Universidade de Lisboa
Faculdade de Farmácia

Genetic variability, disease severity and therapeutic response in a cohort of Angolan Sickle Cell Disease patients

Joana Alexandra Gomes Ferreira

Dissertation supervised by Professor Doctor Rui Miguel Duque de Brito, Principal Coordinating Professor with Habilitation, and co-supervised by Professor Doctor Isabel Maria Antolin Rivera, Assistant Professor with Habilitation.

Master Course in Biopharmaceutical Sciences

2022

Universidade de Lisboa

Faculdade de Farmácia

Genetic variability, disease severity and therapeutic response in a cohort of Angolan Sickle Cell Disease patients

Joana Alexandra Gomes Ferreira

Dissertation supervised by Professor Doctor Rui Miguel Duque de Brito, Principal Coordinating Professor with Habilitation, and co-supervised by Professor Doctor Isabel Maria Antolin Rivera, Assistant Professor with Habilitation.

Master Course in Biopharmaceutical Sciences

2022

The studies presented in this thesis were performed in the Human Genetics Laboratory of Health and Technology Research Center (H&TRC), at Escola Superior de Tecnologia e Saúde de Lisboa, Instituto Politécnico de Lisboa, under the supervision of Professor Miguel Brito, PhD with Habilitation, and the co-supervision of Professor Isabel Rivera, PhD with Habilitation.

ACKNOWLEDGEMENTS

Beforehand, this project would not have been possible without proper authorizations from the Ethics Committee of Centro de Investigação em Saúde de Angola (CISA) and the Ethics Committee of Escola Superior de Tecnologia da Saúde de Lisboa (ESTeSL).

I would especially like to thank Professor Miguel Brito, who transformed what started as an idea into genuine work and made himself and his laboratory available for the development of this work, and most importantly, who believed in me to carry out the project. Thank you for your constant availability, for accepting my ideas and for supporting my work, without ceasing to guide me.

To Dr Adis Cogle, medical doctor at Clínica Girassol, who was responsible for collecting all samples as well as every piece of information required for our study. Your help was indispensable for the realization of this project.

A particular acknowledgement to Mariana, who taught me barely everything I know about laboratory techniques and methods. Thank you for warmly welcoming and integrating me into the lab team. Thank you so much for your support, help and patience throughout this challenging year. Also, I thank Catarina for sharing her knowledge with me and collaborating in building databases with patients' medical records. It was a pleasure to work with both of you.

I thank Professor Isabel Rivera for accepting to co-supervise my work. I appreciate your prompted availability to co-supervise my work.

To Daniela and Raquel, who have been there since the first day of this Master. Thank you for your unconditional support, for always making me laugh and above all, thank you for always being there. I am very grateful for having you as my friends.

Teresa and Vanessa, I can never thank you enough for what you did for me when I most needed someone to just be there and listen. Thank you for brightening my days with your amazingly extraordinary energy.

Sofia and Raquel, who despite being far, were always available when I needed advice or even when I just needed a friend to talk to. Thank you for being my best friends of all time.

Matilde, you know how much you mean to me. Thank you for always being by my side.

Lastly, I must thank my family, particularly, my parents. I would never get so far without your encouragement, support, and motivation. Thank you for pushing me further, especially in the hardest moments.

Genuinely, thank you all so much.

RESUMO

A Drepanocitose ou Anemia das Células Falciformes (ACF) é uma doença hereditária com transmissão recessiva, causada pela mutação c.20A>T no gene *HBB*:. Origina uma variante estrutural da hemoglobina adulta normal denominada de hemoglobina falciforme (HbS). Sob condições de hipoxia, a solubilidade da HbS reduz para um quinto, comparativamente com a hemoglobina adulta. Esta situação causa a polimerização da molécula de HbS com consequente falciformação dos eritrócitos, tornando-os menos flexíveis e mais rígidos, e propensos a eventos de vaso-oclusão e hemólise. As principais manifestações clínicas da ACF são altamente heterogéneas e advém dos eventos de vaso-oclusão e hemólise, causando anemia, isquemia e múltiplas crises de dor que requerem hospitalizações frequentes.

Os indivíduos homozigóticos para a HbS têm o fenótipo mais severo e têm uma prevalência de aproximadamente 70% entre todos os indivíduos com ACF. Estes pacientes desenvolvem uma anemia grave devido à ausência da produção de hemoglobina adulta por falta de síntese de cadeias beta-globina normais. A ACF afeta mais de 300 000 recém-nascidos mundialmente por ano, particularmente na África Subsariana, onde a incidência é de 75%. Esta doença é bastante negligenciada, contabilizando uma taxa de mortalidade entre 50-90% em crianças não diagnosticadas com menos de cinco anos de idade.

Apesar de ser uma doença monogénica, a ACF tem uma heterogeneidade clínica notavelmente elevada na sua expressão fenotípica. Vários fatores demonstraram modular as manifestações clínicas desta doença, nomeadamente moduladores genéticos como a alfa-talassémia e os haplótipos da beta-globina, que modificam parâmetros biológicos como o grau de anemia hemolítica ou os níveis de hemoglobina fetal (HbF). Os genes *ZBTB7A* e *ZNF410* foram propostos como possíveis influenciadores da fisiopatologia da ACF.

O *ZBTB7A* é um regulador importante da transição da expressão de HbF para hemoglobina adulta, atuando como um repressor de transcrição aquando da sua ligação ao promotor do gene da gama-globina, responsável pela expressão de HbF. A potência de repressão do *ZBTB7A* pode originar a perda de 50% da HbF.

Apesar da quantidade reduzida de estudos, sabe-se que o *ZNF410* atua somente num alvo nas células eritroides, que é responsável por silenciar a gama-globina e, consequentemente, inibir a expressão de HbF. Dado que o *ZNF410* e o *ZBTB7A* atuam no complexo co-repressor que desempenha funções cruciais na troca de hemoglobinas, acredita-se que estes genes representam uma influência considerável na síntese de HbF e, subsequentemente, nos fenótipos da drepanocitose. Posto isto, destaca-se a importância de explorar e estudar novos polimorfismos no *ZBTB7A* e *ZNF410*, que possam estar positivamente correlacionados com os níveis de HbF na drepanocitose, severidade da doença e diferenciação de manifestações clínicas.

O tratamento primário da ACF são as transfusões de sangue e estima-se que a maioria dos pacientes com SCA tenham recebido pelo menos uma transfusão. Atualmente, existem quatro medicamentos aprovados para o tratamento da ACF: Hidroxiureia, L-Glutamina, Crizanlizumab-tmca e Voxelotor.

Tecnologias de edição genética como o CRISPR/Cas9 podem alterar permanentemente os genes causadores de doenças, através da correção, exclusão, adição e interrupção de regiões

específicas. Assim, este novo método de tratamento pode providenciar a cura para a maioria dos doentes drepanocíticos. No entanto, questões de segurança necessitam de ser resolvidas para garantir o seu uso clínico seguro.

O objetivo deste estudo é determinar e avaliar a frequência e influência de determinados polimorfismos na severidade da doença e na resposta ao tratamento da ACF, numa população de doentes pediátricos e adultos Africanos, acompanhados na consulta da drepanocitose na Clínica Girassol, em Luanda, Angola. Adicionalmente, pretendemos definir a frequência alélica e genotípica das principais variantes genéticas relacionadas com a drepanocitose, nomeadamente o alelo S, haplótipos da beta-globina e variantes do gene *BCL11A*. Sendo um marcador da ACF, serão analisadas as frequências alélica e genotípica da alfa-talassémia de 3,7kb. Com os dados obtidos, tencionamos realizar estudos de associação entre os marcadores genéticos da drepanocitose e a severidade clínica da doença.

Neste estudo pretendemos identificar moduladores genéticos da ACF, através do estudo de 120 pacientes angolanos drepanocíticos acompanhados na consulta de drepanocitose na Clínica Girassol em Luanda, Angola. Realizaram-se estudos de associação entre fenótipos clínicos e parâmetros bioquímicos e hematológicos, bem como 12 variantes genéticas em 7 genes relacionados com a gravidade da doença (*HBB*, *HBBP1*, *HBA*, *HBE*, *HBG1*, *ZBTB7A* e *ZNF410*). Moduladores genéticos conhecidos da ACF (alfa-talassémia e haplótipos da beta-globina) e genes putativos modificadores de parâmetros hematológicos foram caracterizados, e as diferenças na sua distribuição foram avaliadas. Adicionalmente, tencionamos relacionar a resposta à terapia farmacológica com Hidroxiureia, com os marcadores genéticos estudados, no decorrer de uma análise retrospetiva.

A presença da deleção alfa-talassémia de 3,7kb demonstrou-se protetora em fatores de severidade da doença como a idade de diagnóstico, idade de manifestação dos primeiros sintomas, número de transfusões de sangue recebidas, grau de anemia e ocorrência de crises dolorosas. Neste estudo, observámos um aumento na prevalência do alelo com a deleção relacionada com o envelhecimento, o que se confirmou por uma frequência de 50% de pacientes heterozigóticos e 16,7% de homozigóticos para a deleção com idade superior a 20 anos.

No nosso projeto, verificou-se uma redução da percentagem de indivíduos a receber tratamento com Hidroxiureia diretamente proporcional ao aumento da idade. Além disso, observou-se que a maior percentagem corresponde a pacientes sem deleção. Deste modo, é possível aferir que os pacientes mais velhos e, por consequente, com mais alelos da deleção, não requerem tanto tratamento farmacológico como os indivíduos sem deleção, salientando a componente protetora da alfa-talassémia de 3,7kb em doentes drepanocíticos.

O haplótipo CAR/CAR, predominante em indivíduos angolanos, foi o mais prevalente na nossa população. É caracterizado por um fenótipo mais severo, corroborado no nosso estudo por uma diminuição no número de indivíduos com este haplótipo com o aumento da idade, sugerindo uma redução na sobrevivência. Confirmou-se a severidade nos indivíduos CAR/CAR na idade de diagnóstico, necessidade de tratamento com Hidroxiureia, dactilite, crises de dor e AVCs. Contrariamente ao esperado, pacientes CAR/SEN apresentaram maior severidade na idade de manifestação de primeiros sintomas, necessidade de transfusões de sangue e demonstração de sintomas da ACF.

Relativamente aos polimorfismos, a frequência de mutações no *ZNF410* foi demasiado baixa para tirar elações em análise estatística. Nos polimorfismos do *ZBTB7A*, verificou-se um aumento na idade a que surgiram os primeiros sintomas em indivíduos com a mutação, bem como redução nas crises de dor e anemia e vantagens hematológicas, designadamente o aumento da concentração de HbF e de hemoglobina A2.

Este estudo fornece uma contribuição relevante para o conhecimento genético da população angolana, na qual o haplótipo CAR é indiscutivelmente o mais comum, e a co-herança da alfa-talassémia de 3,7kb tem um grande impacto na severidade clínica da ACF. Neste projeto foram estudados, pela primeira vez, polimorfismos encontrados em genes abordados recentemente e foram realizados estudos de associação entre esses polimorfismos e os parâmetros que caracterizam o fenótipo dos pacientes drepanocíticos. Observaram-se diferenças significativas em vários parâmetros clínicos e em alguns dados hematológicos, em todos os polimorfismos estudados.

Ainda há um longo caminho a percorrer antes de entendermos completamente uma doença tão complexa como a drepanocitose e o motivo pelo qual ela se manifesta de diferentes formas nos pacientes. Esta heterogeneidade é, sem dúvida, influenciada pela herança genética dos pacientes, mas esse não é o único fator. Deste modo, são necessários mais estudos para corroborar os resultados obtidos e confirmar as nossas hipóteses acerca do impacto dos polimorfismos na severidade clínica da ACF.

Palavras-chave: Anemia das Células Falciformes; deleção alfa-talassémia de 3,7kb; *ZBTB7A*; *ZNF410*; hemoglobina fetal.

ABSTRACT

Sickle cell disease (SCD) is an inherited disease with recessive transmission caused by the mutation c.20A>T in the *HBB* gene. It results in a structural variant of normal adult haemoglobin (HbS). The main clinical manifestations of SCD include severe anaemia and multiple pain crises that require regular hospitalizations.

Sickle cell anaemia (SCA) designates homozygosity for the HbS mutation and represents 70% of all SCD cases. SCA affects more than 300,000 newborns worldwide per year. This is a widely neglected disease, accounting for a mortality rate between 50-90% in undiagnosed children under five years old.

Regardless of being a monogenic disease, SCA has a remarkably high clinical heterogeneity in its phenotypic expression. Several factors modulate the clinical manifestations of SCA, namely genetic markers such as alpha-thalassaemia and beta-globin haplotypes. Moreover, *ZBTB7A* and *ZNF410* have been recently proposed as possible influencers of SCA pathophysiology.

This project aims to identify genetic modifiers of the clinical course of SCA by studying 120 Angolan SCA patients followed at Clínica Girassol, in Luanda, Angola. Association studies were performed between the clinical outcomes, and haematological and biochemical parameters of patients, as well as with 12 genetic variants related to disease severity. Known genetic modulators of SCA and putative genetic modifiers of haematological parameters were characterized, and the differences in their distribution were assessed.

The presence of 3.7kb α-thalassaemia deletion was found to be protective in disease severity factors, including the degree of anaemia and occurrence of painful crises. The CAR/CAR haplotype was the most prevalent in our population. It is characterized by a worse phenotype with the most severe disease outcomes. However, CAR/SEN patients had a worse prognosis in some clinical and haematological parameters. Individuals presenting variants in *ZBTB7A* SNPs had a milder phenotype. Notwithstanding, they were strongly associated with higher HbF levels, less symptoms and fewer pain events.

Keywords: Sickle cell anaemia; 3.7kb alpha-thalassaemia deletion; *ZBTB7A*; *ZNF410*; foetal haemoglobin.

TABLE OF CONTENTS

FIGURES

CHAPTER I – INTRODUCTION

CHAPTER IV – RESULTS AND DISCUSSION

TABLES

ABBREVIATIONS

ACS – Acute chest syndrome

bp – base pairs

CISA – Centro de Investigação em Saúde de Angola

CRISPR – Clustered regularly interspaced short palindromic repeats

EMA – European Medicines Agency

ESTeSL – Escola Superior de Tecnologia da Saúde de Lisboa

GVHD – Graft-versus-host disease

Hb – Haemoglobin

HbA – Adult haemoglobin

HbF – Foetal haemoglobin

HbH – Haemoglobin H disease

HbS – Sickle haemoglobin

HPFH – Hereditary persistence of foetal haemoglobin

HPLC – High-performance liquid chromatography

HSCT – Haematopoietic stem cell transplantation

HSPCs – Haematopoietic stem and progenitor cells

HU – Hydroxyurea

HWE – Hardy-Weinberg equilibrium

INDELS – Insertions and deletions

LCR – Locus control region

LDH – Lactate dehydrogenase

LRF – Leukaemia/ lymphoma-related factor

MCV – Mean corpuscular volume

MCH – Mean corpuscular haemoglobin

MRI – Magnetic resonance imaging

NGS – Next Generation Screening

NO – Nitric oxide

NOS – Nitric oxide synthase

NuRD – Nucleosome remodelling and deacetylase

O_2 – Oxygen

PCR – Polymerase chain reaction

RBC – Red blood cell

RFLP – Restriction fragment length polymorphisms

SCA – Sickle Cell Anaemia

SCD – Sickle Cell Disease

SCI – Silent cerebral infarct

SNP – Single nucleotide polymorphism

TCD – Transcranial Doppler (ultrasonography)

TF – Transcription factor

US FDA – United States Food and Drug Administration

UV – Ultra-violet (light)

VOC – Caso-occlusive crisis

WBC – White blood cells

WT – wild-type

CHAPTER I

INTRODUCTION

I. INTRODUCTION

I.1. Human haemoglobin

Haemoglobin (Hb) is a polyfunctional protein of erythrocytes whose primary function is the oxygen (O_2) transport from the lung to tissues.(Ahmed et al., 2020) It is a tetramer constituted by four structurally related and roughly equivalent subunits, each containing a prosthetic group (heme) and a polypeptide chain (globin) consisting of two α-subunits (α1 and α2) and two β-subunits (β1 and β2).(Ahmed et al., 2020) The two αβ dimers (α1β1 and α2β2) are organized along a twofold axis symmetry, resulting in a large central water cavity in the T or unliganded or deoxygenated structure, whereas the R or liganded or oxygenated shape has a narrower cavity.(Ahmed et al., 2020) The α- and β-subunits are composed of seven and eight helices, A to H, respectively, that are connected by non-helical segments, and each subunit includes a binding pocket for the haem molecule that is made up of the E and F helices. (Ahmed et al., 2020) The haem is composed of a ferrous ion held in the centre of porphyrin and coordinated by the porphyrin ring's four nitrogen atoms, which allow the binding to molecular O_2 and subsequently its transport by Hb.(Ahmed et al., 2020)

In humans, there are two clusters of globin chain-coding genes, the α-like chain genes (ξ, α2 and α1) on chromosome 16 and β-like chain genes (ε, $^{G}\gamma$, $^{A}\gamma$, δ and β) on chromosome 11, given that these genes are expressed sequentially in different stages of development (Figure 1).(Guvenc et al., 2012; Razin et al., 2021) During erythropoiesis, a sequence of Hb switches predominantly mediated by changes in the expression of the β-like globin genes results in the sequential expression of embryonic (*HBE1*), foetal (*HBG1*, *HBG2*), and adult (*HBD*, *HBB*) globin genes.(Nandakumar et al., 2016) Critical regulatory components located upstream of the genes are necessary for the coordinated expression of the genes in each cluster at all developmental stages.(HIGGS et al., 2005) The β-globin gene cluster is regulated by the super-enhancer locus control region (LCR), meanwhile the major regulatory element (HS-40) is responsible for the α-globin gene cluster (Figure I.1).(Razin et al., 2021) When erythroid differentiation reaches its terminal stage in healthy people, the synthesis of the α- and β-globin chains is carefully balanced, to ensure the production of well-proportioned red blood cells (RBCs) designated by mean corpuscular volume (MCV), and a good Hb concentration (mean corpuscular haemoglobin, MCH).(Higgs et al., 2005)

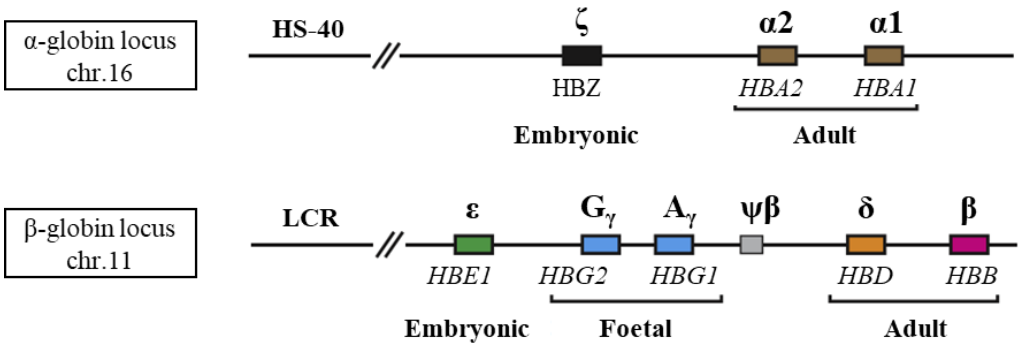


Figure I.1. Schematic representation of the human globin loci. The locations of the α-globin gene on chromosome 16 and the β-globin gene on chromosome 11 are displayed. The α-globin gene includes the ζ-, α2-, and α1-globin genes, mostly regulated by a single upstream strong regulatory region HS-40. The β-globin gene expresses ε-, G_γ-, A_γ-, δ- and β-globin genes under the control of LCR. *Adapted from Wienert B, et al., 2018*.

According to the stage of human development, Hb exhibits heterogeneity throughout the different human development stages.(Khandros & Blobel, 2021) There are several regular human haemoglobins: Hb Gower 1 (ζ2ε2), Hb Gower 2 (α2ε2) and Hb Portland 1 (ζ2γ2) at the embryonic phase; foetal Hb (HbF, α2γ2) is the primary Hb during foetal development and constitutes over 70% of all Hb at birth, before beginning to diminish until approximately 1% during adulthood; and the two haemoglobins of adults are HbA (α2β2) and HbA2 (α2δ2) with HbA making up roughly 97% of the total amount of Hb throughout adult life.(Huehns & Shooter, 1965; Manning et al., 2020) Uncertainty surrounds the start of δ-chain formation; however, residues of these globin chains may be found in cord blood, and by the end of the first year of life, their adult levels are reached.(Weatherall & Clegg, 1976)

There are two switches in the expression of human Hb during development, from embryonic to foetal and from foetal to adult, following the 5' to 3' arrangement of the globin genes in their clusters.(Liu et al., 2021) Synthesis of the primary Hb, schematized in Figure I.2, occurs in the embryonic yolk sac, and after five weeks of gestation, the first Hb switch (primitive to definitive) takes place in the foetal liver, changing embryonic Hb (ζ2ε2) to HbF (α2γ2).(Hariharan & Nadkarni, 2021) The second wave begins throughout pregnancy and is finished postpartum when the major form changes from HbF in the foetal liver to HbA in the bone marrow.(Liu et al., 2021) This complex process is thought to be triggered by a set of models working together, including (i) Gene competition: the γ- and β-globin genes compete with the LCR for interaction in a stage-specific way for their expression, during ontogeny; (ii) Chromosomal looping: the hypersensitive sites in the LCR of the globin cluster loop out to approach the relevant gene inside the cluster and to get near it, for boosting its transcription;

(iii) Autonomous gene silencing: all the components for gene silencing are found inside or close to the target gene that has to be suppressed, making it a self-regulating process; and lastly (iv) Polymerase tracking model: before encountering a promoter region to initiate gene expression, the enhancer-bound protein complex, including RNA polymerase II, scrutinizes the DNA sequence.(Hariharan & Nadkarni, 2021)

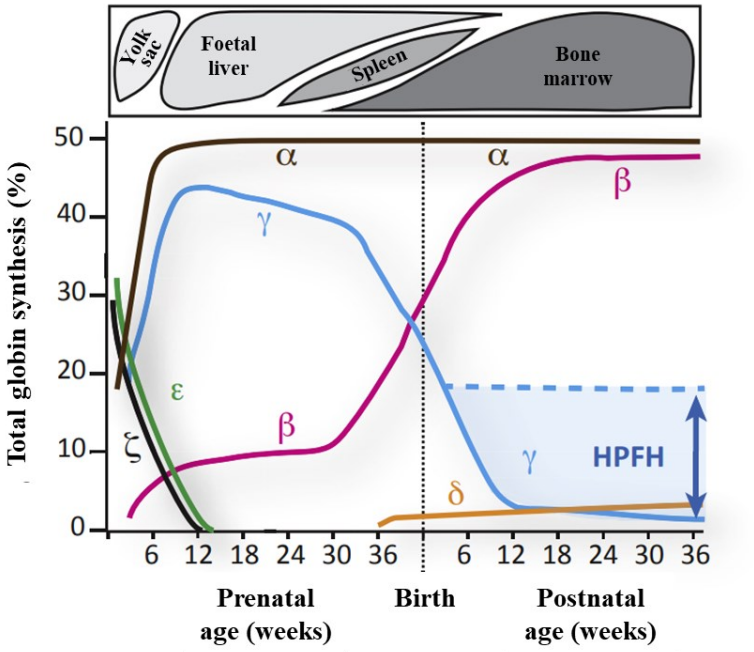


Figure I.2. Developmental regulation of the human β-globin locus. The top represents the location of globin production, and the bottom is the developmental expression of human globin chains. Haemoglobin is initially synthesized in primitive erythroblasts in the embryonic yolk sac. Following the first trimester, the locus of globin expression moves to the foetal liver and then to the bone marrow, as primitive erythropoiesis turns into definitive erythropoiesis. *Adapted from Wienert B, et al., 2018*.

Due to its potential therapeutic applications, understanding the molecular underpinnings of the foetal-to-adult Hb transition has long piqued the interest of researchers in the field.(Nandakumar et al., 2016) Several clinical studies have evidence that increased HbF synthesis lessens the severity of β-haemoglobinopathies.(Liu et al., 2021; Molokie et al., 2021; Nandakumar et al., 2016)

I.1.1. Foetal Haemoglobin

HbF is the predominant type of Hb found in the foetus during pregnancy, produced from 10 to 12 weeks of pregnancy until the first six months of postpartum life, by erythroid precursor cells.(Kaufman et al., 2022) The gamma (γ) subunits that HbF contains induce conformational changes to the Hb protein that provide several physiological differences in O_2 delivery

important in foetal circulation.(Kaufman et al., 2022) The transfer of O_2 from the mother's circulation to the foetal circulation depends critically on HbF, which has a strong O_2 affinity.(Kaufman et al., 2022) HbF plays an extremely important role in haemoglobinopathies, especially in Sickle Cell Disease (SCD).(Kaufman et al., 2022)

The finding that affected infants do not exhibit symptoms until the age of six months is the first indication that HbF performs a protective role in SCD, and the importance of HbF in reducing the sickling process has been established by biochemical and clinical research.(Adekile, 2021) The heterotetramer formed in the presence of γ-globin chains ($\alpha 2\gamma\beta^S$) is considerably more soluble than the homotetramer (α2βS2), allowing an increase in the minimum gelling concentration of mixtures of HbS and HbF.(Adekile, 2021) The delay time to gelation is also increased owing to altered polymerization kinetics.(Adekile, 2021)

In 1961, Jackson et al. performed a clinical study of the influence of HbF on SCD phenotype and concluded that SS patients with HbF levels above 12% had substantially lower pathology indices than individuals with less than 10% of HbF.(Adekile, 2021)

HbF is constrained to a sub-set of RBCs in normal adults so-called F-cells, derived from erythroid cells that do not switch completely from γ- to β-globin chain synthesis during differentiation, and that range from 0.5 to 7% in healthy persons.(Urio et al., 2020) Contrariwise, in SCD patients, the proportion of F-cells may reach up to 80%.(Urio et al., 2020) This percentage depends on the increase in the ratio of erythroid cells that produce γ-globin chains from stress erythropoiesis, and due to the preferential survival of F-cells in the peripheral blood given their resistance to sickling.(Urio et al., 2020) Hence, there is a strong correlation between the F-cell proportion and the percentage of HbF.(Urio et al., 2020)

Given its impact on managing the clinical symptoms and treating SCD, researchers affirm that understanding the mechanisms behind HbF switching and silencing will bring in more targetable and druggable candidates.(Demirci et al., 2021)

I.2. Haemoglobinopathies

Haemoglobinopathies are the most prevalent monogenic inherited diseases and a significant global health issue, that due to immigration, has spread all over the world.(Kohne, 2011) These are caused by mutations, substitutions, and/or deletions in α- or β-globin genes and they can be divided into two groups according to the consequence of the mutation, namely structural variants, and thalassaemia syndromes.(Kohne, 2011) When there are alterations in Hb structure, atypical Hb is produced without affecting its rate of synthesis, whereas in

thalassaemia there is a disorder in the synthesis or stability of the globin protein, altering the proportion of available globin chains.(Kohne, 2011) Hereditary persistence of foetal haemoglobin (HPFH), another diverse class of associated Hb diseases, is brought on by abnormalities in globin switching that hinder the transition from γ- to β-globin, which results in high expression of HbF in adulthood.(Gusmão, 2015; Lin et al., 2013)

I.2.1. Sickle Cell Disease

SCD is a heterogeneous group of haemoglobinopathies, caused by a single nucleotide polymorphism (rs334) at position 17 of the *HBB* gene, which changes thymine to adenine (GAG→GTG), resulting in the substitution of glutamic acid residue (Glu6Val) for valine at the sixth amino acid residue of the β-globin chain.(Kohne, 2011; Williams, 2016) This results in a structural variant of normal adult Hb (HbA; α2β2), where the normal β-globin subunit is replaced by the mutant form (βS), named sickle Hb (HbS).(Williams, 2016) Even though HbS affinity to O_2 is normal, under hypoxia it has about one-fifth of the solubility compared to HbA.(Abdu et al., 2008) This leads to a reversible polymerization of the HbS molecule and the consequent formation of linear elongated fibres that increase cellular rigidity and distort the erythrocyte membrane, resulting in RBC sickling. These sickle erythrocytes are less flexible and stickier and may be secluded into the microcirculation causing vaso-occlusion and local ischaemia.(Sundd et al., 2019) The membrane damaging and RBC dehydration characteristics from HbS polymerization accelerate haemolysis and often cause anaemia. These mechanisms along with the molecular pathophysiology of SCD are represented in Figure I.3.

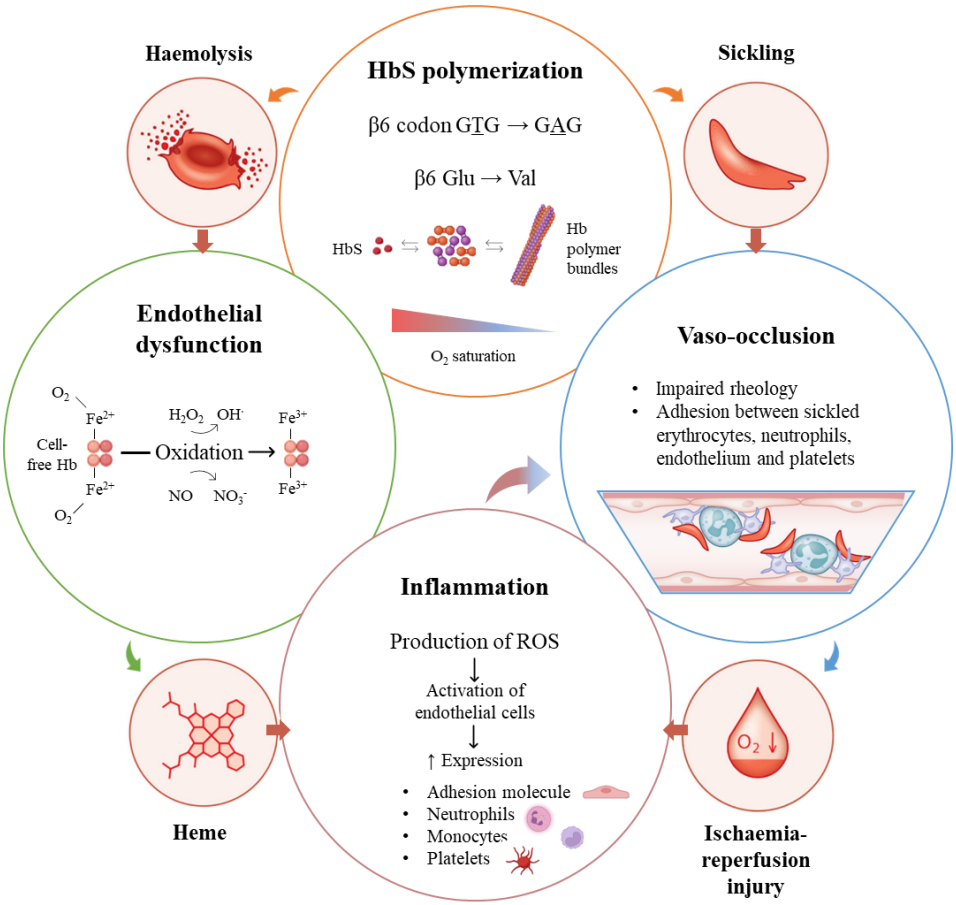


Figure I.3. Molecular pathophysiology of sickle cell disease. In the sixth position of the β-globin chain, a single nucleotide polymorphism causes valine to replace glutamic acid. The mutated haemoglobin (HbS) molecules polymerize upon deoxygenation and form bundles, resulting in impaired rheology of the blood and aggregation of sickle erythrocytes with neutrophils, platelets, and endothelial cells, described as vaso-occlusion. Vaso-occlusion promotes ischaemia-reperfusion (I-R) injury. Hb polymer bundles promote haemolysis or lysis of erythrocytes, releasing cell-free Hb into the blood circulation. Inflammation is caused by reactive O_2 species (ROS) generation, along with the activation of endothelial cells that boost the expression of adhesion molecules, neutrophils, monocytes, and platelets. Following this advent, inflammation also promotes vaso-occlusion through a feedback loop. *Adapted from Sundd P, et al., 2019*.

Epidemiological research proved that several factors are interfering directly with HbS polymerization and, subsequently, with acute systemic painful vaso-occlusive crisis (VOC) which are caused by vaso-occlusion or blood vessel occlusion events resulting in ischaemia.(Sundd et al., 2019) Ischemia is a condition in which there is a blockage of blood vessels, reducing or stopping the blood flow.(Cunha & Davis, 2021) This condition can occur anywhere in the body, including the brain and the heart, in which poor O2 intake results in damage or dysfunction of the affected tissue.(Cunha & Davis, 2021) These are frequently prompted by inflammatory or environmental triggers, such as infection, hypoxia, dehydration, acidosis, and temperature.(Sundd et al., 2019)

I.2.1.1 ***Disease transmission and presentation***

SCD is the umbrella term for a variety of pathologies when HbS predominates, which is present in homozygous individuals (HbS: αA2βS2) or it may be caused by compound heterozygosity for HbS associated with a wide variety of other *HBB* mutations in the second β allele ($HbS\beta^0$-thalassaemia, $HbS\beta^+$-thalassaemia, HbSC).(Williams, 2016) Among all genotypes, the most prevalent and severe is homozygous for HbS mutation, titled sickle cell anaemia (SCA), accounting for 70% of SCD cases.(Onimoe & Rotz, 2020; Sundd et al., 2019)

The HbS gene frequency is strongly correlated with the historical distribution and incidence of malaria due to the carrier-resistance to *Plasmodium falciparum* malaria.(Pule et al., 2017) It is postulated that this geographical coexistence led to local amplification and positive selection of SCD in malaria-endemic regions.(Pule et al., 2017) The *HBB*-β^S distribution outside of Africa largely reflects ancient regional migrations from Africa to the Mediterranean, Middle East, and Indian subcontinent regions, as well as more recent trade routes like the trans-Sahara, Islamic and European, and trans-Atlantic slave trade.(Esoh & Wonkam, 2021) Continuous immigration in the contemporary era is expected to increase the genetic diversity of Africa by enabling the translocation of gene pools, such as mutations associated with malaria, to geographic areas throughout the continent.(Pule et al., 2017)

SCD is a recessive disease, and its severity is determined by sickle cell genotype, being that heterozygosity for HbS provides a milder clinical course although these individuals can exhibit sickle cell trait under certain conditions.(Onimoe & Rotz, 2020; Sundd et al., 2019)

Intermittent vaso-occlusion and persistent haemolytic anaemia are the two hallmarks that characterize the pathophysiology of SCD.(Bender, 1993) These patients usually have lower levels of Hb and RBC, as well as higher counts of lactate dehydrogenase (LDH), white blood cells (WBC) and platelets, possibly related to inflammatory conditions.(Antwi-Boasiako et al., 2018) In cases of hypoxia, sickle erythrocytes can be visualized on a routine peripherical blood smear.(Schnog et al., 2004) The Hb solubility test is based on the formation of a precipitate with O_2 depletion and can be used as a quick screening test for HbS, however, it does not distinguish between the various genotypes.(Schnog et al., 2004)

Among the techniques used to detect the existence of anomalous Hb forms, there are Hb electrophoresis, high-performance liquid chromatography (HPLC), isoelectric focusing, and targeted mutation analysis of the globin genes sequence.(Schnog et al., 2004) It is necessary

to conduct familial or DNA-based research, such as Polymerase chain reaction (PCR) analysis of known mutations using well-designed primers, to distinguish between genotypes.(Schnog et al., 2004)

I.2.1.2 ***Genetic origin and prevalence of sickle cell Hb gene***

The origin of the *HBB*-β^S variant remains uncertain, but it has been described that they are less likely to result from recurrent mutations even because Hb genes mutation rates are too sluggish to account for the β^S variant's several independent occurrences.(Esoh & Wonkam, 2021) There are several mutations in and around *HBB*, represented in Figure I.4, that produce various *HBB* haplotypes and contribute to population variability, reflecting the genetic and evolutionary relationship between this condition in various ethnic environments.(Alsamiri et al., 2021)

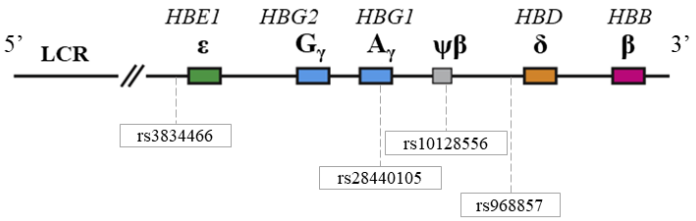


Figure I.4. Representation of SNPs used for classifying HbS most common haplotypes in the β-globin gene cluster. *Adapted from Delgadinho M, et al., 2021.*

Analysis of the β-globin gene cluster haplotypes is determined through restriction fragment length polymorphism (RFLP) in the *HBB* gene cluster, based on detecting the accession of restriction endonucleases, as shown in Table I.1.(Shaikho et al., 2017; Shriner & Rotimi, 2018) There are five common *HBB* haplotypes resulting from the *HBB*-β^S variation, each of which is called after the geographic regions from whence it was initially discovered and was more prevalent: Central African Republic (CAR) or Bantu, Cameroon (CAM), Benin (BEN), Senegal (SEN) and the Arab-Indian (AI).(Esoh & Wonkam, 2021) The five major *HBB* haplotypes are related to different phenotype patterns since they are directly associated with HbF levels, which modulates the clinical course of SCA.(Shaikho et al., 2017) Evidence suggests that patients with AI or SEN haplotypes have the greatest HbF levels and fewer clinical symptoms, whereas those with BEN and CAM haplotypes have intermediate disease severity and individuals with CAR haplotypes have the lowest HbF levels and the worst clinical outcome.(Delgadinho et al., 2021)

Table I.1. Restriction Enzymes and SNP profiles associated with β^S-globin haplotypes. Nucleotides (A = adenosine; T = timine; C = cytosine; G = guanine) after enzymatic recognition, in the five major sickle cell haplotypes.(Shaikho et al., 2017)

SNP	**rs3834466**	**rs28440105**	**rs10128556**	**rs968857**
Gene	ε (*HBE1*)	Aγ (*HBG1*)	ψβ (*HBBP1*)	δ (*HBD*)
Restriction enzyme	*Hinc*II	*Hind*III	*Hinc*II	*Hinc*II
CAR	C	G	C	C
CAM	A	G	T	C
BEN	C	G	T	C
SEN	C	G	T	T
AI	C	GT	T	T

The diversification of the haplotypes around β^S mutation results from the spread of the sickle mutation to various geographic areas because of the protective impact it has against malaria.(Delgadinho et al., 2021)

I.2.1.3 Clinical manifestations

The symptoms of SCD often appear about six months after birth, when the HbF levels start to decline because HbF is responsible for maintaining sickle Hb in solubilized form.(Mangla et al., 2022) Reduced solubility and higher polymerization in HbS lead to haemolysis and vaso-occlusion, two key pathological processes that initiate the clinical manifestations of SCD (Figure I.5).(Booth et al., 2010; Brandow & Liem, 2022) Acute bouts of pain, stroke, priapism, acute chest syndrome (ACS) and chronic organ damage such as osteonecrosis, renal failure, and chronic haemolytic anaemia are the main clinical symptoms.(Bender, 2003; Kato et al., 2007)

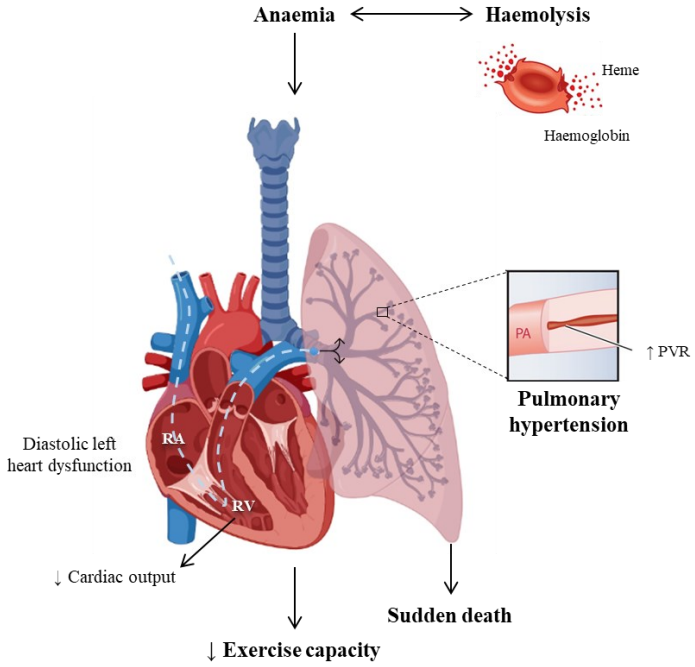


Figure I.5. Endothelial dysfunction in SCD. Anaemia and intravascular haemolysis result in diastolic cardiac failure and pulmonary vascular hypertension, both of which increase the risk of morbidity (due to lower exercise tolerance) and mortality. *Adapted from Sundd P, et al., 2019.*

A heterocellular aggregation made up of sickle cells, non-sickled RBCs, leukocytes, and platelets adhere to the vascular endothelium and clogs the lumen of small blood vessels.(Booth et al., 2010) This microcirculatory obstruction causes multisystemic consequences, especially in the bone, lungs, brain, kidneys, and spleen, as well as acute and chronic tissue ischaemia and infarction.(Booth et al., 2010; Brandow & Liem, 2022) It is the basis of many of the chronic consequences found in SCD, including acute painful episodes, crises, and long-term issues.(Booth et al., 2010)

Sickle cell crisis is a term that refers to a few acute diseases, including the VOC (acute painful crisis), aplastic crisis, splenic sequestration crisis, hyperhaemolytic crisis, hepatic crisis, dactylitis, and ACS.(Borhade & Kondamudi, 2021) Patients often experience moderate to severe pain, with fluctuating frequency and intensity. Young children may have excruciating hand and foot pain and oedema, designated as dactylitis.(Borhade & Kondamudi, 2021)

Although pain in SCD patients is probably caused by VOC, it is advisable to conduct a complete screening for other life-threatening conditions that can be mistaken for sickle cell pain.(Borhade & Kondamudi, 2021) VOC can lead to ACS, an SCD complication with a 25% rate of mortality.(Borhade & Kondamudi, 2021) Commonly, chest hypoventilation brought on by a VOC crisis results in hypoxia, which is the cause of ACS. A vicious cycle is created

when the hypoxia induces sickled erythrocytes to adhere to the pulmonary microvasculature, generating localized hypoxia in the lungs and more RBC sickling.(Borhade & Kondamudi, 2021)

Following the ischaemia-reperfusion event, reactive oxygen species (ROS) are produced, and endothelial cells are activated with enhanced adhesion molecule expression, neutrophils, monocytes, and platelets, and this results in a chronic inflammatory state in SCD.(Nader et al., 2020; Pace et al., 2021) In addition, nitric oxide (NO) is depleted by cell-free Hb due to sickle RBC intravascular haemolysis, which furthers endothelial dysfunction.(Pace et al., 2021) According to growing data, SCD is characterized by ongoing inflammation and oxidative stress, which contribute to the occurrence of chronic vasculopathy, endothelial dysfunction, and several other chronic problems.(Nader et al., 2020)

Before the end of childhood, patients with SCD experience spleen infarction. In this condition, the spleen is impacted due to its constrictive blood arteries and crucial function in the lymphoreticular system causing an abrupt, painful enlargement of the spleen owing to intrasplenic red cell entrapment.(Borhade & Kondamudi, 2021) Hence, recurrent splenic sequestration or splenic dysfunction regularly submits paediatric SCD patients to splenectomy, a surgical procedure to remove the spleen completely or partially.(Al-Balushi et al., 2016)

Cholelithiasis is a medical condition that commonly affects paediatric SCD patients, with prevalence rates ranging from 30 to 70%.(al Talhi et al., 2017) Since its incidence increases with age, it is important to include cholelithiasis and biliary colic in the differential diagnosis for VOC when an individual with SCD experiences an episode of abdominal pain.(al Talhi et al., 2017) Presently, cholecystectomy is the standard treatment for symptomatic patients with gallstones and it consists in removing the gallbladder surgically.(Goodwin et al., 2017)

The reticuloendothelial system, which mediates phagocytosis, destroys sickled RBCs more easily.(Booth et al., 2010) With SCA, this results in chronic anaemia (a steady state Hb of 6 to 8 g/dL), which elevates cardiac workload and output, and results in cardiomegaly and decreased exercise tolerance.(Booth et al., 2010) Haemolytic anaemia dysregulates NO homeostasis, promoting some of the side effects of SCA and other chronic types of anaemia.(Kato et al., 2007) The persistently high rate of haematopoiesis contributes to children's poor development, and people are vulnerable to any condition causing anaemia to worsen, which can lead to circulatory failure.(Booth et al., 2010)

Other complications are involved with SCD physiopathology, as represented in Table I.2, and they negatively impact the quality of life of these patients, as well as their average life expectancy.

Table I.2. Pathophysiology and prevalence of clinical outcomes in SCD. *(Brandow & Liem, 2022)*

Complications in SCD	Pathophysiology	Prevalence
Acute and chronic pain	Tissue ischemia and infarction. Ischemia-reperfusion injury. Haemolysis-induced endothelial dysfunction. Leg ulcers and avascular necrosis. Sensitization of the central or peripherical nervous system.	Acute pain: 70% of acute care visits. Chronic pain: 40% in children; 30% in adults.
Stroke	Cerebral vasculopathy. ↑ sickle RBC adherence. Haemolysis-induced endothelial activation. Altered vasomotor toned.	Silent cerebral infarct: 39% by 18 years old. Overt stroke: 11% by 20 years old.
Nephropathy	Medullary hypoperfusion and ischaemia. Glomerular haemodynamic alterations. Intravascular haemolysis and endothelial dysfunction in the renal cortex. Hypoxia-inducible factor-1α dependent injury.	Hyperfiltration: 40-50% of children. Chronic kidney disease: 20-40% of adults.
Chronic lung disease	Obstructive lung disease: atopy, airway inflammation (↑ leukotrienes). Restrictive lung disease: recurrent acute chest syndrome	Obstructive lung disease: 16% in children, 8% in adults. Restrictive lung disease: 7% in children, 28% in adults.
Pulmonary hypertension	Intravascular haemolysis. NO depletion. Chronic hypoxia. Diastolic dysfunction. Diffuse myocardial fibrosis.	10% in adults by right-heart catheterization.

The most prevalent long-term effects of SCD in children and adults include stroke, silent cerebral infarcts, and cognitive morbidity.(DeBaun et al., 2020) There is a considerable risk of infarct recurrence and clinically severe cognitive deficits following both stroke and silent cerebral infarcts, which indirectly change the quality of life and employment status.(DeBaun et al., 2020)

I.2.1.4 ***Disease management***

SCD is a complicated multisystem disorder defined by acute and chronic consequences.(Kato et al., 2018) The life expectancy of people with SCA has improved significantly due to advancements in general medical care, early diagnosis, and thorough treatment.(Kato et al., 2018)

It is estimated that up to 10 million newborns with SCA might be saved worldwide, most of them in sub-Saharan Africa, by widespread screening and treatment programs.(Kato et al., 2018) The adoption of a universal Newborn Screening Program (NBS) continues to be a significant economic and public health problem; nonetheless, screening for SCD in Africa is essential due to its high incidence.(Kato et al., 2018) African communities and governments should also create culturally appropriate adult screening programs for family planning, to promptly identify symptoms and work to treat them to improve life quality.(Kato et al., 2018)

Given the intricacy of SCD and the variety of potential side effects, healthcare professionals prefer to adopt a multidrug strategy focused on best supportive care, based on penicillin prophylaxis, pneumococcal immunization through vaccination, blood transfusions, and administration of non-steroid anti-inflammatory drugs (NSAIDs) and opioids to alleviate pain crises.(Abboud, 2020; Kato et al., 2018)

The primary treatment for SCD is blood transfusion, and most of the afflicted people are expected to have had at least one transfusion.(Abboud, 2020) In the case of severe anaemia, transfusions can be administered acutely to boost O_2-carrying capacity and improve blood flow, or chronically to avoid long-term complications.(Abboud, 2020) RBCs that contain HbS are diluted during transfusion, which lowers the percentage of abnormal HbS and boosts blood's ability to transport O_2.(Ogu et al., 2021) The main clinical indications for therapeutic transfusions are acute symptomatic anaemia, ACS, stroke, splenic and hepatic sequestration, aplastic crises, sickle hepatopathy, central retinal artery occlusion, and multisystem organ failure.(Abboud, 2020; Ogu et al., 2021) Nevertheless, in primary and secondary stroke prevention, pre-operative management, and pregnancy, blood transfusions are effective

prophylaxis.(Ogu et al., 2021) Despite the advantages, transfusions have drawbacks that prevent them from being used for extended periods, which encompass erythrocyte alloimmunization (an immunological reaction to foreign antigens found in the donor's blood), iron overload and haemolytic transfusion reactions.(Abboud, 2020; Kato et al., 2018)

Presently, there are four medications for SCD authorized by the United States Food and Drug Administration (US FDA), with the majority receiving approval after 2017.(Ogu et al., 2021)

Hydroxycarbamide or Hydroxyurea (HU) was the first drug US FDA and European Medicines Agency (EMA) approved for SCD, in 1998 and 2007, respectively.(Kato et al., 2018) As a strong ribonucleotide diphosphate reductase inhibitor, HU is cytotoxic and myelosuppressive, and its use is linked to reversible decreases in the generation of all blood cell lineages.(Williams, 2016) Its mechanism of action works by inducing HbF synthesis, which reduces sickling inside RBCs by inhibiting intracellular HbS polymerization, and decreasing leukocyte count, thus diminishing inflammatory response.(Ogu et al., 2021; Williams, 2016) Its effects have been demonstrated to reduce the frequency of vaso-occlusive crises, improve the median duration between the first and second crises, and lower the frequency of acute chest syndrome.(Ogu et al., 2021) After two decades of clinical usage, HU has shown to be a safe and effective drug that mostly causes reversible bone marrow suppression, regardless of the need for routine monitoring of peripheral blood counts and chemistries.(Pace et al., 2021) Children with SCA beginning at nine months of age and adults with severe clinical illness are advised to get HU as the standard of therapy.(Pace et al., 2021)

L-Glutamine, commercialized as Endari®, is an essential amino acid required for the synthesis of nicotinamide adenine dinucleotide (NAD), nicotinamide adenine dinucleotide phosphate (NADP), and indirect regulation of the metabolism of glutathione.(Ogu et al., 2021) L-Glutamine increases the rate of reduced NAD (NADH) in sickle RBCs, reducing oxidative stress and thereby, decreasing the number of ACS episodes, pain crises and hospitalizations.(Ogu et al., 2021; Pace et al., 2021) Endari® was approved by FDA for SCA patients aged 5 years old and above.(Ogu et al., 2021; Pace et al., 2021) However, there have been reported concerns regarding patient accessibility and treatment adherence.(Ogu et al., 2021)

Approved in 2019, Crizanlizumab-tmca (Adakveo®) is a humanized immunoglobulin G2 monoclonal antibody that prevents platelets from interacting with erythrocytes and leukocytes through P-selectin on their surface.(Stevens et al., 2021) P-selectin, a protein produced on the

surface of endothelial cells, promotes the aberrant adherence of sickle RBCs to the endothelium, which is a contributing factor to SCD's painful vaso-occlusion.(Ogu et al., 2021) Crizanlizumab dramatically reduced the rate of VOCs and extended the duration between the first and second crises, resulting in its authorization for usage in patients who are at least 16 years old using an intravenous infusion to administer.(Ogu et al., 2021; Stevens et al., 2021)

The most recent disease-modifying drug is an HbS polymerization inhibitor named Voxelotor (Oxbryta®). This small molecule reversibly binds to the α-globin chain and stabilizes the relaxed state (R-state) of Hb via augmenting the O_2-binding capacity of HbS, therefore avoiding RBC sickling and destruction by postponing HbS polymerization.(Ogu et al., 2021; Pace et al., 2021) The improvement of RBC malformation and lengthening of RBC half-life brought along by the reduction of RBC sickling reduces whole blood viscosity, haemolysis, and ensuing anaemia.(Herity et al., 2021) In November 2019, the FDA authorized Voxelotor for the treatment of SCA in individuals 12 years of age and older.(Herity et al., 2021; Ogu et al., 2021)

The existence of new SCA disease-modifying drugs allows the combination of HbF inductors with medicines that target different disease mechanisms.(Pace et al., 2021) A new era of development of personalized combination drug regimens has been ushered in with drug combinations that further reduce clinical severity in an additive and/or synergistic manner, for SCA individuals who are either non-responsive or have contraindications for HU use.(Pace et al., 2021)

With an overall event-free survival rate of 90% to 95% following transplantation from human leukocyte antigen (HLA) matched sibling donors, allogeneic haematopoietic stem cell transplantation (HSCT) is the only treatment that has been recognized as curative.(Doerfler et al., 2021) This procedure consists in entirely swapping out the patient's bone marrow, where RBCs are made, with bone marrow containing RBC-producing stem cells with the proper β-globin gene from an unaffected, tissue-matched sibling donor.(Tisdale et al., 2020) In SCD patients, HSCT can produce donor-derived erythropoiesis and stabilize or even restore the function of the affected organs.(Anurogo et al., 2021) For all children under the age of five and older paediatric patients between the ages of five and eighteen who present with SCD-related problems, matched sibling donor (MSD) HSCT is a curative option.(Anurogo et al., 2021; Tisdale et al., 2020)

Although the outcomes of HSCT using alternate donor sources are encouraging, many patients continue to suffer from immunological issues such as transplant rejection and graft-versus-host disease (GVHD).(Doerfler et al., 2021)

These allogeneic transplantation results have demonstrated the feasibility of genetic modification of the damaged bone marrow stem cells as a potential therapeutic strategy.(Tisdale et al., 2020) As a result, genetic techniques for modifying the patient's stem cells and subsequently performing an autologous transplant on the patient have been actively studied.(Tisdale et al., 2020) The risk for GVHD and transplant rejection can be nearly eradicated by using genetically engineered therapeutic cells that are patient-derived, eliminating the requirement for immunosuppression as a component of the conditioning regimen.(Demirci et al., 2018)

Gene-editing technologies, as opposed to conventional gene therapy methods, have the potential to permanently alter disease-causing genes through the precise correction, deletion, addition, and disruption of specific regions.(Park & Bao, 2021) The therapeutical product for autologous transplantation is a gene-edited haematopoietic stem and progenitor cells (HSPCs) from SCD patients (SCD HSPCs).(Park & Bao, 2021) Recent preclinical studies have indicated the potential of the following gene editing techniques for the treatment of SCD: (i) induction of HbF through gene-disruption of γ-globin repressors, (ii) correction of the HBS point mutation, and (iii) induction of HbF by inducing beneficial HPFH mutation on the β-globin locus.(Park & Bao, 2021)

The clustered regularly interspaced short palindromic repeats (CRISPR)-Cas9 nuclease system emerged as a different option that has shown to be simple to build and capable of delivering highly effective genome editing.(Demirci et al., 2018) Traditional CRISPR-Cas9 editing relies on double-strand breaks (DSBs) and cellular repair of chromosomal breaks using either non-homologous end joining (NHEJ) or homology-directed repair (HDR), being this a precise repair process where demanding mutations can be introduced or corrected, reliant on a donor template availability.(Demirci et al., 2021; Frangoul et al., 2021)

CRISPR-Cas9 gene-editing strategies for SCD treatment are available in a variety of ways, namely introducing beneficial HPFH mutations, targeting HbF transcriptional repressors to produce the necessary concentrations of HbF to reverse sickling, and correcting the sickle mutation in *HBB*.(Park & Bao, 2021)

With the development of CRISPR-Cas9 technology, an autologous transplant of gene-edited HSCs may provide the cure for the majority of SCD patients.(Park & Bao, 2021) However, many outstanding safety concerns need to be resolved to ensure a safe clinical application.(Tisdale et al., 2020)

I.2.1.5 ***Genetic modifiers***

Despite the apparent genetic simplicity of SCA, which is brought on by mutations affecting just one gene, this condition exhibits extraordinary variation in the severity of the disease.(Thein, 2017) Identifying the genetic modifiers through DNA analysis may enable more accurate assessments of the severity of the disease, supporting the principles of precision medicine.(Thein, 2017) Determining the molecular pathways that connect the hereditary components may also point to novel treatment targets.(Thein, 2017)

The discovery of disease severity biomarkers would ideally aid in stratifying patients following their propensity for severe SCD-related outcomes. From clinical genetic research, two significant modifiers have been identified, the co-inheritance of α-thalassaemia and polymorphisms linked to an increased HbF production in adults.(Thein, 2017) These two conditions are also the disease's two primary modulators since they can alter the intracellular concentration of HbS, which in turn affects the rate of polymerization, the primary factor in SCD-related medical issues.(Damanhouri et al., 2015)

Along with genetic factors, social, and economic factors, environmental issues contribute to the clinical course of SCA.(Domingos et al., 2014)

I.2.1.5.1. Beta-globin genotype

The critical stage in the pathophysiology of SCD is deoxygenated HbS polymerization and sickling, which is highly dependent on the intracellular HbS concentration.(Thein, 2013) The different SCD genotypes are associated with distinct levels of intra-erythrocyte HbS and therefore, render diverse phenotypic severities: SCA and SCD-Sβ^0-thalassaemia patients exhibit the most severe forms of the disease since HbS are near all the available Hb, whilst heterozygous states SCD-SC and SCD-Sβ^+-thalassaemia individuals present milder forms of the disease since they have a minor percentage of HbS.(Thein, 2013)

Polymerization of HbS is the main influencer in SCD phenotype, but it has been demonstrated that this process can be delayed by high concentrations of HbF.(Damanhouri et al., 2015) Moreover, HbF can be considered the key regulator of clinical and haematologic aspects of SCD as it is unable to enter the HbS-polymerization phase.(Damanhouri et al., 2015)

The levels of HbF in SCA are derived from the interaction between inherited genetic variables and pathological illness mechanisms.(Menzel & Thein, 2019)

I.2.1.5.2. β^S-globin haplotypes

As was previously mentioned, the sickle cell mutation is linked to particular β^S-globin haplotypes that are representative of the region where the primitive mutation originated (see page 10).(Thein, 2017) These haplotypes have been linked to distinct haematological and clinical variations, notwithstanding the significant variability of clinical manifestations within each haplotype.(Moreira et al., 2016; Steinberg, 2009) Broadly stated, the most severe clinical courses and the highest prevalence of organ damage in Bantu have been attributed to haplotypes; SEN/AI haplotypes have been linked to milder phenotypes, as indicated by lower rates of hospitalization, less painful episodes, and reduced episodes of ACS; and BEN haplotype carriers exhibit intermediate characteristics.(Moreira et al., 2016; Steinberg, 2009)

After many years of genetic admixture, patients in regions where the HbS gene was introduced by gene flow are frequently compound heterozygotes for two different haplotypes. These regions have also been the site of most of the in-depth and larger studies of the clinical and haematological effects of haplotype in SCA.(Steinberg, 2009) Understanding the connection between haplotype and phenotype becomes more difficult with this genetic mixing as well as others.(Steinberg, 2009)

I.2.1.5.3. Hereditary Persistence of Foetal Haemoglobin

HbF is the primary genetic regulator of the haematologic and clinical characteristics of SCD, decreasing HbS polymerization and enhancing haemolytic processes and vaso-occlusion, which are indicative of the patients' clinical symptoms.(Moreira et al., 2016) The two mechanisms responsible for this are the prevention of HbS polymerization due do the inability of hybrid tetramer ($\alpha 2\beta^S\gamma$) to promote so, and the dilution of intracellular HbS concentration.(Steinberg, 2009; Thein, 2013)

SCA (HbSS) individuals whose HbF concentrations are greater than or equal to 20% are compound heterozygotes for HbS and HPFH, also known as HbS-HPFH, and their phenotype is HbSF.(Atinuke, 2022) HPFH is a term defined by deletions or point mutations within the *HBB* gene cluster, being that the most common form of HPFH is caused by large deletions, which result in greater than usual levels of HbF lasting into adulthood.(Steinberg, 2020) Depending on the extent and breakpoints of deletions, the interactions between γ-globin genes, their upstream and downstream regulatory elements, such as the LCR, and

transcription factor complexes may vary.(Steinberg, 2020) This may explain the broad range of HbF levels in compound heterozygotes for the HbS gene.(Steinberg, 2020)

HbF levels between 15 and 30% are occasionally accompanied by mild illness in African descent who are treatment naïve, and concentrations above 20% are characteristic of patients with the AI haplotype.(Steinberg, 2020)

SNPs in the genes *BCL11A* and *MYB*, which code for repressors of HbF gene expression and deletions or single base alterations in the *HBB* gene cluster, are both factors that contribute to abnormally high HbF levels.(Steinberg, 2020)

Recently, new targets presenting the potential for therapeutic reactivation of HbF are being studied.

I.2.1.5.4. BCL11A, ZBTB7A and ZNF410

The DNA-binding transcription factors (TFs) *BCL11A* and zinc-finger and BTB-domain-containing 7A (*ZBTB7A*) along with the nucleosome remodelling and deacetylase (NuRD) chromatin complex are important regulators of the transition from foetal to adult globin gene expression.(Vinjamur et al., 2021) The *ZBTB7A* gene codes for leukaemia/lymphoma-related factor (LRF), a transcription factor that binds DNA via C-terminal C2H2-type zinc fingers and likely attracts a transcriptional repressor complex through its N-terminal BTB domain.(Masuda et al., 2016) Both *BCL11A* and *ZBTB7A* physically engage NuRD by binding to specific locations in the proximal promoters of the duplicated γ-globin genes *HBG1* and *HBG2*, although the mechanisms of interaction diverge.(Tumburu & Thein, 2021; Vinjamur et al., 2021) Additionally, the NuRD subunits *CHD4*, *GATAD2A*, *HDAC2*, *MBD2*, and *MTA2* are required for γ-globin repression.(Lan et al., 2021)

BCL11A is the main modulator of HbF expression but recently, it was found that *ZBTB7A* is the second major regulator of the human γ-globin genes and a major repressor of the mouse embryonic globin genes.(Martyn et al., 2018) *BCL11A* and *ZBTB7A* bind to the sites at -115 and -200bp of the γ-globin gene promoter, respectively, and they are responsible for most of the γ-globin gene repression in adult cells, can bind motifs in the wild-type (WT) γ-globin gene promoter in vitro, and are disrupted by a wide range of mutations related with HPFH.(Martyn et al., 2018) The potency of repression of these proteins is similar, being that individually leads to a loss of almost 50% of HbF, whereas combined knockout represses roughly the whole HbF production.(Orkin & Bauer, 2019)

Regardless of the two main repressor proteins that silence γ-globin genes, *BCL11A* and *ZBTB7A*, individually bind a cognate recognition site within the γ-globin promoter, to exert their repressive activity, both factors must physically associate with members of the NuRD complex, highlighting the critical function of the NuRD corepressor complex in Hb switching.(Tumburu & Thein, 2021) Defining the processes behind how the multiple NuRD subunits interact to produce an active complex for HbF suppression appears to be a good therapeutic approach.(Tumburu & Thein, 2021) Research aiming to find novel regulators for γ-globin expression discovered that the maintenance of γ-globin silencing demands zinc finger 410 (*ZNF410*), a transcription factor (TF) with five tandem canonical C2H2-type ZFs.(Lan et al., 2021)

ZNF410 has only one single target in erythroid cells in the mammalian genome.(Kim & Dean, 2021) It selectively activates the catalytic subunit of the NuRD nucleosome remodelling and deacetylase complex, CHD4, via two dense binding site clusters.(Kim & Dean, 2021; Lan et al., 2021) In erythroid cells, the NuRD complex helps *BCL11A* and LRF perform their γ-globin-repressive roles.(Lan et al., 2021)

Studies assessing *ZNF410*-depleted cells suggest that *CHD4* is the unique *ZNF410*-regulated NuRD subunit once the mRNA levels of the other NuRD subunits were not diminished.(Lan et al., 2021) Following the mRNA analyses, *ZNF410* removal drastically decreased CHD4 protein levels in HUDEP-2 and primary erythroblasts while maintaining *BCL11A*, LRF, HDAC2, and MBD2 protein levels.(Lan et al., 2021) *ZNF410* depletion did not affect any other genes known to control γ-globin silencing, indicating that *CHD4* is the crucial link between *ZNF410* and γ-globin silencing.(Lan et al., 2021; Vinjamur et al., 2021) Because of this fortunate regulatory route, *ZNF410* may be a druggable therapeutic target for severe hemoglobinopathies to raise γ-globin without causing expected off-target consequences.(Kim & Dean, 2021)

Given that *ZNF410* modulates *CHD4*, and that *BCL11A* and *ZBTB7A* both indirectly interact with the CHD4/NuRD complex, it is believed that these genes represent a considerable influence in HbF synthesis and subsequently, in SCA phenotypes.(Shaikho et al., 2016; Tumburu & Thein, 2021) For these reasons, it is important to explore and study new polymorphisms in *ZBTB7A* and *ZNF410* or their promoters and proximal enhancer elements that may be positively correlated with HbF levels in SCA, disease severity and differentiation in clinical outcomes.

Nkya et al. undertook targeted next-generation sequencing to investigate genetic pathways and known and unknown genetic variants linked to excessive HbF levels in SCD patients, in which they discovered several polymorphisms, including single nucleotide polymorphisms (SNPs) and structural variants like insertions and deletions (INDELS), indicating possible modifier effects.(Nkya et al., 2020) While a SNP is a genomic variant at a single base position in the DNA, INDELS refer to a length discrepancy between two alleles where it is not known whether the difference is originally produced by a sequence insertion or deletion.(Gunter, 2022; *INDEL Mutation*, 2022) Results showed that fewer deletions were present in individuals with low HbF levels compared to those with high HbF levels, noting that important newly discovered variations are enriched in the molecular pathways underlying the regulation of HbF, along with previously known variants.(Nkya et al., 2020)

In the same study, Nkya et al. found two SNPs in the *ZBTB7A* gene with functional impact on HbF concentration, either in patients with lower or higher expression.(Nkya et al., 2020) Pursuing this line of thought, the importance of discovering new SNPs that may interfere with the clinical course of the disease is highlighted, not only through modulation of HbF expression but also by improving the haematological indices and clinical symptoms associated with SCA.

Focusing on reducing clinical severity derived from SCA, polymorphisms at both *ZBTB7A* and *ZNF410* were discovered through database research. The main traits in which these SNPs have an influence are haematological SCA effects that worsen the disease's phenotype.

Six SNPs, described in Table I.3, that appear to be connected to the control of these parameters were recently discovered in the *ZBTB7A* and *ZNF410* genes.

Table I.3. Characterization of *ZBTB7A* and *ZNF410* new SNPs and their potential traits.

Gene	SNP	Trait
***ZBTB7A* (chr19)**	rs10415135	Mean corpuscular volume (MCV)
		RBC count
		Mean corpuscular haemoglobin (MCH)
		The whole brain restricted isotropic diffusion
	rs66534382	RBC count
		Reaction time
	rs56356382	MCV
		MCH
		Mean spheric corpuscular volume
		Mean reticulocyte volume
		Body mass index
	rs111929083	Mean spheric corpuscular volume
		Mean reticulocyte volume
***ZNF410* (chr14)**	rs11844552	MCV
		Immature fraction of reticulocytes
	rs144991697	MCV
		MCH

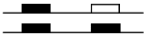
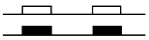
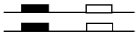
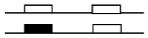
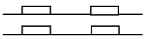
I.2.1.5.5. Alpha-thalassaemia

Alpha-thalassaemia (α-thalassaemia) is characterized by a deficiency in the formation of the α-globin chains of Hb.(Farashi & Harteveld, 2018) It is one of the most prevalent genetic disorders, where carriers' rates can reach up to 80–90% in some places.(Adekile et al., 2020)

The phenotype varies with the degree of α-thalassaemia (Table I.4), being that when one or two α-globin genes are affected (−α/αα, −α/−α, −−/αα), a person is said to have "silent" α-thalassaemia if only one gene is affected or α-thalassaemia trait if both genes are affected, which is regularly associated with mild microcytic and hypochromic anaemia.(Adekile et al., 2020; Farashi & Harteveld, 2018) Loss of three genes (−−/−α) causes a condition known as classical haemoglobin H disease (HbH, β4 tetramer) and it is characterized by moderate to severe anaemia.(Adekile et al., 2020) Upon the deletion of all four genes (−−/−−), Hb Bart's

(γ4) hydrops fetalis originated, and in the absence of early onset of continuous transfusion or stem cell transplantation, is life-threatening.(Adekile et al., 2020)

Table I.4. Classification of α-thalassaemia genotypes and their phenotypic expression.*(Farashi & Harteveld, 2018)*

Phenotype		α-globin genes	Number of functional genes
Normal			4 α-genes
Silent α-thalassaemia			3 α-genes
α-thalassaemia trait	Heterozygous α^0		2 α-genes
	Homozygous α^+		
HbH disease			1 α-gene
Hb Bart's hydrops fetalis syndrome			No functional α-genes

Co-inheritance of α-thalassaemia in SCA is a disease-ameliorating factor since it lowers the amount of Hb in RBCs, which diminishes the propensity of HbS to polymerize, leading to higher Hb levels and slower rates of haemolysis.(Ali Al-Barazanchi et al., 2021) Additionally, the co-inheritance of α-thalassaemia is protective against cerebral vasculopathy, cholelithiasis, leg ulcers, ACS, and chronic renal disease.(Ali Al-Barazanchi et al., 2021)

The 3.7kb α-thalassaemia (-α3.7) is the most common α^+-thalassaemia deletion, derived from the deletion of a single α-globin gene followed by unequal homologous recombination during meiosis.(Farashi & Harteveld, 2018) Previous studies concluded that there is a positive association between -α3.7 and haematological indices, reduced haemolytic rate and reduced anaemia in SCA patients.(Santos et al., 2020)

Furthermore, Santos B. et al observed a lower number of blood transfusions, higher Hb concentration, lower MCV, MCH and number of reticulocytes in these patients, especially in -α3.7 homozygous.(Santos et al., 2020) The most significant aspect contributing to the less severe outcome of this phenotype and perhaps the patient's survival could be the improvement in these haematological indices, namely the reduction of anaemia and the reduction of microcytosis.(Ali Al-Barazanchi et al., 2021; Santos et al., 2020)

CHAPTER II

OBJECTIVES

II. OBJECTIVES

As it was previously mentioned, SCA has a relatively high clinical heterogeneity in its phenotypic expression. According to some studies, genetic factors play a significant role in this phenotypic variability in SCA patients. In the present project, the main objective was to determine and assess the frequency and influence of polymorphisms on disease severity and response to SCA therapy, in a paediatric and adult African patient population followed in the consultation at Clínica Girassol in Luanda, Angola.

This required a few stages:

1. Create a database, containing all relevant demographic, clinical, haematological, biochemical, and imaging information retrospectively collected from hospital records of SCD patients attending Clínica Girassol:
    - Age, sex, age and cause of diagnosis, number of blood transfusions.
    - Steady-state data concerning haemolytic and inflammatory parameters (level of different haemoglobins, LDH, total platelet, reticulocyte, and leukocyte counts).
    - Presence of known phenotypic stroke risk factors, such as high Transcranial Doppler (TCD) velocities, Silent cerebral infarct (SCI) on magnetic resonance imaging (MRI), and previous occurrence of overt stroke.
    - Therapeutic regimen.
2. Identify the sickle cell anaemia genotype (HbSS).
3. Assess the genotypic and allelic frequency of relevant genetic variants:
    - 3.7kb α-thalassaemia
    - Genetic modulators of foetal haemoglobin: β^S-globin haplotype (rs10128556, rs968857, rs3834466, rs28440105)
    - Polymorphisms (SNPs) on *ZBTB7A* (rs10415135, rs111929083, rs56356382, rs66534382) and *ZNF410* (rs144991697, rs11844552).
4. Perform correlation studies between the cited genetic markers of SCA and the haematological, biochemical, and clinical data of patients.
5. Relate the hydroxyurea therapy response with the genetic markers understudy, through retrospective analysis.

CHAPTER III

MATERIALS AND METHODS

III. MATERIALS AND METHODS

III.1. Sample

III.1.1. Population sample

The sample population consisted of 120 SCA Angolan patients selected from a cohort from Clínica Girassol, in Luanda, Angola. As for the eligibility criteria, we expected to have all patients under HU treatment to allow a retrospective study concerning the influence of pharmacological therapy on SCA management. We included individuals of all genders and ages to perform surveillance studies.

In the first medical appointment, a full anamnesis questionnaire was obtained, including family and social context, data regarding the occurrence of clinical manifestations resulting from SCD, as well as parameters under study, specifically the number of hospitalizations, blood transfusions, and strokes. This information was obtained anonymously by the Haematologist in charge of the medical consultation. Clinical, biochemical, and haematological data were also acquired at the medical appointment, along with the remaining information.

This study was approved by Clínica Girassol in Luanda, Angola on 28/05/2021 and by the Ethical Committee of ESTeSL (CE-ESTeSL-Nº51/2021.). Informed consent was obtained and signed by all participants or their caregivers, in the case of children. All the consultations and follow-ups were performed by the project team freely, and with all the human research standards adopted in the Helsinki declaration.

III.1.2. Biological sample

At the selection medical appointment, peripheral blood samples of 3mL were collected in sterile tubes containing EDTA, for laboratory quantitative and genetic analysis. The samples were sent to the Human Genetics Laboratory of H&TRC in Lisbon, Portugal.

III.2. Methods

III.2.1. DNA extraction and Quantification

Genomic DNA was extracted from 200µL of the blood sample using the QIAamp® DNA Blood Mini Kit (50) by Spin Down. This is a very important procedure to ensure that the DNA is purified and does not contain proteins, nucleases, or other contaminants and

inhibitors, to allow an efficient amplification of the fragments during subsequent PCR. The DNA extraction technique was performed according to the QIAamp® DNA Blood Mini Kit (50) manufacturer instructions.

Quality and quantification of the extracted DNA were assessed according to a 260/280 absorbance ratio measured by NanoDrop® (ThermoFisher Scientific) spectrometer. This method allows the determination of DNA concentration in each sample to evaluate whether the amount of DNA in the samples is within the established parameters to enable positive results in subsequent techniques.

DNA samples were stored at -20ºC for posterior analysis.

The molecular identification of sickle cell mutation and the characterization of the eleven remaining β-globin cluster SNPs was performed by PCR followed by RFLP.

Genetic variants under study were selected based on results obtained from previous works on sickle cell anaemia performed in our laboratory (Delgadinho et al., 2021) or were selected from a literature review, using online databases such as NCBI (https://pubmed.ncbi.nlm.nih.gov) and GeneCards (https://www.genecards.org). The molecular methodologies used to study the selected genetic variants are represented in Table III.1.

Table III.1. Molecular techniques are used to analyse each genetic variant.

Category	Gene	Genetic variant	Molecular methodology
Disease confirmation	*HBB*	Sickle cell mutation (*HBB*:c.20A>T)	PCR + RFLP
Haplotype	*HBBP1*	rs10128556 rs968857	RT-PCR
	HBE1 *HBG1*	rs3834466 rs28440105	PCR + RFLP
Genetic modulators	*HBA*	-3.7kb α-thalassaemia	Gap-PCR
	ZBTB7A	rs10415135, rs111929083, rs56356382, rs66534382	RT-PCR
	ZNF410	rs144991697, rs11844552	

III.2.2. Polymerase Chain Reaction

The amplification of genomic DNA extracted was performed through PCR, an essential technique that mimics DNA replication inside the cell to obtain significant amounts of the DNA fragment of interest and subsequently applies other procedures.(Gusmão, 2015) The synthesis of a new strand of DNA complementary to the template strand occurs by using a DNA polymerase and oligonucleotide initiators (primers) that bind to the 3' sequences in each DNA strand of the desired fragment.(*Polymerase Chain Reaction (PCR)*, 2022) The reaction follows three main steps, repeated for several cycles: i) denaturation is the thermal DNA strand separation to make the region available for transcription; ii) annealing allows hybridization of the two primers to the complementary sequences in the DNA strand; iii) extension, through a thermostable DNA polymerase, to synthesize the DNA fragment.(Gusmão, 2015) Each step has a specific temperature, being that the annealing temperature is defined according to the primers' sequence.

In the PCR reaction, each newly synthesized strand works as a new template for the next cycle. Therefore, PCR results in exponential production of the desired fragment, and the number of cycles defines the quantity of final product obtained.(Gusmão, 2015)

Some PCR variants were also performed in this study, namely Gap-PCR and Real-Time (RT)-PCR, allowing direct molecular diagnosis. Gap-PCR is used to detect large known deletions throughout primers flanking the deletion breakpoints.(Gusmão, 2015) The main purpose of using this technique was to detect deletions in the α-globin gene cluster, specifically the 3.7kb α-thalassaemia deletion, according to published literature (see details in Supplementary Material, Table VII.3).(Ferrão et al., 2017)

For every genetic variation studied, genomic DNA was properly amplified in a Thermocycler (Bio-Rad®), using specific primers and PCR conditions, detailed in Supplementary Material, Tables VII.1 to VII.7.

PCR amplification confirmation was performed by electrophoretic migration of the obtained PCR products I agarose gel containing a nucleic acid stain GreenSafe Premium. Electrophoresis was carried out according to each protocol in a Bio-Rad® PowerPacTM Basic Power Supply, with agarose gels of 1% or 2% (w/v), in TBE or TAE 1X, with 3% (v/v) GreenSafe Premium (NZYTech®). Conditions were adapted according to each protocol. Gel revelation was performed in a UV transilluminator (UVP Benchtop UV Transilluminators: 3UV, Fisher Scientific®).

III.2.2.1 ***Molecular identification of the sickle mutation***

Diagnosis confirmation and genotyping for the sickle cell mutation (HbS) were performed by PCR followed by RFLP. Restriction enzymes recognize unique sequences of nucleotides in each DNA molecule strand and cleave them in a specific manner. The enzyme will only hydrolyse DNA molecules that have the nucleotide sequence necessary for recognition in a restriction assay, which is a method for detecting SNPs.

The protocol for this process, including mixture preparation and PCR conditions, is represented in Supplemental Material, in tables VII.1 and VII.2.

III.2.2.2 ***Molecular detection of -3.7kb α-thalassaemia deletion***

Detection of -α3.7kb deletion was evaluated through Gap-PCR, in which two reaction mixtures were prepared for each target fragment. Each mixture had two primers complementary to the sense and antisense chain of DNA regions relative to α2, α1 and

α2/α1. Primer sequence, as well as Gap-PCR conditions, are specified in Supplementary Material, Tables VII.3 and VII.4.

III.2.2.3 ***Haplotype classification***

Classification of HbS haplotypes in our population was based on four previously identified polymorphisms along the β-globin gene cluster.(Shaikho et al., 2017) Two of these SNPs were detected through PCR-RFLP (rs3834466 and rs28440105) whereas the other two SNPs were analysed via RT-PCR (rs10128556 and rs968857).

RT-PCR differs from conventional PCR because it measures the accumulation of amplification product as the reaction progresses, in real-time. This is achieved by the inclusion of a fluorescent reporter molecule in each reaction well that yields increased fluorescence with an increasing concentration of DNA product. Another advantage of RT-PCR is that it does not require gel electrophoresis to evaluate the results, reducing bench time and increasing throughput. Lastly, since RT-PCR reactions occur in a unified and closed-tube qPCR system, the probability of contamination decreases substantially.

All 120 samples were genotyped for the four SNPs to define the five common haplotypes (CAR, CAM, BEN, SEN, and AI). After genotyping all samples for these polymorphisms, they are matched to the predefined genotypic sequence of each HbS haplotype.

The protocol for each SNP detection is detailed in Supplementary Material, including PCR mixtures and reaction conditions, Table VII.5 and Table VII.6.

III.2.2.4 ***Determination of the new SNPs***

ZNF410 and *ZBTB7A* new SNPs were assessed through RT-PCR, following the same principle described for haplotype characterization. Our individuals were genotyped for the 6 new polymorphisms according to the protocols described for each SNP in Supplementary Material, Table VII.7.

III.2.3. Statistical Analysis

It was constructed a database with all the useful information from patient files, including clinical information (age, age at diagnosis, age at first symptoms' manifestation, number of blood transfusions, and clinical outcomes), haematological data (concentration of HbS, HbF and HbA2 pre and post HU treatment), and biochemical data (LDH). The genetic variants analysed throughout the study were all included in this database.

Statistical analysis was performed using SPSS software version 26 (IBM Corp, Armonk, NY, USA) and p-values less than 0.05 were considered significant (95% significance). Allele counts were performed in the overall population (n=120) and the Hardy-Weinberg Equilibrium (HWE) was manually calculated for each genetic variant. For individuals who were in HWE, allele and genotype frequency differences were assessed. Genotypes were also studied under three inheritance modes: dominant (AA), recessive (aa) and overdominant (Aa). For every inheritance mode and every two groups, a new 2x2 contingency table was created.

The normal distribution of the quantitative variables was verified by the Kolmogorov–Smirnov or Shapiro–Wilk normality test. Statistical significance for comparing haematological and biochemical data between genotypes, haplotypes and SNPs was performed using a non-parametric Mann–Whitney U test for single comparisons and Kruskal–Wallis for more than two groups. Chi-square analysis was used for multiple group comparisons between the parameters under study.

CHAPTER IV

RESULTS AND DISCUSSION

IV. Results

IV.1. Genetic characterization

The initial phase of the current project consisted of analysing the DNA samples for each genetic variation understudy. The homozygosity for the HbS mutation was confirmed, the β-globin cluster haplotype was genotyped, and the co-inheritance of α-thalassaemia was evaluated. Afterwards, the potential genetic modulators were investigated.

The 120 samples were examined for the first time in this lab. Most of the techniques have been performed in previous studies, except for the new polymorphisms, which debuted in this project, in our laboratory.

Patients involved in this study were aged from 1 to 69 years old (mean of 15.1 ± 14.2 years), and they comprised both genders with 47.5% females (n = 57) and 52.5% males (n = 63).

IV.1.1. Confirmation of the presence of HbS mutation in the homozygous state

By using PCR-RFLP, it was validated that the HbS mutation was present in homozygosity. The intended DNA fragment was properly amplified for each sample before being submitted to enzymatic restriction by Bsu36I, which recognizes the sequence CC↓TNAGG (nucleotide highlighted is the wild-type at the position of the mutation; "↓" indicates restriction position). When the A>T mutation is present, the sequence is not recognized, not allowing enzymatic restriction to occur. Consequently, as seen in Figure IV.1, the original 375bp PCR product fragment remains unaltered.

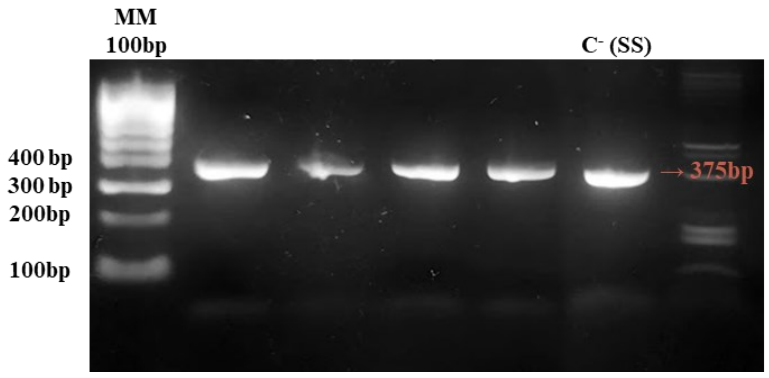


Figure IV.1. Diagnosis of sickle cell mutation by PCR-RFLP. The image represents the revelation by UV light of an agarose gel (2%, w/v) containing GreenSafe Premium to separate products of enzymatic restriction with Bsu36I of samples CG01-CG04. The presence of the HbS mutation eliminates the restriction site. Molecular weight fragments are represented in black (MM 100bp - HyperLadderTM 100bp, Meridian Bioscience®). The molecular weight of restriction-resulting DNA fragments is in red.

All 120 new samples (100%) were homozygous for the HbS mutation, confirming that all subjects suffered from SCA.

IV.1.2. Haplotypes

The β-globin haplotypes can be determined using up to 12 SNPs. However, according to Shaikho EM et al., all the haplotypes spanning the β-globin gene cluster are defined by four SNPs (rs28440105, rs3834466, rs968857, and rs10128556) represented in Table IV.1.(Shaikho et al., 2017)

Table IV.1. Five main haplotypes and genotypes define each haplotype.

Haplotype	rs28440105	rs3834466	rs968857	rs10128556
CAR/CAR	CC	GG	CC	CC
CAR/CAM	CA	GG	CT	CC
CAR/BEN	CC	GG	CT	CC
CAR/SEN	CC	GG	CT	CT
CAR/AI	CC	GGT	CT	CT

All samples were analysed regarding the above-mentioned SNPs. For rs28440105 and rs3834466, the target DNA fragments of each sample were substantially amplified and then subjected to enzymatic restriction by *Hind*III and *Hinc*II, respectively. In both situations, the presence of the recognition sequence causes an enzymatic restriction that is recognized by the detection of fragments with lower molecular weight than the initial one, visible through the agarose gel as exemplified in Figure IV.2.

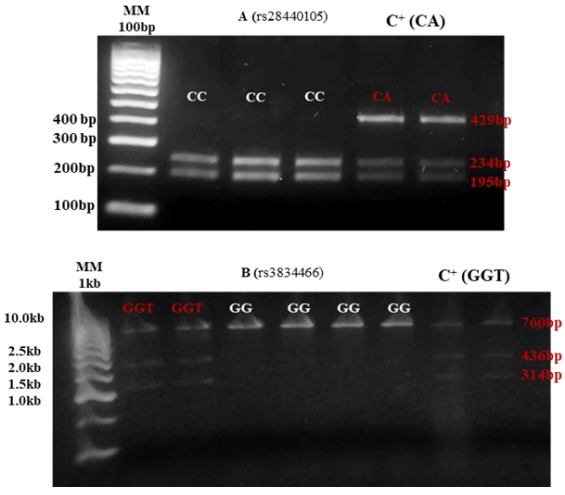


Figure IV.2. Electrophoresis of products of enzymatic restriction for β-globin cluster haplotype. In the image, there are represented two agarose gels (1%, w/v) used to separate the products of enzymatic restrictions with *Hind*III (A) and *Hinc*II (B), after amplification of the fragments of interest. Molecular weight fragments are represented in black (MM 100bp - HyperLadderTM 100bp, Meridian Bioscience®; MM 1kb - GeneRuler 1kb, ThermoFisherTM). The molecular weight of restriction-resulting DNA fragments is in red. "C+" denotes heterozygous positive control for the recognition sequence of each enzyme.

Detection of rs968857 (Figure IV.3) and rs10128556 was performed by RT-PCR.

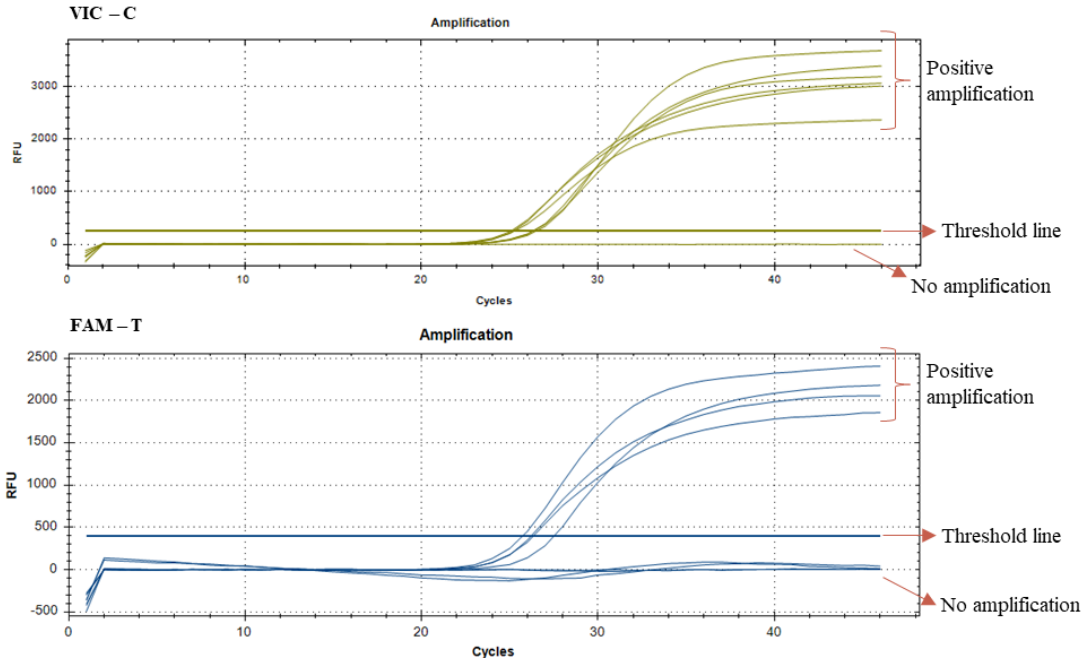


Figure IV.3. Real-Time PCR amplification plots for rs968857. The green curves represent the amplification of the normal allele, while the blue curves represent the amplification of the mutated allele for this SNP.

All SNPs were investigated separately, allowing for simultaneous and non-discriminatory identification of both alleles for each diagnosis. Thus, the combination of SNP profiles for each allele was used to determine the haplotype.

The distribution analysis of haplotype frequencies (Figure IV.4) showed that the most prevalent was the homozygous CAR/CAR, detected in 84.2% (n = 101) of the individuals, followed by the CAR/SEN in 9.2% (n = 11) and CAR/UKN in 4.2% (n = 5). The other haplotypes observed were very rare: 1.7% CAR/CAM (n = 2), 0.8% UKN/UKN (n = 1).

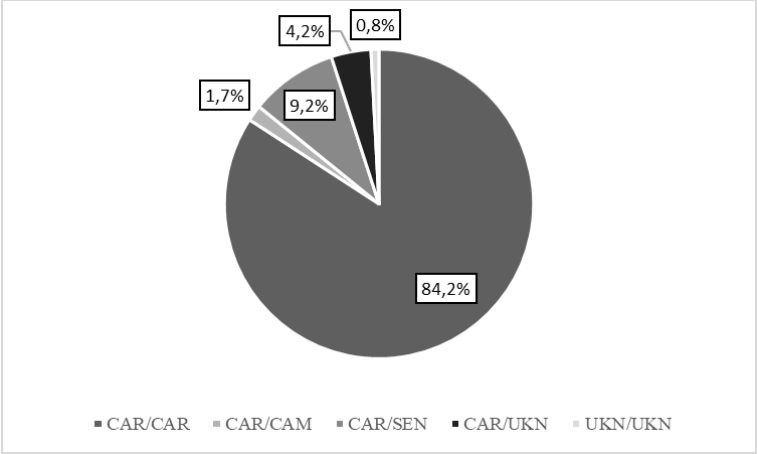


Figure IV.4. Distribution of β-globin haplotypes (%) in the population studied.

Notably, the present study demonstrated a high frequency of the CAR haplotype, with an allelic frequency of 91.7% (220 of 240 alleles). There is a concordance with previous studies that found a high frequency of the homozygous CAR/CAR haplotype and reported the predominance of the CAR haplotype in the Angolan population.

IV.1.3. Diagnosis of α-thalassaemia (-$\alpha^{3.7}$ kb)

The 3.7kb α-thalassaemia deletion was diagnosed by Gap-PCR, using the method reported by Santos et al., 2020, as described in Table VII.3, Supplementary Material, and illustrated in Figure IV.5.

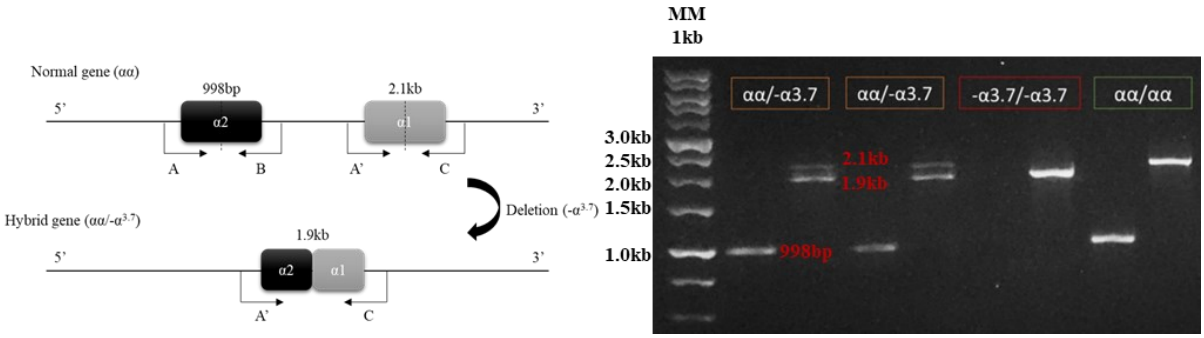


Figure IV.5. Gap-PCR analysis for diagnosis of -α3.7 kb. (A) Schematic representation of breaking points in α2 and α1 globin genes, and the resulting hybrid gene. There are shown the primers used, as well as the size of the fragments. (B) Electrophoresis in agarose gel with the expected results of the three genotypes associated with α-thalassaemia. On the right, homozygous for the non-deletion has both normal α2 and α1 fragments with 998bp and 2.1kb, respectively; homozygous for -α3.7 kb show one single fragment of 1.9kb; and on the left, heterozygous for the deletion have three bands in the gel, corresponding to the three fragments α2, α1 and the hybrid α2α1. Molecular weight fragments are represented in black (MM 1kb - GeneRuler 1kb, ThermoFisher™). The molecular weight of PCR-resulting DNA fragments is in red.

Overall, 44.2% (n = 53) individuals were homozygous for non-deletion, 44.2% (n = 53) of patients were heterozygous (-α3.7/αα), and 11.7% (n = 14) were homozygous for the 3.7kb α-thalassaemia deletion.

The allelic frequency of the 3.7kb α-thalassaemia deletion in our study is 34% (81 of 240 alleles). Our results show that more than 50% of our population has at least one allele with the -α3.7 kb deletion, corroborating previous data concerning the high prevalence of α-thalassaemia in African descent populations.

Based on the three alleles identified in our study and the associated genotypes, our population is in HWE for this variant, with 95% significance ($\chi 2 = 0.018$, $p = 0.892$).

IV.1.4. Analysis of putative genetic modulators

IV.1.4.1 ***ZBTB7A***

Four SNPs in *ZBTB7A*, rs10415135 (Chr.19: 4061546 C>T), rs66534382 (Chr.19: 4056368 G>A), rs56356382 (Chr.19: 4064059 T>C), and rs111929083 (Chr.19: 4079369 A>T), were studied by RT-PCR.

At RT-PCR, the DNA Polymerase cleaves exclusively probes that are hybridized to the target, separating the reporter dye from the quencher dye and therefore, increasing fluorescence by the reporter. The boost in fluorescence only occurs if the amplified target sequence is complementary to the probe, meaning that the fluorescence signal aroused by PCR indicates which alleles are present in the sample.

For SNP rs10415135, the observed allelic frequency was 40% of T alleles and 60% of C alleles. We observed a slight deviation in the HWE ($\chi2 = 6.26$, $p = 0.012$) with a decrease in WT genotype (30%) and SNP homozygous (10.8%), and an increase in heterozygous (59.2%).

The allele distribution for rs66534382 was 40% of A alleles and 60% of G alleles. For this SNP, the population is in HWE ($\chi2 = 3.36$, $p = 0.067$), with 55.8% heterozygous, 11.7% homozygous with the polymorphism, and 32.8% normal homozygous.

For rs56356382, we found 44% C alleles and 56% T alleles in our population. Subsequently, it is not in equilibrium with HW ($\chi2 = 4.90$, $p = 0.027$) due to an increase in heterozygous (59.2%) and a small decrease in WT homozygous (26.7%) and in homozygous with the SNP (14.2%).

Lastly, the observed allelic frequency for rs111929083 was 38% T alleles and 62% A alleles. For this polymorphism, our sample is in HWE ($\chi2 = 0.002$, $p = 0.961$) with 39.2% WT homozygous and 46.7% heterozygous, as expected.

IV.1.4.2 ***ZNF410***

Analysis of two SNPs in *ZNF410*, rs11844552 (Chr.14: 73757874 A>G) and rs144991697 (Chr.14: 73771327 G>A), was made by RT-PCR following the principle described in IV.1.4.1.

For SNP rs11844552, 12% of G alleles and 88% of A alleles were found. The population is in HW equilibrium ($\chi2 = 0.697$, $p = 0.404$) with 85.8% of wild-type homozygous and 14.2% heterozygous for this SNP.

In the population studied for rs144991697, there was only one individual heterozygous for the mutation, being that it was observed with 99.6% G alleles. Our population is in HW equilibrium for this SNP, with 95% significance ($\chi 2 = 0.002$, $p = 0.961$), presenting 99.2% of wild-type homozygous and 0.8% heterozygous.

IV.2. Association studies

For all genetic variants under study, the distribution of genotypes was studied with several factors that have been associated with clinical outcome disparity.

Our study population varied in age from 1 to 69 years, indicating a high survival rate. The average age for SCA diagnosis was 22.4 months and this range was highly variable (0 to 240 months). The mean age for first symptom manifestations was 18.6 months, where severe anaemia was the most common symptom (30%) followed by pain crisis (17.5%) and dactylitis (13.3%). 71.7% of our patients have received at least one blood transfusion, of whom 12.5% have been transfused three or more times.

IV.2.1. Alpha-thalassaemia

In the present study, we observed an increase in the prevalence of the -α3.7 allele related to ageing, confirmed by a frequency of 50% of heterozygous patients above 20 years old and 16.7% homozygous for the deletion. This difference is more evident in -α3.7 homozygous, as the percentage of these individuals increases considerably across age groups, represented in Table IV.2 and Figure IV.6.

Table IV.2. Correlation between age and α-thalassaemia genotype.

		α-thalassaemia					
		αα/αα		**$-\alpha^{3.7}/\alpha\alpha$**		**$-\alpha^{3.7}/-\alpha^{3.7}$**	
		n	%	n	%	n	%
Age (years)	<10	26	48.1%	25	46.3%	3	5.6%
	[10,20[	15	44.1%	14	41.2%	5	14.7%
	≥20	8	33.3%	12	50%	4	16.7%

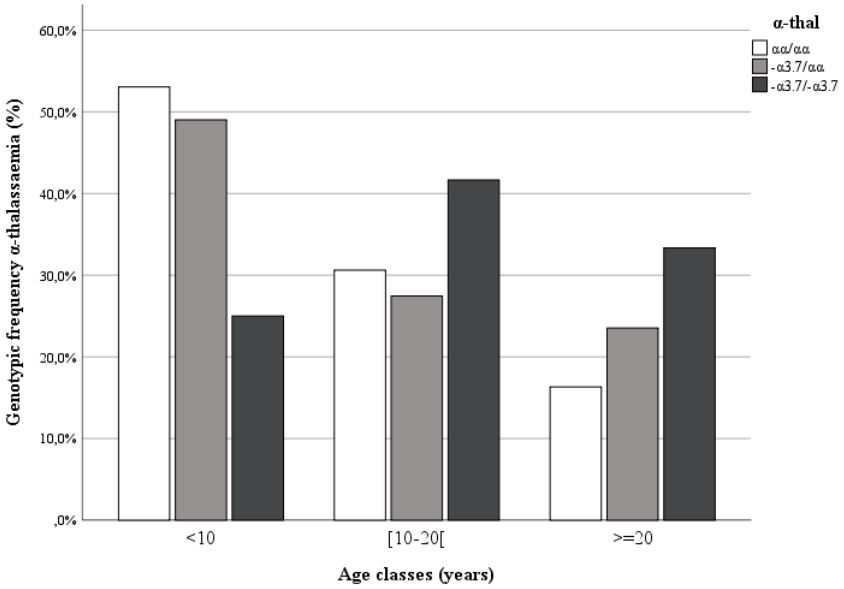


Figure IV.6. Distribution of α-thalassaemia genotypes along age classes.

In addition, the average month for first symptom manifestations also increased with the α-thalassaemia genotype, being that homozygous individuals for this deletion had their first symptoms around 35 months of age, whereas heterozygous manifested symptoms by 17.6 months and non-deletional homozygous began to have symptoms at the age of 15.3 months (represented in Figure IV.7).

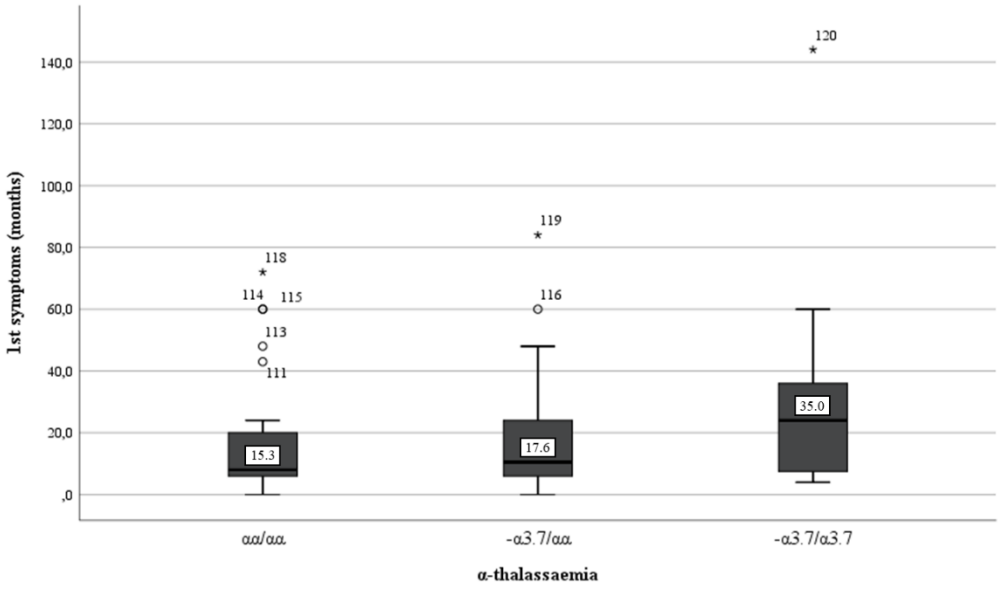


Figure IV.7. Distribution of first symptoms, in months, according to α-thalassaemia genotype.

The age of SCA diagnosis varied according to patients' genotype, following the same principle as the directly above-mentioned parameters. Therefore, normal individuals were

diagnosed by 19.4 months, heterozygous for the deletion at 20.4 months of age and the deletional homozygous were only diagnosed after 45.8 months.

Concerning blood transfusions, 45.3% of both wild-type homozygous and heterozygous required at least one blood transfusion, whilst only 9.3% of homozygous with the α-thalassaemia deletion were transfused. Subsequently, the number of transfusions that deletional homozygous individuals needed was significantly lower than non-deletional homozygous, with a percentage of 10.3% versus 46.2%, respectively for one to two blood transfusions received. Albeit none of the -α3.7 homozygous patients had required more than two transfusions, 28.6% of normal homozygous needed three or more blood transfusions.

Nearly all our individuals (n = 106, 88.3%) were on HU treatment. Among these, 46.5% are children under 10 years old, 32.7% are aged between 10 and 20 years, and 20.8% are adults above 20 years old, indicative of a reduction between pharmacological treatment with HU and ageing. Besides the HU treatment being more prevalent in young children, it is also influenced by the α-thalassaemia deletion. Whilst 45.3% of both SCA non-deletional homozygous and heterozygous individuals were on HU, solely 9.3% of α-thalassaemia homozygous had to undergo treatment with HU.

The occurrence of symptoms increased proportionally to the absence of the -α3.7 allele. In our study, 11.9% of patients with disease symptoms are homozygous for the deletion, while 43.1% were heterozygous and a total of 45% of symptomatic individuals were homozygous for the WT allele. Amongst all symptoms, the biggest disparities related to genotype were related to pain crisis, with 57.1% occurring in normal homozygous, 33.3% in heterozygous, and only 9.5% of -α3.7 homozygous experienced this outcome.

Similarly, we confirmed differences in the prevalence of other clinical outcomes such as dactylitis and severe anaemia, associated with the three α-thalassaemia genotypes.

We observed a great association matching the genotypic frequencies of these SNPs and the α-thalassaemia genotypes. There seems to be a pattern in -α3.7 heterozygotes, as most of them have at least one mutated allele for all the *ZBTB7A* SNPs. The highest frequencies are the α-thalassaemia heterozygous individuals that are heterozygous for the *ZBTB7A* polymorphisms as well; however, more than 60% of -α3.7/αα have leastways one mutated allele from the studied SNPs, knowing that in rs56356382 this frequency is superior to 80%.

Lastly, in our population, two individuals suffered an event of a stroke. One individual was heterozygous for the 3.7kb α-thalassaemia and the other one was homozygous for the WT genotype. Although there was no statistical significance, none of the homozygous patients for the deletion had a stroke event, which may be a beneficial result of the co-inheritance with the α-thalassaemia.

IV.2.2. Beta cluster haplotypes

By assessing the frequency of β-globin haplotypes, summarized in Table IV.3, we found a decrease in the prevalence of CAR/CAR haplotypes linked to ageing, with 49% of CAR/CAR patients being under 10 years old meanwhile the percentage of CAR homozygous aged 20 or more years is 19.8%. Contrary, the presence of an SEN allele seems to be associated with bigger surveillance of SCA patients, grounded by the higher percentage of older individuals with CAR/SEN haplotype (36.4% of CAR/SEN patients over 20 years old versus 27.3% of these patients under 10 years old).

Table IV.3. Correlation between age and β-globin haplotype genotype.

		Haplotype									
		CAR/CAR		**CAR/CAM**		**CAR/SEN**		**CAR/UKN**		**UKN/UKN**	
		n	%	n	%	n	%	n	%	n	%
Age (years)	<10	47	49%	1	50%	3	27.3%	2	100%	1	100%
	[10,20[	30	31.3%	0	0%	4	36.4%	0	0%	0	0%
	≥20	19	19.8%	1	50%	4	36.4%	0	0%	0	0%

The statistical analysis was significant for the age of manifestation of first symptoms in CAR/CAR patients ($p = 0.000$), with a mean age of 16.4 months. The average age of manifestation of first symptoms for individuals with the CAR/SEN haplotype was 38.8 which is more than double comparatively to the CAR/CAR haplotype.

Homozygous CAR patients had a mean age of 20.9 months at diagnosis, CAR/UKN were diagnosed at around 21.5 months and CAR/SEN had the latest diagnosis at 38 months of age (Figure IV.8). The age at diagnosis was statistically significant for CAR/CAR ($p = 0.000$) and CAR/UKN ($p = 0.026$) individuals when using Kolmogorov-Smirnov and Shapiro-Wilk normality tests, respectively.

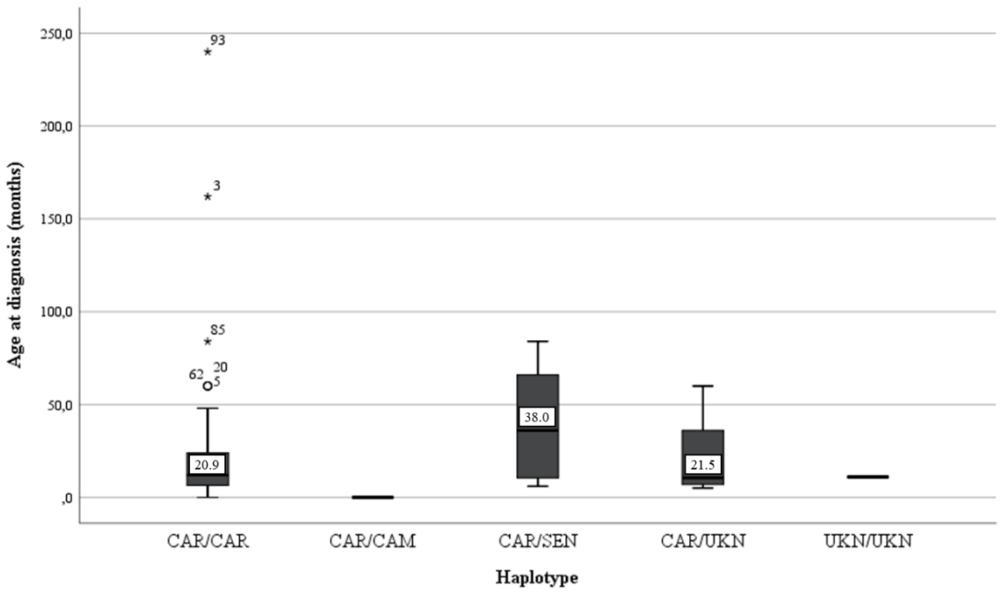


Figure IV.8. Distribution of age at diagnosis, in months, according to β-globin haplotypes.

Despite the previous parameters, the haplotype of individuals who required more transfusions was CAR/SEN with a frequency of 81.8% transfusions. The percentage of transfusions in patients with the remaining haplotypes was also high, with 75% in CAR homozygous, 60% in CAR/UKN and 50% in CAR/CAM individuals. On the other hand, the individuals who required more transfusions were CAR/CAR and 14.6% of them had at least three transfusions, whilst 9.1% of CAR/SEN had the same demand, and no individual of the other haplotypes required the same amount of blood transfusions.

Amongst the studied population, 92 CAR/CAR individuals (94.8%) were on pharmacological treatment with HU whereas the number of CAR/SEN patients receiving this treatment was 7 (77.8%).

Chi-square analysis showed a statistical significance in the percentage of individuals who experienced clinical symptoms derived from SCA according to their β-globin haplotype ($\chi 2 = 31.945$, $p = 0.000$). All CAR/SEN and UKN/UKN patients (100%) had drepanocitosis symptoms, 96.9% of symptomatic individuals were CAR/CAR, and 60% were CAR/UKN. The CAR/CAM haplotype was the only one with just asymptomatic patients.

Out of 16 patients who had dactylitis, 15 were CAR homozygous and there was one CAR/SEN patient. The situation was similar regarding painful crises, with 17 CAR/CAR patients in a total of 21 individuals, 2 CAR/SEN and both CAR/UKN and UKN/UKN had 1 patient suffering from pain crises.

Concerning the stroke incidents, both individuals that suffered a stroke were CAR/CAR. Despite not being statistically significant, this result proves that the CAR/CAR haplotype has a more severe phenotype.

The four *ZBTB7A* new SNPs were analysed and correlated with the β-globin haplotypes. Additionally, we verified that the CAR/CAR haplotype had the greatest genotypic diversity throughout all SNPs. Most of these individuals co-inherited at least one allele of each *ZBTB7A* SNP. The rs111929083 had the highest variability amongst all haplotypes and it was the only SNP having homozygous individuals in all haplotypes.

IV.2.3. *ZBTB7A* polymorphisms

Concerning *ZBTB7A* SNPs, statistical analysis of the age of individuals with the polymorphisms revealed a low frequency of ageing. This data suggests that these SNPs possibly do not interfere with patients' surveillance. Nevertheless, further studies to assess the impact on clinical outcomes were taken.

We noticed that, towards the age of the first symptoms, there was a discrepancy among SNP genotypes. The frequency of heterozygous patients manifesting their first symptom after 12 months was twice the frequency of non-mutated homozygous (22 versus 11 individuals, respectively). This trend has been observed in all the SNPs, evidencing a positive effect of having one mutated allele.

Relatively to clinical outcomes, there was a considerable reduction in pain crisis associated with the presence of the two mutated alleles, meaning that homozygous for the SNPs had less pain crisis than individuals without the polymorphisms. We likewise verified that patients who had the SNPs were less prompted to anaemia events, especially homozygous for rs111929083, represented in Figure IV.9.

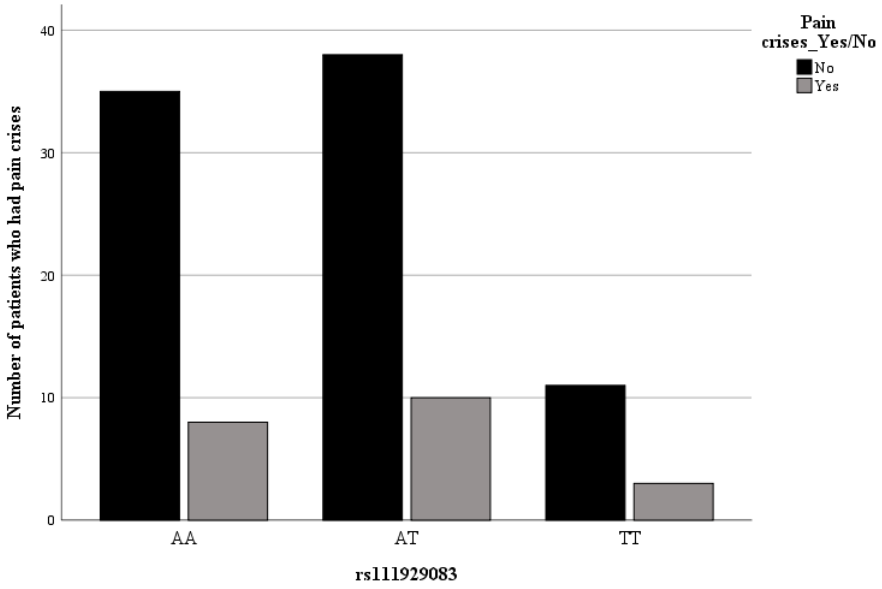


Figure IV.9. Number of patients who had pain crises according to their rs111929083 genotype.

IV.3. Biochemical characterization

IV.3.1. Alpha-thalassaemia

Contrary to expected, HbF levels previous and posterior to HU treatment do not seem to improve in the presence of 3.7kb α-thalassaemia. In our population, both HbS and HbF concentrations got worsened proportionally with the presence of the -α3.7 allele, either before or after pharmacological treatment.

Although not statistically significant, the α-thalassaemia deletion seemed to be positively related to HbA2 levels, since they increased directly with the presence of the -α3.7 allele (Figure IV.10). These values also increased after the treatment with HU, being that the mean HbA2 concentration in WT homozygous ranged from 2.54% to 2.80% pre-HU and post-HU, respectively. In heterozygous it varied from 3.11% to 4.39%, and in deletional homozygous, the concentration went from 4.20% to 4.45% after pharmacological treatment.

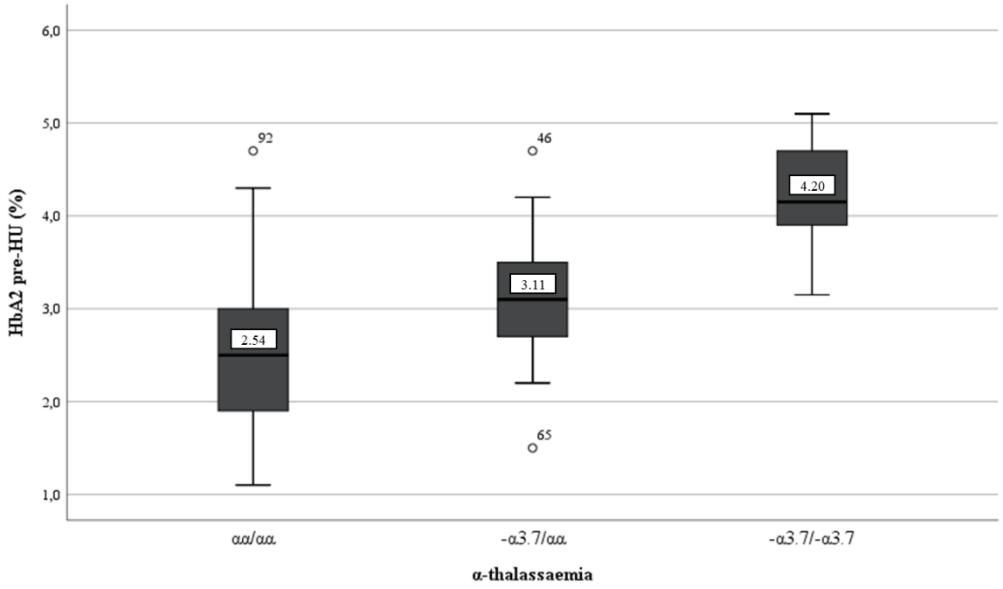


Figure IV.10. The concentration of HbA2 (%) associated with each α-thalassaemia genotype.

LDH levels showed an association with the α-thalassaemia genotypes, and it was statistically significant for -α3.7 heterozygous (p = 0.005) in the Kolmogorov-Smirnov normality test. We observed a decrease in LDH concentration directly proportional to the presence of the -α3.7 allele, where normal homozygous had a mean concentration of 674.82U/L, heterozygous presented 594.41U/L, and -α3.7 homozygous had 472.96U/L mean levels, indicated in Figure IV.11.

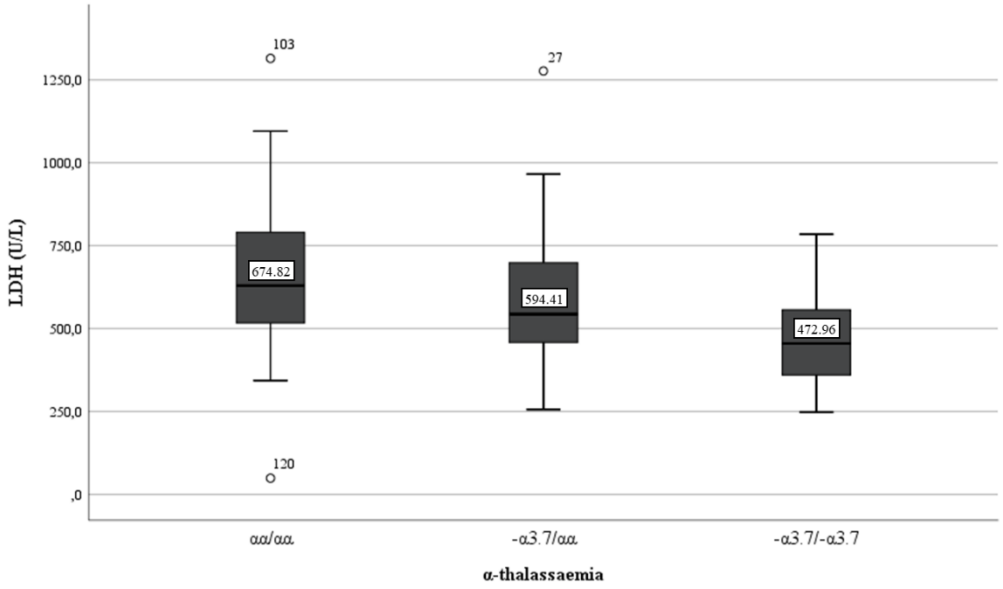


Figure IV.11. LDH concentration, in U/L, according to each α-thalassaemia genotype.

IV.3.2. Beta cluster haplotypes

CAR/CAR haplotype had surprisingly good haematological levels compared to the other haplotypes, being pre-HU HbS mean concentrations of 79.29% and an HbF average was 16.96%. The second less severe haplotype was CAR/UKN with an HbS range of 82.23% and 15.23% of HbF pre-HU. The CAR/SEN was the haplotype with the worse haematological outcomes since HbS mean concentration was 84.34% and HbF was around 12.51% without any treatment.

At the HbA2 level, the most advantageous haplotype was CAR/SEN which showed an average of 3.15% while CAR/CAR had 3.02%, and CAR/UKN presented an HbA2 concentration of 2.75% previously to pharmacological medication. After HU treatment, the levels of HbA2 were 3.91% for CAR/SEN individuals and 3.08% for CAR homozygous.

LDH levels were statistically significant for CAR/CAR patients (p = 0.009) in the Kolmogorov-Smirnov normality test, with a mean of 604.96U/L, whereas CAR/SEN presented 632.92U/L and CAR/UKN individuals had the highest average of 788.70U/L. The lowest LDH levels were observed in CAR/CAM individuals, with a mean range of 544.95U/L (Figure IV.12).

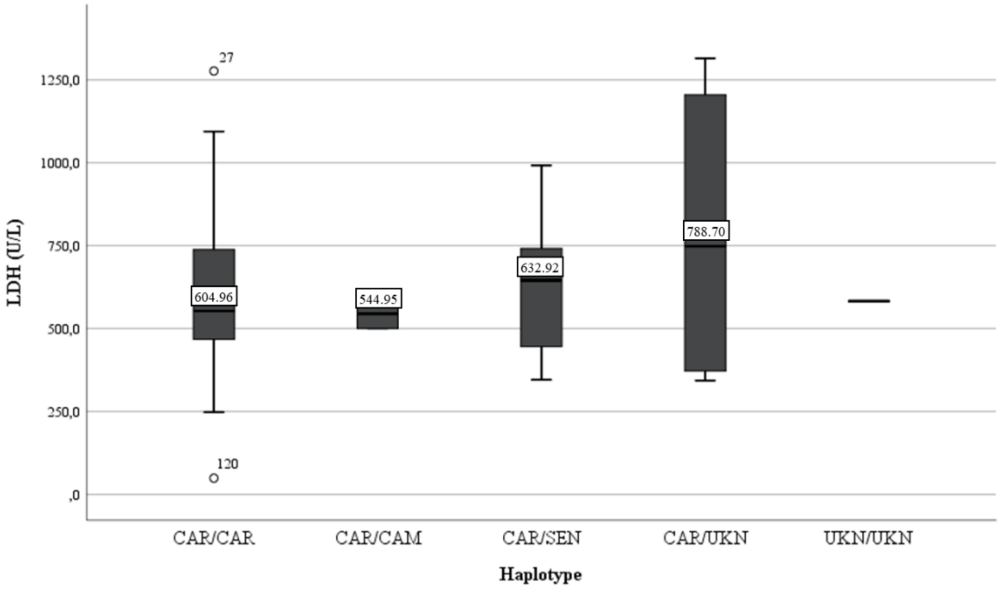


Figure IV.12. LDH concentration, in U/L, according to each β-globin haplotype.

IV.3.3. *ZBTB7A* polymorphisms

In HU-naïve patients, the haematological parameter that showed better results was HbA2, which increased with the presence of mutated alleles for all SNPs.

For rs111929083 there were statistically significant values for HbS and HbF (p = 0.026 and p = 0.038, respectively). In this polymorphism, HbS mean values were reduced by the existence of at least one mutated allele, being that normal homozygous had 81.89% of HbS, heterozygous showed 78.60% and the HbS concentration for mutated homozygous was 79.36% (represented in Figure IV.13). The opposite situation was observed in HbF concentrations, where wild-type homozygous had 14.39% HbF levels, heterozygous had 17.61% and the average concentration in SNP mutated homozygous was 17.72%.

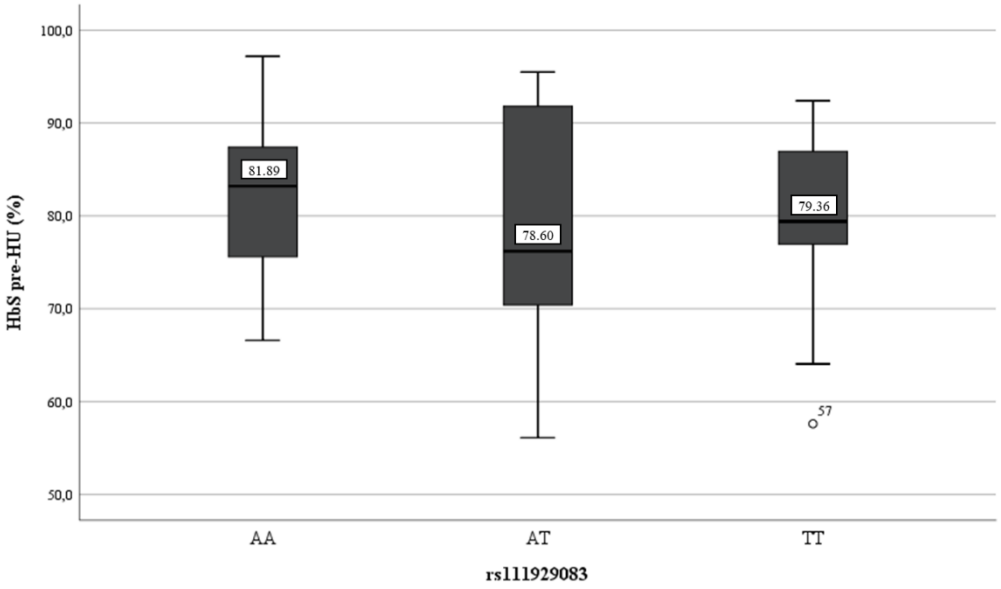


Figure IV.13. HbS concentration (%) in each genotype of rs111929083.

The rs1041535 SNP had statistical significance in association studies between post-HU HbS and HbF, and heterozygous patients (p = 0.003 and p = 0.001, respectively). We also observed statistical significance for these parameters in mutated homozygous individuals (p = 0.018). Concerning the HbS concentrations, even though heterozygous values were higher than WT homozygous (83.59% compared to 79.55%), we observed a reduction of HbS in mutated homozygous for 78.75%. The opposite occurred in HbF levels, shown in Figure IV.14, which were lower in heterozygous (13.18%) and higher in SNP homozygous (18.34%).

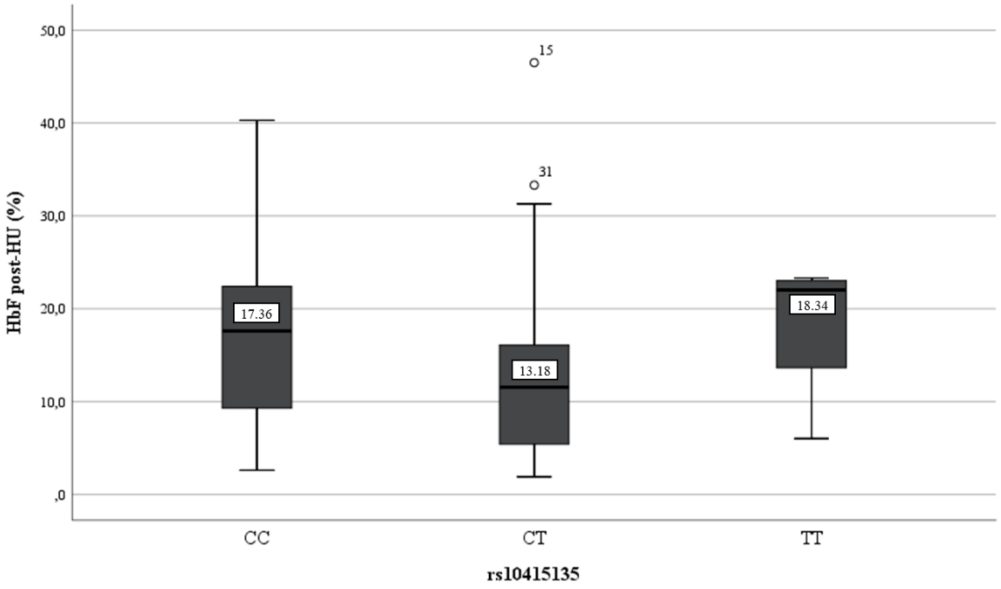


Figure IV.14. HbF concentration (%) in each genotype of rs10415135.

This pattern was observed for all *ZBTB7A* polymorphisms in patients who received HU treatment. Despite heterozygous having worse values than non-mutated homozygous, individuals homozygous for these SNPs have greater results in HbS and HbF values.

In contrast with these results, the HbA2 concentrations post-HU treatment decreases with the presence of mutated alleles, being that WT homozygous show the highest HbA2 levels with statistical significance in rs10415135, rs66534382 and rs56356382 ($p = 0.000$). However, in rs56356382 we obtained statistical significance for polymorphism homozygotes ($p = 0.031$) which is probably related to its substantial reduction in average HbA2 levels (2.87%) compared to the 3.25% in heterozygous and 4.97% in wild-type homozygous.

In rs111929083 the situation reversed and mutated homozygous had statistical significance concerning HbA2 post-HU levels due to its extremely high levels of 5.71% when compared to 3.31% in heterozygous and 3.10% in normal homozygous.

There is statistical significance for LDH levels in rs10415135 with Shapiro-Wilk normality tests for mutated homozygous ($p = 0.005$). Moreover, Mann-Whitney tests showed statistical significance between heterozygous and mutated homozygous ($p = 0.046$) since heterozygous have LDH levels of 635.80 U/L while polymorphism homozygous had an average of 567.16 U/L, which is also lower than WT homozygous mean values (588.05 U/L) (Figure IV.15).

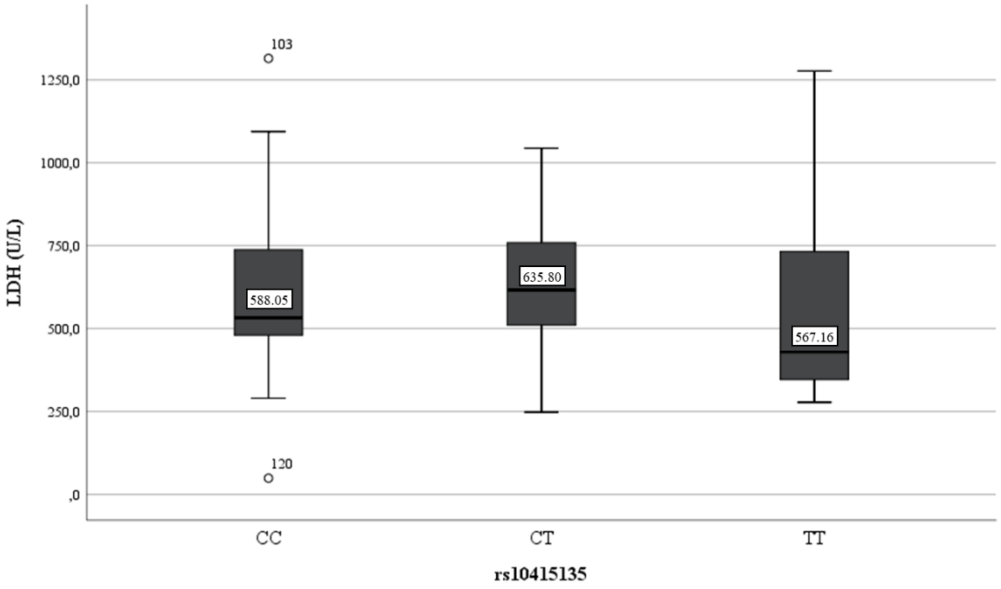


Figure IV.15. LDH concentration (U/L) in each genotype of rs10415135.

For the rs66534382 and rs56356382 SNPs, we also obtained statistically significant results with Shapiro-Wilk normality tests, for homozygous individuals with the polymorphisms ($p = 0.014$ and $p = 0.010$, respectively). In both SNPs, we confirm that mutated patients benefit from the lowest LDH concentrations, whereas heterozygous have the most elevated levels among the tree polymorphism genotypes, as we can confirm in Figure IV.16.

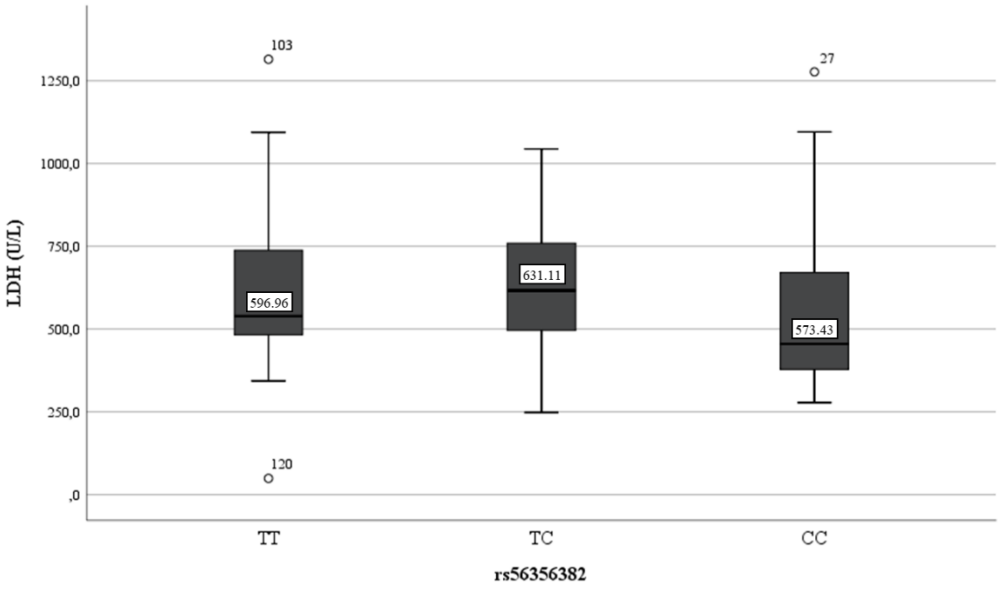


Figure IV.16. LDH concentration (U/L) in each genotype of rs56356382.

For rs111929083 the results are slightly different, and the statistical significance observed was in wild-type homozygous ($p = 0.024$). In this polymorphism, in opposition to the other *ZBTB7A* SNPs, the individuals with the lowest and most favourable LDH levels are the WT

homozygous with an average of 610.49 U/L, compared to the uppermost of 619.34 U/L in mutated homozygous and 614.26 U/L in heterozygous.

IV.4. Discussion

The present study evaluated 120 patients, aged from 1 to 69 years old, with sickle cell anaemia (SCA) for the presence of 12 genetic variants in 7 genes. These included the co-inheritance of α-thalassaemia and β-globin gene cluster haplotypes, which are two well-known modulators of disease severity. Six of the polymorphisms studied were disease modulator candidates, chosen due to their possible role in haematological processes including RBC counting and MCV, MCH, and mean reticulocyte volume. All of these have been suggested as intervening factors for the development of vasculopathy and the worsening of clinical outcomes in drepanocitosis.

Every patient in this Angolan cohort carried at least one CAR allele, which is extremely prevalent in our population, and consequently, it was perhaps challenging to identify a statistical conclusion. The haplotype prevalence in four Sub-Saharan African countries was previously studied and researchers observed a frequency of 92% CAR/CAR homozygous.(McGann et al., 2018) Another study found that the frequency of haplotype carriers was considerably more variable with 82.2% CAR, 11.2% BEN and 6.6% SEN in Angola.(Borges et al., 2019)

The CAR haplotype frequently carries a worse prognosis and seems to be connected to increased haemolysis.(Chamouine et al., 2020) This was unquestionably shown in our community by the existence of the worst clinical haplotypes, as measured by HbA2 levels, HU treatment, number of blood transfusions received, clinical outcomes such as dactylitis and pain crisis, and surveillance rates. Despite the low sample population for the CAR/SEN haplotype, the analysis of the results shows that it presented the worse phenotype in some clinical and haematological parameters, namely the age of manifestation of first symptoms, age at diagnosis, the need for blood transfusions, the occurrence of SCA symptoms and the concentrations of HbS and HbF. This is not in agreement with other studies, which report a less severe disease severity for SEN and AI haplotypes, and an intermediate level for CAM haplotypes.(Piel et al., 2017) However, it is important to mention that the patients we had for the other haplotypes were all heterozygotes, and the variability in our population was quite low.

In addition to the role of haplotypes in regulating this disease, the co-inheritance of α-thalassaemia has also been regarded as a major genetic modulator, providing a milder phenotype when present. This hypothesis was previously proved by our research team in a

study on an Angolan population.(Santos et al., 2020) Our population was in Hardy-Weinberg equilibrium for some genetic variants, of which we emphasize co-inheritance of 3.7kb α-thalassaemia deletion, which was previously described to be implicated as a risk modifier variant in studies performed to assess the impact of this deletion on patients' clinical severity.

In the present study, 11.7% of SCA patients co-inherited the 3.7kb α-thalassaemia deletion in homozygosity. This homozygous frequency is similar to the one observed in Angola (12.5%) (Santos et al., 2020) and Saudi Arabia (14.0%) (El-Hazmi & Warsy, 1993), and it was higher than the results observed in Brazil (1.4%) (Belisário et al., 2010), Georgia, USA (1.7%) (Adams et al., 1994), Senegal (2.3%) (Gueye Tall et al., 2019), Cameroon (6.8%) (Wonkam et al., 2014), and France (8.6%).(Joly et al., 2012) However, it was lower than the frequencies observed in the Democratic Republic of Congo (25.5%) (Mikobi et al., 2018) and New Delhi, India (38.9%).(S. Pandey et al., 2011)

According to preceding research, SCA patients who co-inherit the α-thalassaemia deletion have a different survival rate, with the normal α genotype or heterozygous individuals having greater mortality.(Wonkam et al., 2014) The patients in our study ranged in age from 1 to 69 and we observed a significant increase in the frequency of 3.7kb α-thalassaemia deletion in patients above 20 years old, which is under this hypothesis.

Moreover, in terms of mortality, our findings regarding morbidity have revealed a well-known benefit of the co-inheritance of α-thalassaemia and SCA, especially attending to some clinical adverse events such as painful crisis and severe anaemia, which was confirmed by a noteworthy difference in the number of blood transfusions amongst patients. In accordance, previous researchers have already reported this beneficial effect.(Santos et al., 2020)

Also, the group of 3.7kb α-thalassaemia deletion homozygous patients showed a significantly higher age of first manifestation (15.3 months vs. 35 months) and this correspondingly late illness start had been previously reported in other African populations with the same condition.(Wonkam et al., 2014)

Regarding morbidity, our results have shown a benefit of the co-inheritance of α-thalassaemia with the SCA, specifically attending to some clinical adverse events such as strokes. In our population two individuals experienced an event of a stroke, being that the group presenting homozygosity for α-thalassaemia never experienced a stroke. In agreement, this positive effect has been reported by other authors.(Belisário et al., 2015) Our association studies

revealed that both patients were CAR/CAR which supports previous studies concerning the more severe prognosis characteristic of this haplotype.(Delgadinho et al., 2021)

Our study confirms that there is an improvement in some haematological indices associated with 3.7kb α-thalassaemia deletion, reduced haemolytic rate and reduced anaemia in the population studied. The HbA2 concentration increased with the presence of the deletion, and it has been described to prevent the polymerization of deoxygenated HbS and therefore, reduce clinical events characteristic of SCA. It was speculated that HbA2 can positively influence HbF levels, but it was not confirmed in our study. A lower number of blood transfusions per year, because of less pronounced anaemia, fewer events of pain crisis, and diminished LDH levels were witnessed with the increased number of 3.7kb alpha-thalassaemia deletions, with more significant values in the homozygous.

Several authors believe that the most significant element contributing to the less severe outcome of this phenotype and perhaps the patient's life is the improvement in certain haematological indices, specifically the reduction of anaemia.(Rumaney et al., 2014) In this study, we confirm that the co-inheritance of α-thalassaemia ameliorates the haemolytic rate of SCA, subsequently to an increase in HbA2 concentrations, thus decreasing the tendency of deoxy-HbS to polymerize.(Rumaney et al., 2014)

The effects of α-thalassaemia include a decrease in HbS, which results in a decrease in the rigidity and sickling of red blood, a longer life expectancy, reduced microcytosis, lower reticulocyte counts, MCH, and MCV, as well as an increase in haematocrit and blood viscosity.(Kato et al., 2017; S. K. Pandey et al., 2014) These created the hypothesized biological explanation for the association between SCA and α-thalassaemia and the subsequent clinical beneficial effects on patients. This theory is supported by the current study's finding that the existence of the α-thalassaemia deletion was associated with a decrease in transfusions and pain events.

In summary, our study has revealed the incidence of 3.7kb α-thalassaemia deletion in sickle cell anaemia paediatric and adult patients from Angola, including 11.7% homozygotes and 44.2% heterozygotes, and its relationship to the disease phenotype. In these SCA Angolan patients, there was a less severe phenotype related to the number of deletions. Improved haematological indices, lower blood transfusions, anaemia severity and painful events, contributed to an improvement of the general well-being and improved the survival of SCA homozygous for α-thalassaemia. For the first time in an Angolan population, our results

confirmed the hypothesis of our prior study concerning SCA patients' surveillance in association with the co-inheritance of 3.7kb α-thalassaemia deletion.

Regarding the *ZNF410* SNPs, there was an extremely low genetic variability in both SNPs being that there were not found homozygous for both SNPs. These results did not allow genotype-phenotype association studies, regardless of the theoretically described importance of *ZNF410* in modulating SCA outcomes. This result may be related to a particular disease phenotype, but further studies are needed to confirm this hypothesis.

Lan X et al., through a CRISPR-Cas9 screening study, discovered *ZNF410* to be an HbF repressor with a specialized role in promoting *CHD4* expression and γ-globin silencing, which was confirmed by Vinjamur et al.(Lan et al., 2021; Vinjamur et al., 2021) Besides the previous studies have been performed using the human adult-type erythroid cell line HUDEP-2, mouse erythroid cell line and mouse models, they found that total loss of *CHD4* compromises severely haematopoiesis and erythroid cell growth, but depletion of *ZNF410* does not affect these components because *CHD4* levels remain sufficient to maintain cellular functions.(Lan et al., 2021; Vinjamur et al., 2021) Based on these results, Tumburu L and Thein SL concluded that *ZNF410* deficiency or full absence in mice was not associated with anaemia or haemolysis and was well tolerated throughout erythropoiesis, haematopoiesis, and development.(Tumburu & Thein, 2021)

Bearing this in mind, we were expecting to observe increased concentrations of HbF with no unfavourable significant alterations in the other haematological parameters and clinical outcomes, especially in the degree of anaemia.

In Shaikho EM et al. study assessed eight SNPs in *ZBTB7A* or its promoters and proximal enhancer elements, and in putative binding motifs in and adjacent to the *HBB* gene cluster and compared them with HbF levels in SCA.(Shaikho et al., 2016) The frequencies of SNPs in this study varied between 22% and 56% in the African American cohort.(Shaikho et al., 2016) These authors concluded that SNPs in putative *ZBTB7A* binding sites separate higher levels of HbF in Arab-Indian haplotypes from African haplotypes with lower HbF concentration.

Contrarily, in *ZBTB7A* SNPs, we confirmed a substantial genetic variability with the frequency of the favourable alleles between 38% and 44% in our African population, with an increase in HbF levels in homozygous patients. Moreover, an attenuation in some clinical outcomes was also observed in homozygous for the SNPs, namely in the manifestation of symptoms and pain events.

The Hardy-Weinberg deviations may indicate a natural selection in the African SCA population with *ZBTB7A* polymorphisms, being that the presence of these SNPs provides beneficial haematological data by increasing HbF levels, as well as they reduce clinical events associated with disease severity. Therefore, the differential mortality in individuals with better outcomes is standing out in Angolan SCA patients.

In all SNPs heterozygous the age of manifestation of first symptoms was slightly higher when compared to WT individuals, proving that these SNPs have a beneficial effect on disease severity. Furthermore, we observed a reduction in pain crisis events in mutated homozygous patients, as well as fewer anaemia events, the last one particularly evident in homozygous for rs111929083.

In terms of haematological parameters, the individuals in our study had increased HbA2 levels proportional to the presence of mutated alleles of the four SNPs assessed. In association with rs111929083, beneficial alterations were observed at the level of HbS, which reached its lowest concentrations in heterozygous patients, whilst the homozygous for the polymorphism presented the highest HbF percentage.

Regarding the rs10415135 and haematological indices posterior to HU treatment, solely homozygous individuals for this SNP presented benefits as they had the lowest HbS concentration and the most prominent HbF concentration. Nonetheless, these results were evident in all SNPs but were not as statistically significant as for rs10415135. Ongoing with the subject of post-HU haematological data, patients with the rs111929083 mutation were the only ones showing positive results in HbA2 levels.

Considering LDH, patients with homozygosity for all SNPs beneficiated from the lowest LDH levels, except for the individuals with rs111929083 SNP.

Despite the lack of previous studies on these polymorphisms, we obtained statistically significant in some clinical and haematological parameters, as well as divergences among SNPs. This means that these polymorphisms have an impact on SCA severity and are possible good modulators of the disease pathophysiology.

ZBTB7A SNPs are in chromosome 19 and *ZNF410* SNPs are in chromosome 14, along with the respective gene location. It is known that *ZNF410* main function is to silence HbF expression in adult erythroid cells and that *ZBTB7A* represses the transcription of an extensive range of genes involved in cell proliferation and differentiation, including the switch between

foetal and adult globin expression during erythroid cells maturation. Given this information and the described effects of the polymorphisms analysed in this study, there is a possibility that they may alter the function of the protein, namely in terms of its activity or expression.

It is known that *ZBTB7A* acts as one of the major regulators' of γ-globin genes, causing its repression in adults.(Martyn et al., 2018) Additionally, it has been published that *ZNF410* has a single target in erythroid cells of the human genome, and it selectively activates *CHD4* to promote γ-globin gene silencing.(Lan et al., 2021) Therefore, the repression of both genes can be very profitable for SCA patients. Our results are innovative since it is the first time these SNPs are being studied. Therefore, we highlight the importance of more studies with a bigger population and assessing more haematological parameters, to corroborate our results.

CHAPTER V

CONCLUDING REMARKS AND FUTURE PERSPECTIVES

V. Concluding Remarks and Future Perspectives

This study provides a relevant contribution to the Angolan population's genetic background, where the CAR haplotype is unquestionably the most common *HBB* haplotype, and the co-inheritance of the 3.7kb α-thalassaemia has a major impact on modulating the clinical course of SCA. In this project, we studied new SNPs found in recently approached genes for the first time, and we performed association studies between these polymorphisms and the parameters that characterize SCA patients' phenotypes. Significant differences were observed in several clinical parameters as well as in some haematological data, in all polymorphisms studied.

SCA has a high heterogeneity in its phenotypic expression. Several polymorphisms are being discovered every day that could explain the variation in HbF levels according to different geographic regions, as well as the severity of clinical events affected by the genotype of individuals.

There is still a long way to go before we completely understand a disease as complex as SCA and why it manifests in patients in such a variety of ways. This heterogeneity is undoubtedly influenced by patients' genetic heritage, but it is not the only factor. We consider that using more advanced techniques such as Next Generation Screening (NGS) could expand our knowledge of SCA heterogeneity and correlated severity since it allows the study of the effect of multiple variants.

To support our study, a population of older patients should be carefully assessed to reduce potential confounding factors or to build distinct protective characteristics unique to each specific population or subgroup. Additionally, it would be important to collect more information regarding clinical manifestations or events related to disease severity. In future studies, it should also be included more haematological variables, namely erythrocyte, white blood cell, neutrophil platelet, and reticulocyte count, as well as mean corpuscular volume and mean corpuscular haemoglobin to precisely infer results towards our hypotheses.

The results of this study emphasize the importance of personalized healthcare for SCA patients, particularly in Angola. We hope to contribute to the progress in SCA research globally with our results.

VI. Literature Cited

Abboud, M. R. (2020). Standard management of sickle cell disease complications. *Hematology/Oncology and Stem Cell Therapy*, *13*(2), 85–90. https://doi.org/10.1016/j.hemonc.2019.12.007

Abdu, A., Gómez-Márquez, J., & Aldrich, T. K. (2008). The oxygen affinity of sickle hemoglobin. *Respiratory Physiology & Neurobiology*, *161*(1), 92–94. https://doi.org/10.1016/j.resp.2007.12.005

Adams, R. J., Kutlar, A., McKie, V., Carl, E., Nichols, F. T., Liu, J. C., McKie, K., & Clary, A. (1994). Alpha thalassemia and stroke risk in sickle cell anemia. *American Journal of Hematology*, *45*(4), 279–282. https://doi.org/10.1002/ajh.2830450402

Adekile, A. (2021). The Genetic and Clinical Significance of Fetal Hemoglobin Expression in Sickle Cell Disease. *Medical Principles and Practice*, *30*(3), 201–211. https://doi.org/10.1159/000511342

Adekile, A., Sukumaran, J., Thomas, D., D'Souza, T., & Haider, M. (2020). Alpha thalassemia genotypes in Kuwait. *BMC Medical Genetics*, *21*(1), 170. https://doi.org/10.1186/s12881-020-01105-y

Ahmed, M. H., Ghatge, M. S., & Safo, M. K. (2020). *Hemoglobin: Structure, Function and Allostery* (pp. 345–382). https://doi.org/10.1007/978-3-030-41769-7_14

al Talhi, Y., Shirah, B. H., Altowairqi, M., & Yousef, Y. (2017). Laparoscopic cholecystectomy for cholelithiasis in children with sickle cell disease. *Clinical Journal of Gastroenterology*, *10*(4), 320–326. https://doi.org/10.1007/s12328-017-0750-3

Al-Balushi, Z., Bhatti, K., Ehsan, M., Al-Shaqsi, Y., Al-Sharji, N., Mady, H., & Sherif, M. (2016). Laparoscopic Splenectomy Alone for Sickle Cell Disease: Account of 50 paediatric cases. *Sultan Qaboos University Medical Journal*, *16*(4), e482-486. https://doi.org/10.18295/squmj.2016.16.04.013

Ali Al-Barazanchi, Z., Abdulateef, S., & Hassan, M. (2021). Co-Inheritance of α-thalassemia gene mutation in patients with sickle cell Disease: Impact on clinical and hematological variables. *Nigerian Journal of Clinical Practice*, *24*(6), 874. https://doi.org/10.4103/njcp.njcp_11_20

Alsamiri, A., Alzahrani, F., Filimban, N., Khojah, A., Felimban, R., & Qadah, T. (2021). A Cross-Sectional Study of Detection of Beta Globin (HBB) Haplotypes Among Beta Thalassemia Patients. *Cureus*. https://doi.org/10.7759/cureus.13367

Antwi-Boasiako, C., Ekem, I., Abdul-Rahman, M., Sey, F., Doku, A., Dzudzor, B., Dankwah, G., Hager Otu, K., Ahenkorah, J., & Aryee, R. (2018). Hematological parameters in Ghanaian sickle cell disease patients. *Journal of Blood Medicine*, *Volume 9*, 203–209. https://doi.org/10.2147/JBM.S169872

Anurogo, D., Yuli Prasetyo Budi, N., Thi Ngo, M.-H., Huang, Y.-H., & Pawitan, J. A. (2021). Cell and Gene Therapy for Anemia: Hematopoietic Stem Cells and Gene Editing. *International Journal of Molecular Sciences*, *22*(12), 6275. https://doi.org/10.3390/ijms22126275

Atinuke, F. (2022). The Effect Of Alpha Thalassemia, HbF And HbC On Haematological Parameters Of Sickle Cell Disease Patients In Ibadan, Nigeria. *Mediterranean Journal of Hematology and Infectious Diseases*, *14*(1), e2022001. https://doi.org/10.4084/MJHID.2022.001

Belisário, A. R., Nogueira, F. L., Rodrigues, R. S., Toledo, N. E., Cattabriga, A. L. M., Velloso-Rodrigues, C., Duarte, F. O. C., Silva, C. M., & Viana, M. B. (2015). Association of alpha-thalassemia, TNF-alpha (-308G>A) and VCAM-1 (c.1238G>C) gene polymorphisms with cerebrovascular disease in a newborn cohort of 411 children with sickle cell anemia. *Blood Cells, Molecules, and Diseases*, *54*(1), 44–50. https://doi.org/10.1016/j.bcmd.2014.08.001

Belisário, A. R., Rodrigues, C. V., Martins, M. L., Silva, C. M., & Viana, M. B. (2010). Coinheritance of α-Thalassemia Decreases the Risk of Cerebrovascular Disease in a Cohort of Children with Sickle Cell Anemia. *Hemoglobin*, *34*(6), 516–529. https://doi.org/10.3109/03630269.2010.526003

Bender, M. (1993). *Sickle Cell Disease*.

Bender, M. (2003). *Sickle Cell Disease* (M. P. Adam, D. B. Everman, G. M. Mirzaa, R. A. Pagon, S. E. Wallace, L. J. Bean, K. W. Gripp, & A. Amemiya, Eds.). University of Washington, Seattle, Seattle (WA).

Booth, C., Inusa, B., & Obaro, S. K. (2010). Infection in sickle cell disease: A review. *International Journal of Infectious Diseases*, *14*(1), e2–e12. https://doi.org/10.1016/j.ijid.2009.03.010

Borges, E., Tchonhi, C., Couto, C. S. B., Gomes, V., Amorim, A., Prata, M. J., & Brito, M. (2019). Unusual β-Globin Haplotype Distribution in Newborns from Bengo, Angola. *Hemoglobin*, *43*(3), 149–154. https://doi.org/10.1080/03630269.2019.1647230

Borhade, M. B., & Kondamudi, N. P. (2021). *Sickle Cell Crisis*. StatPearls Publishing, Treasure Island (FL).

Brandow, A. M., & Liem, R. I. (2022). Advances in the diagnosis and treatment of sickle cell disease. *Journal of Hematology & Oncology*, *15*(1), 20. https://doi.org/10.1186/s13045-022-01237-z

Chamouine, A., Saandi, T., Muszlak, M., Larmaraud, J., Lambrecht, L., Poisson, J., Balicchi, J., Pissard, S., & Elenga, N. (2020). High fetal hemoglobin level is associated with increased risk of cerebral vasculopathy in children with sickle cell disease in Mayotte. *BMC Pediatrics*, *20*(1), 302. https://doi.org/10.1186/s12887-020-02187-6

Cunha, J. P., & Davis, C. P. (2021). *Ischemia Symptoms, Causes, Treatment, and Prognosis*.

Damanhouri, G. A., Jarullah, J., Marouf, S., Hindawi, S. I., Mushtaq, G., & Kamal, M. A. (2015). Clinical biomarkers in sickle cell disease. *Saudi Journal of Biological Sciences*, *22*(1), 24–31. https://doi.org/10.1016/j.sjbs.2014.09.005

DeBaun, M. R., Jordan, L. C., King, A. A., Schatz, J., Vichinsky, E., Fox, C. K., McKinstry, R. C., Telfer, P., Kraut, M. A., Daraz, L., Kirkham, F. J., & Murad, M. H. (2020). American Society of Hematology 2020 guidelines for sickle cell disease: prevention,

diagnosis, and treatment of cerebrovascular disease in children and adults. *Blood Advances*, *4*(8), 1554–1588. https://doi.org/10.1182/bloodadvances.2019001142

Delgadinho, M., Ginete, C., Santos, B., Miranda, A., & Brito, M. (2021). Genotypic Diversity among Angolan Children with Sickle Cell Anemia. *International Journal of Environmental Research and Public Health*, *18*(10), 5417. https://doi.org/10.3390/ijerph18105417

Demirci, S., Leonard, A., Essawi, K., & Tisdale, J. F. (2021). CRISPR-Cas9 to induce fetal hemoglobin for the treatment of sickle cell disease. *Molecular Therapy - Methods & Clinical Development*, *23*, 276–285. https://doi.org/10.1016/j.omtm.2021.09.010

Demirci, S., Uchida, N., & Tisdale, J. F. (2018). Gene therapy for sickle cell disease: An update. *Cytotherapy*, *20*(7), 899–910. https://doi.org/10.1016/j.jcyt.2018.04.003

Doerfler, P. A., Sharma, A., Porter, J. S., Zheng, Y., Tisdale, J. F., & Weiss, M. J. (2021). Genetic therapies for the first molecular disease. *Journal of Clinical Investigation*, *131*(8). https://doi.org/10.1172/JCI146394

Domingos, I. F., Falcão, D. A., Hatzlhofer, B. L., Cunha, A. F., Santos, M. N., Albuquerque, D. M., Fertrin, K. Y., Costa, F. F., Azevedo, R. C., Machado, C. G., Araújo, A. S., Lucena-Araujo, A. R., & Bezerra, M. A. (2014). Influence of the βs haplotype and α-thalassemia on stroke development in a Brazilian population with sickle cell anaemia. *Annals of Hematology*, *93*(7), 1123–1129. https://doi.org/10.1007/s00277-014-2016-1

El-Hazmi, M. A. F., & Warsy, A. S. (1993). On the Molecular Interactions Between -thalassaemia and Sickle Cell Gene. *Journal of Tropical Pediatrics*, *39*(4), 209–213. https://doi.org/10.1093/tropej/39.4.209

Esoh, K., & Wonkam, A. (2021). Evolutionary history of sickle-cell mutation: implications for global genetic medicine. *Human Molecular Genetics*, *30*(R1), R119–R128. https://doi.org/10.1093/hmg/ddab004

Farashi, S., & Harteveld, C. L. (2018). Molecular basis of α-thalassemia. *Blood Cells, Molecules, and Diseases*, *70*, 43–53. https://doi.org/10.1016/j.bcmd.2017.09.004

Ferrão, J., Silva, M., Gonçalves, L., Gomes, S., Loureiro, P., Coelho, A., Miranda, A., Seuanes, F., Reis, A. B., Pina, F., Maia, R., Kjöllerström, P., Monteiro, E., Lacerda, J. F., Lavinha, J., Gonçalves, J., & Faustino, P. (2017). Widening the spectrum of deletions and molecular mechanisms underlying alpha-thalassemia. *Annals of Hematology*, *96*(11), 1921–1929. https://doi.org/10.1007/s00277-017-3090-y

Frangoul, H., Altshuler, D., Cappellini, M. D., Chen, Y.-S., Domm, J., Eustace, B. K., Foell, J., de la Fuente, J., Grupp, S., Handgretinger, R., Ho, T. W., Kattamis, A., Kernytsky, A., Lekstrom-Himes, J., Li, A. M., Locatelli, F., Mapara, M. Y., de Montalembert, M., Rondelli, D., … Corbacioglu, S. (2021). CRISPR-Cas9 Gene Editing for Sickle Cell Disease and β-Thalassemia. *New England Journal of Medicine*, *384*(3), 252–260. https://doi.org/10.1056/NEJMoa2031054

Goodwin, E. F., Partain, P. I., Lebensburger, J. D., Fineberg, N. S., & Howard, T. H. (2017). Elective cholecystectomy reduces morbidity of cholelithiasis in pediatric sickle cell disease. *Pediatric Blood & Cancer*, *64*(1), 113–120. https://doi.org/10.1002/pbc.26179

Gueye Tall, F., Martin, C., Ndour, E. H. M., Renoux, C., Ly, I. D., Connes, P., Gueye, P. M., Diallo, R. N., Diagne, I., Diop, P. A., Cissé, A., Lopez Sall, P., & Joly, P. (2019). Combined and differential effects of alpha-thalassemia and HbF-quantitative trait loci in Senegalese hydroxyurea-free children with sickle cell anemia. *Pediatric Blood & Cancer*. https://doi.org/10.1002/pbc.27934

Gunter, C. (2022, September). *Single Nucleotide Polymorphisms (SNPS)*.

Gusmão, S. (2015). *Cerebral Vasculopathy in Children with Sickle Cell Disease: A study of genetic modulators of the disease*. Faculdade de Ciências e Tecnologia.

Guvenc, B., Canataroglu, A., Unsal, C., Yildiz, S. M., Turhan, F. T., Bozdogan, S. T., Dincer, S., & Erkman, H. (2012). β-Globin chain abnormalities with coexisting α-thalassemia mutations. *Archives of Medical Science*, *4*, 644–649. https://doi.org/10.5114/aoms.2012.28723

Hariharan, P., & Nadkarni, A. (2021). Insight of fetal to adult hemoglobin switch: Genetic modulators and therapeutic targets. *Blood Reviews*, *49*, 100823. https://doi.org/10.1016/j.blre.2021.100823

Herity, L. B., Vaughan, D. M., Rodriguez, L. R., & Lowe, D. K. (2021). Voxelotor: A Novel Treatment for Sickle Cell Disease. *Annals of Pharmacotherapy*, *55*(2), 240–245. https://doi.org/10.1177/1060028020943059

Higgs, D. R., Garrick, D., Anguita, E., Gobbi, M., Hughes, J., Muers, M., Vernimmen, D., Lower, K., Law, M., Argentaro, A., Deville, M. A., & Gibbons, R. (2005). Understanding α-Globin Gene Regulation: Aiming to Improve the Management of Thalassemia. *Annals of the New York Academy of Sciences*, *1054*(1), 92–102. https://doi.org/10.1196/annals.1345.012

Huehns, E. R., & Shooter, E. M. (1965). Human Haemoglobins. *Journal of Medical Genetics*, *2*(1), 48–90. https://doi.org/10.1136/jmg.2.1.48

INDEL Mutation. (2022). MeSH Database.

Joly, P., Pondarré, C., Bardel, C., Francina, A., & Martin, C. (2012). The alpha-globin genotype does not influence sickle cell disease severity in a retrospective cross-validation study of the pediatric severity score. *European Journal of Haematology*, *88*(1), 61–67. https://doi.org/10.1111/j.1600-0609.2011.01705.x

Kato, G. J., Gladwin, M. T., & Steinberg, M. H. (2007). Deconstructing sickle cell disease: Reappraisal of the role of hemolysis in the development of clinical subphenotypes. *Blood Reviews*, *21*(1), 37–47. https://doi.org/10.1016/j.blre.2006.07.001

Kato, G. J., Piel, F. B., Reid, C. D., Gaston, M. H., Ohene-Frempong, K., Krishnamurti, L., Smith, W. R., Panepinto, J. A., Weatherall, D. J., Costa, F. F., & Vichinsky, E. P. (2018). Sickle cell disease. *Nature Reviews Disease Primers*, *4*(1), 18010. https://doi.org/10.1038/nrdp.2018.10

Kato, G. J., Steinberg, M. H., & Gladwin, M. T. (2017). Intravascular hemolysis and the pathophysiology of sickle cell disease. *Journal of Clinical Investigation*, *127*(3), 750–760. https://doi.org/10.1172/JCI89741

Kaufman, D. P., Khattar, J., & Lappin, S. L. (2022). *Physiology, Fetal Hemoglobin*.

Khandros, E., & Blobel, G. A. (2021). Heterogeneity of fetal hemoglobin production in adult red blood cells. *Current Opinion in Hematology*, *28*(3), 164–170. https://doi.org/10.1097/MOH.0000000000000640

Kim, J., & Dean, A. (2021). A Zinc Finger Transcription Factor Faithfully Dedicated to Only a Single Target Gene in Erythroid Cells. *Molecular Cell*, *81*(2), 218–219. https://doi.org/10.1016/j.molcel.2020.12.023

Kohne, E. (2011). Hemoglobinopathies. *Deutsches Ärzteblatt International*. https://doi.org/10.3238/arztebl.2011.0532

Lan, X., Ren, R., Feng, R., Ly, L. C., Lan, Y., Zhang, Z., Aboreden, N., Qin, K., Horton, J. R., Grevet, J. D., Mayuranathan, T., Abdulmalik, O., Keller, C. A., Giardine, B., Hardison, R. C., Crossley, M., Weiss, M. J., Cheng, X., Shi, J., & Blobel, G. A. (2021). ZNF410 Uniquely Activates the NuRD Component CHD4 to Silence Fetal Hemoglobin Expression. *Molecular Cell*, *81*(2), 239-254.e8. https://doi.org/10.1016/j.molcel.2020.11.006

Lin, M., Wen, Y.-F., Wu, J.-R., Wang, Q., Zheng, L., Liu, G.-R., Huang, Y., Yang, H., Lin, F., Zhan, X.-F., Lin, C.-P., Yang, H.-T., Weng, Q.-Q., Huang, F.-T., Wang, Y., Yao, M.-Q., Chen, H.-Z., Wu, D.-H., Zeng, J.-B., … Yang, L.-Y. (2013). Hemoglobinopathy: Molecular Epidemiological Characteristics and Health Effects on Hakka People in the Meizhou Region, Southern China. *PLoS ONE*, *8*(2), e55024. https://doi.org/10.1371/journal.pone.0055024

Liu, N., Xu, S., Yao, Q., Zhu, Q., Kai, Y., Hsu, J. Y., Sakon, P., Pinello, L., Yuan, G.-C., Bauer, D. E., & Orkin, S. H. (2021). Transcription factor competition at the γ-globin promoters controls hemoglobin switching. *Nature Genetics*, *53*(4), 511–520. https://doi.org/10.1038/s41588-021-00798-y

Mangla, A., Ehsan, M., Agarwal, N., Maruvada, S., & Doerr, C. (2022). *Sickle Cell Anemia (Nursing)*.

Manning, J. M., Manning, L. R., Dumoulin, A., Padovan, J. C., & Chait, B. (2020). *Embryonic and Fetal Human Hemoglobins: Structures, Oxygen Binding, and Physiological Roles* (pp. 275–296). https://doi.org/10.1007/978-3-030-41769-7_11

Martyn, G. E., Wienert, B., Yang, L., Shah, M., Norton, L. J., Burdach, J., Kurita, R., Nakamura, Y., Pearson, R. C. M., Funnell, A. P. W., Quinlan, K. G. R., & Crossley, M. (2018). Natural regulatory mutations elevate the fetal globin gene via disruption of BCL11A or ZBTB7A binding. *Nature Genetics*, *50*(4), 498–503. https://doi.org/10.1038/s41588-018-0085-0

Masuda, T., Wang, X., Maeda, M., Canver, M. C., Sher, F., Funnell, A. P. W., Fisher, C., Suciu, M., Martyn, G. E., Norton, L. J., Zhu, C., Kurita, R., Nakamura, Y., Xu, J., Higgs, D. R., Crossley, M., Bauer, D. E., Orkin, S. H., Kharchenko, P. v., & Maeda, T. (2016). Transcription factors LRF and BCL11A independently repress expression of fetal hemoglobin. *Science*, *351*(6270), 285–289. https://doi.org/10.1126/science.aad3312

McGann, P. T., Williams, T. N., Olupot-Olupot, P., Tomlinson, G. A., Lane, A., Luís Reis da Fonseca, J., Kitenge, R., Mochamah, G., Wabwire, H., Stuber, S., Howard, T. A., McElhinney, K., Aygun, B., Latham, T., Santos, B., Tshilolo, L., & Ware, R. E. (2018). Realizing effectiveness across continents with hydroxyurea: Enrollment and baseline characteristics of the multicenter REACH study in Sub-Saharan Africa. *American Journal of Hematology*, *93*(4), 537–545. https://doi.org/10.1002/ajh.25034

Menzel, S., & Thein, S. L. (2019). Genetic Modifiers of Fetal Haemoglobin in Sickle Cell Disease. *Molecular Diagnosis & Therapy*, *23*(2), 235–244. https://doi.org/10.1007/s40291-018-0370-8

Mikobi, T. M., Lukusa, P. T., Aloni, M. N., Lumaka, A., Akilimali, P. Z., Devriendt, K., Matthijs, G., Mbuyi Muamba, J.-M., & Race, V. (2018). Association between sickle cell anemia and alpha thalassemia reveals a high prevalence of the $\alpha^{3.7}$ triplication in congolese patients than in worldwide series. *Journal of Clinical Laboratory Analysis*, *32*(1), e22186. https://doi.org/10.1002/jcla.22186

Molokie, R., DeSimone, J., & Lavelle, D. (2021). Epigenetic regulation of hemoglobin switching in non-human primates. *Seminars in Hematology*, *58*(1), 10–14. https://doi.org/10.1053/j.seminhematol.2020.12.001

Moreira, J. A., Machado, R. P. G., Laurentino, M. R., Lemes, R. P. G., Barbosa, M. C., Santos, T. E., Bandeira, I. C. J., & Martins, A. M. C. (2016). Influence of β S-Globin Haplotypes and Hydroxyurea on Arginase I Levels in Sickle Cell Disease. *Disease Markers*, *2016*, 1–5. https://doi.org/10.1155/2016/9172726

Nader, E., Romana, M., & Connes, P. (2020). The Red Blood Cell—Inflammation Vicious Circle in Sickle Cell Disease. *Frontiers in Immunology*, *11*. https://doi.org/10.3389/fimmu.2020.00454

Nandakumar, S. K., Ulirsch, J. C., & Sankaran, V. G. (2016). Advances in understanding erythropoiesis: evolving perspectives. *British Journal of Haematology*, *173*(2), 206–218. https://doi.org/10.1111/bjh.13938

Nkya, S., Mwita, L., Mgaya, J., Kumburu, H., van Zwetselaar, M., Menzel, S., Mazandu, G. K., Sangeda, R., Chimusa, E., & Makani, J. (2020). Identifying genetic variants and pathways associated with extreme levels of fetal hemoglobin in sickle cell disease in Tanzania. *BMC Medical Genetics*, *21*(1), 125. https://doi.org/10.1186/s12881-020-01059-1

Ogu, U. O., Badamosi, N. U., Camacho, P. E., Freire, A. X., & Adams-Graves, P. (2021). Management of Sickle Cell Disease Complications Beyond Acute Chest Syndrome. *Journal of Blood Medicine*, *Volume 12*, 101–114. https://doi.org/10.2147/JBM.S291394

Onimoe, G., & Rotz, S. (2020). Sickle cell disease: A primary care update. *Cleveland Clinic Journal of Medicine*, *87*(1), 19–27. https://doi.org/10.3949/ccjm.87a.18051

Orkin, S. H., & Bauer, D. E. (2019). Emerging Genetic Therapy for Sickle Cell Disease. *Annual Review of Medicine*, *70*(1), 257–271. https://doi.org/10.1146/annurev-med-041817-125507

Pace, B. S., Starlard-Davenport, A., & Kutlar, A. (2021). Sickle cell disease: progress towards combination drug therapy. *British Journal of Haematology*, *194*(2), 240–251. https://doi.org/10.1111/bjh.17312

Pandey, S. K., Pandey, S., Ranjan, R., Shah, V., Mishra, R. M., Sharma, M., & Saxena, R. (2014). Phenotypic Effect of α-Globin Gene Numbers on Indian Sickle β-Thalassemia Patients. *Journal of Clinical Laboratory Analysis*, *28*(2), 110–113. https://doi.org/10.1002/jcla.21652

Pandey, S., Pandey, S., Mishra, R. M., Sharma, M., & Saxena, R. (2011). Genotypic influence of α-deletions on the phenotype of Indian sickle cell anemia patients. *The Korean Journal of Hematology*, *46*(3), 192. https://doi.org/10.5045/kjh.2011.46.3.192

Park, S. H., & Bao, G. (2021). CRISPR/Cas9 gene editing for curing sickle cell disease. *Transfusion and Apheresis Science*, *60*(1), 103060. https://doi.org/10.1016/j.transci.2021.103060

Piel, F. B., Steinberg, M. H., & Rees, D. C. (2017). Sickle Cell Disease. *New England Journal of Medicine*, *376*(16), 1561–1573. https://doi.org/10.1056/NEJMra1510865

Polymerase Chain Reaction (PCR). (2022). NCBI.

Pule, G. D., Chimusa, E. R., Mnika, K., Mhandire, K., Kampira, E., Dandara, C., & Wonkam, A. (2017). Beta-globin gene haplotypes and selected Malaria-associated variants among black Southern African populations. *Global Health, Epidemiology and Genomics*, *2*, e17. https://doi.org/10.1017/gheg.2017.14

Razin, S. v., Ioudinkova, E. S., Kantidze, O. L., & Iarovaia, O. v. (2021). Co-Regulated Genes and Gene Clusters. *Genes*, *12*(6), 907. https://doi.org/10.3390/genes12060907

Rumaney, M. B., Ngo Bitoungui, V. J., Vorster, A. A., Ramesar, R., Kengne, A. P., Ngogang, J., & Wonkam, A. (2014). The Co-Inheritance of Alpha-Thalassemia and Sickle Cell Anemia Is Associated with Better Hematological Indices and Lower Consultations Rate in Cameroonian Patients and Could Improve Their Survival. *PLoS ONE*, *9*(6), e100516. https://doi.org/10.1371/journal.pone.0100516

Santos, B., Delgadinho, M., Ferreira, J., Germano, I., Miranda, A., Arez, A. P., Faustino, P., & Brito, M. (2020). Co-Inheritance of alpha-thalassemia and sickle cell disease in a cohort of Angolan pediatric patients. *Molecular Biology Reports*, *47*(7), 5397–5402. https://doi.org/10.1007/s11033-020-05628-8

Schnog, J. B., Duits, A. J., Muskiet, F. A. J., Cate, H. ten, Rojer, R. A., & Brandjes, D. P. M. (2004). Sickle cell disease; a general overview. *The Netherlands Journal of Medicine*, *62*(10), 364–374.

Shaikho, E. M., Farrell, J. J., Alsultan, A., Qutub, H., Al-Ali, A. K., Figueiredo, M. S., Chui, D. H. K., Farrer, L. A., Murphy, G. J., Mostoslavsky, G., Sebastiani, P., & Steinberg, M. H. (2017). A phased SNP-based classification of sickle cell anemia HBB haplotypes. *BMC Genomics*, *18*(1), 608. https://doi.org/10.1186/s12864-017-4013-y

Shaikho, E. M., Habara, A. H., Alsultan, A., Al-Rubaish, A. M., Al-Muhanna, F., Naserullah, Z., Alsuliman, A., Qutub, H. O., Patra, P. K., Sebastiani, P., Baltrusaitis, K., Farrell, J. J.,

Jiang, Z., Luo, H., Chui, D. H. K., Al-Ali, A. K., & Steinberg, M. H. (2016). Variants of ZBTB7A (LRF) and its β-globin gene cluster binding motifs in sickle cell anemia. *Blood Cells, Molecules, and Diseases*, *59*, 49–51. https://doi.org/10.1016/j.bcmd.2016.04.001

Shriner, D., & Rotimi, C. N. (2018). Whole-Genome-Sequence-Based Haplotypes Reveal Single Origin of the Sickle Allele during the Holocene Wet Phase. *The American Journal of Human Genetics*, *102*(4), 547–556. https://doi.org/10.1016/j.ajhg.2018.02.003

Steinberg, M. H. (2009). Genetic Etiologies for Phenotypic Diversity in Sickle Cell Anemia. *The Scientific World JOURNAL*, *9*, 46–67. https://doi.org/10.1100/tsw.2009.10

Steinberg, M. H. (2020). Fetal Hemoglobin in Sickle Hemoglobinopathies: High HbF Genotypes and Phenotypes. *Journal of Clinical Medicine*, *9*(11), 3782. https://doi.org/10.3390/jcm9113782

Stevens, D. L., Hix, M., & Gildon, B. L. (2021). Crizanlizumab for the Prevention of Vaso-Occlusive Pain Crises in Sickle Cell Disease. *Journal of Pharmacy Technology*, *37*(4), 209–215. https://doi.org/10.1177/87551225211008460

Sundd, P., Gladwin, M. T., & Novelli, E. M. (2019). Pathophysiology of Sickle Cell Disease. *Annual Review of Pathology: Mechanisms of Disease*, *14*(1), 263–292. https://doi.org/10.1146/annurev-pathmechdis-012418-012838

Thein, S. L. (2013). Genetic association studies in β-hemoglobinopathies. *Hematology*, *2013*(1), 354–361. https://doi.org/10.1182/asheducation-2013.1.354

Thein, S. L. (2017). *Genetic Basis and Genetic Modifiers of β-Thalassemia and Sickle Cell Disease* (pp. 27–57). https://doi.org/10.1007/978-1-4939-7299-9_2

Tisdale, J. F., Thein, S. L., & Eaton, W. A. (2020). Treating sickle cell anemia. *Science*, *367*(6483), 1198–1199. https://doi.org/10.1126/science.aba3827

Tumburu, L., & Thein, S. L. (2021). Targeting ZNF410 as a potential β-hemoglobinopathy therapy. *Nature Genetics*, *53*(5), 589–590. https://doi.org/10.1038/s41588-021-00817-y

Urio, F., Nkya, S., Rooks, H., Mgaya, J. A., Masamu, U., Zozimus Sangeda, R., Mmbando, B. P., Brumat, M., Mselle, T., Menzel, S., Luzzatto, L., & Makani, J. (2020). F cell numbers are associated with an X-linked genetic polymorphism and correlate with haematological parameters in patients with sickle cell disease. *British Journal of Haematology*, *191*(5), 888–896. https://doi.org/10.1111/bjh.17102

Vinjamur, D. S., Yao, Q., Cole, M. A., McGuckin, C., Ren, C., Zeng, J., Hossain, M., Luk, K., Wolfe, S. A., Pinello, L., & Bauer, D. E. (2021). ZNF410 represses fetal globin by singular control of CHD4. *Nature Genetics*, *53*(5), 719–728. https://doi.org/10.1038/s41588-021-00843-w

Weatherall, D. J., & Clegg, J. B. (1976). Molecular Genetics Of Human Hemoglobin. *Annual Review of Genetics*, *10*(1), 157–178. https://doi.org/10.1146/annurev.ge.10.120176.001105

Williams, T. N. (2016). Sickle Cell Disease in Sub-Saharan Africa. *Hematology/Oncology Clinics of North America*, *30*(2), 343–358. https://doi.org/10.1016/j.hoc.2015.11.005

Wonkam, A., Rumaney, M. B., Ngo Bitoungui, V. J., Vorster, A. A., Ramesar, R., & Ngogang, J. (2014). Coinheritance of sickle cell anemia and α-thalassemia delays disease onset and could improve survival in cameroonian's patients (Sub-Saharan Africa). *American Journal of Hematology*, *89*(6), 664–665. https://doi.org/10.1002/ajh.23711

SUPPLEMENTAL MATERIAL

VII. Supplemental Material

Table VII.1. PCR conditions for the amplification of the HBB fragment for detection of the sickle cell mutation.

Gene	Genetic variant	Primers				Reaction mixture		PCR Conditions			Fragment size (bp)
		Primer ID	Sequence	Length (nt)	Tm (ºC)	Reagent	Volume (µL)	T (ºC)	Δt		
HBB **(Chr. 11)**	GAG>GTG Mutation at codon 6 (*HBB*:c.20A>T)	1 (Fw) 67 (Rev)	5' – ACCTCACCCTGTGGAGCCAC – 3' 5' – ACCAGCAGCCTAAGGGTGGGAAAATACACC – 3'	20 30	69.2 74	Platinum™ II Hot-Start PCR Master Mix (2X) Primer 1 Primer 67 H_2O DNA (50ng/ µL)	12.5 0.5 0.5 10.5 1	94 94 60 68 4	2' 15'' 15'' 15'' ∞	x30	375

Table VII.2. Reaction mixture for enzymatic restriction for detection of rs334. Electrophoresis conditions for visualization of results.

Gene	Genetic variant	Reaction mixture (per sample)		Reaction Conditions		Fragment size (bp)		Electrophoresis		
		Reagents	Volume (µL)	T (ºC)	Δt	(-) T (mutation)	(+) A (WT)	Reagents	Volume	Conditions
HBB **(Chr. 11)**	GAG>GTG Mutation at codon 6 (*HBB*:c.20A>T)	CutSmart® Buffer	2	37	60'	375	199 176	2% Agarose gel with TBE (1X)	100mL	100V – 35'
		BSU36I enzyme	1					GreenSafe Premium (NZYTech®)	3µL	
		H_2O	2	Recognition Sequence				HyperLadder 100bp (Bioline®)	5µL	
		DNA (PCR product)	15	5' C C↓T N A G G 3' 3' G G A N T↑C C 5'				DNA Loading Buffer (ThermoFisher®) + PCR Product	2µL + 8µL	

Table VII.3. Reaction mixtures for Gap-PCR for 3.7kb α-thalassaemia deletion detection.

Gene	Genetic variant	Primers				Reaction mixture		PCR Conditions			Fragment size (bp)
		Primer ID	Sequence	Length (nt)	Tm (ºC)	Reagent	Volume (µL)	T (ºC)	Δt		
***HBA* (Chr. 16)**	–α3.7kb deletion	A	5' - GGA CTC CCC TGC GGT CCA GG - 3'	20	62	Platinum™ II Hot-Start PCR Master Mix (2X)	12.5	94	5'		998bp (α1) 1.9kb (α2α1) 2.1kb (α2)
		B	5' - CTC CAT TGT TGG CAC ATT CCG GG - 3'	23	58.8	GC Enhancer	5	94	1'	x30	
		A'	5' - GGG ATG CAC CCA CTG GCA CT - 3'	20	57.9	MgCl2	2	64	1'		
		C	5' - CTG CTG TCC ACG CCC ATG CC - 3'	20	60.0	10X PCR	1	72	1'		
						Primer A/A'	1.25	72	10'		
						Primer B/C	1.25				
						H2O	1				
						DNA (~50ng)	1				

Table VII.4. Electrophoresis conditions for visualization of the results of 3.7kb α-thalassaemia deletion detection.

Gene	Genetic variant	Fragment size (bp)	Electrophoresis		
			Reagents	Volume	Conditions
***HBA* (Chr. 16)**	–α3.7kb deletion	998bp (α1) 1.9kb (α2α1) 2.1kb (α2)	1% Agarose gel with TAE (1X)	100mL	80V – 90'
			GreenSafe Premium (NZYTech®)	3µL	
			GeneRuler 1kb DNA Ladder (ThermoFisher®)	5µL	
			DNA Loading Buffer (ThermoFisher®) + PCR Product	2µL + 8µL	

Table VII.5. PCR conditions for the amplification of fragments with the polymorphisms of HBB haplotypes.

Gene	Genetic variant	Primers				Reaction mixture		PCR Conditions			Fragment size (bp)
		Primer ID	Sequence	Length (nt)	Tm (ºC)	Reagent	Volume (µL)	T (ºC)	Δt		
HBBP1 **(Chr. 11)**	rs10128556 rs968857		20X Assay Mix rs101 20X Assay Mix rs96			Platinum™ II Hot-Start PCR Master Mix (2X) 20X Assay Mix (101/96) H2O DNA (~50ng)	10 1 8 1	95 95 60	10’ 15’’ 1’	x45	
HBE1 **(Chr. 11)**	rs3834466	Fw Rev	5’ – AGT CAT TGG TCA AGG CTG ACC – 3’ 5’ – TCT CTG TTT GAT GAC AAA TTC – 3’	21 21	65.4 57.4	Platinum™ II Hot-Start PCR Master Mix (2X) Primer Fw Primer Rev H2O DNA (~50ng)	12.5 0.5 0.5 9.5 2	94 94 57 68	2’ 15’’ 15’’ 15’’	x30	429
HBG1 **(Chr. 11)**	rs28440105	Fw Rev	5’ – GCT CTG AAT CAT GGG CAG TG – 3’ 5’ – GTG TGT CAG CGT GTG TTT CT – 3’	20 20	64.1 63.9	Platinum™ II Hot-Start PCR Master Mix (2X) Primer Fw Primer Rev H2O DNA (~50ng)	12.5 0.5 0.5 9.5 2	94 94 57 68	2’ 15’’ 15’’ 15’’	x30	760

Table VII.6. Electrophoresis conditions for visualization of the results of rs3834466 and rs28440105.

Gene	Genetic variant	Reaction mixture (per sample)		Reaction Conditions		Fragment size (bp)		Electrophoresis		
		Reagents	Volume (µL)	T (ºC)	Δt	(-) GT (mutation)	(+) G (WT)	Reagents	Volume	Conditions
HBE1 **(Chr. 11)**	rs3834466	NEBuffer™ 2	12.5	37	15’	429	234 195	1% Agarose gel with TBE (1X)	100mL	100V – 45’
		*Hinc*II enzyme	1	80	20’			GreenSafe Premium (NZYTech®)	3µL	
		H_2O	29	Recognition Sequence				HyperLadder 100bp (Bioline®)	5µL	
		DNA (PCR product)	15	5' G T Y↓R A C 3' 3' C A R↑Y T G 5'				DNA Loading Buffer (ThermoFisher®) + PCR Product	2µL + 8µL	
HBG1 **(Chr. 11)**	rs28440105	NEBuffer™ 2	12.5	T (ºC)	Δt	(-) A (mutation)	(+) C (WT)	1% Agarose gel with TAE (1X)	100mL	100V – 45’
		*Hind*III enzyme	1	37	60’	635	327 308	GreenSafe Premium (NZYTech®)	3µL	
				80	20’			HyperLadder 100bp (Bioline®)DNA Loading Buffer (ThermoFisher®) + PCR Product	5µL	
		H_2O	29	Recognition Sequence					2µL + 8µL	
		DNA (PCR product)	15	5' A↓A G C T T 3' 3' T T C G A↑A 5'						

Table VII.7. RT-PCR conditions for the detection of the candidate polymorphisms in the genes of interest.

Gene	Genetic variant	Polymorphism	Reaction mixture		PCR Conditions		
			Reagent	Volume (µL)	T (ºC)	Δt	
	rs10415135	C>T					
ZBTB7A	rs66534382	G>A					
(Chr.19)	rs56356382	T>C					
	rs111929083	A>T	NZYSupreme qPCR Probe Master Mix (2x)	12.5	95	10’	
			20X Assay Mix	1.25	95	15’’	x45
			H2O	10.25	60	15’	
ZNF410	rs144991697	G>A	DNA (~50ng)	1			
(Chr. 14)	rs11844552	A>G					