



INSTITUTO DE HIGIENE E  
MEDICINA TROPICAL  
DESDE 1902

**Universidade Nova de Lisboa**  
**Instituto de Higiene e Medicina Tropical**

Influência de Variantes Genéticas em *CD36*, *ICAM-1* e  
*HBA* na Variabilidade Fenotípica da Drepanocitose e da  
Malária

**Irina Sofia Gouveia Matos**

**DISSERTAÇÃO PARA A OBTENÇÃO DO GRAU DE MESTRE EM CIÊNCIAS BIOMÉDICAS  
ESPECIALIDADE BIOLOGIA MOLECULAR EM SAÚDE TROPICAL E INTERNACIONAL**

(MARÇO, 2024)





INSTITUTO DE HIGIENE E  
MEDICINA TROPICAL  
DESDE 1902

**Universidade Nova de Lisboa**  
**Instituto de Higiene e Medicina Tropical**

Influência de variantes genéticas em *CD36*, *ICAM-1* e  
*HBA* na Variabilidade Fenotípica da Drepanocitose e da  
Malária

**Autor:** Irina Sofia Gouveia Matos

**Orientadora:** Doutora Paula Faustino, Instituto Nacional de Saúde Doutor Ricardo Jorge

**Coorientadora:** Doutora Ana Paula Arez, Instituto de Higiene e Medicina Tropical,  
Universidade NOVA de Lisboa; Doutor Miguel Brito, Instituto Politécnico de Lisboa/Escola  
Superior de Tecnologia da Saúde de Lisboa

Dissertação apresentada para cumprimento dos requisitos necessários à obtenção do grau de  
Mestre em Ciências Biomédicas

Apoio financeiro de *FCT/Aga Khan Development Network, projeto SCAfoldChild,*  
*330842553*



## **Agradecimentos**

Em primeiro lugar, agradecer ao Dr. Fernando de Almeida, presidente do conselho diretivo do Instituto Nacional de Saúde Doutor Ricardo Jorge (INSA) e à Dra. Glória Isidro, coordenadora do Departamento de Genética Humana (DGH) do INSA, por permitirem a realização deste projeto, respetivamente, na instituição e departamento que dirigem, que contribuiu muito para a minha formação tanto profissional como académica.

À minha orientadora Doutora Paula Faustino por me ter dado a oportunidade de desenvolver este projeto sob a sua orientação no Grupo de Investigação em Hemoglobinopatias, Metabolismo do Ferro e Patologias Associadas, da Unidade de Investigação e Desenvolvimento do DGH/INSA. O meu enorme obrigada por todos os ensinamentos, paciência, disponibilidade e partilha.

À minha coorientadora Inv<sup>a</sup> Ana Paula Arez por me ter possibilitado embarcar neste projeto, por me ter ouvido e ensinado e pelo apoio no decorrer de todo o mestrado.

Ao coorientador Prof. Miguel Brito do Instituto Politécnico de Lisboa/Escola Superior de Tecnologia da Saúde de Lisboa, pela ajuda nos dados necessários à realização deste projeto.

Aos profissionais de saúde que, previamente a mim, participaram e permitiram o desenvolvimento deste projeto, muito especialmente à médica pediatra Dra. Brígida Santos do Hospital Pediátrico David Bernardino em Luanda, à colega Isabel Germano e ao Pedro Lopes por toda a disponibilidade e acompanhamento diário imprescindível na minha formação pessoal e académica no DGH/INSA.

O meu agradecimento sincero a todos os membros da Unidade de Tecnologia e Inovação (UTI) do DGH/INSA, pelo apoio técnico prestado na metodologia de Sequenciação de Sanger e Análise de Fragmentos.

À Rita, Catarina, Joana; Madalena, Mariana e Fátima, pela amizade e apoio incondicional.

Ao Gonçalo pela paciência e por me mostrar que até os dias mais difíceis se podem viver com leveza.

À minha família pelo apoio e motivação, em especial à minha mãe, avó e irmã por me mostrarem sempre que há uma luz ao fundo do túnel e que quando existe vontade e um sonho existe concretização.



## Resumo

A drepanocitose é uma doença hereditária monogénica, causada pela mutação c.20A>T no gene *HBB*, caracterizada por eritrócitos falciformes, anemia hemolítica crónica e elevada morbidade e mortalidade, sobretudo no continente africano onde tem elevada prevalência.

A malária humana é causada pelo parasita do género *Plasmodium* e é considerada um problema de saúde pública, nomeadamente na região da África Subsariana e, tal como a drepanocitose, encontra-se associada a uma elevada taxa de morbidade e mortalidade infantil.

O objetivo principal deste estudo foi investigar a relação entre a variabilidade fenotípica da drepanocitose e da malária e a variabilidade genética dos genes *CD36*, *ICAM-1* e *HBA*, numa população sujeita a ambas as patologias.

O estudo incidiu sobre 65 crianças com drepanocitose residentes em Angola. Genotiparam-se cinco regiões polimórficas no gene *CD36*, seis no gene *ICAM-1* e uma em *HBA*, recorrendo a técnicas moleculares de PCR, Gap-PCR, sequenciação de Sanger e análise de fragmentos.

Os resultados revelaram uma elevada prevalência da deleção  $\alpha$ -talassémica de 3,7kb (73%) na população drepanocítica estudada. A deleção talassémica revelou-se benéfica para os doentes no que toca à gravidade da anemia hemolítica.

Os estudos de associação demonstraram que o alelo\_C do SNP rs3211891 T>C, no gene *CD36*, se encontra associado negativamente à concentração média de hemoglobina ( $p=0,013$ ) pois os doentes com genótipos contendo o alelo\_C nesse SNP apresentaram uma concentração significativamente mais baixa de hemoglobina que os restantes, portanto um maior grau de anemia.

No gene *ICAM-1*, os genótipos com o alelo\_T no SNP rs5491 A>T demonstraram estar associados a um número de eritrócitos significativamente mais baixo ( $p=0,044$ ) e a níveis de LDH superiores ( $p=0,020$ ), que os genótipos *wild-type*. Concluimos assim que se encontra associado a um agravamento da anemia e da taxa de hemólise. Ainda em *ICAM-1*, os genótipos com o alelo\_A no SNP rs5496 G>A revelaram uma associação com uma concentração de hemoglobina ( $p=0,009$ ) e número de eritrócitos ( $p=0,036$ ) superior, bem como com uma percentagem de reticulócitos ( $p=0,044$ ) inferior aos demais, concluindo-se que melhora os fenótipos da drepanocitose, nomeadamente a anemia e a taxa de hemólise. Em sentido contrário, os genótipos contendo o alelo\_G no SNP rs5498 A>G revelaram associação com uma elevada percentagem de reticulócitos ( $p=0,015$ ).

Em relação à malária, os doentes drepanocíticos com genótipos que contêm o alelo\_G no SNP rs5494 C>T revelaram ter uma probabilidade 5,63 vezes maior de apresentarem malária em comparação com aqueles que têm os genótipos *wild-type*, na população estudada, OR=5,63, IC95% de 1,07-29,73,  $p=0,028$ .

Este estudo permitiu desvendar o papel modulador de algumas variantes genéticas em relação aos fenótipos da drepanocitose e da malária.

**Palavras-chave:** Drepanocitose, Malária, *Plasmodium falciparum*, Modificadores genéticos



## Abstract

Sickle cell anemia is a hereditary disease caused by c.20A>T mutation in the *HBB* gene. It is characterized by sickle-shaped erythrocytes, chronic hemolytic anemia, and high morbidity and mortality rates. The disease is particularly prevalent in the African continent.

Human malaria, caused by the *Plasmodium* parasite, is a public health issue, especially in sub-Saharan Africa. Like sickle cell disease, malaria is associated with high morbidity and infant mortality rates.

The aim of this study was to investigate the relationship between the phenotypic variability of sickle cell disease and malaria and the genetic variability of *CD36*, *ICAM-1*, and *HBA* genes in a population subject to both pathologies.

The study included 65 children with sickle cell disease living in Angola. Five polymorphic regions were genotyped in the *CD36* gene, six in the *ICAM-1* gene, and one in *HBA* using molecular techniques like PCR, Gap-PCR, Sanger sequencing, and fragment analysis.

The results showed a high prevalence of  $\alpha$ -thalassemia deletion of 3.7kb (73%) in the studied population. The thalassemia deletion was shown to be beneficial for patients concerning the severity of hemolytic anemia.

The association studies revealed that the C allele of the rs3211891 T>C SNP in the *CD36* gene is negatively associated with mean hemoglobin concentration ( $p=0.013$ ). Patients with genotypes containing the allele C in that SNP had a significantly lower hemoglobin concentration than the others, indicating a higher degree of anemia.

In the *ICAM-1* gene, genotypes with the allele T in the rs5491 A>T SNP were associated with a significantly lower number of erythrocytes ( $p=0.044$ ) and higher LDH levels ( $p=0.020$ ) than wild-type genotypes. This suggests that it worsens anemia and hemolysis rates. Conversely, genotypes containing the allele A in the rs5496 G>A SNP were associated with a higher hemoglobin concentration ( $p=0.009$ ) and erythrocyte count ( $p=0.036$ ), as well as a percentage of reticulocytes ( $p=0.044$ ) lower than others, indicating an improvement in the phenotypes of sickle cell anemia, namely in anemia level and hemolysis rate. On the other hand, genotypes containing the allele G in the rs5498 A>G SNP were associated with a high percentage of reticulocytes ( $p=0.015$ ).

Regarding malaria, sickle cell anemia patients with genotypes containing the allele G in the rs5494 C>T SNP showed a 5.63-fold higher risk of malaria compared to those with wild-type genotypes (OR=5.63, 95%CI 1.07-29.73,  $p=0.028$ ).

Overall, this study sheds light on the role of some genetic variants in modulating sickle cell disease and malaria phenotypes.

**Keywords:** Sickle cell anemia, Malaria, *Plasmodium falciparum*, Genetic modifiers



# Índice

1	Introdução.....	1
1.1	Malária.....	1
1.2	Espécie <i>Plasmodium falciparum</i> .....	2
1.2.1	Ciclo de Vida.....	2
1.2.2	Fatores de virulência.....	4
1.3	Eritrócito .....	6
1.3.1	Hemoglobina e Hemoglobinopatias.....	6
1.4	Drepanocitose .....	9
1.5	Epidemiologia da malária e drepanocitose em África .....	10
1.6	Modeladores genéticos de variabilidade fenotípica da malária e da drepanocitose.....	12
1.6.1	Os genes da $\alpha$ -globina, a $\alpha$ -talassémia e as suas consequências.....	13
1.6.2	CD36 – O gene, a proteína e a sua função.....	14
1.6.3	ICAM-1 – O gene, a proteína e a sua função.....	17
2	Justificação e Objetivos.....	21
3	Materiais e Métodos .....	23
3.1	Ética e consentimento informado .....	23
3.2	Caraterização da amostra .....	23
3.2.1	Amostra Populacional.....	23
3.2.2	Amostra Biológica .....	24
3.3	Desenho experimental e técnicas aplicadas .....	25
3.3.1	Amplificação <i>in vitro</i> de DNA ( <i>Polymerase Chain Reaction</i> , PCR).....	26
3.3.2	Gap-PCR para o estudo dos genes <i>HBA</i> .....	32
3.3.3	Eletroforese em gel de agarose .....	35
3.3.4	Sequenciação de Sanger.....	36
3.3.5	Análise de microssatélites In3(TG)n no gene <i>CD36</i> .....	38
3.4	Tratamento dos dados <i>in silico</i> .....	40
3.4.1	varSeak.....	40
3.4.2	PolyPhen-2.....	41
3.4.3	SPSS.....	42
3.4.4	Globin Gene Server, Ithantet, Uniprot, Ensembl e Pubmed .....	43
4	Resultados e Discussão.....	45

4.1	Caracterização demográfica, hematológica, bioquímica e clínica da população estudada.....	45
4.2	Co-herança de $\alpha$ -talassémia .....	50
4.3	Pesquisa e caracterização de variantes nos genes <i>CD36</i> e <i>ICAM-1</i> na população drepanocítica .....	54
4.3.1	Variantes genéticas detetadas no gene <i>CD36</i> .....	58
4.3.2	Variantes genéticas detetadas no gene <i>ICAM-1</i> .....	72
4.4	Estudos de Associação .....	76
4.4.1	Variantes genéticas e a variabilidade fenotípica da drepanocitose .....	76
4.4.2	Variantes genéticas detetadas e a variabilidade fenotípica da malária.....	89
5	Conclusões e Perspetivas Futuras.....	93
6	Referências Bibliográficas.....	97
7	Anexos.....	107
	Anexo 1-Documento de consentimento informado para os tutores das crianças participantes .....	107
	Anexo 2-Códigos União Internacional de Química Pura e Aplicada (IUPAC) .....	108
	Anexo 3-Material Suplementar ao capítulo Materiais e Métodos .....	110
	Anexo 4-Valores de referência dos parâmetros hematológicos para a população Portuguesa em idade pediátrica.....	112

## Índice de Figuras

<b>Figura 1.1-</b> Ciclo de vida do parasita <i>Plasmodium falciparum</i> . .....	3
<b>Figura 1.2-A)</b> Esquema da estrutura molecular da família de proteínas. ....	6
<b>Figura 1.3-A)</b> Imagem 3D da estrutura quaternária da hemoglobina humana. ....	7
<b>Figura 1.4-</b> Representação das diferentes cadeias globínicas e sua expressão ao longo das diferentes etapas de desenvolvimento, desde a fase embrionária à adulta e respetivos locais de eritropoiese. ....	8
<b>Figura 1.5-</b> Representação esquemática da alteração falciforme: substituição do aminoácido Ácido Glutâmico por Valina na proteína $\beta$ -globina como consequência da alteração do tripleto GAG por GUG na sequência codificante do gene HBB. ....	9
<b>Figura 1.6-A)</b> Distribuição dos novos casos de malária, causada por <i>P. falciparum</i> , no continente africano no espaço de um ano. B) Distribuição do alelo $\beta$ S no continente africano. ....	11
<b>Figura 1.7-</b> Esquema representativo da estrutura do gene <i>CD36</i> de acordo com a base de dados NCBI - NG_008192.1. ....	14
<b>Figura 1.8-</b> Sequência aminoacídica (452aa) da proteína CD36. ....	15
<b>Figura 1.9-</b> Resíduo de fenilalanina da proteína CD36 como local conservado de ligação à bolsa hidrofóbica, situada no domínio CIDR $\alpha$ 2 da proteína PfEMP1. ....	16
<b>Figura 1.10-</b> Esquema representativo da estrutura do gene <i>ICAM-1</i> de acordo com a base de dados NCBI- NG_012083.1. ....	17
<b>Figura 1.11-</b> Representação esquemática da molécula de ICAM-1 com todos os seus domínios extracelulares D1 a D5 e cauda citoplasmática. ....	18
<b>Figura 3.1-</b> Equação de Beer-Lambert. ....	25
<b>Figura 3.2-</b> Representação esquemática dos genes <i>HBA2</i> e <i>HBA1</i> e do gene híbrido resultante da deleção talassémica de 3,7kb, bem como dos locais de ligação dos primers (C10, C3 e C2). ....	33
<b>Figura 4.1-</b> Representação gráfica da frequência dos genótipos referentes à deleção $\alpha$ -talassémica de 3,7kb na população drepanocítica analisada (n=63). ....	51
<b>Figura 4.2-</b> Representação gráfica das frequências dos genótipos relativos à deleção de 3,7kb ocorrida nos genes <i>HBA2</i> e <i>HBA1</i> na população drepanocítica angolana analisada no nosso estudo, na população drepanocítica nigeriana visada no artigo de Olatunya et al. (93) e na população drepanocítica residente em Lisboa e com ascendência africana referida no artigo de Silva et al (88). ....	52
<b>Figura 4.3-</b> Resultado parcial da sequenciação de Sanger do décimo exão do gene <i>CD36</i> revelando a alteração c.975T>G em heterozigotia. ....	58

<b>Figura 4.4</b> -Resultado parcial da sequenciação de Sanger do quarto intrão do gene <i>CD36</i> revelando a alteração c.282-31 T>C. ....	60
<b>Figura 4.5</b> -Resultado parcial da análise <i>in silico</i> do impacto no <i>splicing</i> da variante <i>CD36</i> : c.282-31T>C pelo <i>software</i> varSeak. ....	61
<b>Figura 4.6</b> -Eletroforese em gel de agarose Seakem e NuSieve (1:3; 4% p/v), dos produtos de PCR obtidos na genotipagem da variante In3(TG)n em diferentes casos de homozigotia. ....	62
<b>Figura 4.7</b> -Perfil de casos homozigóticos para a variante In3(TG)n após análise de marcadores de DNA. ....	63
<b>Figura 4.8</b> -Genótipos encontrados aquando do estudo da variante In3(TG)n, na população angolana drepanocítica. ....	64
<b>Figura 4.9</b> -Representação gráfica das frequências genótípicas para a variante In3(TG)n na população angolana drepanocítica (estudada neste trabalho). ....	65
<b>Figura 4.10</b> -Representação gráfica das frequências genótípicas dos genótipos TG12/TG12, TG12/TG13 e TG13/TG13 para a variante In3(TG)n nas populações angolana drepanocítica (estudada), africana e ioruba. ....	66
<b>Figura 4.11</b> -Resultado parcial da sequenciação de Sanger da região promotora do gene <i>CD36</i> revelando a alteração c.-184+11225A>G em heterozigotia (A) e a alteração c.-184+11308G>C em heterozigotia (B) e em homozigotia (C). ....	68
<b>Figura 4.12</b> -Representações gráficas das frequências alélicas (rs1413661) nas populações africana (AFR), ioruba (YRI) e angolana drepanocítica (população estudada). ....	69
<b>Figura 4.13</b> -Resultado parcial da sequenciação de Sanger do quinto intrão do gene <i>CD36</i> revelando a alteração c.429+2T>C. ....	70
<b>Figura 4.14</b> -Resultado parcial da análise <i>in silico</i> do impacto no <i>splicing</i> da variante <i>CD36</i> : c.429+2T>C pelo <i>software</i> varSeak. ....	70
<b>Figura 4.15</b> -Representações gráficas das frequências alélicas (rs3211893) nas populações africana (AFR), ioruba (YRI) e angolana drepanocítica (população estudada). ....	71
<b>Figura 4.16</b> -Resultado parcial da sequenciação de Sanger do sexto exão do gene <i>ICAM-1</i> revelando a alteração c.1190G>A em heterozigotia. ....	72
<b>Figura 4.17</b> -Representação gráfica dos resultados obtidos através do estudo <i>in silico</i> (Poly-Phen) para a variante <i>ICAM-1</i> , p.Arg397Gln. ....	73
<b>Figura 4.18</b> -Resultado parcial da sequenciação de Sanger do segundo exão do gene <i>ICAM-1</i> revelando a alteração c.167A>T. ....	74

<b>Figura 4.19</b> -Representações gráficas das frequências alélicas (rs5491) nas populações africana (AFR), ioruba (YRI) e angolana drepanocítica (população estudada). .....	76
<b>Figura 4.20</b> -Representação gráfica da distribuição média de Hb por genótipo em rs3211891 no gene <i>CD36</i> . .....	79
<b>Figura 4.21</b> -Representação gráfica da distribuição média da concentração de Hb por genótipo <i>wild-type versus</i> genótipos contendo a variante rs5496_A. ....	80
<b>Figura 4.22</b> -Representação gráfica da distribuição média de eritrócitos por genótipo <i>wild-type versus</i> genótipos contendo a variante rs5496_A. ....	80
<b>Figura 4.23</b> -Representação gráfica da distribuição média de eritrócitos por genótipo <i>wild-type versus</i> genótipos contendo a variante rs5491_T. ....	81
<b>Figura 4.24</b> -Representação gráfica da distribuição média de eritrócitos por genótipo da deleção talassémica de 3,7kb. ....	82
<b>Figura 4.25</b> -Representações gráficas dos valores médios de VGM (A) e HGM (B) por grupo de genótipo da deleção talassémica de 3,7kb. ....	83
<b>Figura 4.26</b> -Representação gráfica da distribuição média do número de episódios de malária por genótipo <i>wild-type versus</i> genótipos contendo a variante rs5496_A. ....	90



## Índice de Tabelas

<b>Tabela 1.1</b> -Fenótipos e genótipos das possíveis deleções associadas à $\alpha$ -talassémia....	13
<b>Tabela 3.1</b> -Metodologias aplicadas para o estudo das diversas variantes genéticas analisadas .....	26
<b>Tabela 3.2</b> -Mistura de PCR para o estudo da região promotora do gene <i>CD36</i> .....	28
<b>Tabela 3.3</b> -Condições de PCR para o estudo da região promotora do gene <i>CD36</i> .....	28
<b>Tabela 3.4</b> -Mistura de PCR para o estudo do quinto exão do gene <i>CD36</i> .....	28
<b>Tabela 3.5</b> -Condições de PCR para o estudo do quinto exão do gene <i>CD36</i> .....	29
<b>Tabela 3.6</b> -Mistura de PCR para o estudo do décimo exão do gene <i>CD36</i> .....	29
<b>Tabela 3.7</b> -Condições de PCR para o estudo do décimo exão do gene <i>CD36</i> .....	29
<b>Tabela 3.8</b> -Mistura de PCR para o estudo da variante In3(TG)n no gene <i>CD36</i> , utilizando <i>primers</i> marcados com FAM .....	30
<b>Tabela 3.9</b> -Condições de PCR para o estudo da variante In3(TG)n no gene <i>CD36</i> , utilizando <i>primers</i> marcados com FAM .....	30
<b>Tabela 3.10</b> -Mistura de PCR para o estudo da variante In3(TG)n no gene <i>CD36</i> .....	30
<b>Tabela 3.11</b> -Condições de PCR para o estudo da variante In3(TG)n no gene <i>CD36</i> ....	31
<b>Tabela 3.12</b> -Mistura de PCR para o estudo do segundo exão do gene <i>ICAM-1</i> .....	31
<b>Tabela 3.13</b> -Condições de PCR para o estudo do segundo exão do gene <i>ICAM-1</i> .....	32
<b>Tabela 3.14</b> -Mistura de PCR para o estudo do quarto ao sexto exões do gene <i>ICAM-1</i> 32	
<b>Tabela 3.15</b> - Condições de PCR para o estudo do quarto ao sexto exões do gene <i>ICAM-1</i> .....	32
<b>Tabela 3.16</b> -Mistura II de Gap-PCR para o estudo da deleção $\alpha$ -talassêmica 3,7kb que afeta os genes <i>HBA2</i> e <i>HBA1</i> .....	34
<b>Tabela 3.17</b> -Condições de Gap-PCR para o estudo da deleção $\alpha$ -talassêmica de 3,7 kb .....	34
<b>Tabela 3.18</b> -Condições de gel e eletroforese aplicadas para o estudo das variantes genéticas estudadas .....	35
<b>Tabela 3.19</b> -Mistura reacional para a realização da Sequenciação de Sanger.....	37
<b>Tabela 3.20</b> -Condições de PCR para a realização de Sequenciação de Sanger.....	37
<b>Tabela 3.21</b> -Sequências <i>RefSeq</i> utilizadas para comparação das sequências obtidas por sequenciação de Sanger .....	38

<b>Tabela 4.1</b> -Estatística descritiva dos parâmetros demográficos, hematológicos e bioquímicos da população em estudo .....	47
<b>Tabela 4.2</b> -Variantes genéticas detetadas no gene <i>CD36</i> na população drepanocítica em estudo .....	54
<b>Tabela 4.3</b> -Variantes genéticas detetadas no gene <i>ICAM-1</i> na população drepanocítica em estudo .....	55
<b>Tabela 4.4</b> -Variantes detetadas no gene <i>CD36</i> e respetivas frequências genótípicas e alélicas, na população estudada .....	55
<b>Tabela 4.5</b> -Variantes detetadas no gene <i>ICAM-1</i> e respetivas frequências genótípicas e alélicas, na população estudada .....	56
<b>Tabela 4.6</b> -Variantes genéticas cujas frequências genótípicas diferem entre a população drepanocítica estudada (angolana) e a população africana em geral .....	57
<b>Tabela 4.7</b> -Variantes genéticas cujas frequências genótípicas diferem entre a população drepanocítica estudada (angolana) e a população ioruba em geral.....	57
<b>Tabela 4.8</b> -Associações significativas entre a média da idade da primeira manifestação drepanocítica e os diferentes genótipos em rs5491 .....	77
<b>Tabela 4.9</b> -Associações entre a média da concentração de Hb e contagem de eritrócitos e os diferentes genótipos em rs3211891, rs5496, rs5491 e a deleção $\alpha$ -talassémica de 3,7kb .....	78
<b>Tabela 4.10</b> -Associações entre a média de HGM e VGM e os diferentes genótipos em rs3211891, rs5496, rs5491 e a deleção $\alpha$ -talassémica de 3,7kb.....	83
<b>Tabela 4.11</b> -Associações entre os biomarcadores de hemólise (LDH e percentagem de reticulócitos) e os diferentes genótipos em rs5491, rs5498, rs5496 e a deleção $\alpha$ -talassémica de 3,7kb .....	85
<b>Tabela 4.12</b> -Associações entre o número médio de plaquetas e os diferentes genótipos em rs3211938 e rs5497 .....	87
<b>Tabela 4.13</b> -Associações significativas entre a média de episódios de malária e os diferentes genótipos em rs5496 .....	89
<b>Tabela 4.14</b> -Associação dos genótipos em rs5494 e os casos de malária.....	91
<b>Tabela 7.1</b> -Códigos IUPAC para nucleótidos .....	108
<b>Tabela 7.2</b> -Código IUPAC para os diferentes aminoácidos .....	109
<b>Tabela 7.3</b> -Composição dos diferentes reagentes utilizados ao longo do trabalho prático .....	110

<b>Tabela 7.4</b> -Descrição dos <i>primers</i> utilizados na genotipagem das diferentes regiões do gene <i>CD36</i> .....	110
<b>Tabela 7.5</b> -Descrição dos <i>primers</i> utilizados na genotipagem do terceiro intrão do gene <i>CD36</i> , para o estudo do STR In3(TG)n.....	111
<b>Tabela 7.6</b> -Descrição dos <i>primers</i> utilizados na genotipagem das diferentes regiões do gene <i>ICAM-1</i> .....	111
<b>Tabela 7.7</b> -Descrição dos <i>primers</i> utilizados na genotipagem da deleção talassêmica de 3,7kb ocorrida nos genes <i>HBA2</i> e <i>HBA1</i> .....	111



## Lista de Abreviaturas e Acrónimos

AGEZ– <i>Adenine Guanine Exclusion Zone</i>	INSA– Instituto Nacional de Saúde Doutor Ricardo Jorge
AVCs– Acidentes Vasculares Cerebrais	IUPAC- União Internacional de Química Pura e Aplicada
BSA- Albumina Bovina-Sigma	kb- Kilobases
CD36- <i>Cluster of differentiation 36</i>	LDH– Desidrogenase láctica
ddNTPs- didesoxirribonucleótidos	ng- Nanogramas
DNA- Ácido desoxirribonucleico	nm– Nanometros
dNTPs– desoxirribonucleótidos	NMD- <i>Nonsense Mediated Decay</i>
EDTA– Ácido etilenodiamino tetra-acético	OMS- Organização Mundial de Saúde
ESTeS-L– Escola Superior de Tecnologia da Saúde de Lisboa	OR– <i>Odds Ratio</i>
EtBr- Brometo de Etídeo	pb- Par de bases
Exo– Exonuclease I	PCR- <i>Polymerase Chain Reaction</i>
Hb- Hemoglobina	PfEMP1- Proteína de Membrana do <i>Plasmodium falciparum</i> 1
HbA- Hemoglobina Adulta	pmol- Picomole
HBA- Hemoglobina Alfa	RIFINS- <i>Repetitive interspersed family</i>
HBB- Hemoglobina Beta	SAP– <i>Shrimp Alkaline Phosphatase</i>
HbF- Hemoglobina Fetal	SCD- <i>Sickle Cell Disease</i>
HbS- Hemoglobina S	SNP– <i>Single Nucleotide Polymorphism</i>
HGM– Hemoglobina Globular Média	STEVOR– <i>Subtelomeric variable open reading frame</i>
HPDB– Hospital Pediátrico David Bernardino	STR– <i>Short Tandem Repeat</i>
HU- Hidroxiureia	TBE– Tris-Borato-EDTA
HW– Hardy-Weinberg	TSP- Trombospondina
ICAM-1- Molécula de adesão intercelular 1	UTI– Unidade de Tecnologia e Inovação
IHMT– Instituto de Higiene e Medicina Tropical	UV– Ultravioleta
	VGM- Volume Globular Médio
	µL– Microlitro



# 1 Introdução

## 1.1 Malária

A malária é uma parasitose causada pelo parasita do género *Plasmodium*, é considerada uma das doenças infecciosas mais antigas do mundo e uma das parasitoses que maior número de pessoas afeta anualmente. É um problema de saúde pública, nomeadamente na região da África Subsariana onde apresenta maior incidência e prevalência (endemicidade). É também a doença tropical associada a maior mortalidade infantil, de crianças até aos 5 anos de idade, sendo estas um dos grupos de risco mais vulnerável em áreas endémicas para a malária. De acordo com o último relatório da organização mundial de saúde (OMS), em 2022 cerca de metade da população mundial encontrava-se exposta à malária e foram estimados 249 milhões de casos. Nesse mesmo relatório, a OMS relatou ainda que 94% dos casos e 96% das mortes ocorreram em África (1).

A malária humana pode ser causada por seis espécies do género *Plasmodium*, sendo estas *Plasmodium falciparum*, *Plasmodium vivax*, *Plasmodium ovalecurtisi* e *Plasmodium ovalewallikeri*, *Plasmodium malariae* e *Plasmodium knowlesi*, porém é na infeção por *Plasmodium falciparum* que as consequências se configuram mais severas (2–5). O parasita em questão apresenta mecanismos ao nível da infeção e da evasão ao sistema imune do hospedeiro e fatores de virulência, responsáveis pelos fenómenos de sequestro e citoadesão dos eritrócitos infetados. É nesta espécie de *Plasmodium* que se concentra o presente trabalho.

A malária é uma infeção que tem contribuído, ao longo dos anos, para a seleção de variantes polimórficas do hospedeiro Humano. Desta forma, exerce um mecanismo de pressão seletiva das variantes que são vantajosas para o hospedeiro perante a malária, resultando no aumento das suas frequências populacionais (6). Deste modo, fala-se de uma “Hipótese da malária” que tenta explicar esta seleção. A “Hipótese da malária” foi primeiramente proposta por Haldane em 1949, que inferiu sobre a existência de uma pressão seletiva positiva exercida sobre os portadores de talassémia que conferia algum grau de proteção contra a malária em áreas endémicas para esta parasitose (7).

Mais tarde, em 1954, Allison sugeriu que a anemia das células falciformes estaria igualmente associada a proteção contra a malária e sujeita a esta pressão que seleciona

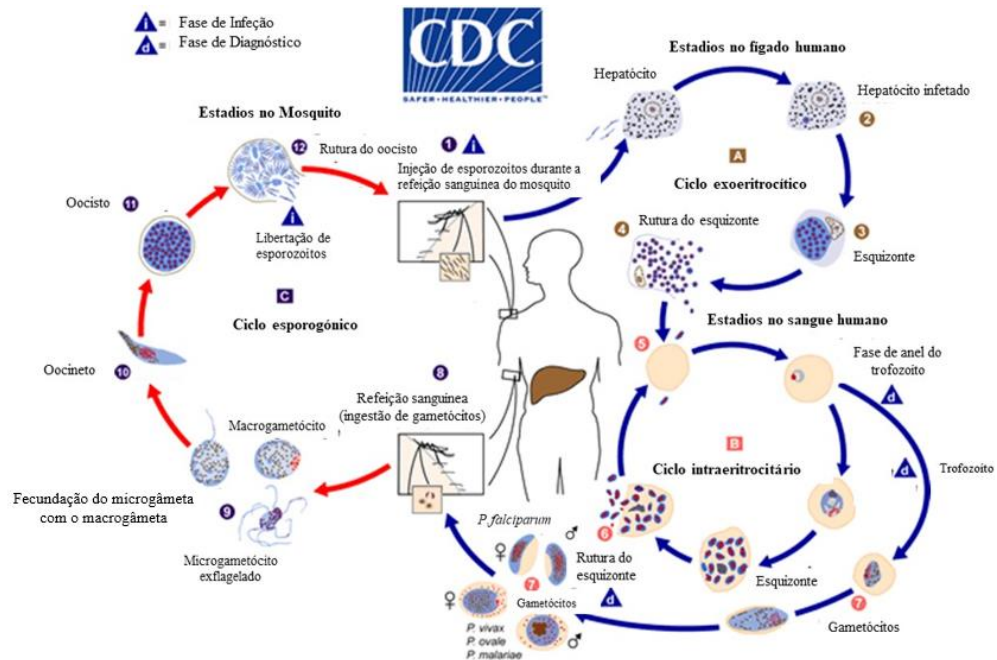
polimorfismos/mutações e variantes genéticas que conferem proteção às populações que nestas áreas habitam (8).

## **1.2 Espécie *Plasmodium falciparum***

O parasita da espécie *Plasmodium falciparum* é um protozoário unicelular pertencente ao filo Apicomplexa. Este filo é caracterizado pela presença de complexo apical importante no momento de invasão celular. É um parasita intracelular, tendo um desenvolvimento e reprodução dependente das vias enzimáticas das células-alvo do hospedeiro. É a espécie de parasita *Plasmodium* associado ao maior número de mortes por malária e ao maior número de casos de malária cerebral. Configura-se como um dos agentes etiológicos mais perigosos, uma vez que apresenta um lato espectro de resistências a antimaláricos, pelo que o seu estudo se torna essencial para uma tentativa de contenção de infeções (9,10).

### **1.2.1 Ciclo de Vida**

No caso da malária humana, o parasita possui um ciclo esporogónico (reprodução sexuada) no mosquito do género *Anopheles* (hospedeiro definitivo) e um ciclo esquizogónico no hospedeiro intermediário, Homem. O ciclo de vida do plasmódio no hospedeiro intermediário compreende duas fases de infeção distintas, uma primeira fase de infeção de hepatócitos e a seguinte de infeção de eritrócitos e encontra-se representado na figura 1.1.



**Figura 1.1-** Ciclo de vida do parasita *Plasmodium falciparum*. 1- Durante uma refeição sanguínea, dá-se a injeção de esporozoítos na corrente sanguínea do Homem (por parte do mosquito-vetor). 2- São infectados os hepatócitos, onde se inicia um ciclo exoeritrocitário (A). 3- Esporozoítos sofrem esquizogonia que dá origem a esquizontes, células multinucleadas que posteriormente se dividem em merozoítos. 4- Dá-se a rutura do esquizonte e libertação de merozoítos para a corrente sanguínea. 5- Os merozoítos vão infectar eritrócitos, dando início ao ciclo intraeritrocitário (B). 6- Esquizogonia (intraeritrocitária) em trofozoito, etapas de diferenciação deste em esquizonte (ou em gametócito) e rutura do esquizonte. 7- Estrutura do gametócito (microgametócito ou macrogametócito). 8- Refeição sanguínea e ingestão de gametócitos (infecção do mosquito), início do ciclo esporogónico (C). 9- Diferenciação dos gametócitos em gâmetas e fecundação do micro com o macrogâmeta. 10- A fecundação dá origem ao oocineto, que se vai alojar na matriz peritrófica do intestino médio do mosquito. 11- Reprodução sexuada, meiose, e origem do oocisto que permite o desenvolvimento de esporozoítos. 12- Rutura do oocisto e novo início de ciclo. Adaptado de Centers for Disease Control and Prevention (<https://www.cdc.gov/parasites/>).

O ciclo inicia-se pela injeção de esporozoítos na corrente sanguínea, no decorrer de uma refeição sanguínea por parte do mosquito fêmea (vetor). Os esporozoítos vão infectar hepatócitos, no fígado, iniciando uma fase hepática e assintomática (Fig.1.1-nº2). Aqui ocorre uma reprodução assexuada, esquizogonia exoeritrocitária, dos esporozoítos em merozoítos que são posteriormente libertados na corrente sanguínea. De seguida, dá-se o início da fase intraeritrocitária (com duração de 48 horas), sintomática, onde ocorre uma adesão do merozoíto à membrana do eritrócito (Fig.1.1-nº5)(11).

A internalização do parasita é feita a partir de uma reorientação da extremidade apical para a zona de contacto com o eritrócito, com recurso a filamentos de actina e miosina

presentes no complexo apical do plasmódio. Após esta reorientação entram em ação as róptrias e micronema do parasita, são libertadas proteases, fosfolipases e outras enzimas que perturbam a estabilidade da membrana do eritrócito e provocam uma disrupção membranar. Esta cascata de acontecimentos permite que se suceda uma invaginação da membrana originando o vacúolo parasitóforo, onde o merozoíto se desenvolve (12,13).

No eritrócito, o desenvolvimento do parasita compreende os estádios de merozoíto, trofozoíto (anel e maturo) e esquizonte, sendo que este último leva ao rebentamento do eritrócito infetado e libertação de merozoítos na corrente sanguínea para infeção de mais células sanguíneas.

O ciclo continua, por um lado com a invasão de mais eritrócitos sãos e de vários picos de parasitémia e, por outro lado, com a diferenciação dos merozoítos em gametócitos (Fig.1.1-nº7). A ingestão destes pela picada de um mosquito, origina a sua infeção e, posteriormente, permite a fecundação do macro com o microgâmeta e formação do oocineto (Fig.1.1-nº8 a 10). O oocineto através de meiose e, posteriormente, de mitoses sucessivas origina o oocisto que alberga os esporozoítos, que migram para as glândulas salivares do mosquito para nova infeção (Fig.1.1- nº11 e 12).

## 1.2.2 Fatores de virulência

A espécie *P. falciparum* apresenta mecanismos que auxiliam a evasão ao sistema imunitário do hospedeiro, através dos fenómenos de citoadesão e *rosetting* (14).

No decorrer da fase sanguínea da infeção por *P. falciparum*, este exporta para o folheto externo da membrana do eritrócito, proteínas que apresentam variação antigénica, como a proteína de membrana 1 (PfEMP1) (ver ponto 1.2.2.1), as proteínas da *repetitive interspersed family* (RIFINS) e as proteínas da região variável subtelomérica, *subtelomeric variable open reading frame* (STEVOR), entre outras. Estas proteínas codificadas por famílias génicas, respetivamente *var*, *rif* e *stevor*, permitem a já mencionada evasão ao sistema imune através da variação antigénica e consequentemente, das proteínas apresentada às células do hospedeiro (15). A PfEMP1 especificamente, encontra-se já cientificamente provado ser capaz de estabelecer ligação com recetores do endotélio vascular humano, sendo associada aos fenómenos de citoadesão e sequestro (importantes para a virulência de *P. falciparum*) (16). Por estas razões, a PfEMP1 será a

proteína sobre a qual a informação será desenvolvida daqui para a frente.

O fenómeno de citoadesão, ou sequestro, consiste na adesão dos eritrócitos infetados ao endotélio vascular. Desta forma, os eritrócitos evitam a sua eliminação pelo fígado e baço. Por outro lado, o *rosetting* traduz-se na capacidade dos eritrócitos infetados se cercarem por outros constituintes sanguíneos, por exemplo eritrócitos sãos e plaquetas. Desta forma, escapam ao reconhecimento pelas células do sistema imunitário como corpos estranhos e evitam a sua sinalização e posterior eliminação (17). Para além disso, esta capacidade de se rodearem por outros eritrócitos não infetados permite um maior rendimento e infeção localizada quando se dá a libertação de merozoítos, com objetivo de nova invasão (16).

### 1.2.2.1 A família PfEMP1

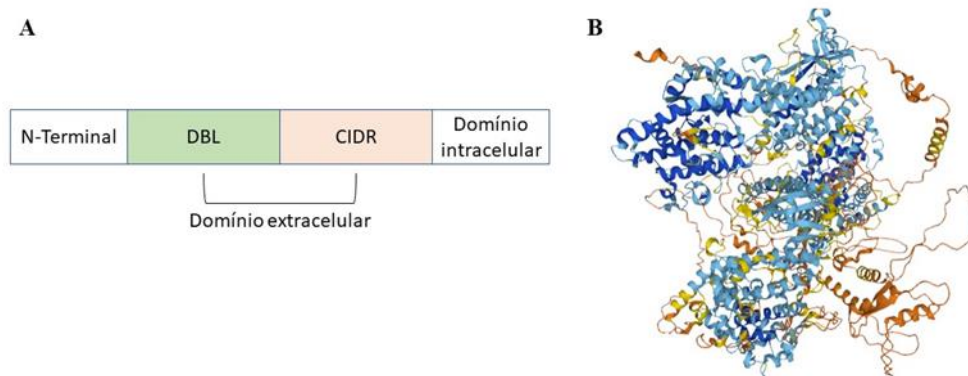
A PfEMP1 é uma família de proteínas codificadas por genes do parasita e que se expressam à superfície do eritrócito do hospedeiro, sendo o mais importante fator de virulência do parasita pois permite o reconhecimento de recetores humanos e o estabelecimento de ligações ao endotélio vascular humano (18,19).

A PfEMP1 tem a sua expressão regulada pelos genes de variação antigénica (*var gene*). Estes são cerca de 60 genes que através de *switching* transcricional alteram os antigénios apresentados às células hospedeiras (11,20). Deste modo, quando o sistema imunitário do hospedeiro adquire imunidade contra os antigénios apresentados ocorre o *switching* do gene expresso e o processo de resposta imunitária inicia-se novamente, levando a picos de parasitémia constantes (21–23).

A proteína em tradução ativa é exportada pelo parasita para a membrana externa do eritrócito infetado, na fase sanguínea da infeção, e acumula-se em estruturas denominadas de *knobs* (14,16,24).

No que diz respeito à sua estrutura molecular, a PfEMP1 é constituída por um domínio transmembranar e um extracelular constituído por dois tipos de subdomínios, os *Duffy binding- like* (DBL) e os *Cystein-rich interdomain region* (CIDR) (Fig.1.2) (25). Cada

um dos subdomínios estabelece ligação com recetores humanos específicos (16,19).



**Figura 1.2-**A) Esquema da estrutura molecular da família de proteínas PfEMP1. B) Estrutura 3D da PfEMP1.Uniprot- C6KT15. Informação recolhida a 10/07/2023.

### 1.3 Eritrócito

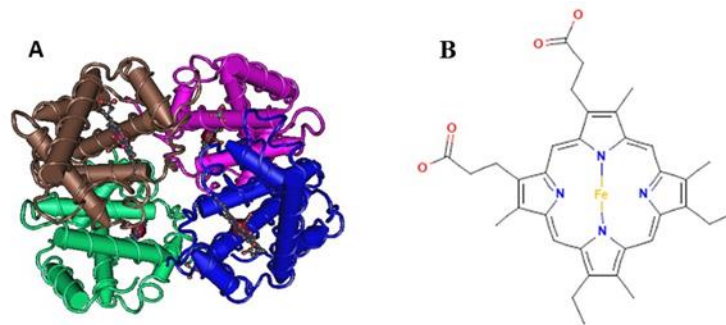
Os eritrócitos, ou glóbulos vermelhos, são células anucleadas, circulantes na corrente sanguínea, responsáveis pelo transporte de oxigénio e consequentes trocas gasosas. Apresentam forma bicôncava e uma membrana flexível que permite não só o aumento da área de troca dos gases, mas também a manutenção da estrutura e função da célula. A flexibilidade da membrana facilita a passagem pelos inúmeros capilares. Esta membrana é constituída por uma bicamada fosfolipídica e um citoesqueleto composto por proteínas, tais como espectrina, anquirina, actina e banda 3 (17,26,27).

Os eritrócitos têm um tempo de vida de 120 dias e apresentam vias enzimáticas imprescindíveis à sua manutenção. Possuem uma cor vermelha devido à presença de hemoglobina (Hb), proteína citoplasmática abundante sem a qual o transporte de oxigénio seria impossível, uma vez que é esta que alberga o ião ferro que apresenta afinidade para as moléculas de oxigénio (28).

#### 1.3.1 Hemoglobina e Hemoglobinopatias

A hemoglobina é a principal proteína presente nos eritrócitos. A sua morfologia e variações moleculares desde cedo foram campo de investigação e de interesse para o Homem, não só pela função que desempenha a nível biológico, mas também pela sua associação com diversas patologias humanas. Foi em 1970 que Perutz descreveu, pela primeira vez, a estrutura molecular da hemoglobina humana (29).

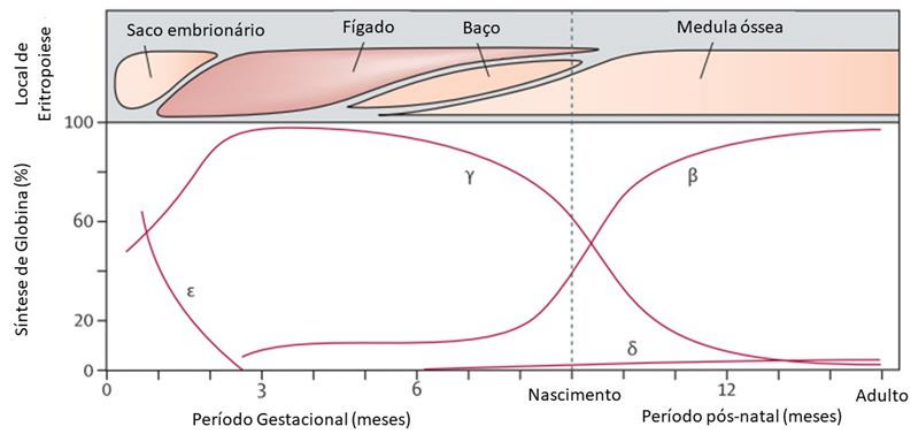
A hemoglobina é uma proteína tetramérica de estrutura globular, isto é, com quatro subunidades estruturais, duas cadeias globínicas alfa e duas beta (Fig.1.3-A), no caso da hemoglobina A, cada uma com ligação a um grupo prostético heme (Fig.1.3-B) (29). O íon ferro presente no grupo heme, no seu estado ferroso, é o responsável pela elevada afinidade e transporte de moléculas de oxigénio, desempenhando desta forma uma das funções fisiológicas mais importantes, o transporte de oxigénio e trocas gasosas, nos tecidos e órgãos (28,30).



**Figura 1.3-A)** Imagem 3D da estrutura quaternária da hemoglobina humana. NCBI – MMDB ID:184277/PDB ID:6LCX. **B)** Estrutura 2D do grupo heme. PubChem SID:144080007. Informação recolhida a 03/02/2023.

Ao longo do desenvolvimento fisiológico são sintetizados diferentes tipos de hemoglobinas que dependem do estadió de desenvolvimento (embrionário, fetal ou adulto) e da expressão ativa de diferentes cadeias globínicas. No período adulto, observa-se a síntese maioritária da hemoglobina adulta (HbA), e das hemoglobinas minoritárias, hemoglobina A<sub>2</sub> (HbA<sub>2</sub>) e hemoglobina fetal (HbF).

A alteração da expressão de cada tipo desta proteína é acompanhada pela mudança de local onde ocorre a eritropoiese (Fig.1.4), sendo que a partir do nascimento a eritropoiese tem lugar na medula óssea (31).



**Figura 1.4**-Representação das diferentes cadeias globínicas e sua expressão ao longo das diferentes etapas de desenvolvimento, desde a fase embrionária à adulta e respectivos locais de eritropoiese. Adaptado de (31).

As patologias genéticas relacionadas com a hemoglobina têm o nome de hemoglobinopatias. As hemoglobinopatias são doenças monogênicas de transmissão autossômica recessiva que se caracterizam por variantes patogênicas nos genes globínicos. Estas alterações genéticas podem afetar o nível de expressão do gene globínico, causando diminuição ou ausência de síntese da proteína correspondente ou, por outro lado, podem dar origem à síntese de uma proteína estruturalmente diferente, as chamadas variantes da hemoglobina. Deste modo, as hemoglobinopatias podem ser divididas em dois grupos: as hemoglobinopatias quantitativas e as qualitativas.

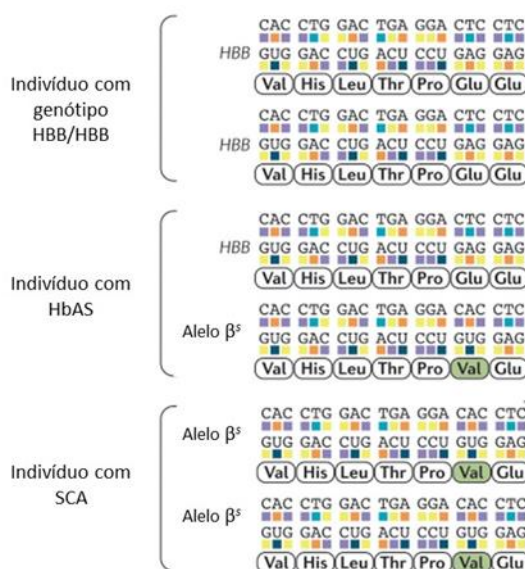
As hemoglobinopatias quantitativas são denominadas talassémias. Estas tomam o nome correspondente ao gene globínico cuja expressão está diminuída (por exemplo,  $\beta$ -talassémia e  $\alpha$ -talassémia) devido a mutações pontuais, a deleções ou inserções, localizadas no próprio gene ou nas suas regiões regulatórias (32).

As hemoglobinopatias qualitativas, ou variantes de hemoglobina, ocorrem quando se está na presença de alterações genéticas ao nível das sequências codificantes dos genes globínicos (exões) e que dão origem à alteração de um ou mais aminoácidos na proteína correspondente (geralmente mutações *missense*).

A simples alteração de um aminoácido origina uma hemoglobina estruturalmente diferente e que pode ter propriedades bioquímicas bastante diversas e nefastas, em relação à hemoglobina normal. O exemplo mais frequente e conhecido deste tipo de hemoglobinopatia é a drepanocitose (33).

## 1.4 Drepanocitose

A drepanocitose, ou anemia das células falciformes, é uma hemoglobinopatia causada por uma variante da hemoglobina, uma vez que tem na sua gênese uma mutação que altera o 6º aminoácido da proteína  $\beta$ -globina constituinte da HbA passando esta a denominar-se HbS (30). Os processos e características moleculares desta patologia foram, pela primeira vez, descritas por Linus Pauling em 1949, dando a conhecer os princípios desta mutação que afeta milhares de pessoas (34). Assim, a drepanocitose é uma doença hereditária monogénica, caracterizada pela mutação de um único nucleótido c.20A>T (substituição de adenina por timina, tabela 7.1 do anexo 2) no gene *HBB*, este que codifica as cadeias da  $\beta$ -globina. Esta mutação (rs334) é considerada *missense*, tendo como consequência a alteração do aminoácido ácido glutâmico (Glu) para valina (Val) (Tabela 7.2, do anexo 2) no sexto codão da referida proteína (Fig.1.5) (35).



**Figura 1.5**-Representação esquemática da alteração falciforme: substituição do aminoácido Ácido Glutâmico por Valina na proteína  $\beta$ -globina como consequência da alteração do tripleto GAG por GUG na sequência codificante do gene *HBB*. Adaptado de (36).

As diferentes características bioquímicas entre os dois aminoácidos levam a uma alteração no modo de ligação estrutural entre as cadeias alfa e beta, o que provoca uma modificação morfológica do eritrócito de uma forma bicôncava para uma forma de foice (falciforme) (26). Esta alteração confere maior rigidez, menor flexibilidade e tendência para a desidratação eritrocitária, culminando numa maior tendência para a ocorrência de

hemólise dos eritrócitos e, conseqüentemente, diminuição do seu número e ocorrência de anemia (anemia hemolítica crônica) (36). Outras conseqüências da drepanocitose, igualmente relevantes do ponto de vista clínico, são a obstrução vascular, ocorrência de acidentes vasculares cerebrais (AVCs), crises dolorosas, dactilite, entre outras (37).

Um indivíduo é doente drepanocítico se for homocigótico para o alelo mutante  $\beta^s$  (genótipo HbSS), enquanto se o indivíduo for heterocigótico (genótipo HbAS) diz-se que apenas apresenta o traço drepanocítico/falciforme. Os indivíduos com traço-falciforme são clinicamente assintomáticos e apresentam proteção contra a malária (8,35,38).

A drepanocitose apresenta uma elevada prevalência em populações que habitam regiões endêmicas da malária, o que se explica pela hipótese da malária previamente descrita (ver ponto 1.1).

Assim, devido às alterações morfológicas dos eritrócitos, origina-se um ambiente desfavorável à sobrevivência do parasita, devido a stresse celular e radicais oxidativos que não permitem a continuidade do ciclo intraeritrocitário do plasmódio. Por outro lado, a alteração do citoesqueleto do eritrócito não permite uma eficiente exportação da PfEMP1 para a membrana celular e diminui conseqüentemente a sua capacidade de citoadesão (39,40).

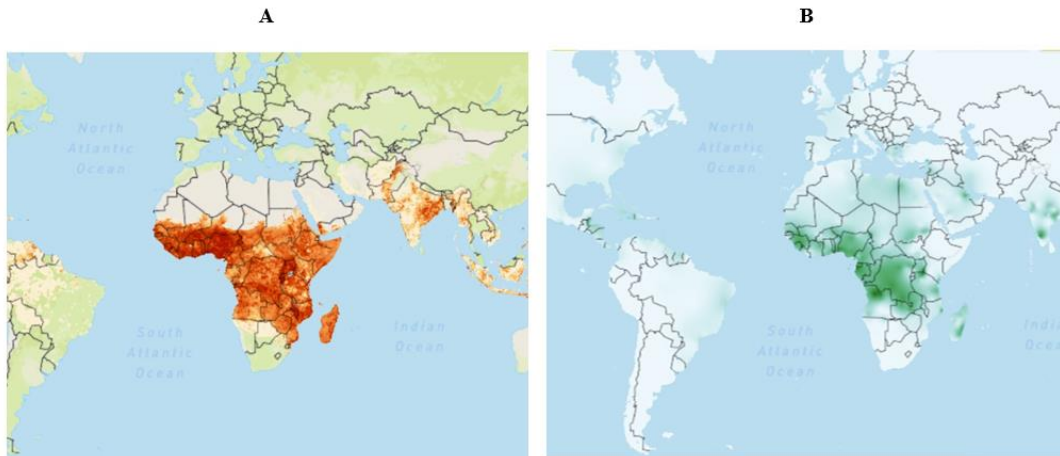
## 1.5 Epidemiologia da malária e drepanocitose em África

De acordo com a OMS, a anemia, no seu sentido lato, é um problema de saúde pública que afeta 1,8 bilhões de pessoas a nível global e dessas 269 milhões são crianças até aos 5 anos. As regiões mais afetadas são a África Subsariana e o Sudeste Asiático. Estima-se também que cerca de 75% dos nascimentos de crianças com drepanocitose (globalmente) ocorrem na África Subsariana (41,42).

A anemia pode dever-se a deficiências nutricionais (por exemplo, carência nutricional em ferro), a doenças infecciosas que alterem a absorção de nutrientes e metabolismo no eritrócito e a doenças hereditárias do sangue. Nas regiões endêmicas para a malária, a anemia possui duas causas prevalentes, as doenças infecciosas e as doenças hereditárias do sangue, nomeadamente, a prevalência do alelo  $\beta^s$  (43,44).

A figura 1.6 (A e B), mostra a prevalência do alelo responsável pela drepanocitose em regiões com maior incidência de malária, causada por *P. falciparum*. Por outro lado,

mostra também que a região africana é a fortemente afetada por ambas as condições. Isto, evidencia a pressão seletiva da malária (ao longo dos anos) sobre variantes polimórficas que confirmam algum nível de proteção, aumentando a sua frequência alélica na população.



**Figura 1.6-**A) Distribuição dos novos casos de malária, causada por *P. falciparum*, no continente africano no espaço de um ano. B) Distribuição do alelo  $\beta S$  no continente africano. Mapas gerados pelo *software* MalariaAtlas project (<https://malariaatlas.org/>). Informação recolhida a 11/07/2023.

Sabe-se que Angola é um dos países da África Subsariana com maior número de casos e mortes associadas à malária, este facto é associado às condições climatéricas e também à falta de recursos por parte do sistema de saúde (1).

A nível epidemiológico as províncias do Bengo e Luanda consideram-se meso-endémicas para a malária, com taxas de incidência da infeção entre os 10 e 30%, estando associadas a climas tropicais e a malária sazonal, sendo a época de maior transmissão entre novembro e maio. Diz-se que a infeção por malária nestes locais é sazonal devido às variações climatéricas do país, tendo alturas do ano de seca e outras de humidade e chuva constante. É no decorrer desta última que se dá o pico de transmissão do parasita, facto este resultante de um maior número de criadouros estáveis, o que facilita a reprodução do vetor (45).

Por outro lado, segundo a ANPAF (Associação Angolana de Apoiantes dos Portadores de Falciformação), 20% da população adulta angolana é portadora do alelo falciforme e anualmente cerca de dez mil crianças nascem com a doença (46).

É, portanto, notória a coexistência entre a ocorrência de drepanocitose e a endemicidade de malária.

## 1.6 Modeladores genéticos de variabilidade fenotípica da malária e da drepanocitose

Polimorfismos genéticos são variações que ocorrem no DNA e que originam diversidade genética nas espécies. Uma variante genética é considerada um polimorfismo quando a sua frequência populacional ultrapassa 1% dos alelos presentes na população em estudo (47). Os *Single Nucleotide Polymorphism* (SNP), dizem respeito a alterações de um só nucleótido. Estas variantes pontuais podem ter diversas consequências. Quando ocorrem na região codificante de um gene e levam à alteração de um aminoácido na proteína denominam-se alterações *missense*. Por outro lado, se a alteração introduz um codão de stop prematuro denomina-se de alteração *nonsense*. Neste caso, origina-se uma proteína truncada com perda da sua função biológica ou o RNA mensageiro que contém o codão de stop prematuro é degradado por um mecanismo de defesa da célula denominado *Nonsense Mediated Decay* (NMD), não havendo assim síntese da proteína correspondente (47–49). Outro tipo de variantes genéticas são os *Short Tandem Repeat* (STR), que dizem respeito a pequenas sequências nucleotídicas que se repetem em *tandem* e que variam em número no que diz respeito ao motivo de repetição (47). Para além dos acima mencionados, outro tipo de polimorfismo genético corresponde a deleções ou inserções, de tamanho variável, e que podem afetar a expressão de um ou mais genes (47).

Alguns dos polimorfismos referidos podem ser encontrados nos genes da  $\alpha$ -globina (*HBA1* e *HBA2*), nos genes *ICAM-1* (*intercellular adhesion molecule 1*) e *CD36* (*cluster of differentiation 36*) e têm vindo a ser amplamente estudados, uma vez que variantes polimórficas ocorridas nestes genes parecem ser relevantes, quer por aumento da suscetibilidade ou por proteção a múltiplas manifestações da malária e/ou por serem modificadores de alguns subfenótipos da drepanocitose.

Os genes *ICAM-1* e *CD36* codificam proteínas de membrana, expressas em inúmeras células humanas (descritas nos pontos 1.6.2 e 1.6.3), que se apresentam como elos de ligação entre os eritrócitos e o endotélio vascular humano através da PfEMP1, de origem parasitária (26,50,51). Deste modo, estas moléculas humanas podem ser consideradas moduladores genéticos da infeção por *P. falciparum*. Por outro lado, encontram-se também associados a manifestações de drepanocitose por modificarem a gravidade de

hemólise e vaso-occlusão, por exemplo (52,53).

### 1.6.1 Os genes da $\alpha$ -globina, a $\alpha$ -talassémia e as suas consequências

A  $\alpha$ -talassémia, hemoglobinopatia quantitativa, resulta de uma síntese diminuída das cadeias de  $\alpha$ -globina. A nível molecular esta patologia pode resultar de deleções totais ou parciais dos genes  $\alpha$ -globina (*HBA1* e *HBA2*). Deleções que envolvam apenas um gene  $\alpha$ -globina (genótipo  $-\alpha/\alpha\alpha$ ) são geralmente silenciosas ou de expressão muito ligeira; quando envolvem a eliminação de dois genes (genótipos  $-\alpha/-\alpha$  ou  $--/\alpha\alpha$ ) estão associadas a fenótipos ligeiros de microcitose e hipocromia, com ou sem anemia, considerados como  $\alpha$ -talassémia. Quando três genes  $\alpha$ -globina estão afetados (genótipo  $--/-\alpha$ ) dá-se origem à produção de HbH (tetramerização das cadeias  $\beta$ -globina) e surge uma anemia microcítica e hipocrômica de gravidade variável. A total ausência de genes  $\alpha$ -globina origina a síndrome da Hb *Bart Hydrops Fetalis*, incompatível com a vida (Tabela 1.1) (33). A nível hematológico/clínico a  $\alpha$ -talassémia encontra-se associada a casos de hipocromia (valores de hemoglobina globular média (HGM) abaixo da referência) e microcitose (volume globular médio (VGM) abaixo do esperado). Para além dessas alterações nos eritrócitos, a diminuição do nível da hemoglobina origina anemia nas situações de gravidade intermédia a grave (32).

**Tabela 1.1**-Fenótipos e genótipos das possíveis deleções associadas à  $\alpha$ -talassémia

Fenótipo da $\alpha$ -talassémia	Genótipos $\alpha$ -globínicos
Normal	$\alpha\alpha/\alpha\alpha$
Portador silencioso / $\alpha$ -talassémia menor	$-\alpha/\alpha\alpha$
$\alpha$ -talassémia	$-\alpha/-\alpha$
	$--/\alpha\alpha$
Doença da HbH	$--/-\alpha$
Síndrome da Hb <i>Bart Hydrops Fetalis</i>	$--/--$

A deleção  $\alpha$ -talassémica mais frequente em África tem um tamanho de 3,7 kb e ocorre entre o gene *HBA2* e o *HBA1*, dando origem a um gene híbrido funcional (54). Os portadores desta deleção que apresentam o genótipo  $-\alpha/\alpha\alpha$  ou  $-\alpha/-\alpha$  têm  $\alpha^+$ -talassémia e, segundo vários estudos, apresentam uma vantagem protetora contra a malária. Num estudo de caso-controlo realizado por Allen e colegas, os resultados revelaram uma vantagem significativa dos indivíduos homocigóticos para a  $\alpha^+$ -talassémia em relação à

malária grave comparando com indivíduos controlo. O mecanismo pelo qual esta proteção é conferida pensa-se ser ao nível da resposta imunitária, uma vez que eritrócitos  $\alpha^+$ -talassémicos parasitados são mais rapidamente fagocitados por monócitos (55).

Por outro lado, vários estudos apontam o benefício da co-herança da  $\alpha$ -talassémia com a drepanocitose, na medida em que esta última parece mitigar algumas das manifestações da drepanocitose (54). Na drepanocitose, em condições de desoxigenação, ocorre a polimerização das cadeias HbS no interior do eritrócito o que origina a deformação da célula e todas as consequências nefastas a ela associadas. Contudo, se coexistir  $\alpha$ -talassémia vai haver um menor número de cadeias  $\alpha$ -globina disponíveis para a formação do tetrâmero  $\alpha_2\beta^S_2$  resultando, assim, numa deficiência na síntese de HbS e consequentemente uma menor taxa de polimerização de HbS intraeritrocitária.

Deste modo, ocorre também a diminuição dos efeitos nefastos dessa polimerização e falciformização, tais como diminuição da gravidade da anemia hemolítica, uma vez que se irão encontrar em circulação eritrócitos falciformizados em menor número (56,57).

### 1.6.2 CD36 – O gene, a proteína e a sua função

O gene *CD36* codifica uma glicoproteína de membrana expressa em reticulócitos, plaquetas, macrófagos, células do endotélio, entre outras e é também recetor da trombospondina (TSP) humana, uma importante molécula de adesão. A glicoproteína CD36 desempenha funções ao nível do metabolismo de ácido gordos, metabolismo de lípidos oxidados de baixa densidade, fagocitose e angiogénese. É uma das moléculas de adesão recetoras para a PfEMP1 do *P. falciparum*, associação já descrita na literatura (51,58,59).

Em termos estruturais, o gene *CD36* é constituído por 15 exões sendo que destes apenas do terceiro ao 14º exão é que se consideram codificantes (Fig.1.7).



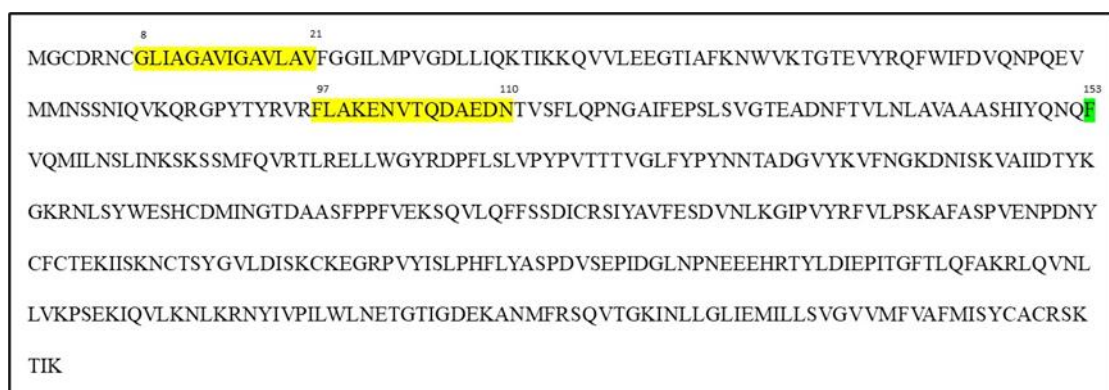
**Figura 1.7**-Esquema representativo da estrutura do gene *CD36* de acordo com a base de dados NCBI - NG\_008192.1. As barras verticais, preenchidas a cor roxa, correspondem aos exões. Informação recolhida a 02/03/2023.

A região N-terminal da proteína é codificada pelo terceiro exão e a C-terminal pelo 14º,

ambas as regiões se constituem como transmembranares e citoplasmáticas. É dos genes que apresenta maior probabilidade de sofrer alterações nucleotídicas e apresenta inúmeros locais de *splicing* alternativo (pelo menos 11 locais de *splice junction*, que se encontram em regiões perto ou dentro dos exões) com promotores independentes (60,61).

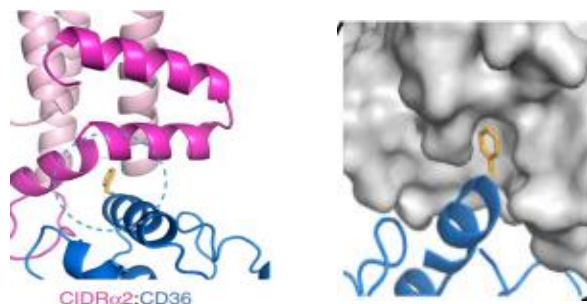
A proteína codificada pelo gene *CD36*, apresenta uma estrutura em *hairpin* e pertence à superfamília de recetores que reconhecem padrões moleculares associados a patógenos (PAMPs) sendo de igual modo considerada um recetor de tipo B. Ao gene *CD36* e proteína resultante são atribuídas diferentes funções biológicas que, aquando do estudo da malária, podem levar a conclusões contraditórias e a incertezas quanto ao papel deste gene na suscetibilidade ou proteção contra as diferentes manifestações da infeção. A diversidade de resultados e conclusões poderá ser explicada porque a glicoproteína CD36 é expressa em macrófagos, contribuindo para uma melhor adesão e fagocitose do eritrócito parasitado. Por outro lado, é também expressa em células do endotélio e plaquetas, assistindo aos fenómenos de citoadesão e *rosetting* beneficiando o parasita (62).

O local de ligação entre a CD36 e a PfEMP1, é uma região conservada em ambas as proteínas, isto é, apesar de modificações que possam ocorrer nestas proteínas, esta região mantém-se sempre inalterável. Na proteína CD36, os locais de ligação são codificados pelos aminoácidos (aa) nas posições 8 a 21 e 97 a 110, correspondentes no gene, respetivamente, aos exões 3 e 5. No entanto o resíduo de fenilalanina (aa 153) é apontado como o local de ligação fixo à proteína PfEMP1(Fig.1.8) (63,64).



**Figura 1.8**-Sequência aminoacídica (452aa) da proteína CD36. Regiões destacadas a amarelo indicam locais de ligação à PfEMP1-codificadas no gene pelo exão 3 e 5- e o aminoácido destacado a verde representa o resíduo conservado de fenilalanina que estabelece ligação.

De igual forma, a PfEMP1 possui também uma região conservada (comum a todas as variantes dos *var genes* que possam ser expressas). Esta região forma uma bolsa hidrofóbica com a qual o resíduo de fenilalanina estabelece ligação e que pertence aos subdomínios CIDR $\alpha$ 2-6 (Fig.1.9) (24,61,64).



**Figura 1.9**-Resíduo de fenilalanina da proteína CD36 como local conservado de ligação à bolsa hidrofóbica, situada no domínio CIDR $\alpha$ 2 da proteína PfEMP1. Adaptado de (64).

Em relação à drepanocitose, sabe-se que a proteína CD36 é expressa nos eritrócitos em formação, incluindo nos reticulócitos, sendo perdida após maturação destes em eritrócitos. A drepanocitose ao provocar um aumento do nível de hemólise dos eritrócitos (hemólise crónica) leva ao aumento da taxa de eritropoiese por um mecanismo de stresse hematopoiético. Assim, um maior número de reticulócitos circulantes é sinónimo de uma maior quantidade de células eritrocitárias que expressam CD36 e, conseqüentemente, de um aumento do fenómeno de adesão ao endotélio e vaso-oclusão (49).

Alguns polimorfismos no gene *CD36* já foram descritos como modificadores da gravidade da hemólise e dos níveis de hemoglobina em doentes drepanocíticos. Foi descrita a associação do SNP rs1984112\_alelo G, ocorrido na região promotora do gene (5'UTR), com o aumento do número de reticulócitos, um dos biomarcadores de hemólise (52,53). Também o SNP rs1413661\_alelo C foi relacionado com baixos níveis de hemoglobina, o que agrava o estado de anemia (49,65).

### 1.6.3 ICAM-1 – O gene, a proteína e a sua função

O gene *ICAM-1* localiza-se no cromossoma 19 e possui sete exões (Fig.1.10).



**Figura 1.10**-Esquema representativo da estrutura do gene *ICAM-1* de acordo com a base de dados NCBI- NG\_012083.1. As barras verticais, preenchidas a cor roxa, correspondem aos exões. Informação recolhida a 02/03/2023.

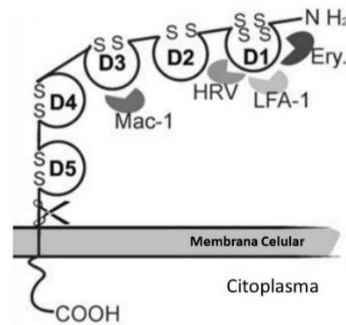
Codifica uma glicoproteína de membrana de 532 aminoácidos, que permite a interação célula-célula e possui um papel importante no que toca a processos imunitários tanto inatos como adaptativos.

Ao nível da expressão génica, o gene *ICAM-1* é regulado por fatores de transcrição nos inúmeros locais de ligação na região promotora. É também regulado, através de vias de transdução de sinal, por citocinas e por níveis de estímulos específicos, isto é, a expressão aumenta aquando de processos inflamatórios e diminui quando o contrário ocorre.

A proteína ICAM-1 é expressa em células endoteliais, células epiteliais e células do sistema imunitário, evidenciando o importante papel que possui no que diz respeito ao contacto célula-célula. Para além disso, encontra-se envolvida no processo de transmigração dos leucócitos, na fase crucial de adesão firme ao endotélio, facilitando processos de adesão endotelial (26,66,67).

Esta glicoproteína pertence à superfamília das imunoglobulinas apresentando cinco domínios *Ig-like* extracelulares, codificados do segundo ao sexto exão, e um domínio transmembranar codificado pelo sétimo exão (quando na sua isoforma acoplada). O gene pode igualmente dar origem a uma forma solúvel da proteína, resultante de processos de *splicing* alternativo e ação proteolítica por parte de enzimas localizadas na membrana ou através de enzimas bacterianas (68). A expressão desta forma solúvel é dependente da célula em que ocorre a sua expressão ativa (67).

Esta molécula de adesão é também recetora humana do rinovírus humano (HRV), do antígeno macrofágico 1 (Mac-1) e do antígeno associado à função linfocitária 1 (LFA-1), cada uma destas estabelecendo ligações com domínios extracelulares da proteína (Fig.1.11).



**Figura 1.11**-Representação esquemática da molécula de ICAM-1 com todos os seus domínios extracelulares D1 a D5 e cauda citoplasmática. Os ligandos da molécula estão representados nos seus locais de ligação, Mac-1, HRV, LFA-1 e Ery (eritrócito infetado). Adaptado de (67).

O primeiro domínio reconhece a PfEMP1, o LFA-1 e HRV e o terceiro domínio reconhece o Mac-1.

Em relação à malária, este recetor humano revela-se importante não só por reconhecer a família PfEMP1, permitindo a citoadesão de eritrócitos parasitados, mas também porque a proteína ICAM-1, tanto na sua forma acoplada como solúvel, são expressas em grandes quantidades no endotélio cerebral (69). Assim, a associação desta proteína ICAM-1 à malária cerebral tem sido estudada na tentativa de encontrar uma relação entre variantes genéticas e a suscetibilidade ou proteção para a malária cerebral (26). A ligação ICAM-1 à PfEMP1 ocorre ao nível do subdomínio DBL reconhecido pelo primeiro domínio da proteína ICAM-1. Assim, qualquer alteração que ocorra ao nível da sequência codificante do domínio 1 de ICAM-1 irá afetar diretamente a afinidade de ligação desta molécula à proteína virulenta (69,70)

Sendo uma molécula de adesão com elevada expressão em células do endotélio, o seu estudo revela-se importante em relação à drepanocitose, uma vez que a obstrução vascular é a maior causa de morbidade e mortalidade desta patologia (52).

Um estudo de Kucukal e colaboradores demonstrou, *in vitro*, que eritrócitos provenientes de doentes drepanocíticos (HbSS) apresentam maior afinidade para proteína ICAM-1 do

que eritrócitos com apenas um dos alelos mutados (provenientes de indivíduos heterozigóticos, HbAS) ou eritrócitos normais (provenientes de indivíduos normais, HbAA), suportando a ideia de que a ICAM-1 está diretamente relacionada com a maior adesão e a ocorrência de vaso-oclusão em indivíduos drepanocíticos (71).



## 2 Justificação e Objetivos

Face às consequências tanto da malária como da drepanocitose, é importante perceber que influência podem as variantes genéticas ter no desenvolvimento destas e em que medida podem ser consideradas modificadores e biomarcadores genéticos. É, igualmente, importante perceber qual o papel das variantes genéticas humanas na suscetibilidade para desenvolver, no caso da malária, infeções mais graves e, por outro lado, desenvolver uma proteção contra a infeção da malária. A aquisição de conhecimentos, neste contexto, pode colmatar lacunas tanto, nos métodos desenvolvidos para combater a infeção por *P. falciparum* de modo que a terapêutica seja direcionada às células do hospedeiro e não do parasita, por exemplo. Por outro lado, estudar as interações celulares na presença de drepanocitose pode ser benéfico para melhor entender como mitigar crises de vaso-occlusão e ocorrências de AVCs, as duas principais consequências que levam à morte dos portadores homozigóticos desta mutação.

Sendo a malária considerada um mecanismo de pressão seletiva de variantes polimórficas nos países endémicos para a parasitose, e uma vez que estes são, igualmente, países com elevada prevalência da mutação drepanocítica é interessante estudar a sua interação e perceber que vantagens/desvantagens a coexistência desta parasitose e patologia podem trazer e poder aplicá-los no desenvolvimento de terapêutica.

Com este trabalho pretende-se contribuir para o conhecimento de possíveis interações entre a malária e a drepanocitose, ambas de elevada prevalência, morbidade e mortalidade em África, mais especificamente em Angola. Para além disso, pretende-se aumentar a compreensão de como o *background* genético dos doentes drepanocíticos pode influenciar a variabilidade dos fenótipos relacionados com a drepanocitose e com as infeções por malária.

A presente dissertação tem como objetivos principais a investigação da relação entre a variabilidade genética dos genes *CD36*, *ICAM-1* e *HBA* e a variabilidade fenotípica da malária e drepanocitose, assim como o estudo de possíveis interações entre as duas doenças. Neste sentido, foram definidos objetivos específicos:

- 1) Estudar os genes *CD36*, *ICAM-1* e *HBA* em DNAs de crianças angolanas com drepanocitose, recorrendo a metodologias de biologia molecular;
- 2) Determinar as frequências alélicas e genotípicas das variantes genéticas detetadas;

- 3) Determinar a potencial patogenicidade das variantes detetadas, através de estudos *in silico*;
- 4) Investigar uma possível relação entre as variantes genéticas identificadas e os fenótipos (hematológico, bioquímico e clínico) apresentados pelas crianças estudadas, relacionados com a drepanocitose e com infeção de malária;
- 5) Investigar o papel modulador das variantes genéticas encontradas na interação da drepanocitose com a malária.

### 3 Materiais e Métodos

#### 3.1 Ética e consentimento informado

Este estudo inseriu-se no projeto de investigação clínica de âmbito mais lato intitulado “SCAFfoldChild – Sickle Cell Anemia and Fetal Hemoglobin. Genetic modifiers in an Angolan Children Cohorte” que envolveu o Hospital Pediátrico David Bernardino (HPDB), em Luanda, o Centro de Investigação em Saúde de Angola, localizado no Bengo, a Escola Superior de Tecnologia da Saúde de Lisboa (ESTeS-L), o Instituto de Higiene e Medicina Tropical (IHMT NOVA) e o Instituto Nacional de Saúde Doutor Ricardo Jorge (INSA), Lisboa. Este projeto obteve as devidas aprovações pelas comissões de ética das instituições envolvidas e pelo Ministério da Saúde de Angola. Acrescenta-se, ainda, que a participação dos doentes drepanocíticos no projeto de investigação clínica foi voluntária e que os tutores legais das crianças participantes foram devidamente esclarecidos sobre os objetivos do projeto e assinaram um documento de consentimento informado (Anexo 1).

#### 3.2 Caraterização da amostra

##### 3.2.1 Amostra Populacional

A população estudada envolveu 65 crianças previamente diagnosticadas com drepanocitose, com uma média de idades de 7 anos (mínimo de 3 e máximo de 12 anos), sendo constituída por 33 (50,8%) indivíduos do sexo feminino e 32 (49,2%) do sexo masculino. Estas crianças foram acompanhadas clinicamente, num estudo longitudinal, pela Dra. Brígida Santos, médica pediátrica, no HPDB em Luanda e no Hospital Geral do Bengo. Na primeira consulta foi realizado um inquérito de anamnese e recolhidos dados demográficos e clínicos, incluindo a idade e o tipo da primeira manifestação drepanocítica. Para além disso, o acompanhante da criança foi inquirido acerca do historial de episódios de malária que, eventualmente, a criança já tivesse tido (42).

Ao longo das consultas, programadas semestralmente, foram recolhidos dados clínicos e hematológicos, tais como: Hb, VGM, HGM, contagem de eritrócitos, reticulócitos e plaquetas, e número de episódios de malária. Todos estes dados foram recolhidos antes do início do tratamento com hidroxiureia (HU).

Os critérios aplicados para a participação no estudo foram:

- Critérios de inclusão: Crianças entre os 3 e os 12 anos com drepanocitose confirmada a nível laboratorial por método de focagem isoelétrica ou eletroforese de hemoglobinas e posterior confirmação molecular.
- Critérios de exclusão: Transfusões recentes (nos três meses anteriores ao início do estudo), administração prévia de HU, tuberculose ativa, infecção por HIV, neoplasias e evidências de disfunção medular.

### **3.2.2 Amostra Biológica**

Foi colhido, pela Dra. Brígida Santos, sangue periférico em EDTA para a avaliação hematológica dos doentes. Posteriormente, foram aplicadas gotas de sangue em *guthrie cards* (WHAT10531018 Cartão 903 *Protein Saver*) com fim ao seu transporte para o IHMT NOVA. O DNA humano e do parasita foram extraídos pelo método de Chelex, de acordo com o protocolo descrito por Plowe *et al* (72).

#### **3.2.2.1 Análise quantitativa e qualitativa do DNA**

A quantificação do DNA foi realizada por método espectrofotométrico. Utilizou-se o equipamento NanoDrop One (Thermo Fisher Scientific) para determinar a concentração dos DNAs em estudo, em nanogramas por microlitro (ng/ $\mu$ L), e para avaliar a sua qualidade. O espectrofotómetro NanoDrop One permite quantificar o DNA em volumes muito pequenos de amostra (1  $\mu$ L) e indica os rácios de pureza, permitindo inferir sobre a integridade/qualidade da amostra. Desta forma, é realizada uma análise aos possíveis contaminantes, solventes e reagentes utilizados num pré-tratamento da amostra, quer estes sejam proteicos ou outros. A concentração do DNA é determinada com base na equação de Beer-Lambert (Fig.3.1), a uma absorvância de 260 nm.

Os rácios de pureza, A<sub>260</sub>/A<sub>280</sub> e A<sub>260</sub>/A<sub>230</sub>, são calculados após a leitura das absorvâncias dos ácidos nucleicos a 230, 260 e 280 nm (73).

Valores de A260/A280 entre 1,8 e 2,0 e de A260/A230 entre 2,0 e 2,2 indicam pureza adequada do DNA, enquanto valores fora destes intervalos indicam a presença de contaminantes.

Concentração do analito (M)

Coefficiente de extinção

$$c = A / (\epsilon \times b)$$

Absorvância

Largura da cuvette

O diagrama mostra a equação  $c = A / (\epsilon \times b)$  no centro. Uma seta aponta de 'c' para cima para 'Concentração do analito (M)'. Outra seta aponta de 'A' para cima para 'Coefficiente de extinção'. Uma seta aponta de 'A' para baixo para 'Absorvância'. Uma seta aponta de 'b' para baixo para 'Largura da cuvette'.

**Figura 3.1**-Equação de Beer-Lambert. O cálculo da concentração de DNA é realizado de acordo com a equação de Beer-Lambert onde se substitui o valor de absorvância resultante da leitura, o coeficiente de extinção que diz respeito à opacidade do material em leitura (líquido ou sólido) e a largura da cuvette (se utilizada) na leitura.

### 3.3 Desenho experimental e técnicas aplicadas

O desenho experimental deste trabalho foi realizado partindo de conhecimentos e fundamentos básicos de biologia e genética molecular e também recorrendo a bibliografia científica procurada em diversas bases de dados, como por exemplo, o PubMed. Por outro lado, foi também projetado com base nos objetivos do projeto de investigação. Na tabela 3.1 são identificadas as variantes genéticas estudadas, os respetivos genes e as metodologias de genética molecular aplicadas.

**Tabela 3.1**-Metodologias aplicadas para o estudo das diversas variantes genéticas analisadas

Gene	Identificação da Variante Genética	Metodologia	
<b>CD36</b>	rs1413661	PCR* + Sequenciação de Sanger	
	rs1984112		
	rs3211891		
	rs3211892		
	rs3211893		
	rs3138813 – STR In3(TG)n	PCR + Análise de microssatélites	
	rs3211938	PCR + Sequenciação de Sanger	
	rs5491		
	<b>ICAM-1</b>	rs5030352	PCR + Sequenciação de Sanger
		rs1799969	
rs5498			
rs1801714			
rs5496			
rs5494			
<b>HBA</b>	rs5497	Gap-PCR	
	$\alpha$ - talassémia (del 3,7kb)		

\*PCR - *Polymerase Chain Reaction*.

### 3.3.1 Amplificação *in vitro* de DNA (*Polymerase Chain Reaction*, PCR)

A *Polymerase Chain Reaction* (PCR) é uma metodologia que permite a amplificação *in vitro* de uma sequência específica de DNA. A reação depende da ação enzimática de uma DNA-polimerase (TaqDNA polimerase, estável a temperaturas elevadas) que sintetiza novas cadeias de DNA partindo de um DNA molde. A atuação correta da TaqDNA polimerase depende de um tampão estabilizador da reação, que contém a concentração correta de sais e de alguns adjuvantes, tais como, albumina sérica bovina (BSA) e cloreto de magnésio (MgCl<sub>2</sub>) (composição dos tampões e adjuvantes na tabela 7.3, apresentada no anexo 3). A amplificação é também dependente da existência, no meio reacional, dos quatro desoxirribonucleótidos trifosfatados (dNTPs): adenina (A), citosina (C), timina (T) e guanina (G). Estes são ligados entre si, através de ligações fosfodiéster, dando origem às novas cadeias de DNA a partir dos oligonucleótidos iniciadores, os *primers*.

Os *primers* são sequências curtas de DNA, de cadeia simples, específicas para a região-

alvo da amplificação e que, por isso, determinam o início (*primer forward*, direto ou de sentido *sense*) e o fim (*primer reverso* ou *anti-sense*) da sequência a amplificar (74).

Os programas de PCR caracterizam-se por, geralmente, 30 ciclos sucessivos de aumento e descida de temperaturas específicas e compreendem algumas etapas obrigatórias: 1) uma desnaturação inicial ( $\approx 94^{\circ}\text{C}$ ) que permite que cada cadeia de DNA original sirva de molde para a síntese de novas cadeias e que a região alvo fique acessível; 2) hibridação ( $\approx 55\text{-}65^{\circ}\text{C}$ ) que diz respeito ao momento de reconhecimento e ligação dos *primers* ao local alvo; e 3) extensão da reação ( $72^{\circ}\text{C}$ ) que permite a ação da TaqDNA polimerase e, portanto, a síntese de novas cadeias de DNA (75).

Neste trabalho, as PCRs, e as metodologias nelas baseadas, foram realizadas no termociclador T Gradient Thermocycler da marca Biometra. As sequências e características dos *primers* bem como condições de PCR e mistura reacional utilizada para cada região génica estudada encontram-se descritas nos pontos seguintes (3.3.1.1 a 3.3.2).

A constituição das misturas de reação para as diversas PCRs realizadas, bem como as respectivas condições reacionais foram, inicialmente, obtidas a partir de bibliografia científica. Contudo, as condições das reações foram otimizadas do ponto de vista laboratorial e, por vezes, ajustadas às condições utilizadas no laboratório. Para além disto, em todas as PCRs realizadas foi incluído um controlo negativo para monitorização de contaminações e validação da técnica.

Em algumas amostras, devido à falta de integridade do DNA original e para que fosse possível obter produto amplificado em quantidade suficiente, optou-se por realizar duas reações de PCR consecutivas (equivalente à realização de *nested-PCR*). Na segunda reação de PCR, o DNA molde consistia na utilização de um  $\mu\text{L}$  do produto amplificado resultante da primeira reação de PCR, por vezes diluído a 1:10. Para além disso, optou-se pela utilização dos *primers* numa concentração de 15 picomoles por  $\mu\text{L}$  ( $\text{pmol}/\mu\text{L}$ ).

### **3.3.1.1 PCRs para o estudo do gene *CD36***

Nas tabelas seguintes (tabelas 3.2 a 3.11) são apresentadas as misturas reacionais utilizadas nas reações de PCR, bem como as condições dessas mesmas reações, para cada região génica estudada, no gene *CD36*.

As sequências e caracterização dos *primers* utilizados, para as reações de PCR deste gene, encontram-se descritas na tabela 7.4 e 7.5, presentes no anexo 3.

**Tabela 3.2-**Mistura de PCR para o estudo da região promotora do gene *CD36*

Mistura de PCR	
Reagente	Volume por reação (µL)
Tampão B	22,8
CD36_33137_Fw (25 pmol/µL)	0,5
CD36_33137_Rv (25 pmol/µL)	0,5
Taq AmpliTaq*(5U/µL)	0,2
DNA**	2

\*Taq AmpliTaq™ da Applied Biosystems;\*\*2 µL de DNA molde contém cerca de 70 ng de DNA.

**Tabela 3.3-**Condições de PCR para o estudo da região promotora do gene *CD36*

Condições de PCR			Fragmento amplificado (pb)
T (°C)	Tempo	Nº de ciclos	
94	10 min	x1	314
94	30 seg		
56	30 seg	x30	
72	40 seg		
72	10 min	x1	

**Tabela 3.4-**Mistura de PCR para o estudo do quinto exão do gene *CD36*

Mistura de PCR	
Reagente	Volume por reação (µL)
dH <sub>2</sub> O	19,55
Tampão David (10x)	2,5
BSA (10 mg/mL)	0,35
dNTPs (100 mM)	0,5
CD36_Ex5_Fw (25 pmol/µL)	0,5
CD36_Ex5_Rv (25 pmol/µL)	0,5
Taq AmpliTaq* (5U/µL)	0,1
DNA**	2

\*Taq AmpliTaq™ da Applied Biosystems;\*\*2 µL de DNA molde contém cerca de 70 ng de DNA.

**Tabela 3.5**-Condições de PCR para o estudo do quinto exão do gene *CD36*

Condições de PCR			Fragmento amplificado (pb)
T (°C)	Tempo	Nº de ciclos	
94	5 min	x1	707
94	45 seg		
55	45 seg	x35	
72	35 seg		
72	5 min	x1	

**Tabela 3.6**-Mistura de PCR para o estudo do décimo exão do gene *CD36*

Mistura de PCR	
Reagente	Volume por reação (µL)
Tampão B	23,4
CD36_SNPex10_Fw (25 pmol/µL)	0,5
CD36_SNPex10_Rv (25 pmol/µL)	0,5
Taq AmpliTaq* (5U/µL)	0,1
DNA**	2

\*Taq AmpliTaq™ da Applied Biosystems; \*\*2 µL de DNA molde contém cerca de 70 ng de DNA.

**Tabela 3.7**-Condições de PCR para o estudo do décimo exão do gene *CD36*

Condições de PCR			Fragmento amplificado (pb)
T (°C)	Tempo	Nº de ciclos	
94	4 min	x1	390
94	45 seg		
56	45 seg	x30	
72	45 seg		
72	4 min	x1	

Para analisar o STR In3(TG)<sub>n</sub> presente no gene *CD36*, foram realizadas duas reações de PCR distintas para posterior análise por metodologias diferentes. Assim, numa das reações de PCR foram utilizados *primers* marcados com fluoróforo FAM para que fosse possível, posteriormente, realizar a técnica de análise de STRs (descrita no ponto 3.3.5). Na outra reação de PCR foram utilizados *primers* desenhados no laboratório e não marcados com fluoróforo com o objetivo de posterior análise do produto da amplificação por sequenciação de Sanger (descrita no ponto 3.3.4).

**Tabela 3.8-** Mistura de PCR para o estudo da variante In3(TG)n no gene *CD36*, utilizando *primers* marcados com FAM

Mistura de PCR	
Reagente	Volume por reação (µL)
Tampão B	23
CD36_TGn_Fw (25 pmol/µL)	0,5
CD36_TGn_Rv (25 pmol/µL)	0,5
Taq AmpliTaq* (5U/µL)	0,1
DNA**	2

\*Taq AmpliTaq™ da Applied Biosystems; \*\*2 µL de DNA molde contém cerca de 70 ng de DNA.

**Tabela 3.9-** Condições de PCR para o estudo da variante In3(TG)n no gene *CD36*, utilizando *primers* marcados com FAM

Condições de PCR			Fragmento amplificado (pb)
T (°C)	Tempo	Nº de ciclos	
94	5 min	x1	120
94	45 seg		
45	45 seg	x38	
72	30 seg		
72	5 min	x1	

**Tabela 3.10-** Mistura de PCR para o estudo da variante In3(TG)n no gene *CD36*

Mistura de PCR	
Reagente	Volume por reação (µL)
dH <sub>2</sub> O	19,55
Tampão David (10x)	2,5
BSA (10 mg/mL)	0,35
dNTPs (100 mM)	0,5
CD36_IVS3_FW (25 pmol/µL)	0,5
CD36_IVS3_Rv (25 pmol/µL)	0,5
Taq AmpliTaq* (5U/µL)	0,1
DNA**	2

\*Taq AmpliTaq™ da Applied Biosystems; \*\*2 µL de DNA molde contém cerca de 70 ng de DNA.

**Tabela 3.11**-Condições de PCR para o estudo da variante In3(TG)n no gene *CD36*

Condições de PCR			Fragmento amplificado (pb)
T (°C)	Tempo	N° de ciclos	
94	5 min	x1	578
94	45 seg		
53	45 seg	x35	
72	45 seg		
72	5 min	x1	

### 3.3.1.2 PCRs para o estudo do gene *ICAM-1*

As seguintes tabelas (Tabela 3.12 a 3.15) apresentam as misturas reacionais utilizadas nas reações de PCR, bem como as condições dessas mesmas reações, para cada região génica estudada no gene *ICAM-1*.

As sequências e a caracterização dos *primers* utilizados para as reações de PCR efetuadas para o estudo do gene *ICAM-1* encontram-se descritas na tabela 7.6, presente no anexo 3.

**Tabela 3.12**-Mistura de PCR para o estudo do segundo exão do gene *ICAM-1*

Mistura de PCR	
Reagente	Volume por reação (µL)
Tampão B	23
ICAM1_K29M_Fw (25 pmol/µL)	0,5
ICAM1_K29M_Rv (25 pmol/µL)	0,5
Taq AmpliTaq* (5U/µL)	0,1
DNA**	2

\*Taq AmpliTaq™ da Applied Biosystems; \*\*2 µL de DNA molde contém cerca de 70 ng de DNA.

**Tabela 3.13**-Condições de PCR para o estudo do segundo exão do gene *ICAM-1*

Condições de PCR			Fragmento amplificado (pb)
T (°C)	Tempo	Nº de ciclos	
94	5 min	x1	416
94	45 seg		
54	45 seg	x30	
72	45 seg		
72	5 min	x1	

**Tabela 3.14**-Mistura de PCR para o estudo do quarto ao sexto exão do gene *ICAM-1*

Mistura de PCR	
Reagente	Volume por reação (µL)
dH <sub>2</sub> O	19,55
Tampão David (x10)	2,5
BSA (10 mg/mL)	0,35
dNTPs (100 mM)	0,5
ICAM1_G241R_Fw (25 pmol/µL)	0,5
ICAM1_K469E_Rv (25 pmol/µL)	0,5
Taq AmpliTaq* (5U/µL)	0,1
DNA**	2

\*Taq AmpliTaq™ da Applied Biosystems; \*\*2 µL de DNA molde contém cerca de 70 ng de DNA.

**Tabela 3.15**- Condições de PCR para o estudo do quarto ao sexto exão do gene *ICAM-1*

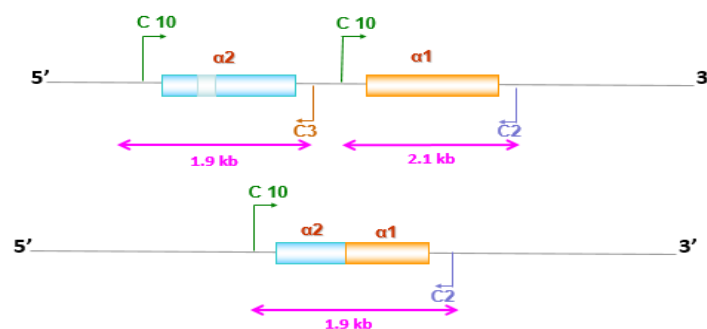
Condições de PCR			Fragmento amplificado (pb)
T (°C)	Tempo	Nº de ciclos	
94	5 min	x1	1053
94	45 seg		
62	45 seg	x30	
72	45 seg		
72	5 min	x1	

### 3.3.2 Gap-PCR para o estudo dos genes *HBA*

O Gap-PCR é uma variante da PCR convencional, que se destina ao estudo de sequências de DNA com deleções ou inserções conhecidas. Esta técnica foi utilizada, neste trabalho,

para determinar a presença da deleção  $\alpha$ -talassêmica de 3,7kb que envolve ambos os genes  $\alpha$ -globina (*HBA2* e *HBA1*).

Para a sua realização são utilizados *primers* que flanqueiam externamente as regiões da deleção. Geralmente, para a determinação do genótipo de cada amostra são realizadas duas reações de PCR utilizando três *primers* (um *forward* e dois *reverse*), para que se consiga amplificar tanto os genes normais como o gene híbrido, que advém da presença da deleção, tal como esquematizado na figura 3.2 e descrito por Dodé e seus colegas (76,77).



**Figura 3.2**-Representação esquemática dos genes *HBA2* e *HBA1* e do gene híbrido resultante da deleção talassêmica de 3,7kb, bem como dos locais de ligação dos *primers* (C10, C3 e C2). O gene *HBA1* está representado por  $\alpha 1$  (cor laranja), o gene *HBA2* está representado por  $\alpha 2$  (cor azul) e o gene híbrido por  $\alpha 2\alpha 1$  (ambas as cores).

Na realização do Gap-PCR para a pesquisa da deleção  $\alpha$ -talassêmica de 3,7kb são realizadas, geralmente, duas misturas reacionais (mistura I e II). Estas têm a mesma constituição reacional, contêm o mesmo *primer Fw* mas, cada uma delas, contêm um dos *primers* reverse (respetivamente, C3 ou C2).

Pela observação do esquema apresentado na figura 3.2 é possível concluir que: 1) na situação normal, ou seja, ausência de deleção ( $\alpha\alpha/\alpha\alpha$ ), há a amplificação de dois fragmentos com 1,9kb (mistura I) e 2,1kb (mistura II); 2) na presença da deleção em homocigotia ( $-\alpha^{3,7}/-\alpha^{3,7}$ ) ocorre a amplificação de um fragmento com 1,9kb (mistura II) e ausência de amplificação na mistura I; 3) quando a deleção se encontra em heterocigotia ( $\alpha\alpha/-\alpha^{3,7}$ ) há a amplificação de fragmento de 1,9kb (mistura I) e de dois fragmentos de 1,9kb e 2,1kb (mistura II).

Neste trabalho, a metodologia de Gap-PCR para determinar a presença da deleção  $\alpha$ -

talassêmica de 3,7kb foi apenas realizada para dois dos DNA em estudo, uma vez que os restantes já tinham sido genotipados pela colega Isabel Germano para realização da dissertação intitulada “Anemia das Células Falciformes - influência de fatores genéticos na gravidade da doença em idade pediátrica”(49,78). Para além disso, foi apenas realizada a mistura II, uma vez que o objetivo era detetar o gene híbrido. A constituição desta mistura reacional encontra-se descrita na tabela 3.16 e as condições da Gap-PCR na tabela 3.17.

As sequências e caracterização dos *primers* utilizados para as reações deste Gap-PCR encontram-se descritas na tabela 7.7, presente no anexo 3.

**Tabela 3.16-**Mistura II de Gap-PCR para o estudo da deleção  $\alpha$ -talassêmica 3,7kb que afeta os genes *HBA2* e *HBA1*

Mistura II de Gap-PCR	
Reagente	Volume por reação ( $\mu$ L)
dH <sub>2</sub> O	15,3
Tampão $\alpha$ + $\beta$ ME (10x)	2,5
DMSO	2,5
MgCl <sub>2</sub> (0,1 M)	0,5
BSA (10 mg/mL)	0,4
dNTPs (100 mM)	0,5
C10- A_5'UTR_Fw (25 pmol/ $\mu$ L)	0,5
C2- A1_3'UTR_Rv (25 pmol/ $\mu$ L)	0,5
Taq AmpliTaq* (5U/ $\mu$ L)	0,3
DNA**	2

\*Taq AmpliTaq<sup>TM</sup> da Applied Biosystems; \*\*2 $\mu$ L de DNA molde contém cerca de 70 ng de DNA.

**Tabela 3.17-**Condições de Gap-PCR para o estudo da deleção  $\alpha$ -talassêmica de 3,7 kb

Condições de Gap-PCR			Fragmento amplificado (kb)
T (°C)	Tempo	Nº de ciclos	
94	5 min	x1	1,9
94	1 min		
64	1 min	x34	
72	90 seg		2,1
72	10 min	x1	

### 3.3.3 Eletroforese em gel de agarose

Após cada reação de PCR, ou de Gap-PCR, a verificação do sucesso da amplificação foi realizada por eletroforese em gel de agarose.

A eletroforese é um método de separação de moléculas pelo seu peso molecular através da aplicação de corrente elétrica. O DNA apresenta carga negativa devido ao grupo fosfato, desse modo quando lhe é aplicada uma corrente elétrica/voltagem as moléculas migram do respetivo ponto de aplicação em direção ao polo positivo, permitindo uma separação pelo número de pares de bases (pb) (79).

Esta técnica requer a utilização de uma tina de eletroforese contendo tampão estabilizador, Tris-Borato-EDTA concentrado uma vez (TBE 1x) e que permite a distribuição da corrente elétrica gerada por um equipamento, neste caso pelo Power-Pac 300 da BioRad.

O gel usado na eletroforese, onde se aplicam os produtos resultantes da amplificação (8  $\mu$ L) aos quais é adicionado azul de bromofenol ou *loading buffer* (5  $\mu$ L), foi preparado utilizando agarose, tampão TBE 1x e brometo de etídeo (EtBr) (3  $\mu$ L para 100 mL de gel). Este último funciona como agente intercalante de marcação de ácidos nucleicos, o que permite a sua visualização sob UV após eletroforese.

A concentração dos diversos géis de agarose efetuados neste estudo, o tipo de agarose e as condições de eletroforese para as várias regiões génicas estudadas encontram-se descritas na tabela 3.18.

**Tabela 3.18**-Condições de gel e eletroforese aplicadas para o estudo das variantes genéticas estudadas

Gene	Regiões génicas amplificadas	Gel/Tipo de Agarose*	Gel de agarose (% p/v)	Voltagem e tempo de eletroforese
<i>ICAM-1</i>	Exão 2		2%	70V~45min
	Exão 4 a 6		1%	70V~50min
<i>CD36</i>	Região promotora	SeaKem @LE Agarose		70V~45min
	Exão 5		2%	70V~50min
	Exão 10			70V~45min
	STR In3(TG)n	NuSieve™ 3:1 Agarose	4%	60V~120min
<i>HBA</i>	<i>HBA2</i> e <i>HBA1</i>	SeaKem @LE Agarose	1%	70V~90min

\*Da marca Lonza.

Para se verificar o tamanho dos produtos de amplificação, aplicaram-se nos diversos géis de agarose e em paralelo com as amostras, marcadores de peso molecular, como o marcador de 100pb plus DNA Ladder (Bioron) para géis a 2%, marcador 1kb DNA Ladder (BioLabs) para géis a 1% e 50bp DNA Ladder (Invitrogen) no caso da eletroforese para o fragmento correspondente ao STR In3(TG)<sub>n</sub> no gene *CD36*.

O resultado da eletroforese foi acompanhado de um registo fotográfico obtido aquando da exposição do gel à luz UV em câmara Uvitec. Devido à presença do EtBr, as moléculas de DNA, tanto das amostras como dos marcadores de peso molecular, apresentam fluorescência o que permite a sua visualização e registo fotográfico e consequente validação do sucesso da PCR ou Gap-PCR.

### 3.3.4 Sequenciação de Sanger

A técnica de sequenciação de Sanger foi desenvolvida por Sanger e pela sua equipa na década de 1970 e permite a determinação da sequência nucleotídica de uma cadeia de DNA. Atualmente, a metodologia encontra-se automatizada e, neste caso, o princípio da técnica baseia-se na incorporação de dNTPs na síntese de uma nova cadeia de DNA e a incorporação aleatória de didesoxirribonucleótidos trifosfatados (ddNTPs) marcados com um fluoróforo, que após excitação se diferencia na sua cor de emissão, de acordo com o ddNTP que é incorporado (A, C, T ou G). Os ddNTPs não possuem o grupo 3'-OH e impedem a ligação fosfodiéster entre nucleótidos terminando a síntese da cadeia. Desta forma, a polimerização ocorre até que seja incorporado um ddNTP, resultando na síntese de cadeias de DNA que diferem em tamanho apenas num nucleótido entre si (80). Estas são, posteriormente, separadas por capilaridade. Os resultados da sequenciação apresentam-se sob a forma de eletroferogramas que, após leitura, nos indicam a sequência do fragmento de DNA em análise.

Neste trabalho recorreu-se à técnica de sequenciação automática de Sanger partindo de DNA amplificado por PCR. Previamente à execução desta metodologia é necessário executar um passo de purificação dos produtos amplificados, para se obter uma maior eficiência na sequenciação.

No processo de purificação utilizou-se o kit ExoSAP-IT PCR *Product Clean-up* (Applied Biosystems). Este baseia-se na ação enzimática conjunta de uma exonuclease I (Exo) e

de uma fosfatase alcalina (SAP) que hidrolisam *primers* e dNTPs que não foram incorporados no decorrer da PCR. Assim sendo, após a reação de amplificação foi preparada uma mistura contendo 5  $\mu\text{L}$  do DNA amplificado e 2  $\mu\text{L}$  do ExoSAP (conservado a  $-20^{\circ}\text{C}$ ), que foi depois submetida à temperatura ótima de ação enzimática,  $37^{\circ}\text{C}$ , por 15 minutos e em seguida a outros 15 minutos de cessação da ação enzimática, a  $80^{\circ}\text{C}$  (81). A reação foi executada no termociclador T Gradient Thermocycler da Biometra.

Os produtos purificados foram de seguida submetidos a sequenciação, realizada no termociclador de acordo com a mistura reacional descrita na tabela 3.19 e as condições presentes na tabela 3.20. Para tal, 1,5  $\mu\text{L}$  de produto purificado foi adicionado a uma mistura reacional que continha tampão Big Dye (tampão estabilizador da reação), o Big Dye (ambos do kit de terminação Big Dye © Terminator V1.1 Cycle Kit da Applied Biosystems) e o *primer forward* (utilizado na reação PCR) caso o tamanho do fragmento fosse inferior a cerca de 400pb ou ambos os *primers* (*forward* e *reverse*), em misturas singulares para cada, caso o fragmento tivesse um tamanho superior a 700 pb. O Big Dye consiste na solução que contém os dNTPs e ddNTPs necessários à síntese das cadeias (82).

**Tabela 3.19-**Mistura reacional para a realização da Sequenciação de Sanger

Mistura de Sequenciação	
Reagente	Volume por reação ( $\mu\text{L}$ )
dH <sub>2</sub> O	5,25
Tampão Big Dye (5x)	1,75
<i>Primer</i> (2 pmol/ $\mu\text{L}$ )	1
Big Dye	0,5

**Tabela 3.20-**Condições de PCR para a realização de Sequenciação de Sanger

Condições de Sequenciação		
T ( $^{\circ}\text{C}$ )	Tempo	Nº de ciclos
96	4 min	x1
96	10 seg	
55	5 seg	x25
60	4 min	
60	8 min	x1

Após a reação de sequenciação, o produto foi submetido a separação das cadeias por peso molecular efetuada em eletroforese por capilaridade no equipamento 3500 Genetic Analyser, ABI Prism da Applied Biosystems. A separação capilar foi realizada na Unidade de Tecnologia e Inovação (UTI) do INSA.

Os eletroferogramas contendo as sequências resultantes foram analisados recorrendo ao *software* FinchTV v.1.4.0 (Geospiza). Para a análise das sequências, para além do uso do referido *software*, efetuou-se um BLAST para comparação com as *RefSeq* dos genes estudados (Tabela 3.21), disponíveis no Ensembl.

**Tabela 3.21**–Sequências *RefSeq* utilizadas para comparação das sequências obtidas por sequenciação de Sanger

Gene	NG	NM	NP
<i>ICAM-1</i>	NG_012083.1	NM_000201.3	NP_000192.2
<i>CD36</i>	NG_008192.1	NM_000072.3	NP_000063.2

NG–sequência de referência com base na região genómica; NM–sequência de referência com base no RNA codificante da proteína; NP–sequência de referência com base na proteína.

### 3.3.5 Análise de microssatélites In3(TG)n no gene *CD36*

Uma das metodologias realizadas para o estudo do microssatélite (TG)n, presente no terceiro intrão do gene *CD36*, consistiu na análise de marcadores STR a partir do produto amplificado por PCR. Esta técnica permite a determinação do número de repetições dinucleotídicas TG (que podem variar entre oito e 18 repetições). Assim, obtêm-se fragmentos amplificados que irão apresentar tamanhos diferentes consoante o número de repetições em cada alelo (por exemplo, 118 bp equivale a 12 repetições). Sabendo-se o tamanho previsto do produto amplificado e comparando com a *RefSeq* (NM\_000072.3), conclui-se o número de repetições de TG.

A amplificação desta região génica (parte do terceiro intrão) foi realizada com recurso a um *primer forward* marcado com FAM (tabela 7.5, presente no anexo 3), um fluoróforo que permite a emissão de fluorescência aquando excitação, para que fosse possível a análise posterior do fragmento amplificado (condições de PCR presentes nas tabelas 3.9 e 3.11).

Após PCR, 5 µL de cada produto de amplificação foi analisado por eletroforese capilar, no equipamento 3500 *Genetic Analyser da Applied Biosystems*, e em paralelo, foi

aplicado um marcador de peso molecular, neste caso o *GeneScan* TM 500 LIZ® *Size Standard* (leitura de 35 a 500 pb) (83). Este marcador molecular foi o adequado ao tamanho do fragmento alvo e fluoróforo utilizado na amplificação, a sua migração gerou uma reta de calibração a partir da qual o *software* de análise (*GeneMapper Software 6*) calculou o peso molecular do produto amplificado. Os resultados são apresentados em forma de picos que indicam o peso do fragmento e a área do pico que é proporcional ao número de cópias do DNA com o mesmo peso. A análise de separação dos fragmentos por eletroforese capilar foi realizada na UTI.

Para se confirmar que o tamanho molecular que se associou a um número de repetições TG era o correto realizou-se a sequenciação de Sanger em algumas amostras e controlos normais homocigóticos para um determinado número de repetições. Contudo, para ser possível realizar esta validação foi necessário efetuar outra PCR que incluía a região em estudo, mas desta vez fazendo uso de *primers* não marcados com FAM. Assim, foram desenhados novos *primers*, como abaixo descrito, e as condições da PCR utilizadas encontram-se enumeradas nas tabelas 3.10 e 3.11.

### **3.3.5.1 Desenho de *primers* para reações PCR que incluem o STR In3(TG)<sub>n</sub>**

Como foi já mencionado e após tentativa de sequenciar os produtos resultantes da PCR, para amostras homocigóticas quanto ao número de repetições TG (no decorrer da metodologia acima descrita) concluiu-se que o facto do *primer forward* ser marcado com FAM resultava em sequências com muito ruído, tornando impossível a leitura e validação do tamanho do fragmento em relação ao número de repetições. Para resolver este problema, optou-se pelo desenho de um novo par de *primers* que permitiram amplificar um fragmento de 578 pb contendo a região em apreço.

Para o desenho dos *primers* recorreu-se ao *software Primer Blast Tool* do NCBI (<https://www.ncbi.nlm.nih.gov/tools/primer-blast/>).

As condições pretendidas de funcionamento dos *primers* foram: a) Tamanho dos *primers* entre 18 e 22 nucleótidos, com tamanho ótimo de 20 nucleótidos; b) Temperatura de *melting* entre os 52 e 63°C, com uma temperatura ótima de 58°C; c) Diferença máxima entre as temperaturas de *melting* de 5°C; d) Conteúdo GC entre os 40 e os 60%; e) Um

tamanho máximo de produto de amplificação de 700 pb; f) Um número máximo de três bases C/G nos últimos cinco nucleótidos na sequência a 3'. Após apresentação pelo programa dos pares de *primers* resultantes, avaliaram-se as características de cada par e submeteu-se o par escolhido a um BLAST para confirmar a sua especificidade para a região alvo pretendida.

O par de *primers* desenhado e escolhido foi o seguinte:

*Primer forward*: 5'-TGCAACTCTATATGATGCG-3' e *Primer reverse*: 5'-AGCTAATTCGAATAAGCTAGGG-3'.

### 3.4 Tratamento dos dados *in silico*

Sendo um dos objetivos deste projeto o estudo de variantes que possam desempenhar um papel modulador no que toca à suscetibilidade para malária e na drepanocitose, era esperado que a análise de cada variante fosse realizada por forma a averiguar se estas têm impacto na proteína (variantes localizadas nas sequências codificantes dos genes) ou no *splicing* (variantes localizadas em intrões) e que tipo de consequência manifestam. Portanto, para efeitos de estudo *in silico* das variantes estudadas recorreu-se a ferramentas bioinformáticas que permitem avaliar as possíveis consequências patogénicas destas alterações.

#### 3.4.1 varSeak

A ferramenta varSeak (<https://varseak.bio/>), permitiu a análise do impacto de variantes ao nível do *splicing*. De acordo com a variante a analisar, o varSeak fornece informação quanto à localização dessa mesma variante em relação à sequência de referência (HGVS), quanto à sua distância em relação aos locais canónicos de *splicing* e até mesmo se se configuram locais crípticos de *splicing*.

Para além disto, permite uma comparação exata no que toca à funcionalidade destes locais, uma vez que calcula as probabilidades de locais de *splicing* serem funcionais e a diferença entre o local de *splicing* na sequência referência e no local na sequência variante, respetivamente – *score* e *delta score*. Estas informações são apresentadas sob a forma de esquema das sequências onde se apresentam os locais canónicos (aceitadores ou dadores de *splicing*), local de *Branch point* e região de exclusão de AG

(adenina/guanina) (AGEZ). É apresentada também uma tabela, onde se indica a classe de previsão da variante em análise, o *score* e o *delta score*.

O *score* diz respeito à probabilidade prevista de um local de *splicing* ser funcional (valores positivos) ou não (valores negativos), configura valores que podem ir de -100% a +100%. Locais de *splicing* com funcionalidade desconhecida admitem um *score* de 0%. Por outro lado, o *delta score* ( $\Delta score$ ) avalia a diferença entre o *score* do local de *splicing* canônico e o *score* do local de *splicing* apresentado na sequência variante.

A classe de previsão de afetação do *splicing* varia de um a cinco, cada classe é representada por cores diferentes e encontra-se associada à probabilidade de apresentar um efeito no *splicing*. Assim, a classe um significa que a variante não tem efeito, a classe dois informa que provavelmente a variante não apresenta um efeito, a classe três informa que o efeito é desconhecido, a classe quatro é atribuída a variantes que provavelmente apresentam efeito e, por fim, a classe cinco atribui-se a variantes com efeito claro no *splicing*.

O *branch point* diz respeito a um nucleótido – adenina (A) - localizado numa região localizada entre 18 e 40 nucleótidos a montante da extremidade 3' do intrão. Este nucleótido de adenina é seguido de uma sequência polipirimidínica e de um dinucleótido AG a 3'.

A AGEZ, localiza-se entre o *branch point* e o local de *splicing* a 3'. E é uma região caracterizada pela ausência do dinucleótido AG, tal como a própria identificação indica (84).

### 3.4.2 PolyPhen-2

O PolyPhen (<http://genetics.bwh.harvard.edu/pph2/>) analisa o impacto das variantes a nível proteico, isto é, prevê o efeito das alterações aminoacídicas na função e estabilidade das proteínas humanas. Esta análise é realizada com base na estrutura e bioquímica do aminoácido e com base na sua conservação entre as espécies (que possuem homologia para a proteína em estudo) permitindo estimar a probabilidade de uma alteração se configurar prejudicial. Esta ferramenta aplica-se então a variantes que ocorram em regiões codificantes dos genes (exões) e utiliza-se para avaliar, fundamentalmente, alterações *missense*.

O *software* utiliza dois modelos de análise, HumDiv e HumVar:

- i. O modelo HumDiv, é o modelo padrão na avaliação de alelos raros e na análise de seleção natural. Consiste no conjunto de todos os alelos prejudiciais conhecidos como causadores de doenças mendelianas humanas. Apresenta ainda as diferenças entre proteínas humanas e as suas homólogas não prejudiciais, presentes noutros mamíferos.
- ii. O modelo HumVar é o adequado para diagnóstico de doenças mendelianas e requer a distinção entre mutações com efeitos drásticos e outras variantes humanas. Este modelo vai então reunir todas as mutações que causam doença humana, juntamente com SNPs humanos comuns e não prejudiciais.

Aquando da análise é apresentado ao operador um *score* que prevê o impacto de uma substituição na estrutura e função de uma proteína. O *score* varia entre zero e um, sendo que entre zero e 0,5 as variantes são consideradas como benignas e entre 0,5 e um, são consideradas possivelmente prejudiciais (85).

### 3.4.3 SPSS

O *software* IBM SPSS *Statistics* 27 utilizou-se para avaliar a possível associação entre os genótipos e os dados clínicos e hematológicos da população em estudo. Para tal, foi construída uma base de dados, no programa Excel, e esta foi exportada e processada no programa estatístico SPSS. Testou-se primeiramente a distribuição das variáveis pelos testes de normalidade e homogeneidade.

Uma vez que todas as variáveis quantitativas analisadas não seguiam uma distribuição normal, os testes estatísticos de associação realizados posteriormente foram não-paramétricos. Realizou-se o teste de Mann-Whitney-Wilcoxon para avaliar a associação entre duas variáveis e o teste de Kruskal-Wallis para avaliar a associação de mais do que duas variáveis. No caso do estudo de associação entre uma variável quantitativa e uma variável qualitativa recorreram-se a testes de Qui-Quadrado quando todas as frequências esperadas foram superiores a cinco. Quando este pressuposto não se verificou, utilizou-se como alternativa o Teste Exato de Fisher. A significância estatística foi atribuída quando  $p\text{-value} < 0,05$ .

### 3.4.4 Globin Gene Server, Ithanel, Uniprot, Ensembl e Pubmed

O Globin Gene Server (<https://globin.bx.psu.edu/>) e o Ithanel (<https://www.ithanel.eu/db/ithagenes>) são duas bases de dados que incidem sobre sequências e variantes de DNA, respetivamente envolvidas na produção de hemoglobina ou que dão origem a hemoglobinopatias (quer sejam talassémias ou variantes de hemoglobina).

A base de dados Uniprot (<https://www.uniprot.org/>) permite o acesso a informações sobre sequências genéticas e proteicas, informação em relação a modelos funcionais e estrutura dessas mesmas proteínas. Fornece uma informação detalhada de cada molécula. Neste trabalho foi utilizada para a caracterização dos genes e das proteínas resultantes, CD36 e ICAM-1.

O Ensembl (<https://www.ensembl.org/index.html>) é um *browser* para genomas de vertebrados. Fornece informações sobre genómica comparativa, evolução, variação de sequências e regulação transcricional. Possui anotação de genes, permite o acesso a alinhamentos múltiplos, contém informação sobre doenças e prevê funções regulatórias.

O PubMed (<https://pubmed.ncbi.nlm.nih.gov/>) apresenta o estado-de-arte de uma panóplia de temáticas ligadas às ciências biomédicas e ciências da vida, entre outras. Apresenta os fundamentos necessários à compreensão da temática e assuntos abordados neste projeto/dissertação.



## 4 Resultados e Discussão

Os dados hematológicos, bioquímicos e clínicos da população pediátrica, cujos representantes legais consentiram que participassem no estudo, foram recolhidos em consulta, protocolada de seis em seis meses ao longo de dois anos, concretizando assim um estudo longitudinal. Na primeira consulta, foram recolhidos dados demográficos (como idade e sexo) e informação retrospectiva, como idade da primeira manifestação drepanocítica e número de episódios de malária até ao momento dessa primeira consulta. Este último dado foi obtido através dos tutores das crianças, sem confirmação molecular, pelo que constitui um dado subjetivo.

Durante as consultas, para além do exame clínico que envolvia a observação da eventual presença de esplenomegalia, hepatomegalia e icterícia, foram recolhidas amostras biológicas (sangue em EDTA) às crianças participantes que se encontravam em período estável (sem crise drepanocítica). Essas amostras foram utilizadas para a realização da caracterização dos parâmetros hematológicos (Hb, VGM, HGM, contagem de eritrócitos, reticulócitos, leucócitos e plaquetas) e bioquímicos [desidrogenase láctica (LDH) e bilirrubina total]. Posteriormente, e partindo do sangue periférico colhido, foram aplicadas gotas de sangue nos *guthrie cards* (ponto 3.2.2 do capítulo Materiais e Métodos) para posterior deteção do parasita (por diagnóstico molecular por PCR) e extração do DNA humano para a realização da genotipagem dos genes de interesse por PCR.

Os dados hematológicos, bioquímicos e clínicos utilizados neste trabalho referem-se aos que foram recolhidos antes dos participantes iniciarem o tratamento da drepanocitose com HU.

### 4.1 Caracterização demográfica, hematológica, bioquímica e clínica da população estudada

A população estudada é constituída por 65 crianças, 33 (50,8%) do sexo feminino e 32 (49,2%) do sexo masculino, com idades compreendidas entre os 3 e os 12 anos (média de  $7 \pm 2,8$  anos). Todas estas crianças residem em Luanda ou no Bengo, Angola. Uma vez que não foram encontradas diferenças genéticas significativas entre as crianças residentes nas diferentes localidades, estas foram tratadas como um todo.

Todos os participantes apresentam drepanocitose, confirmada através de focagem isoelétrica ou eletroforese de hemoglobinas, ou seja, são homozigóticos para a mutação c.20A>T (rs334) no gene da beta-globina (*HBB*).

Em média, na população estudada, a primeira manifestação drepanocítica revelou-se aos 16 meses de idade. No entanto, a idade da primeira manifestação revelou-se bastante variável pois esta tanto foi observada logo aos 2 meses de idade como apenas aos 96 meses (8 anos). A dactilite foi a primeira manifestação mais frequente (66,2%), seguida de crises dolorosas (27,7%) e de anemia (6,2%). Resultados estes expectáveis, uma vez que a dactilite é, geralmente, a mais precoce e comum manifestação nos doentes com anemia falciforme (37,86).

A drepanocitose caracteriza-se por ser uma anemia hemolítica crónica. Assim, na drepanocitose ocorre eliminação aumentada de eritrócitos (hemólise), tanto a nível intravascular como extravascular (86). Esta última afeta especialmente alguns órgãos como o baço e o fígado, originando, respetivamente, esplenomegalia e hepatomegalia. Para além disso, a taxa elevada de hemólise intravascular origina aumento da bilirrubina sérica e a consequente deposição em tecidos e observação de icterícia (pele e olhos de cor amarelada). Estes indicadores clínicos de hemólise anormalmente elevada foram observados nos nossos participantes. No final do estudo longitudinal concluiu-se que das 65 crianças, nove (13,8%) apresentaram esplenomegalia, 27 (41,5%) hepatomegalia e 32 (49,2%) icterícia. De entre as crianças que apresentaram icterícia, em 31 esta foi classificada, pela pediatra, como ligeira a moderada e numa como intensa.

Uma outra consequência da hemólise aumentada pode ser observada a nível hematológico, pela presença de anemia (baixa concentração de hemoglobina) e diminuição da concentração de eritrócitos no sangue. Estes fenótipos anómalos foram observados e encontram-se apresentados na tabela 4.1.

**Tabela 4.1**-Estatística descritiva dos parâmetros demográficos, hematológicos e bioquímicos da população em estudo

	N*	Mínimo	Máximo	Mediana	Média	Desvio padrão
<b>Idade ao integrar o estudo (anos)</b>	65	3	12	6	7	2,8
<b>Idade da 1ª manifestação drepanocítica (meses)</b>	65	2	96	10	16	17,0
<b>Hb (g/dL)</b>	65	5,5	9,9	7,2	7,3	0,9
<b>VGM (fL)</b>	65	60,3	97,3	78,3	77,6	7,2
<b>HGM (pg)</b>	65	18,8	30,6	25,8	25,6	2,6
<b>Eritrócitos (x10<sup>12</sup>/L)</b>	65	2,1	4,8	2,8	2,9	0,5
<b>Reticulócitos (%)</b>	65	3,1	25,6	9,6	10,0	3,8
<b>Plaquetas (x10<sup>9</sup>/L)</b>	65	171,2	770,7	451	444,1	152,6
<b>Leucócitos (x10<sup>9</sup>/L)</b>	65	7,2	29,7	13,2	13,9	4,3
<b>LDH (U/L)</b>	57	246	683	424	428	123
<b>Bilirrubina Total (mg/dL)</b>	63	0,2	5,1	1,1	1,3	0,9

\*N=número de doentes analisados em cada parâmetro.

Sendo a nossa população de estudo constituída por crianças em idade pediátrica seria importante confrontar os dados hematológicos com valores de referência para uma população pediátrica de preferência angolana, no entanto estes valores não se encontram estipulados. Dada esta lacuna os dados hematológicos apresentados na tabela 4.1 foram confrontados com o descrito como referência para a população portuguesa em geral, numa faixa etária semelhante. Foram usados os valores de referência publicados na Norma nº 063/2011 da Direção-Geral de Saúde (DGS), apresentados no anexo 4.

Assim, todas as crianças analisadas apresentam anemia, pois os níveis de Hb observados (o valor máximo é de 9,9 g/dL) encontram-se bastante abaixo dos considerados no intervalo de referência na referida norma (11-14 g/dL, na idade dos 6 meses aos 11 anos). Todas as crianças apresentam baixa contagem de eritrócitos (média de  $2,9 \pm 0,5 \times 10^{12}/L$ ) quando este parâmetro é confrontado com o intervalo de referência publicado na referida norma ( $3,80-5,40 \times 10^{12}/L$ , na idade dos 6 meses aos 11 anos).

Assim podemos concluir, tal como esperado para doentes drepanocíticos, que as crianças estudadas apresentam anemia, que é observada pelo nível baixo de Hb assim como pela baixa contagem de eritrócitos.

No que toca aos índices hematimétricos, a população estudada apresenta um valor médio de HGM ( $25,6 \pm 2,6$  pg) abaixo do esperado, quando confrontado com o intervalo de referência da norma da DGS, 26,0 a 34,0 pg, para uma idade superior a 5 anos. Este valor é sugestivo da presença de uma hipocromia (geralmente, caracterizada por valores inferiores a 27 pg). A hipocromia é uma consequência hematológica comum em doentes portadores de talassémia, podendo aqui ser justificada pela elevada co-herança de  $\alpha$ -talassémia, nomeadamente da deleção de 3,7kb nos genes *HBA2* e *HBA1* (33). O estudo desta população relativamente à presença e frequência da deleção  $\alpha$ -talassémica será abordado no ponto 4.2.

Contrariamente aos valores da HGM, o valor médio do VGM ( $77,6 \pm 7,2$  fL) para a população estudada encontra-se inserido no intervalo de valores de referência (entre 72,0 e 86,6 fL) para uma população pediátrica com idades entre os 6 meses e os 11 anos (estipulado pela DGS). No entanto, o valor do VGM mínimo determinado para a população estudada é de 60,3 fL, este abaixo do valor mínimo de referência. Valores abaixo dos 70 fL são associados a microcitose, fenótipo este que poderá estar relacionado com talassémia. Assim, os valores mais baixos de VGM poderão ser justificados, tal como para o HGM, pela co-herança de  $\alpha$ -talassémia.

A população estudada apresenta um elevado número de reticulócitos ( $10,0 \pm 3,8\%$ ). O número de reticulócitos é um biomarcador indireto de hemólise, desta forma podemos concluir que nesta população existe uma anemia hemolítica, tal como esperado para uma população de doentes drepanocíticos e coerentemente com o observado clinicamente (esplenomegalia, hepatomegalia e icterícia). Para além disso, podemos afirmar que há uma resposta da medula óssea à anemia pois observa-se um aumento da taxa de eritropoiese libertando um maior número de reticulócitos (eritrócitos imaturos) para a corrente sanguínea.

A nossa população apresenta, em média, uma contagem elevada de plaquetas ( $444,1 \pm 152,6 \times 10^9/L$ ) e de leucócitos ( $13,9 \pm 4,3 \times 10^9/L$ ), quando comparadas com os intervalos de referência descritos na referida norma (respetivamente, 140-440  $\times 10^9/L$  na idade superior a dois meses; e 4,5-13,0  $\times 10^9/L$ , na faixa etária entre os 5 e os 11 anos). Os valores elevados de plaquetas e leucócitos podem dever-se a processos inflamatórios, uma vez que a drepanocitose é caracterizada por uma inflamação crónica conjuntamente com

a anemia hemolítica. Por outro lado, as plaquetas e os leucócitos, são constituintes celulares associados à vaso-oclusão, muito comum na drepanocitose (36).

Quanto aos biomarcadores da hemólise, LDH e bilirrubina total, os valores determinados para a população estudada foram comparados com os valores de referência estipulados pela *Roche Diagnostics GmbH*, os quais consideram que um valor de LDH superior a 332 U/L e um valor de bilirrubina total superior a 1,1 mg/dL são indicadores de hemólise grave, em doentes em idade pediátrica até aos 12 anos (87). Assim sendo, a população estudada apresenta valores médios para esses parâmetros (LDH =  $428 \pm 123$  U/L e bilirrubina total =  $1,3 \pm 0,9$  mg/dL) que caracterizam a ocorrência de hemólise grave, como seria esperado sendo uma população drepanocítica caracterizada por anemia hemolítica, tal como foi já mencionado.

Um estudo de Silva. *et al* reuniu dados demográficos, clínicos, hematológicos e bioquímicos de 70 crianças drepanocíticas numa faixa etária semelhante à da população aqui estudada (3 aos 16 anos), residentes em Portugal, mais precisamente, na área metropolitana de Lisboa, mas com ancestralidade Africana (na sua maioria, com progenitores provenientes de Angola) (88). Relativamente aos parâmetros hematológicos, como Hb e contagem de eritrócitos, a nossa população estudada neste trabalho apresenta valores inferiores aos determinados para a população apresentada no referido artigo: Hb = 7,3 *versus* 8,0 g/dL e número de eritrócitos = 2,9 *versus* 3,0  $\times 10^{12}$ /L. O mesmo sucede ao compararmos os valores dos índices hematimétricos determinados na nossa população *versus* o descrito por Silva *et al*: VGM = 77,6 *versus* 81,3 fL e HGM = 25,6 *versus* 26,9 pg. Assim, podemos concluir que a nossa população de estudo, residente em Angola, apresenta uma anemia mais acentuada, mais microcítica e mais hipocrômica do que uma população de crianças residentes em Portugal com a mesma patologia. Estas diferenças entre duas populações drepanocíticas dentro de uma mesma faixa etária podem ser atribuídas a fatores ambientais. A escassez de alimentos que forneçam as quantidades necessárias de ferro, por exemplo, podem levar ao agravamento da anemia na população que reside em Angola, para além disto, a carência de ferro pode dever-se igualmente à permanente exposição a coinfeções por parasitas, como por exemplo, o parasita da espécie *P. falciparum*, agente etiológico da malária. Este último metaboliza o ferro presente nos eritrócitos em cristais de hemozoína, agravando a carência de ferro dos doentes anémicos e drepanocíticos (89). Também o acesso facilitado a cuidados médicos

e de saúde e acompanhamento por parte de profissionais poderá ter um impacto no tratamento da anemia e até mesmo da drepanocitose no que toca às crianças residentes em Lisboa, melhorando o quadro clínico e os valores hematológicos dos doentes.

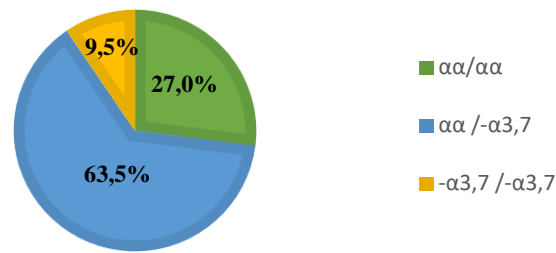
Por outro lado, a população estudada no presente estudo apresenta, em média, valores mais elevados de contagem de plaquetas e leucócitos do que a população de ascendência africana residente em Lisboa, descrita por Silva *et al*: plaquetas = 444,1 *versus* 404,1  $\times 10^9/L$  e leucócitos = 13,9 *versus* 12,6  $\times 10^9/L$  (88). Sendo estes parâmetros hematológicos associados a processos inflamatórios como foi já mencionado, pode-se prever que estes valores derivem, não apenas da inflamação crónica que caracteriza a drepanocitose, mas também de respostas do sistema imunitário a coinfeções. A população pediátrica residente em Angola encontra-se frequentemente exposta a infeções parasitárias, não só de infeções por malária, mas também por outras doenças tropicais (como esquistossomose e filarioses, por exemplo). Para além disso encontra-se também sujeita a uma elevada taxa de coinfeção por HIV e tuberculose (90). Encontram-se, portanto, mais suscetíveis a um agravamento das respostas inflamatórias, o que justifica os elevados valores de plaquetas e leucócitos.

#### 4.2 Co-herança de $\alpha$ -talassémia

O diagnóstico da presença da deleção talassémica de 3,7kb, que abrange parcialmente os genes *HBA2* e *HBA1*, foi realizado através da técnica de Gap-PCR, tal como descrito no ponto 3.3.2 do capítulo Materiais e Métodos.

A genotipagem desta deleção nos DNAs em estudo foi executada previamente, no âmbito da dissertação da colega Isabel Germano (78). Os resultados por ela obtidos foram, no presente projeto, adequados aos objetivos propostos e tratados estatisticamente. É também importante mencionar que para dois DNAs não houve sucesso ao realizar a técnica de diagnóstico da deleção talassémica em causa.

Assim das 63 crianças estudadas com sucesso, 17 (27,0%) apresentam o genótipo *wild-type* ( $\alpha\alpha/\alpha\alpha$ ; ausência da deleção), 40 (63,5%) apresentam a deleção em heterozigotia ( $\alpha\alpha/-\alpha^{3,7}$ ) e seis (9,5%) apresentam a deleção em homozigotia ( $-\alpha^{3,7}/-\alpha^{3,7}$ ). Na figura 4.1 encontram-se representados graficamente estes resultados.



**Figura 4.1**-Representação gráfica da frequência dos genótipos referentes à deleção  $\alpha$ -talassêmica de 3,7kb na população drepanocítica analisada (n=63).

Observa-se nesta população pediátrica, com drepanocitose, que a  $\alpha$ -talassémia é muito prevalente, uma vez que 46 (73,0%) participantes apresentam pelo menos um alelo com a deleção de 3,7kb. A frequência alélica para  $-\alpha^{3,7}$  revelou-se ser de 41,3% (52 alelos apresentam a deleção e 74 são normais).

Através do cálculo do equilíbrio de Hardy-Weinberg (HW), pode-se concluir que a proporção dos diferentes genótipos referentes à deleção  $\alpha$ -talassêmica de 3,7kb não se encontra em equilíbrio [ $\chi^2$  (gl=1, N=63) = 6,04; p=0,0139]. Seria de esperar que o número de doentes com o genótipo heterozigótico ( $\alpha\alpha/-\alpha^{3,7}$ ) fosse menor, pois o número observado foi igual a 40 e o número de doentes esperado era de 30,5 (de acordo com o cálculo do equilíbrio de HW). Este resultado poderá ser devido à dimensão reduzida da amostra ou a um efeito de seleção positiva da deleção em heterozigotia, para a população drepanocítica estudada. No sentido de esclarecer este assunto, pretendeu-se comparar os nossos resultados com outros reportados na literatura.

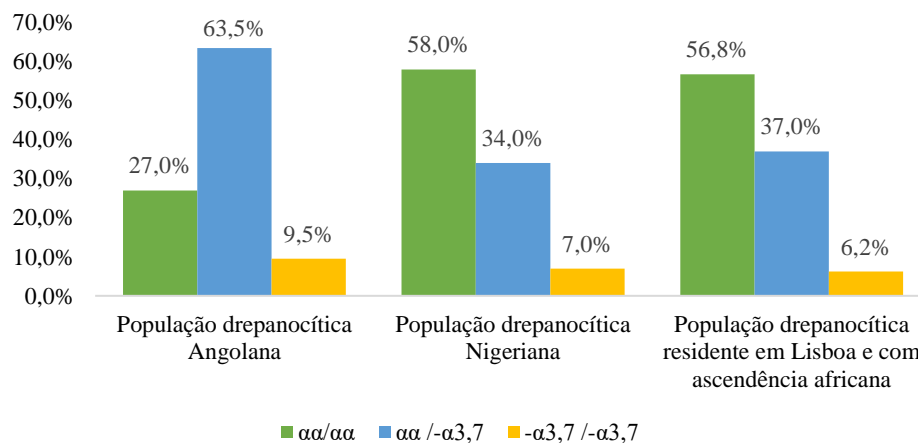
Estes resultados de elevada prevalência de  $\alpha$ -talassémia conjuntamente com a drepanocitose eram esperados, uma vez que a nossa população estudada é de ascendência africana e residente em África, continente este com elevada frequência do alelo mutado resultante da deleção talassêmica. Na literatura encontra-se reportado que, pelo menos, 30,0% das populações de doentes com drepanocitose apresentem as mutações  $\alpha$ -talassémicas (pelo menos um dos alelos mutados para esta hemoglobinopatia) (91). Para além disto sabe-se que a presença de pelo menos um dos alelos mutados em países tropicais pode ser igual ou superior a 80,0% (92).

Um estudo de Olatunya *et al* que incidiu sobre uma população de 100 crianças e adolescentes drepanocíticos (com idades entre os 2 e os 21 anos e uma média de 9 anos)

e de origem nigeriana, concluiu que destes 34 (34,0%) eram heterozigóticos para a mutação  $-\alpha^{3,7}$  e sete (7,0%) eram homozigóticos (93).

Noutro estudo que envolveu 81 crianças drepanocíticas, com uma faixa etária semelhante à da população estudada, de ascendência africana mas residentes em Lisboa [estudo este ainda não publicado mas englobado no projeto que deu origem ao artigo já mencionado de Silva *et al* (88)] concluiu-se que 46 (56,8%) apresentavam o genótipo *wild-type* para a deleção  $\alpha$ -talassémica de 3,7kb, enquanto 30 (37,0%) apresentavam o genótipo heterozigótico para a mutação e cinco (6,2%) tratavam-se de doentes homozigóticos para a deleção.

Na figura 4.2 pode-se observar a diferença entre o número de participantes drepanocíticos que apresentam cada genótipo para a deleção talassémica de 3,7kb nas três populações com drepanocitose acima mencionadas, a angolana analisada neste estudo, a nigeriana estudada no artigo de Olatunya *et al* (93) e a de ascendência africana residente em Lisboa.



**Figura 4.2**-Representação gráfica das frequências dos genótipos relativos à deleção de 3,7kb ocorrida nos genes *HBA2* e *HBA1* na população drepanocítica angolana analisada no nosso estudo, na população drepanocítica nigeriana visada no artigo de Olatunya *et al.* (93) e na população drepanocítica residente em Lisboa e com ascendência africana referida no artigo de Silva *et al* (88).

Tanto a população drepanocítica nigeriana, estudada no artigo de Olatunya *et al*, como a população drepanocítica residente em Lisboa apresentam uma maior frequência do genótipo *wild-type* em  $\alpha$ -globina em comparação com a nossa população angolana (respetivamente, 58,0% e 56,8% *versus* 27,0%). O contrário ocorre quando são comparados os números referentes ao genótipo heterozigótico para a deleção talassémica:

na nossa população observa-se em 63,5% dos participantes, enquanto na população nigeriana em 34,0% e na população drepanocítica residente em Lisboa e com ascendência africana em 37,0%.

Relativamente ao número de participantes homozigóticos ( $-\alpha^{3,7}/-\alpha^{3,7}$ ), as três populações apresentam semelhanças, sendo que na nossa população estudada 9,5% dos participantes são homozigóticos, na população nigeriana estes representam 7,0% e na população residente em Lisboa representam 6,2% dos doentes.

No caso da população drepanocítica nigeriana, a proporção entre os diversos genótipos referentes à deleção  $\alpha$ -talassémica de 3,7kb encontra-se em equilíbrio de HW [ $\chi^2_{(gl=1, N=100)} = 0,42$ ;  $p=0,518$ ]. Também para a população de ascendência africana e residente em Lisboa, o mesmo é verificado [ $\chi^2_{(gl=1, N=81)} = 0,001$ ;  $p=0,971$ ].

No entanto, pode-se afirmar que todas as populações mencionadas (drepanocítica angolana, nigeriana e residente em Lisboa, mas com ascendência africana) vão ao encontro da frequência alélica (do alelo mutado  $-\alpha^{3,7}$ ) reportada na literatura para populações drepanocíticas (respetivamente, 41,3%, 24,0% e 24,7%). Como foi já mencionado, é estimado que pelo menos cerca de 30% das populações drepanocíticas apresentem, pelo menos, um dos alelos mutados pela deleção talassémica de 3,7kb (91).

Assim, fica por explicar a discrepância entre a frequência dos genótipos da  $\alpha$ -globina entre as três populações de crianças com drepanocitose referidas. Uma das razões poderá ser o pequeno número da amostra nos três estudos.

Sabe-se que, em termos gerais, a co-herança da  $\alpha$ -talassémia é benéfica para os doentes com drepanocitose (54,94). No entanto, são poucos os estudos epidemiológicos e clínicos publicados que se foquem no estudo da co-herança da  $\alpha$ -talassémia (deleção de 3,7kb) em populações pediátricas com essa patologia. Os estudos abordados apresentam um número de participantes reduzido o que implica elevada incerteza na determinação das frequências alélicas e genotípicas e em relação à observação da vantagem/desvantagem desta co-herança. Assim, deveriam ser realizadas mais investigações neste âmbito, com um maior número de amostra populacional, para ser possível esclarecer este assunto.

### 4.3 Pesquisa e caracterização de variantes nos genes *CD36* e *ICAM-1* na população drepanocítica

O estudo genético efetuado nos genes *CD36* e *ICAM-1* foi executado através da técnica de PCR, sequenciação de Sanger e análise de microssatélites, tal como se encontra descrito, respetivamente, nos pontos 3.3.1., 3.3.4 e 3.3.5 (Materiais e Métodos).

No estudo do gene *CD36* foram analisadas as regiões consideradas de maior interesse: região promotora, intrões três, quatro e cinco, e exão dez. No estudo do gene *ICAM-1* foram analisados os exões dois, quatro, cinco e seis, e os intrões dois e cinco. A análise das respetivas sequências permitiu detetar 15 variantes diferentes apresentadas nas tabelas 4.2 e 4.3.

**Tabela 4.2-**Variantes genéticas detetadas no gene *CD36* na população drepanocítica em estudo

Gene	Região no Gene	ID*	Nomenclatura HGVS**			Tipo
			Genómica	Genética	Proteica	
<i>CD36</i>	Promotor	rs1984112	g.80613604 A>G	n.230+1122 5 A>G	-	Intrónica
		rs1413661	g.80613687 G>C	c.- 184+11308 G>C	-	Intrónica
	Intrão 3	rs3138813	g.80647283_80 647284 ins(TG)n	In3(TG)n	-	<i>Short Tandem Repeat</i>
	Intrão 4	rs3211891	g.80661032 T>C	c.282- 31T>C	-	Intrónica
		rs3211892	g.80661053 A>G	c.282- 10A>G	-	<i>Splice polypyrimidin e tract variant</i>
	Intrão 5	rs3211893	g.80661212 T>C	c.429+2T>C	-	Intrónica
	Exão 10	rs3211938	g.80671133 T>G	c.975T>G	p.Tyr325 *	<i>Nonsense/ stop gained</i>

\*ID, identificação; \*\*HGVS, nomenclatura de acordo com a *Human Genome Variation Society*.

**Tabela 4.3-** Variantes genéticas detetadas no gene *ICAM-1* na população drepanocítica em estudo

Gene	Região no Gene	ID*	Nomenclatura HGVS**			Tipo
			Genômica	Genética	Proteica	
<i>ICAM-1</i>	Exão 2	rs5491	g.10274864 A>T	c.167 A>T	p.Lys56Met	<i>Missense</i>
	Intrão 2	rs5030352	g.10275067 C>G	c.331+39 C>G	-	<i>Intrónica</i>
	Exão 4	rs1799969	g.10284116 G>A	c.721 G>A	p.Gly241Arg	<i>Missense</i>
		rs5494	g.10284241 C>T	c.846 C>T	p.Asp282=	<i>Sinónima</i>
	Exão 5	rs1801714	g.10284532 C>T	c.1055 C>T	p.Pro352Leu	<i>Missense</i>
	Intrão 5	rs5496	g.10284771 G>A	c.1181-12 G>A	-	<i>Intrónica</i>
	Exão 6	rs5497	g.10284792 G>A	c.1190 G>A	p.Arg397Gln	<i>Missense</i>
		rs5498	g.10285007 A>G	c.1405 A>G	p.Lys469Glu	<i>Missense</i>

\*ID, identificação; \*\*HGVS, nomenclatura de acordo com a *Human Genome Variation Society*.

O número de casos detetados na população estudada, bem como as frequências genotípicas e alélicas para cada variante, encontram-se apresentadas na tabela 4.4 para as variantes detetadas no gene *CD36* e na tabela 4.5 para as variantes detetadas no gene *ICAM-1*.

**Tabela 4.4-** Variantes detetadas no gene *CD36* e respetivas frequências genotípicas e alélicas, na população estudada

Gene	Variante genética	Genótipos		Alelos
		Nº de casos detetados	Frequência genotípica (%)	Frequência alélica (%)
<i>CD36</i>	c.-184+11225A>G (rs1984112)	AA - 35	53,8	A - 74,6
		AG - 27	41,5	G - 25,4
		GG - 3	4,6	
	c.-184+11308G>C (rs1413661)	GG - 3	4,6	G - 30,8
		GC - 34	52,3	<b>C - 69,2</b>
		CC - 28	43,0	
	c.282-31T>C (rs3211891)	TT - 49	75,0	T - 85,4
		TC - 13	20,0	C - 14,6
		CC - 3	5,0	
	c.282-10A>G (rs3211892)	AA - 8	12,3	A - 26,9
		AG - 19	29,2	<b>G - 73,1</b>
		GG - 38	58,5	
	c.429+2T>C (rs3211893)	TT - 60	92,3	T - 94,6
		TC - 3	4,6	C - 5,4
		CC - 2	3,0	
c.975T>G (rs3211938)	TT - 48	74,0	T - 86,9	
	TG - 17	26,0	G - 13,0	
	GG - 0	0		

Os valores a negrito representam casos em que a frequência alélica do alelo variante é superior à do alelo *wild-type*.

No caso das variantes encontradas no gene *CD36*, pode-se constatar que as variantes para as quais o alelo variante apresenta uma frequência alélica superior ao alelo *wild-type* são a rs1413661 e a rs3211892. No caso da variante rs1413661, a maioria da população estudada revelou um genótipo heterozigótico (GC), 52,3%, enquanto no caso da variante rs3211892, a maioria da população apresentou um genótipo homozigótico para o alelo variante (GG), 58,5%.

**Tabela 4.5**-Variantes detetadas no gene *ICAM-1* e respectivas frequências genotípicas e alélicas, na população estudada

Gene	Variante genética	Genótipos		Alelos
		Nº de casos detetados	Frequência genotípica (%)	Frequência alélica (%)
<i>ICAM-1</i>	c.167A>T (rs5491)	AA - 33	50,8	A - 70,8 T - 29,2
		AT - 26	40,0	
		TT - 6	9,2	
	c.331+39C>G (rs5030352)	CC - 0	0	C - 18,7 <b>G - 81,3</b>
		CG - 24	37,0	
		GG - 41	63,0	
	c.721G>A (rs1799969)	GG - 65	100	<b>G - 100</b>
	c.846C>T (rs5494)	CC - 51	78,5	C - 88,5 T - 11,5
		CT - 13	20,0	
		TT - 1	1,5	
	c.1055C>T (rs1801714)	CC - 63	96,9	C - 98,5 T - 1,5
		CT - 2	3,1	
		TT - 0	0	
	c.1181-12 G>A (rs5496)	GG - 56	86,2	G - 92,3 A - 7,7
		GA - 8	12,3	
AA - 1		1,5		
c.1190G>A (rs5497)	GG - 52	80,0	G - 90 A - 10	
	GA - 13	20,0		
	AA - 0	0		
c.1405A>G (rs5498)	AA - 55	84,7	A - 91,5 G - 8,5	
	AG - 9	13,8		
	GG - 1	1,5		

Os valores a negrito representam casos em que a frequência alélica do alelo variante é superior à do alelo *wild-type*.

No caso das variantes encontradas no gene *ICAM-1*, a variante rs5030352 revelou para além de uma frequência do alelo variante superior ao alelo *wild-type* (81,30% *versus* 18,70%), uma frequência genotípica nula para o alelo *wild-type* em homozigotia. Ainda sobre a variante rs5030352, a frequência do genótipo com o alelo variante em homozigotia (GG), 62,5%, representa a maioria da população estudada.

Em nenhum dos doentes estudados foi detetada a alteração c.721G>A (rs1799969). Esta observação vai ao encontro do que se encontra descrito pelo Ensembl (1000 Genomes

Project Phase 3), onde a frequência do alelo variante é de apenas 1,0% para a população africana em geral e para a população ioruba a frequência é nula.

Pretendeu-se investigar se as frequências genotípicas observadas na nossa população drepanocítica angolana seriam semelhantes às descritas para a população africana em geral e para a população ioruba (no caso destas duas últimas, recorrendo aos dados presentes no Ensembl, fornecidos pelo 1000 Genomes Project Phase 3). Foi usado o modelo dominante que confronta as frequências dos genótipos que contêm a variante com a frequência do genótipo *wild-type*. A análise estatística de comparação das referidas frequências permitiu observar que as variantes cujas frequências genotípicas diferem significativamente entre a nossa população de estudo e pelo menos uma das outras populações analisadas são a rs3211891 e a rs3211938, no gene *CD36*, e a rs5030352 e a rs5497, no gene *ICAM-1*, tal como se encontra apresentado nas tabelas 4.6 e 4.7.

**Tabela 4.6**-Variantes genéticas cujas frequências genotípicas diferem entre a população drepanocítica estudada (angolana) e a população africana em geral

Gene	ID	Genótipos*	Frequência (%)		<i>p-value</i> ***
			População Estudada	População Africana**	
<b><i>CD36</i></b>	rs3211891	TT	75	88	<b>0,028</b>
		TC+CC	25	12	
<b><i>ICAM-1</i></b>	rs5497	GG	80	93	<b>0,012</b>
		GA+AA	20	7	
	rs5030352	CC	0	7	<b>0,014</b>
		CG+GG	100	93	

A negrito apresentam-se os valores de *p* que se revelaram estatisticamente significativos. \*De acordo com o modelo dominante;\*\*Dados apresentados no Ensembl, consultado a 5/2/2023; \*\*\*Teste exato de Fisher.

**Tabela 4.7**-Variantes genéticas cujas frequências genotípicas diferem entre a população drepanocítica estudada (angolana) e a população ioruba

Gene	ID	Genótipos *	Frequência (%)		<i>p-value</i> ***
			População Estudada	População Ioruba**	
<b><i>CD36</i></b>	rs3211938	TT	74	52	<b>0,002</b>
		TG+GG	26	48	
	rs3211891	TT	75	93	<b>&lt;0,001</b>
<b><i>ICAM-1</i></b>	rs5497	TC+CC	25	7	<b>0,012</b>
		GG	80	93	
	rs5030352	GA+AA	20	7	<b>&lt;0,001</b>
		CC	0	12	
		CG+GG	100	88	

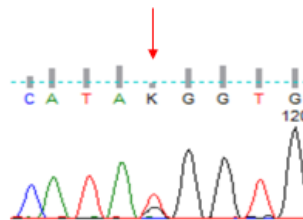
A negrito apresentam-se os valores de *p* que se revelaram estatisticamente significativos. \*De acordo com o modelo dominante;\*\*Dados apresentados no Ensembl, consultado a 5/2/2023; \*\*\*Teste exato de Fisher.

Assim, apesar de terem sido detetadas na nossa população drepanocítica angolana, 15 variantes diferentes nos dois genes em análise, verificou-se que apenas quatro diferiam significativamente, em termos de frequências genótípicas, entre a nossa população e a população africana ou a população ioruba. Este facto foi considerado relevante para serem estas variantes as selecionadas para um estudo mais aprofundado, como a seguir se descreve. Por outro lado, optou-se também por efetuar um estudo mais aprofundado para as variantes que apresentam elevada frequência em populações de ascendência africana, como a rs5491 e, do mesmo modo, a variantes que podem revelar um impacto significativo na expressão dos genes estudados, rs3138813, rs1984112 e rs1413661.

#### 4.3.1 Variantes genéticas detetadas no gene *CD36*

##### 4.3.1.1 Variante *CD36*: c.975T>G; (p.Tyr325\*)

A variante *CD36*: c.975T>G (rs3211938) traduz-se na alteração de uma timina (T) por uma guanina (G) no 975º nucleótido da região codificante do referido gene (Fig.4.3). Ocorre no décimo exão do gene e resulta na alteração de TAT>TAG, ou seja, no ganho de um codão de stop prematuro no mRNA.



**Figura 4.3**-Resultado parcial da sequenciação de Sanger do décimo exão do gene *CD36* revelando a alteração c.975T>G em heterozigotia.

Este mRNA contendo um codão de stop prematuro dá origem a uma proteína truncada que é eliminada após a síntese (95). Assim, ocorre uma perda quantitativa de função da proteína, ou seja, uma deficiência em *CD36* do tipo I, caracterizada por ausência de expressão em células como plaquetas, monócitos e macrófagos (62,96,97). Esta deficiência apenas é notória quando o alelo variante se encontra em homozigotia (5).

Pensa-se que este tipo de deficiência em *CD36* esteja envolvida em mecanismos de proteção contra a infeção por malária, apesar das opiniões serem controversas(11,95,98). Como foi já mencionado, a variante rs3211938, em homozigotia, ao proporcionar a

eliminação pós-síntese da proteína CD36, apresenta relativamente à malária um impacto positivo para o hospedeiro. A deficiência de CD36 em monócitos e macrófagos implica a ausência de um dos principais ligandos humanos da PfEMP1 (fator de virulência do parasita da espécie *P. falciparum*) (99). No entanto, as opiniões relativas a esta variante são contrárias entre autores pois a ausência de expressão da proteína CD36 em macrófagos compromete o mecanismo de fagocitose macrofágica. Sendo assim, tanto apresenta consequências negativas para o hospedeiro relativamente à resposta imunitária que não será eficiente, não ocorrendo a eliminação dos eritrócitos sinalizados após reconhecimento do epitopo estranho do parasita, PfEMP1, como se poderá configurar vantajosa para o plasmódio precisamente pela inibição da fagocitose de eritrócitos infetados (24,62,98,100). Pelas diferentes perspectivas de estudo mencionadas, a variante não foi ainda associada nem à proteção nem ao risco de manifestações graves ou duradouras da malária, nem à própria suscetibilidade de desenvolver a infeção.

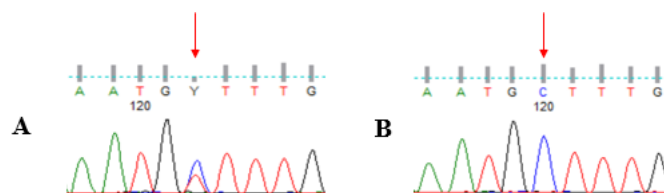
Por outro lado, é referido na literatura que as variantes do gene *CD36* se encontram associadas à variação do risco para doenças cardiometabólicas e para a síndrome metabólica (101,102). A função transmembranar e lipídica da proteína CD36 implica que possíveis alterações na sua estrutura ponham em causa o normal funcionamento do metabolismo lipídico (60,103). As doenças cardiometabólicas definem patologias que afetam o sistema circulatório, como a diabetes, hipercolesterolemia, hipertrigliceridemia e hipertensão arterial. Já a síndrome metabólica é definida como o conjunto de fatores de risco para as doenças cardiovasculares, como a coexistência de resistência à insulina, obesidade, diabetes do tipo II, hipertensão e dislipidemia (104).

A variante em questão (rs3211938), expressa em heterozigotia, foi já associada a um aumento nos níveis da lipoproteína de alta densidade (101,105,106), efeito protetor em relação à síndrome metabólica e capacidade, como biomarcador em conjunto com a proteína reativa C, de determinar a propensão para doenças cardiovasculares (107). Desta forma, as hipóteses de estudo não se encontram diretamente relacionadas com a drepanocitose, mas sim com outros fatores que podem agravar algumas manifestações clínicas desta. A desregulação cardiometabólica, por exemplo, pode agravar a ocorrência de AVCs.

Um outro resultado que obtivemos diz respeito à comparação das frequências genotípicas da variante *CD36*: c.975T>G (rs3211938) na população drepanocítica angolana estudada com a população africana e a ioruba (resultados apresentados nas tabelas 4.3 e 4.4). Daqui conclui-se que: *i*) a nossa população apresenta tendencialmente uma maior frequência dos genótipos que contêm a variante em comparação com a população africana em geral; e *ii*) contrariamente, a nossa população apresenta uma frequência estatisticamente menor ( $p = 0,002$ ) dos genótipos que contêm a variante em comparação com a população ioruba. Era esperada uma elevada frequência desta variante uma vez que esta alteração genética é descrita como a variante *nonsense* mais comum em indivíduos de ancestralidade africana (100). Pensa-se que a variante seja selecionada não só pela endemicidade da malária, mas também devido a outras patologias relacionadas com desregulação metabólica como doenças cardiovasculares e a síndrome metabólica. Por outro lado, a maior expressão do alelo variante na população ioruba, foi igualmente mencionada na literatura, pondo-se a questão do porquê da pressão seletiva desta variante (100). Tal como mencionado, existem outras patologias que podem selecionar esta alteração e que podem, portanto, explicar a maior expressão do alelo na população ioruba, apesar de não ser ainda certo o que exercerá essa pressão.

#### 4.3.1.2 Variante *CD36*: c.282-31T>C

Esta variante (rs3211891) ocorre no quarto intrão do gene *CD36* e consiste na substituição de uma timina (T) por uma citosina (C), alteração esta que ocorre 31 nucleótidos antes do início do quinto exão (c.282-31 T>C). Na figura 4.4 são apresentados dois exemplos da sua deteção em dois dos indivíduos analisados.



**Figura 4.4**-Resultado parcial da sequenciação de Sanger do quarto intrão do gene *CD36* revelando a alteração c.282-31 T>C. Em (A) é apresentado um caso de heterozigotia; em (B) é apresentado um caso de homozigotia.

Uma vez que se trata de uma variante localizada na extremidade 3' do intrão, pretendeu-se investigar se esta teria algum efeito nefasto ao nível do *splicing*. Assim, foi realizado o seu estudo *in silico* recorrendo ao *software* varSeak (Fig.4.5).



**Figura 4.5**-Resultado parcial da análise *in silico* do impacto no *splicing* da variante *CD36*: c.282-31T>C pelo *software* varSeak. U2-Branch point; AGEZ- Zona de exclusão AG; 1-local canônico de *splicing*.

O resultado da análise *in silico* revelou que a variante é classificada como benigna e de classe 1, ou seja, não é previsível que afete o normal processo de *splicing*. No entanto, encontra-se apenas a quatro nucleótidos (a montante) de uma zona *branch point*, U2. O *branch point* é constituído por um nucleótido (adenina) imprescindível para a ligação da maquinaria do spliceossoma e, deste modo, imprescindível para a sinalização do local de aceitação de *splicing* no intrão (108,109).

Deste modo, aparentemente, a variante rs3211891 não possui impacto no *splicing*, nem altera o seu local canônico. No entanto, ocorrendo tão próxima do local U2 responsável pela sinalização da excisão intrónica, poderá levantar a hipótese de não ocorrer a normal sinalização e como consequência este intrão poderá ficar retido no mRNA e ser codificante, alterando a proteína CD36. Esta hipótese poderia ser confirmada prosseguindo, por exemplo, para estudos de mRNA de indivíduos homocigóticos.

A frequência dos genótipos com o alelo variante (TC+CC) é significativamente superior na nossa população drepanocítica estudada do que na população africana ( $p = 0,028$ ) e ioruba ( $p < 0,001$ ) (tabelas 4.3 e 4.4). Devido a estas diferenças significativas levantamos a hipótese que possa existir um fator ambiental que exerce uma seleção sobre esta variante na nossa população estudada. Alternativamente, a variante pode também ter sido selecionada pela drepanocitose.

Através de uma revisão bibliográfica apenas foi detetado um trabalho que envolveu o estudo da variante rs3211891 no âmbito de uma outra patologia. No trabalho de Farook *et al*, foram estudados indivíduos de ascendência mexicana em risco de desenvolver

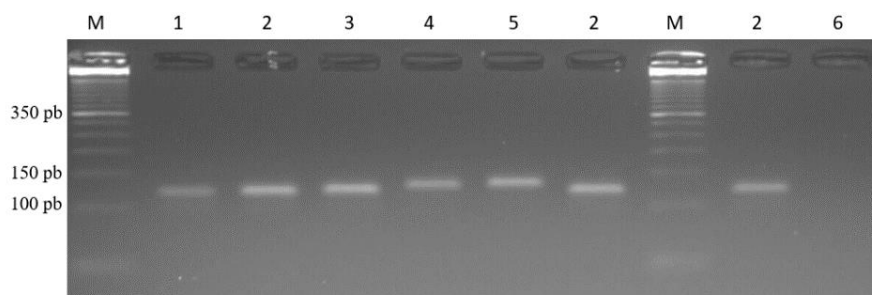
síndrome metabólica, com o objetivo de estudar possíveis determinantes genéticas de risco para a patologia mencionada (104). Esta variante foi uma das analisadas e não foi associada significativamente com essa síndrome.

Não existem atualmente mais estudos publicados que investiguem o papel desta variante nos indivíduos que a expressam, seria por isso interessante prosseguir com os estudos funcionais acima referidos, por análise da estrutura e quantificação do mRNA correspondente ao alelo com a variante, e também ao nível da quantificação da proteína CD36.

#### 4.3.1.3 Variante *CD36*: STR In3(TG)*n*

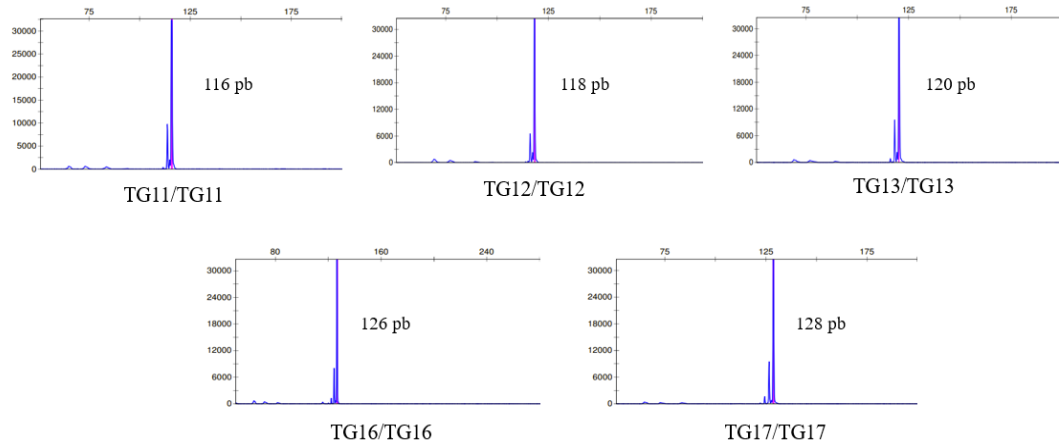
Esta variante In3(TG)*n* (rs3138813) ocorre no terceiro intrão do gene *CD36*, caracteriza-se pelo diferente número de repetições do dinucleótido TG e é classificado como um *Short Tandem Repeat* (STR). O número de repetições encontra-se descrito no NCBI ([https://jul2023.archive.ensembl.org/Homo\\_sapiens/Variation/Explore?r=7:80646759-80647783;v=rs3138813;vdb=variation;vf=730172513](https://jul2023.archive.ensembl.org/Homo_sapiens/Variation/Explore?r=7:80646759-80647783;v=rs3138813;vdb=variation;vf=730172513)) como variando entre oito e 20, sendo que 12 repetições é considerado o número *wild-type*.

As técnicas executadas para o estudo do STR In3(TG)*n* encontram-se descritas no ponto 3.3.5 dos Materiais e Métodos. Em primeiro lugar, foi realizada a amplificação por PCR de várias amostras para identificar, por sequenciação de Sanger, futuros controles homocigóticos a usar posteriormente na análise por microssatélites. Na figura 4.6 são mostrados alguns exemplos dos resultados obtidos.



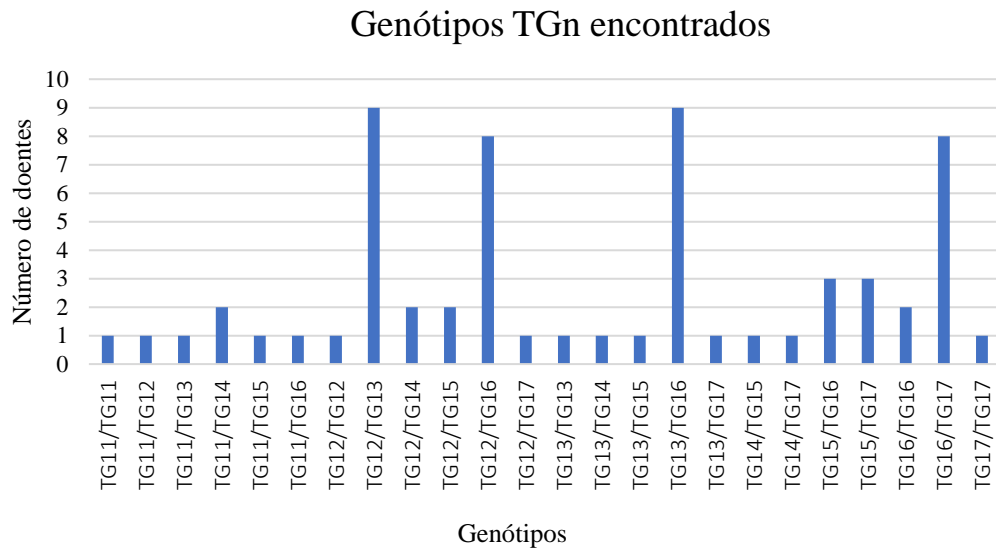
**Figura 4.6**-Eletroforese em gel de agarose Seakem e NuSieve (1:3; 4% p/v), dos produtos de PCR obtidos na genotipagem da variante In3(TG)*n* em diferentes casos de homocigotia. M-marcaador de peso molecular (50bp) da marca Invitrogen; 1-Homocigotia TG11/TG11; 2-Homocigotia TG12/TG12; 3-Homocigotia TG13/TG13; 4-Homocigotia TG16/TG16; 5-Homocigotia TG17/TG17; 6-Controlo negativo.

Após a identificação de casos homocigóticos para diferentes números de repetições TG, submetem-se os produtos da amplificação à técnica de análise de marcadores de DNA (descrita também no ponto 3.3.5 dos Materiais e Métodos). Os resultados obtidos através da realização desta metodologia para os casos homocigóticos encontram-se apresentados na figura 4.7.



**Figura 4.7**-Perfil de casos homocigóticos para a variante  $In3(TG)_n$  após análise de marcadores de DNA. Genótipos representados por TG seguido do número de repetições do dinucleótido.

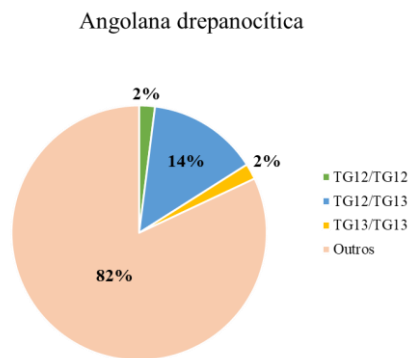
Deste modo cada genótipo e número de repetições TG foi associado a um pico com um peso molecular específico. Após esta análise primária, procedeu-se então à genotipagem das restantes amostras de DNA da população estudada. A diversidade dos vários genótipos encontrados na população estudada encontra-se representado na figura 4.8, bem como o número de doentes que os possuem.



**Figura 4.8**-Genótipos encontrados aquando do estudo da variante In3(TG)n, na população angolana drepanocítica. Os genótipos são representados pelo número de repetições do dinucleótido TG.

Interpretando a figura 4.8, na nossa população drepanocítica, os genótipos mais frequentes são TG12/TG13 (n=9, 13,8%), TG13/TG16 (n=9, 13,8%), TG12/TG16 (n=8, 12,3%) e TG16/TG17 (n=8, 12,3%). Relativamente à frequência alélica, o alelo mais frequente é o TG16 (25,0%), seguido do alelo TG13 (21,0%) e do alelo TG12 (19,0%). O grande número de diferentes genótipos na população estudada (grande variabilidade) não era expectável, sendo que apenas 2,0% da população apresenta o genótipo considerado *wild-type* (TG12/TG12).

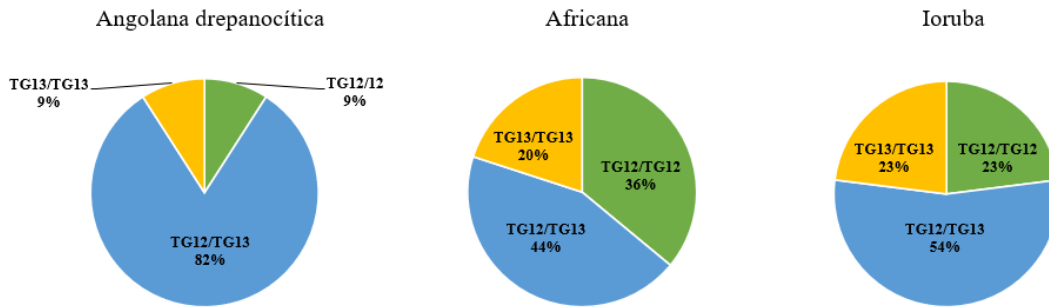
Na base de dados Ensembl apenas existe informação relativa às frequências dos genótipos TG13/TG13, TG12/TG12 e TG12/TG13, pelo que na figura 4.9 encontram-se representadas as frequências genotípicas destes para a população angolana drepanocítica.



**Figura 4.9**-Representação gráfica das frequências genóticas para a variante In3(TG)n na população angolana drepanocítica (estudada neste trabalho).

Através da figura 4.9 pode-se perceber que os genótipos considerados como os mais comuns pelo Ensembl, não o são na nossa população de estudo. Dado o resultado apresentado, seria importante confrontar os dados com uma população angolana normal [sem a mutação drepanocítica (rs344)] e perceber se numa população drepanocítica genótipos que originam um transcrito anormal de CD36 (isoforma de CD36 sem o quarto e quinto exões transcritos) representam algum tipo de benefício e até mesmo se sofrem seleção. Sendo que apenas três doentes (4,6% da população estudada) possuem genótipos que originam o transcrito normal, TG11/TG11, TG11/TG12 e TG12/TG12.

Por forma a comparar a representatividade dos genótipos ditos mais comuns (TG12/TG12, TG12/TG13 e TG13/TG13) nas três populações (drepanocítica, africana e ioruba), calcularam-se as percentagens destes na população estudada excluindo os restantes genótipos. As proporções dos três genótipos para as diferentes populações encontram-se representadas na figura 4.10.



**Figura 4.10**-Representação gráfica das frequências genótípicas dos genótipos TG12/TG12, TG12/TG13 e TG13/TG13 para a variante In3(TG)n nas populações angolana drepanocítica (estudada), africana e ioruba.

Comparando as frequências dos três genótipos agora em análise concluímos que a população angolana drepanocítica apresenta frequências mais baixas para os genótipos TG12/TG12 (9,0%) e TG13/TG13 (9,0%) quando comparadas com as frequências destes nas outras populações (africana e ioruba). Contrariamente, a frequência do genótipo TG12/TG13 na população angolana drepanocítica (tendo em conta que se consideraram os 11 doentes com estes três genótipos como 100%) revelou-se superior à das restantes populações, 82,0% *versus* 44,0% e 54,0%.

Curiosamente, nas três populações comparadas, o genótipo com maior representatividade populacional é o TG12/TG13, isto é, em que o alelo *wild-type* com 12 repetições de TG se encontra em heterozigotia com um alelo com 13 repetições que dá origem a um transcrito anormal.

Na literatura, esta variante STR In3(TG)n encontra-se relacionada com a deficiência do tipo II da proteína CD36. Esta deficiência caracteriza-se pela ausência de expressão da proteína em plaquetas e é mais frequente em indivíduos asiáticos e de ascendência africana (61,97). Consoante o número de repetições pode ocorrer o *skipping* do quarto e do quinto exões (no momento de *splicing*), dando origem a uma proteína truncada. De acordo com a literatura, indivíduos com 12 repetições, portanto com um número canónico de repetições, apresentam o transcrito normal, isto é, contendo o quarto e quinto exões para além dos restantes. Pelo contrário, indivíduos com o número de repetições diferente de 12 irão apresentar o transcrito resultante do *skipping* mencionado, com falta das sequências do quarto e quinto exões, o que dá origem a uma isoforma da proteína CD36 que se traduz, posteriormente, numa deficiência na sua expressão em plaquetas (62,95).

Assim, as variantes do STR In3(TG)n afetam o transcrito do gene *CD36* o que poderá ter consequências no processo de ligação entre a proteína CD36, expressa nas células do hospedeiro e a PfEMP1, expressa à superfície dos eritrócitos infetados pelo plasmódio, uma vez que o quinto exão desse gene codifica umas das regiões reconhecidas pela proteína do parasita. As variantes In3(TG)n podem então, em teoria, configurar-se como protetoras do risco de infecção por malária.

Num estudo conduzido por Omi e colegas foi efetuada a genotipagem desta variante numa população tailandesa com malária (62). O estudo revelou que, à semelhança da população angolana drepanocítica analisada neste trabalho, os alelos mais frequentes foram o TG12 (32,0%), o TG13 (31,5%) e o TG16 (24,6%). Estes resultados podem sugerir que perante a alta suscetibilidade para a infecção por malária de ambas as populações [a Tailândia é também uma região endémica para a malária (1)], os alelos mencionados poderão ser benéficos e conferir proteção.

Para além disto, Omi e colegas concluíram que o alelo TG11 em homozigotia ou heterozigotia com o alelo TG12 originava o transcrito normal, sem o *skipping* de exões (62). Tendo em conta esta informação, concluiu-se que apenas 6,0% da nossa população estudada expressa o transcrito normal da proteína CD36. Sendo assim, 94,0% dos doentes apresentam genótipos que originam transcritos que não apresentam o quarto e o quinto exões e consequentemente dão origem à isoforma de CD36 e à sua deficiência em certo tipo de células.

No estudo mencionado a redução do risco na infecção por malária não se terá mostrado significativa, no entanto, isso não significará que esta variante não possa apresentar um impacto benéfico na infecção malárica para a nossa população angolana em análise. Isto porque a população tailandesa encontra-se exposta a uma espécie diferente de plasmódio (*P. vivax*) o que significa que se tem de ter em conta os locais de ligação deste parasita, preferências celulares (no hospedeiro), entre outros fatores.

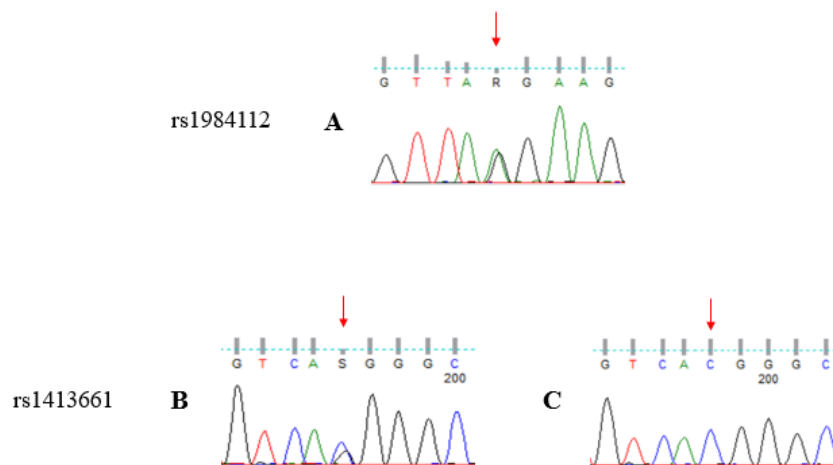
Existem ainda poucos estudos publicados que abordem esta variante In3(TG)n no gene *CD36* apesar desta apresentar inúmeros impactos, tanto a nível de expressão celular devido à deficiência do tipo II a que se encontra associada, como ao nível da ligação com a PfEMP1. Será, por isso, importante a continuação do seu estudo, sobretudo ao nível funcional.

#### 4.3.1.4 Região promotora do gene *CD36*: Variantes c.-184+11225A>G e c.-184+11308G>C

O estudo das regiões promotoras dos genes possui grande importância uma vez que albergam locais de ligação de fatores de transcrição. Alterações na sequência de referência localizadas nesta região podem definir e alterar a expressão génica.

Neste sentido, foi estudada a região promotora do gene *CD36*. O estudo prático desta região foi executado recorrendo a reações de PCR e posterior sequenciação de Sanger, tal como se encontra descrito nos pontos 3.3.1 e 3.3.1.1 dos Materiais e Métodos.

A variante c.-184+11225A>G (rs1984112) consiste na substituição de uma adenina (A) por guanina (G), numa região não codificante do gene *CD36*. Já a variante c.-184+11308G>C (rs1413661) consiste na substituição de uma guanina (G) por uma citosina (C), a apenas 82pb da variante rs1984112. As alterações mencionadas encontram-se representadas na figura 4.11.



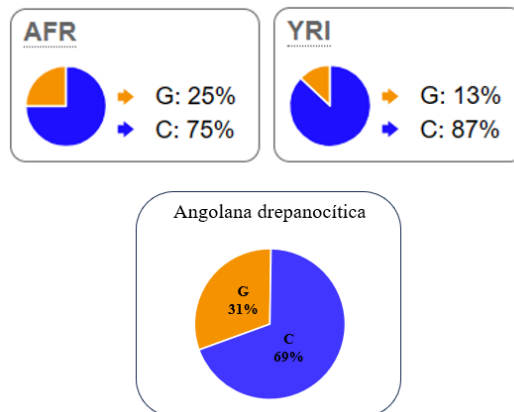
**Figura 4.11**-Resultado parcial da sequenciação de Sanger da região promotora do gene *CD36* revelando a alteração c.-184+11225A>G em heterozigotia (A) e a alteração c.-184+11308G>C em heterozigotia (B) e em homozigotia (C).

No sentido de investigar o papel modulador destas variantes na drepanocitose e na malária e o seu impacto a nível da regulação da expressão do gene *CD36*, o estudo destas variantes foi efetuado recorrendo a literatura e comparando as frequências alélicas e genotípicas entre as populações de ascendência africana.

Relativamente à frequência genotípica para os três genótipos possíveis de cada uma destas variantes, tem-se que para a variante c.-184+11225A>G a frequência de genótipos com o alelo\_G é semelhante em todas as populações (angolana drepanocítica, africana e ioruba), sendo quase 50%.

Relativamente à variante rs1413661, c.-184+11308G>C, a frequência genotípica do genótipo GG é baixa na nossa população angolana com drepanocitose (5%), assim como na população africana (6%) e ioruba (1%).

Comparando as frequências alélicas para este SNP nas três populações, representadas na figura 4.12, pode-se observar que a população angolana drepanocítica apresenta uma menor frequência alélica para o alelo\_C, levantando a hipótese de ocorrer uma pressão seletiva negativa da drepanocitose relativamente a este alelo. Hipótese esta que seria importante estudar e confrontar com dados para uma população angolana sem drepanocitose.

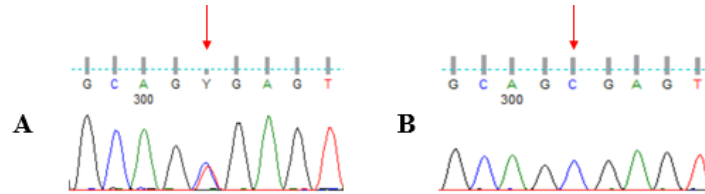


**Figura 4.12-**Representações gráficas das frequências alélicas (rs1413661) nas populações africana (AFR), ioruba (YRI) e angolana drepanocítica (população estudada).

Estes nossos resultados e a hipótese levantada vem ao encontro do anteriormente publicado, uma vez que o SNP rs1413661\_alelo C, foi associado com níveis de hemoglobina mais baixos em doentes com drepanocitose, o que se traduz por uma anemia mais grave (49).

#### 4.3.1.5 Variante *CD36*: c.429+2T>C

A variante c.429+2T>C (rs3211893) consiste na alteração do nucleótido de timina (T) por citosina (C), no segundo nucleótido do quinto intrão do gene *CD36* (Fig.4.13).



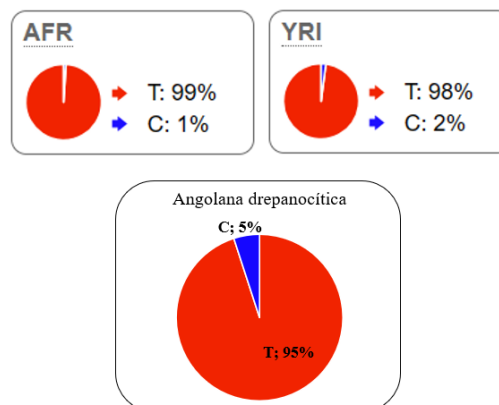
**Figura 4.13**-Resultado parcial da sequenciação de Sanger do quinto intrão do gene *CD36* revelando a alteração c.429+2T>C. Em (A) é apresentado um caso de heterozigotia; em (B) é apresentado um caso de homozigotia.

Sendo uma variante localizada apenas a dois nucleótidos do local de aceitação do *splicing* do quinto exão, colocou-se a hipótese de que esta pudesse afetar o mecanismo normal de *splicing* e, por consequência, pudesse alterar a estrutura da proteína. Para averiguar este possível impacto, recorreram-se a estudos *in silico* realizados no varSeak. Os resultados mostraram que esta variante é prejudicial para o *splicing*, tendo sido classificada de classe 5 (Fig.4.14). A substituição afeta precisamente o dinucleótido que constitui o local canônico de *splicing*, o que resulta numa perda de função deste último. A consequência da perda de função é o *skipping* do quinto exão do gene, isto é, este exão deixa de integrar o mRNA acabando por não ser traduzido. A proteína *CD36* é então traduzida sem a presença do quinto exão, isto é, ocorre a junção *in frame* do quarto e do sexto exões.



**Figura 4.14**-Resultado parcial da análise *in silico* do impacto no *splicing* da variante *CD36*: c.429+2T>C pelo software varSeak.

Tal como foi já mencionado (ver ponto 1.6.2), o quinto exão codifica o local de ligação entre a proteína PfEMP1 e a proteína CD36, deste modo esta variante poderá ter impacto nesta ligação. Em teoria, seria expectável que esta alteração conferisse alguma proteção contra a infeção por malária e, portanto, pudesse surgir em elevada frequência nas populações africanas. No entanto, e pelo contrário, a frequência alélica do alelo variante\_C, nas populações de ascendência africana é baixa (figura 4.15, informação presente no Ensembl e resultante da genotipagem efetuada). Uma explicação poderá ser a seleção negativa como consequência da deficiência em proteína CD36 que ocorre, uma vez que a proteína perde um dos exões, é sintetizada na sua forma truncada e posteriormente poderá ser eliminada. A eliminação afeta a sua expressão em células como plaquetas e macrófagos (61). A proteína CD36 auxilia no mecanismo de fagocitose macrofágica (como foi já mencionado, ponto 1.6.2) e, portanto, na resposta do sistema imunitário, não sendo benéfica a sua ausência e, por isso, não sendo esta variante selecionada pela malária.



**Figura 4.15**-Representações gráficas das frequências alélicas (rs3211893) nas populações africana (AFR), ioruba (YRI) e angolana drepanocítica (população estudada).

O impacto e papel modulador desta variante ainda está por desvendar. Existe, até à data, apenas um artigo publicado que associa esta variante à deficiência da proteína CD36, quando em heterozigotia composta com outra variante detetada no gene *CD36* (detetada em apenas um indivíduo) (110). A ausência de estudos sobre esta variante deve-se, possivelmente, à baixa frequência em que ocorre nas populações. De facto, segundo o Ensembl, apenas populações africanas apresentam o alelo\_C.

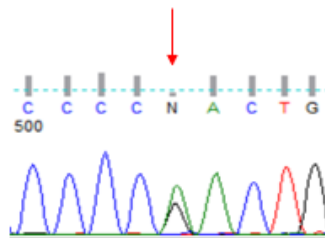
Seria importante o estudo do papel modulador da variante rs3211893, uma vez que esta

apresenta um grande impacto no *splicing*. Estudos funcionais poderiam desvendar as consequências provocadas ao nível proteico que esta variante representa.

### 4.3.2 Variantes genéticas detetadas no gene *ICAM-1*

#### 4.3.2.1 Variante *ICAM-1*: c.1190G>A; (p.Arg397Gln)

A variante c.1190G>A (rs5497) ocorre no sexto exão do gene *ICAM-1*, é uma variante *missense* e caracteriza-se pela alteração de uma guanina (G) por uma adenina (A) (Fig.4.16). Ao nível proteico a alteração traduz-se numa substituição aminoacídica de arginina (Arg) por glutamina (Gln) no resíduo 397.

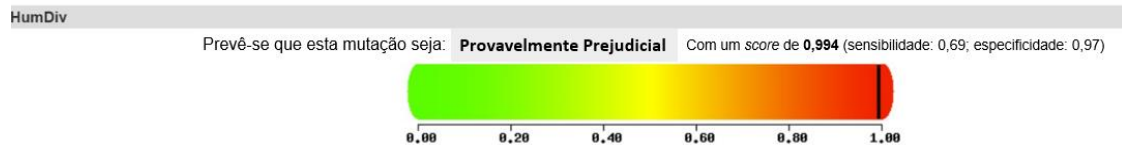


**Figura 4.16**-Resultado parcial da sequenciação de Sanger do sexto exão do gene *ICAM-1* revelando a alteração c.1190G>A em heterozigotia.

Tanto a Arg como a Gln são aminoácidos polares (hidrofilicos), no entanto apresentam propriedades ácido/base distintas. A Arg (PubChem CID 6322) é um aminoácido alcalino, por outro lado, a Glu (PubChem CID 5961) é um aminoácido ácido. Esta substituição pode alterar a carga da proteína sintetizada e, conseqüentemente o seu ponto isoelétrico, o que pode alterar a solubilidade da proteína. Como consequência pode ocorrer desnaturação proteica ou um rearranjo da sua estrutura nativa e, conseqüentemente, perda da sua função e de interações com outras moléculas e células (111,112).

Com o objetivo de se estudar o impacto proteico desta variante, recorreu-se a ferramentas bioinformáticas para realizar os estudos *in silico*.

Recorrendo à ferramenta bioinformática PolyPhen, para o estudo do impacto a nível proteico, a variante rs5497 foi classificada como provavelmente prejudicial (*probably damaging*), com um *score* de 0,994 (Fig.4.17), quando avaliada de acordo com o modelo HumDiv (descrito no capítulo dos Materiais e Métodos, no ponto 3.4.2).



**Figura 4.17**-Representação gráfica dos resultados obtidos através do estudo *in silico* (Poly-Phen) para a variante ICAM-1, p.Arg397Gln.

A classificação atribuída à variante pelo *software* Poly-Phen vai ao encontro da análise bioquímica que se realizou anteriormente. Será, por isso, relevante perceber o impacto e de que forma esta alteração na proteína afeta a sua estrutura e função.

O sexto exão do gene *ICAM-1* codifica o quinto domínio extracelular da proteína ICAM-1. Não são conhecidos ligandos específicos para este domínio, no entanto sabe-se que este se encontra associado a uma função estabilizadora e de dimerização da estrutura da proteína ICAM-1 (67,68). Assim, dada a localização da variante p.Arg397Gln neste domínio, podemos levantar a hipótese de que se poderá originar a destabilização da estrutura da proteína e afetar a sua dimerização e a sua capacidade de estabelecer ligações intercelulares. A dimerização nesta proteína aumenta a capacidade de adesão, pelo que se esta for comprometida irá não só interferir na sua função recetora como na sua intervenção no processo de transmigração leucocitária (67).

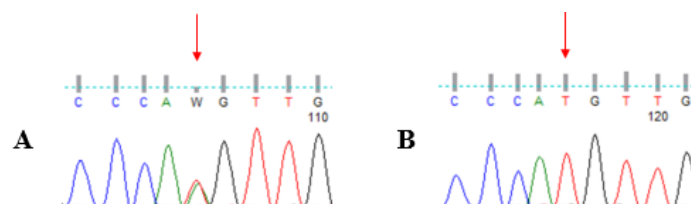
Na nossa população drepanocítica estudada a frequência de génotipos com a variante é superior à frequência na população africana em geral e da população ioruba (sendo esta diferença entre populações estatisticamente significativa, tabelas 4.6 e 4.7). É necessário salientar que na população estudada não foi identificado nenhum indivíduo homozigótico para a variante. Sabendo que as populações africanas se encontram expostas à infeção pelo plasmódio, tal como a nossa população estudada, esta diferença significativa e seleção da variante pode prender-se com outros fatores para além da exposição à malária. Sendo a nossa população selecionada pela presença de uma patologia, isto é, constituída apenas por doentes drepanocíticos, a maior frequência deste alelo variante (em

heterozigotia) nesta população sugere que este poderá traduzir-se num efeito benéfico relativamente à drepanocitose ou, até mesmo, relativamente à drepanocitose acoplada com a constante exposição à malária. A destabilização mencionada da estrutura da proteína e consequente diminuição da capacidade de adesão pode traduzir-se num benefício relativamente à ocorrência de crises dolorosas e vaso oclusão na drepanocitose ou diminuir a ocorrência de sequestro de eritrócitos parasitados aquando da infeção pelo plasmódio.

Tanto quanto sabemos existem poucos estudos publicados que reportem os efeitos prejudiciais (ou outros) desta variante. De facto, esta variante apenas foi mencionada num estudo realizado em gémeos monozigóticos, onde se pretendeu investigar os perfis genéticos possivelmente associados a imunodeficiência (21). Foi também mencionada num estudo que pretendia investigar haplótipos do gene *ICAM-1* relacionados com a exposição à malária a um nível global, neste estudo os autores apenas concluíram (tal como no presente projeto) que a variante é classificada como provavelmente prejudicial pelo Poly-Phen (22). Assim seria de todo o interesse prosseguir a investigação desta variante de *ICAM-1* mutada para validar os resultados obtidos *in silico* e compreender os seus possíveis efeitos patogénicos e funcionais.

#### 4.3.2.2 Variante *ICAM-1*: c.167A>T (p.Lys56Met)

A variante c.167A>T (rs5491), ocorre no segundo exão do gene *ICAM-1* e consiste na alteração de uma adenina (A) por uma timina (T) (Fig.4.18), o que se traduz (ao nível proteico) na substituição aminoacídica do resíduo 56 de uma lisina (Lys) por uma metionina (Met). É conhecida, comumente, pela variante Kilifi. Nome atribuído por ter sido primeiramente descrita, em 1997, na população de Kilifi, no Quênia (113).

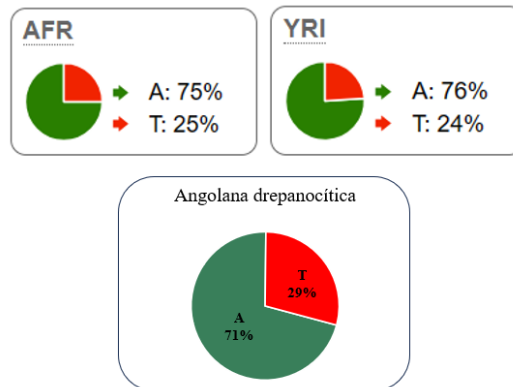


**Figura 4.18**-Resultado parcial da sequenciação de Sanger do segundo exão do gene *ICAM-1* revelando a alteração c.167A>T. Em (A) é apresentado um caso de heterozigotia; em (B) é apresentado um caso de homozigotia.

No gene *ICAM-1*, o segundo exão codifica o primeiro domínio extracelular da proteína e, portanto, esta alteração afeta esse domínio. Sabe-se que a Lys (PubChem CID 5962) é um aminoácido básico e polar (hidrofílico), no entanto a Met (PubChem CID 6137) é um aminoácido apolar (hidrofóbico). Aminoácidos hidrofóbicos localizam-se, geralmente, no interior das estruturas proteicas. A substituição de Lys por Met no primeiro domínio extracelular da proteína ICAM-1 poderá traduzir-se numa mudança conformacional da sua estrutura apresentando um impacto nas suas interações celulares e com ligandos (111,112). Este domínio é o recetor humano para o rinovírus humano e para a PfEMP1, deste modo a alteração na sequência aminoacídica pode interferir na afinidade das ligações que se estabelecem entre esta molécula de adesão e outras células ou recetores (67). Nomeadamente, interferir na ligação da PfEMP1 à ICAM-1, afetando, por sua vez, o mecanismo de adesão ao endotélio pelos eritrócitos parasitados (aquando de uma infeção por malária).

A investigação dos efeitos desta variante tem revelado resultados controversos e até aos dias de hoje não é ainda consensual o seu efeito (114). No artigo de Fernandez-Reyes em 1997, no qual foi feita a primeira descrição desta variante, o alelo kilifi foi descrito como responsável pela suscetibilidade para o desenvolvimento de malária cerebral grave. Por outro lado, em 1999, Kun estudou grupos-caso constituídos por indivíduos originários de outras regiões africanas e também asiáticas (115). A conclusão do referido artigo foi a de que a variante kilifi não teria qualquer efeito significativo na suscetibilidade para a malária grave. Apesar desta conclusão, os autores referiram que as ambiguidades nas conclusões referentes a este alelo teriam de ser analisadas tendo, também, em conta fatores ambientais e a estirpe de plasmódio prevalente em cada região estudada.

Relativamente às frequências alélicas do alelo *wild-type* e do alelo kilifi (de acordo com os resultados da genotipagem e-também de acordo com os dados do Ensembl) (Fig.4.19), pode-se observar que as três populações apresentam frequências muito semelhantes, como seria de esperar, uma vez que este alelo se encontra amplamente presente em populações de ascendência africana (115).



**Figura 4.19**-Representações gráficas das frequências alélicas (rs5491) nas populações africana (AFR), ioruba (YRI) e angolana drepanocítica (população estudada).

#### 4.4 Estudos de Associação

Com o objetivo de se perceber se alguma das variantes genéticas detetadas nos genes *CD36*, *ICAM-1* e *HBA* se encontrava associada à variabilidade fenotípica tanto da malária como da drepanocitose, foram realizados estudos de associação recorrendo ao SPSS (tal como explicado no ponto 3.4.3 dos Materiais e Métodos).

Serão apresentadas associações significativas ( $p < 0,05$ ) e associações que apresentem tendência para melhorar ou piorar os fenótipos.

Não foram encontradas associações entre o sexo dos doentes e os diferentes parâmetros estudados.

##### 4.4.1 Variantes genéticas e a variabilidade fenotípica da drepanocitose

A drepanocitose caracteriza-se, como foi já mencionado, por uma inflamação e anemia hemolítica crónica. Como tal, alterações genéticas que melhorem os parâmetros hematológicos, índices hematimétricos e dados clínicos serão fatores modificadores benéficos para o prognóstico da doença.

Nas tabelas 4.8 a 4.12 encontram-se apresentadas as associações estatisticamente significativas que foram determinadas entre as variantes genéticas e os dados clínicos e parâmetros hematológicos/hematimétricos e bioquímicos no âmbito da drepanocitose.

**Tabela 4.8**-Associações significativas entre a média da idade da primeira manifestação drepanocítica e os diferentes genótipos em rs5491

Fenótipo clínico	Genótipo em rs5491 no gene <i>ICAM-1</i>			<i>p</i> *
	AA (N=33)	AT (N=26)	TT (N=6)	
Média da idade da primeira manifestação da drepanocitose (meses)	15	12	27	<b>0,023</b>

A negrito apresenta-se o valor de *p* que se revelou estatisticamente significativo. \*Teste de Kruskal-Wallis.

Relativamente à idade da primeira manifestação drepanocítica constatou-se que, no caso da variante rs5491 A>T no gene *ICAM-1*, os doentes com o alelo variante em homozigotia (genótipo TT) têm a manifestação mais tardiamente do que os doentes com o alelo variante em heterozigotia ou com o genótipo *wild-type*.

Segundo os estudos efetuados na nossa população, dos doentes homozigóticos para o alelo\_T (seis doentes) a primeira manifestação só se revelou a partir dos 12 meses de idade. A idade média da primeira manifestação drepanocítica para estes doentes é aos 27 meses. Adicionalmente, observou-se que em dois doentes homozigóticos (33,3%), a manifestação revelou-se só aos 24 meses. Por outro lado, no caso dos doentes heterozigóticos para o alelo variante, a idade média da primeira manifestação é aos 12 meses e no caso dos indivíduos homozigóticos *wild-type*, a primeira manifestação apresenta-se, em média, aos 15 meses de vida, sendo que a maioria revelou a manifestação aos seis meses.

De facto, as diferenças entre pares são significativas quando confrontando as médias de idades da primeira manifestação drepanocítica: dos doentes com o genótipo AA *versus* TT com  $p=0,023$  e o mesmo se verifica quando confrontado o genótipo AT com TT,  $p=0,006$ .

Concluimos então que a variante T em rs5491 no gene *ICAM-1* é um fator genético modificador do fenótipo clínico com efeitos benéficos para os doentes drepanocíticos uma vez que atrasa o aparecimento das primeiras manifestações da doença. Esta associação não terá sido ainda mencionada na literatura e pode justificar a elevada

frequência do alelo variante não só na nossa população estudada como também noutras populações de ascendência africana.

Na tabela 4.9 encontram-se as associações determinadas para os parâmetros hematológicos concentração média de Hb e contagem média de eritrócitos segundo os diferentes genótipos em rs3211891 no gene *CD36*, rs5496, rs5491 no gene *ICAM-1* e mutação  $\alpha$ -talassémica de 3,7kb nos genes *HBA2* e *HBA1*.

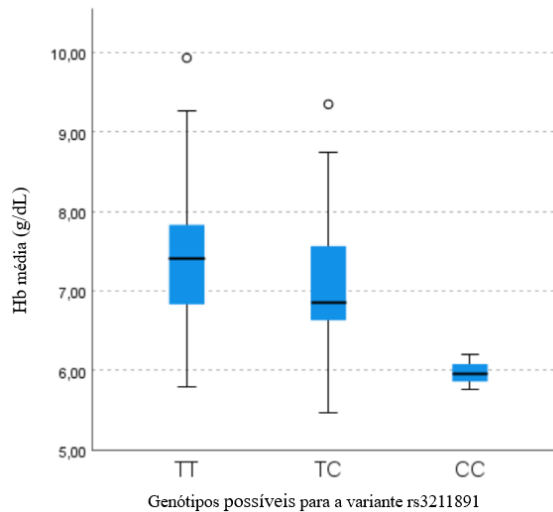
**Tabela 4.9**-Associações entre a média da concentração de Hb e contagem de eritrócitos e os diferentes genótipos em rs3211891, rs5496, rs5491 e a deleção  $\alpha$ -talassémica de 3,7kb

Fenótipo hematológico (média)		N	Hb (g/dL)	<i>p</i> *	Eritrócitos ( $\times 10^{12}/L$ )	<i>p</i> *
Genótipos em rs3211891 no gene <i>CD36</i>	TT	49	7,43	<b>0,013</b>	2,93	0,074
	TC	13	7,20			
	CC	3	5,98			
Genótipos em rs5496 no gene <i>ICAM-1</i> **	GG	56	7,21	<b>0,009</b>	2,84	<b>0,036</b>
	GA+AA	9	7,99			
Genótipos em rs5491 no gene <i>ICAM-1</i> **	AA	33	7,54	0,098	3,05	<b>0,044</b>
	AT+TT	32	7,08			
Deleção talassémica de 3,7kb (Genes <i>HBA2</i> e <i>HBA1</i> )	$\alpha\alpha/\alpha\alpha$	17	7,31	0,119	2,66	<b>&lt;0,001</b>
	$\alpha\alpha/-\alpha^{3,7}$	40	7,11			
	$-\alpha^{3,7}/-\alpha^{3,7}$	6	8,10			

A negrito apresentam-se os valores de *p* que se revelaram estatisticamente significativos. \*Teste de Kruskal Wallis nos casos em que o modelo dominante não foi utilizado e teste de Mann-Whitney nos casos em que o modelo dominante foi utilizado.\*\*Modelo dominante.

No caso da variante rs3211891T>C no gene *CD36*, o alelo variante\_C em homozigotia foi detetado em doentes com níveis anormalmente baixos de Hb,  $5,98 \pm 0,22$  g/dL, em comparação com os doentes nos quais a variante foi detetada em heterozigotia,  $7,20 \pm 1,03$  g/dL, e onde não foi detetada,  $7,43 \pm 0,81$  g/dL.

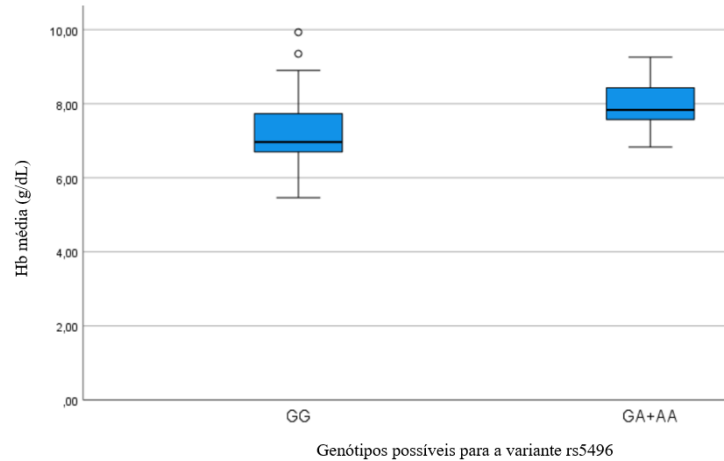
De facto, na figura 4.20 é notório o decrescer dos valores médios de Hb para os três genótipos.



**Figura 4.20**-Representação gráfica da distribuição média de Hb por genótipo em rs3211891 no gene *CD36*.

Para os doentes com drepanocitose esta variante não se revela vantajosa. A drepanocitose é por si só caracterizada por níveis de Hb abaixo do que é considerado normal. A presença desta variante, quer em homozigotia quer em heterozigotia parece agravar a carência de Hb, ou seja, aumenta o grau de anemia, o que se torna prejudicial.

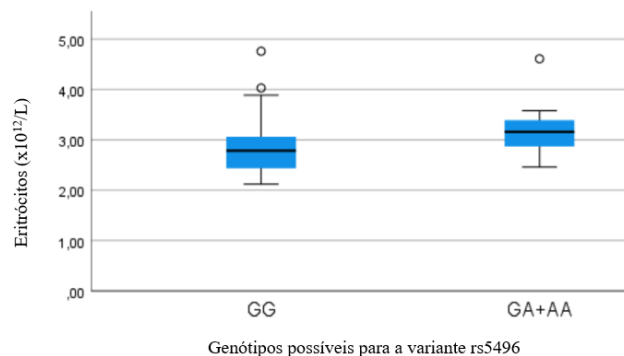
Em relação à variante rs5496 G>A no gene *ICAM-1*, associada de forma significativa com os valores de Hb média ( $p=0,009$ ), observou-se que os doentes com o alelo variante *\_A* em homozigotia e em heterozigotia apresentam valores superiores de Hb,  $7,99 \pm 0,72$  g/dL, quando comparados com os doentes que apresentam genótipo *wild-type*, para os quais o valor médio de Hb é de  $7,21 \pm 0,87$  g/dL (Figura 4.21).



**Figura 4.21**-Representação gráfica da distribuição média da concentração de Hb por genótipo *wild-type* versus genótipos contendo a variante rs5496\_A.

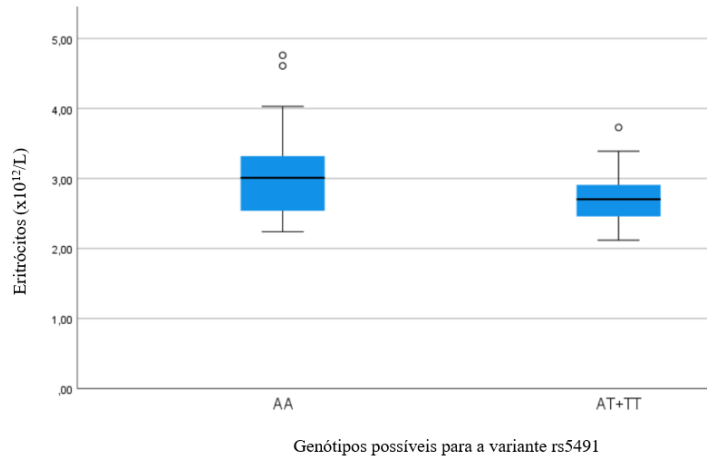
Desta forma, a presença do alelo variante\_A poderá revelar-se benéfica para os doentes drepanocíticos, na medida em que aumenta os níveis médios da Hb e, portanto, diminui a anemia.

Relativamente ao número de eritrócitos, verificou-se que a variante rs5496 G>A, no gene *ICAM-1*, se encontra associada de forma significativa ( $p=0,036$ ) com este parâmetro hematológico. Os estudos revelaram que os doentes com presença do alelo variante (em homozigotia e em heterozigotia) apresentam valores médios superiores de eritrócitos,  $3,20 \pm 0,66 \times 10^{12}/L$ , quando comparados com as contagens médias dos doentes com genótipo *wild type*,  $2,84 \pm 0,51 \times 10^{12}/L$  (Fig.4.22). Resultado este expectável, uma vez que se verificou previamente uma associação com o aumento na concentração média de Hb.



**Figura 4.22**-Representação gráfica da distribuição média de eritrócitos por genótipo *wild-type* versus genótipos contendo a variante rs5496\_A.

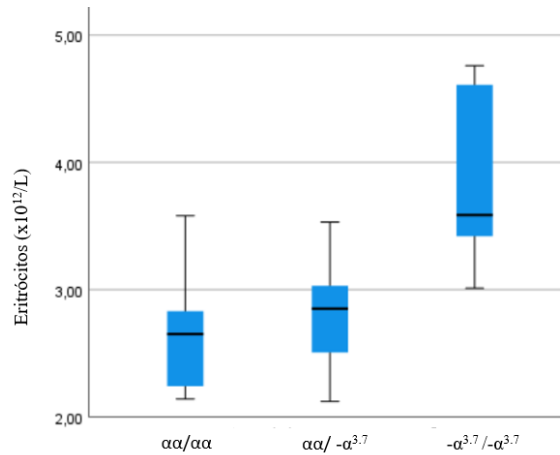
Já para a variante rs5491 A>T (gene *ICAM-1*) concluiu-se que a presença de genótipos com o alelo variante\_T se encontra significativamente associada ( $p=0,044$ ) com um número médio de eritrócitos inferior,  $2,83 \times 10^{12}/L$  em comparação com os doentes que não possuem genótipos com o alelo variante, média de  $3,05 \times 10^{12}/L$  (Fig.4.23).



**Figura 4.23**-Representação gráfica da distribuição média de eritrócitos por genótipo *wild-type* versus genótipos contendo a variante rs5491\_T.

O resultado aferido para a variante rs5491 A>T levanta a hipótese de que os doentes que possuem genótipos com o alelo variante apresentem uma maior taxa de eliminação dos eritrócitos, quando comparados com os doentes com genótipo *wild-type*.

A deleção  $\alpha$ -talassémica de 3,7kb (ocorrida nos genes *HBA2* e *HBA1*) encontra-se, de igual modo, associada a um maior número de eritrócitos ( $p < 0,001$ ) determinados nos doentes. De facto, os doentes com a mutação talassémica em homocigotia apresentam uma contagem média de eritrócitos significativamente superior em relação aos portadores da mutação em heterocigotia e aos não portadores,  $3,83 \pm 0,70 \times 10^{12}/L$  versus  $2,80 \pm 0,34 \times 10^{12}/L$  e  $2,66 \pm 0,42 \times 10^{12}/L$  (Fig.4.24).



**Figura 4.24**-Representação gráfica da distribuição média de eritrócitos por genótipo da deleção talassémica de 3,7kb.

Estes resultados relativos à deleção talassémica de 3,7kb podem relacionar-se com o já comprovado aumento do hematócrito e da viscosidade sanguínea na sua presença (54). O hematócrito diz respeito à percentagem de sangue que é constituída por eritrócitos e influencia também a viscosidade do sangue, quanto maior o hematócrito (acima de valores normais) maior a viscosidade. Na nossa população estudada a presença desta mutação em homozigotia aumenta os valores de eritrócitos para valores que se encontram dentro do intervalo de referência estipulado para uma população pediátrica (entre os seis meses e os 11 anos, anexo 4), o que se pode revelar benéfico, melhorando a capacidade de transporte de oxigénio ao organismo.

Sabe-se que as alterações morfológicas da membrana do eritrócito (nos doentes com drepanocitose), que advém da polimerização da HbS, aumentam a eliminação dos mesmos por hemólise precoce. Portanto, a mutação talassémica em causa, é benéfica no sentido em que diminui o número de cadeias de  $\alpha$ -globina que estarão disponíveis para se juntarem com as cadeias  $\beta^S$  para a formação da HbS. A diminuição da concentração de HbS no interior do eritrócito leva à diminuição da taxa de falciformização e consequentemente diminuição de hemólise dos eritrócitos e da sua fagocitose e normalização do número de eritrócitos nos doentes. Esta hipótese foi comprovada (como descrito mais abaixo) uma vez que também detetámos associação entre a presença de  $\alpha$ -talassémia e menor percentagem de reticulócitos (um biomarcador do nível de hemólise).

As associações com os parâmetros hematimétricos VGM e HGM inferem sobre o tipo de anemia a que os doentes se encontram sujeitos, isto é, se as variantes melhoram ou pioram

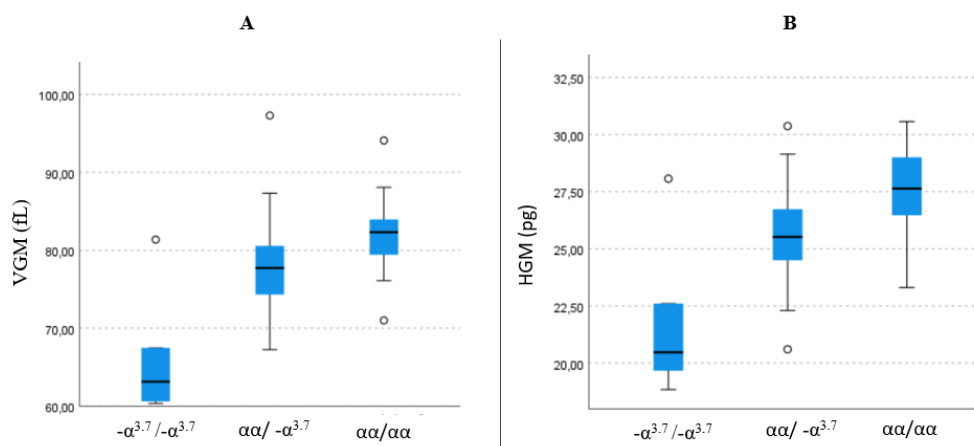
os fenótipos de microcitose ou hipocromia. As associações entre estes parâmetros e as variantes genéticas detetadas estão representadas na tabela 4.10.

**Tabela 4.10**-Associações entre a média de HGM e VGM e os diferentes genótipos em rs3211891, rs5496, rs5491 e a deleção  $\alpha$ -talassémica de 3,7kb

Fenótipo hematológico (média)		N	HGM	$p^*$	VGM	$p^*$
Genótipos em rs3211891 no gene <i>CD36</i>	TT	49	25,76	0,378	77,53	0,476
	TC	13	24,91			
	CC	3	26,10			
Genótipos em rs5496 no gene <i>ICAM-1</i> **	GG	56	25,73	0,442	78,10	0,247
	GA+AA	9	24,85			
Genótipos em rs5491 no gene <i>ICAM-1</i> **	AA	33	25,10	0,177	76,24	0,153
	AT+TT	32	26,12			
Deleção talassémica de 3,7kb (Genes <i>HBA2</i> e <i>HBA1</i> )	$\alpha\alpha/\alpha\alpha$	17	27,44	<b>&lt;0,001</b>	81,87	<b>&lt;0,001</b>
	$\alpha\alpha/-\alpha^{3,7}$	40	25,55			
	$-\alpha^{3,7}/-\alpha^{3,7}$	6	21,68			

A negrito apresentam-se os valores de  $p$  que se revelaram estatisticamente significativos. \*Teste de Kruskal Wallis nos casos em que o modelo dominante não foi utilizado e teste de Mann-Whitney nos casos em que o modelo dominante foi utilizado. \*\*Modelo dominante.

Com os resultados dos estudos de associação apresentados na tabela 4.10, pode dizer-se que como esperado as associações significativas com os parâmetros VGM e HGM foram detetadas para a deleção talassémica ocorrida nos genes *HBA2* e *HBA1*. A distribuição dos valores médios de VGM e HGM pelos genótipos possíveis da deleção de 3,7kb encontram-se apresentados graficamente na figura 4.25.



**Figura 4.25**-Representações gráficas dos valores médios de VGM (A) e HGM (B) por grupo de genótipo da deleção talassémica de 3,7kb.

Como se pode observar na figura 4.25 a presença da mutação  $\alpha$ -talassémica em questão, tanto para os valores de VGM (Fig.4.25- A) como para os valores de HGM (Fig.4.25- B) encontra-se associada a valores mais baixos.

No caso dos valores médios de VGM, a média para os doentes homozigóticos foi de  $66,00 \pm 7,95$  fL, valor este que se encontra abaixo do esperado para uma população pediátrica entre os seis meses e os 11 anos de idade (Anexo 4). Contrariamente a esta observação, os valores médios de VGM para os doentes heterozigóticos para a mutação ( $77,88 \pm 5,86$  fL) e homozigóticos *wild-type* ( $81,87 \pm 5,24$  fL) encontram-se compreendidos no intervalo de valores normais. Estes resultados vão ao encontro dos obtidos no estudo de Olatunya *et al.*, no qual os doentes com a mutação em homozigotia apresentaram valor médio de VGM de 74,00 fL enquanto que os doentes com a mutação em heterozigotia apresentaram um valor médio de 78,00 fL e com genótipo *wild-type* um valor de 84,00 fL (93). Do mesmo modo, no estudo de Al-Barazanchi *et al.* os valores de VGM correspondem aos obtidos no presente projeto, os doentes drepanocíticos na presença da deleção talassémica de 3,7kb apresentaram um valor médio de VGM de 68,06 fL e os doentes sem a deleção apresentaram um valor de VGM de 85,72 fL (116). Comprovou-se, desta forma, que os valores mais baixos de VGM e associados a um fenótipo de microcitose se devem à presença da deleção  $\alpha$ -talassémica de 3,7kb.

Relativamente aos valores de HGM, o valor médio determinado para os doentes homozigóticos para a deleção foi de  $21,68 \pm 3,38$  pg e para os doentes com genótipo heterozigótico o valor determinado foi de  $25,55 \pm 2,06$  pg. Valores estes que se encontram abaixo do valor de referência estipulado na norma nº 063/2011 da DGS. Já no caso dos doentes com um genótipo *wild-type* o valor médio de HGM determinado,  $27,44 \pm 1,90$  pg, encontra-se dentro do intervalo de referência. Nos estudos previamente mencionados e à semelhança do concluído para os valores de VGM, os valores mais baixos de HGM encontram-se significativamente associados à homozigotia da deleção (93,116). Comprovando, também, que o fenótipo de hipocromia na população estudada se encontra associado à presença da deleção  $\alpha$ -talassémica de 3,7kb.

O estudo dos biomarcadores diretos de hemólise, LDH e bilirrubina total, e indireto, percentagem de reticulócitos, revela-se importante para avaliar a gravidade da drepanocitose uma vez que esta é caracterizada por uma anemia hemolítica crónica.

No decorrer dos estudos de associação verificou-se que nenhuma das variantes genéticas detetadas apresentava uma relação significativa com os valores de bilirrubina total.

**Tabela 4.11**-Associações entre os biomarcadores de hemólise (LDH e percentagem de reticulócitos) e os diferentes genótipos em rs5491, rs5498, rs5496 e a deleção  $\alpha$ -talassémica de 3,7kb

Biomarcador de hemólise		N	LDH (U/L)	<i>p</i> *	Reticulócitos (%)	<i>p</i> *
Genótipos em rs5491 no gene <i>ICAM-1</i> **	AA	33	389	<b>0,020</b>	9,13	0,109
	AT+TT	32	465		10,82	
Genótipos em rs5498 no gene <i>ICAM-1</i> **	AA	55	420,16	0,251	9,57	<b>0,015</b>
	AG+GG	10	476,05		12,09	
Genótipos em rs5496 no gene <i>ICAM-1</i> **	GG	56	425,73	0,694	10,31	<b>0,044</b>
	GA+AA	9	440,13		7,79	
Deleção talassémica de 3,7kb (Genes <i>HBA2</i> e <i>HBA1</i> )	$\alpha\alpha/\alpha\alpha$	17	452,07	0,522	10,68	<b>0,018</b>
	$\alpha\alpha/-\alpha^{3,7}$	40	420,63		10,35	
	$-\alpha^{3,7}/-\alpha^{3,7}$	6	410,34		6,53	

A negrito apresentam-se os valores de *p* que se revelaram estatisticamente significativos. \*Teste de Kruskal Wallis nos casos em que o modelo dominante não foi utilizado e teste de Mann-Whitney nos casos em que o modelo dominante foi utilizado.\*\*Modelo dominante.

Quanto aos valores de LDH, houve apenas associação significativa com a variante rs5491 A>T (no gene *ICAM-1*), com  $p=0,020$ . Concluiu-se que, para a população estudada, os indivíduos com genótipo homozigótico ou heterozigótico para o alelo variante\_T apresentavam valores médios de LDH mais elevados,  $465 \pm 125$  U/L. Os doentes com genótipo *wild-type* revelaram níveis de LDH perto dos normais (332 U/L de acordo com o estipulado pela *Roche Diagnostics GmbH*), de  $389 \pm 111$  U/L. Pode-se dizer que este resultado corrobora a hipótese apresentada aquando da apresentação da associação desta variante com o número médio de eritrócitos. A presença do alelo variante tanto em homozigotia como em heterozigotia encontra-se associada a um menor número de eritrócitos e a valores elevados de LDH. A LDH encontra-se contida nos eritrócitos e é libertada quando ocorre eliminação dos mesmos, o aumento da sua concentração na corrente sanguínea indica um aumento da hemólise e, por consequente, uma diminuição do número de eritrócitos.

Assim, a presença da variante rs5491 A>T não se revela benéfica para a população drepanocítica estudada, uma vez que muito provavelmente se encontra relacionada com um aumento da taxa de hemólise.

A avaliação da percentagem de reticulócitos fornece informações sobre a resposta da medula óssea à anemia hemolítica crónica característica da drepanocitose, neste sentido uma maior percentagem de reticulócitos é sinónimo de uma eritropoiese elevada por stress devido à grande eliminação de eritrócitos.

As associações significativas com a percentagem de reticulócitos encontradas estão representadas na tabela 4.11.

A variante genética rs5498 A>G, localizada no gene *ICAM-1*, encontra-se associada significativamente,  $p=0,021$ , com a percentagem de reticulócitos. A presença do alelo variante\_G, em homozigotia e em heterozigotia, revelou uma percentagem de reticulócitos superior à percentagem determinada para os doentes com genótipo *wild-type* ( $12,09 \pm 3,18\%$  versus  $9,57 \pm 3,80\%$ ). O que indica que para os doentes com genótipos que possuem esta variante, a resposta da medula óssea é, possivelmente, exacerbada.

A variante rs5498 A>G no gene *ICAM-1*, é apenas mencionada na literatura como estando relacionada com a ocorrência do cancro (na sua generalidade), com doenças vasculares e com patologias ligadas à carência de ferro não tendo sido ainda reportada nenhuma associação significativa com manifestações da drepanocitose, para além da relação com a anemia (117,118). Esta relação com a anemia relatada na literatura vai ao encontro do que se verifica na nossa população estudada. Os doentes com drepanocitose apresentam anemia hemolítica crónica e consequentemente carência em ferro, o que estimula a resposta por parte da medula óssea, desta forma pode-se dizer que os doentes com genótipos variantes para a variante rs5498 A>G apresentam uma resposta eritropoética mais elevada que os doentes com genótipo *wild-type*.

Relativamente à variante rs5496 G>A no gene *ICAM-1* averiguou-se que a presença do alelo variante\_A, em homozigotia e em heterozigotia se encontra associado ( $p=0,044$ ) a uma percentagem média de reticulócitos inferior ( $7,79 \pm 2,31\%$ ), quando comparado com a percentagem de reticulócitos nos doentes *wild-type* ( $10,31 \pm 3,89\%$ ). Este resultado vai ao encontro do aferido para a concentração média de Hb e número de eritrócitos, na qual a presença de um genótipo com o alelo variante\_A se encontra associado a concentrações superiores de Hb e maior número médio de eritrócitos. Levanta-se, deste modo, a hipótese de que a presença da variante seja benéfica para a população uma vez que normaliza os valores de Hb e eritrócitos e estabiliza a taxa de hemólise.

A mutação  $\alpha$ -talassémica de 3,7kb (genes *HBA2* e *HBA1*) em homozigotia foi detetada em indivíduos com uma média de reticulócitos,  $6,53 \pm 2,36\%$ , significativamente inferior em relação aos doentes com a mutação em heterozigotia,  $10,35 \pm 3,94\%$ , e aos doentes *wild-type*,  $10,68 \pm 3,20\%$ . De facto, esta observação de que a deleção mencionada diminui a taxa de hemólise foi já apontado na literatura como sendo um dos benefícios apontados para as populações com drepanocitose e como uma das razões para a elevada prevalência da co-herança (54). Este resultado está em concordância com os resultados que obtivemos de associação entre a presença de  $\alpha$ -talassémia e níveis mais elevados de eritrócitos, fazendo antever uma menor taxa de hemólise nos doentes drepanocíticos que co-herdam  $\alpha$ -talassémia.

A avaliação do número de plaquetas pode-se revelar importante para perceber que impacto as variantes genéticas apresentam em relação aos processos inflamatórios, por exemplo. As associações determinadas para este parâmetro encontram-se na tabela 4.12.

**Tabela 4.12**-Associações entre o número médio de plaquetas e os diferentes genótipos em rs3211938 e rs5497

Fenótipo hematológico (média)		N	Plaquetas ( $\times 10^9/L$ )	<i>p</i> *
Genótipos em rs3211938 gene CD36**	TT	48	468,94	<b>0,030</b>
	TG+GG	17	373,78	
Genótipos em rs5497 no gene ICAM-1**	GG	52	425,51	0,051
	GA+AA	13	518,21	

A negrito apresenta-se o valor de *p* que se revelou estatisticamente significativo. \*Teste de Mann-Whitney.\*\*Modelo dominante.

No que toca aos valores de contagem de plaquetas, a variante genética associada de forma significativa foi a rs3211938 T>G ocorrida no gene *CD36* com o *p*= 0,030. A presença do alelo variante\_G em homozigotia e em heterozigotia parece diminuir o número de plaquetas nos doentes, estes apresentam uma média de  $373,78 \times 10^9/L$  plaquetas *versus* uma média de  $468,94 \times 10^9/L$  plaquetas nos doentes com o genótipo *wild-type*. O número médio de plaquetas nos indivíduos drepanocíticos com o alelo variante é considerado normal, segundo a norma nº 063/2011 da DGS (Anexo 4). Contrariamente a esta observação, o número médio de plaquetas nos doentes com genótipo *wild-type* encontra-se elevado e acima do intervalo normal apresentado pela norma.

Põem-se a hipótese desta variante modular as respostas inflamatórias e processos inflamatórios exacerbados a que a população angolana estudada esteja exposta. Como foi previamente mencionado, a população estudada encontra-se exposta não só à malária, mas também a outras parasitoses e bacterioses que desencadeiam respostas inflamatórias. Caso o papel desta variante se prenda com a modulação e desencadeamento das respostas imunitárias e inflamatórias, de forma que estas não sejam exacerbadas, isto é, respostas controladas e direcionadas ao corpo estranho, então será considerada benéfica.

Por outro lado, e como foi já mencionado no ponto 4.2.1.1 da caracterização genética das variantes detetadas, pensa-se que a variante rs3211938 T>G resulte numa deficiência da proteína CD36 em monócitos e macrófagos (deficiência de tipo I), como consequência a resposta imunitária não se espera que seja tão eficiente como na ausência da variante (62,98,100). Deste modo pode-se pôr a hipótese de que o número de plaquetas reflita a incapacidade de resposta imunitária e inflamatória na presença de corpos estranhos.

Relativamente à variante rs5497 G>A, ocorrida no gene *ICAM-1*, verificou-se que os doentes drepanocíticos com genótipo que contém o alelo variante\_A apresentam um número médio de plaquetas superior ao número de plaquetas determinado para os indivíduos com genótipo *wild-type*, respetivamente  $518,2 \times 10^9/L$  e  $425,5 \times 10^9/L$ . A variante revelou uma forte tendência de associação com o aumento do número de plaquetas, com  $p=0,051$ .

Esta variante foi já mencionada em estudos de imunodeficiência e devido à sua localização no gene *ICAM-1* (c.1190G>A) sabe-se que pode interferir com a estabilização da proteína ICAM-1 traduzida (ponto 4.3.2.1, dos Resultados e Discussão). A perda de função da proteína ICAM-1 pode resultar na ausência de interações moleculares importantes para a resposta imunitária e interferir também no processo de transmigração leucocitária (67,119).

Deste modo, pode levantar-se a hipótese de que a associação (ainda que não seja significativa) de genótipos que contém o alelo variante\_A com um número elevado de plaquetas, possa estar associado à ineficiência da resposta imunitária que resulta num aumento da produção de plaquetas e consequente resposta imunitária exacerbada.

#### 4.4.2 Variantes genéticas detetadas e a variabilidade fenotípica da malária

Para se efetuarem os estudos de associação entre as variantes detetadas nos genes de estudo e a malária recorreu-se apenas à informação subjetiva dos tutores relativa ao número de episódios de malária e aos resultados do PCR (deteção do DNA do plasmódio). Infelizmente, as lâminas observadas não se encontravam legíveis pelo que os resultados que dessa observação resultaram não foram utilizados no corrente projeto por não serem fiáveis.

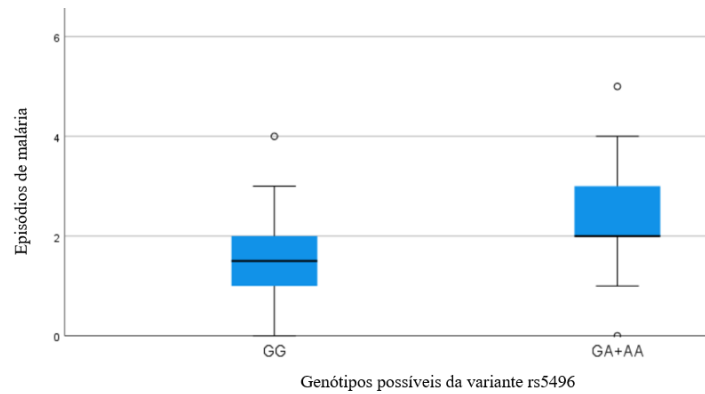
Não foram observadas associações significativas entre os episódios de malária e as variantes detetadas. Somente a variante rs5496 G>A, que ocorre no gene *ICAM-1*, revelou uma tendência para estar associada com um maior número de episódios de malária,  $p=0,060$  (tabela 4.13).

**Tabela 4.13**-Associações significativas entre a média de episódios de malária e os diferentes genótipos em rs5496

Fenótipo da malária		N	Número de episódios de malária	$p^*$
Genótipos em rs5496 no gene <i>ICAM-1</i> **	GG	56	1,46	0,060
	GA+AA	9	2,33	

\*Teste de Mann-Whitney.\*\*Modelo dominante.

Os doentes com genótipos que contém o alelo variante A apresentaram uma média de  $2,33 \pm 1,50$  episódios, enquanto os doentes com genótipo *wild-type* reportaram uma média de  $1,46 \pm 1,21$  episódios (Fig.4.26).



**Figura 4.26-**Representação gráfica da distribuição média do número de episódios de malária por genótipo *wild-type* versus genótipos contendo a variante rs5496\_A.

Assim, a presença desta variante rs5496 G>A (gene *ICAM-1*) parece proporcionar um maior número de episódios de malária nestes doentes drepanocíticos. Concluimos previamente que esta variante se encontra associada, de forma significativa, com valores superiores de Hb e contagem de eritrócitos em relação aos doentes com genótipo *wild-type*. Com uma maior abundância de eritrócitos (células-alvo do plasmódio) e maior concentração de Hb, esta metabolizada pelo parasita para efeitos nutricionais e destoxificação, o plasmódio apresenta melhores condições para se desenvolver e propagar (120). Pode-se então, levantar a hipótese de que estas observações se encontrem relacionadas. Desta forma, a variante rs5496 pode representar, simultaneamente, um benefício relativamente ao fenótipo de drepanocitose e uma desvantagem para os doentes expostos à malária.

Quanto aos resultados obtidos na PCR para deteção de DNA do plasmódio apenas foram conseguidos resultados para 54 dos 65 doentes drepanocíticos que constituíam a população estudada. Os resultados em questão foram utilizados nos estudos de associação com o objetivo de se perceber se alguma das variantes detetadas apresentaria uma relação significativa com o resultado positivo e/ou negativo de malária.

Neste sentido, foram realizados o teste de Qui-Quadrado e o teste exato de Fisher, este último quando uma contagem menor que cinco era observada em mais do que 25% das células. Posteriormente, calculou-se também o valor de *odds ratio* (OR).

Na tabela 4.14 encontra-se a associação significativa observada entre a presença do DNA do parasita (diagnóstico positivo de malária) e as variantes genéticas detetadas.

**Tabela 4.14**-Associação dos genótipos em rs5494 e os casos de malária

Casos de malária		Casos positivos para malária (N)	Casos negativos para malária (N)	<i>p</i> *	OR	95%IC
Genótipos em rs5494 no gene <i>ICAM-1</i> **	CC	3 (7,3%)	38 (92,7%)	<b>0,028</b>	5,63	1,07-29,73
	CT+TT	4 (30,8%)	9 (69,2%)			

A negrito apresenta-se o valor de *p* que se revelou estatisticamente significativo. \*Teste de Qui-Quadrado.\*\*Modelo dominante.

Concluiu-se haver uma relação significativa entre a presença da variante rs5494 C>T e os resultados positivos de malária (confirmados por PCR) e aferiu-se um valor de OR de 5,63, com intervalo de confiança a 95% de 1,07 - 29,73,  $p=0,028$ .

A correlação entre os genótipos contendo o alelo variante *T*, quer em homozigotia, quer em heterozigotia, e a malária (detetada por PCR) é positiva. A presença de genótipos que contêm o alelo variante encontra-se, portanto, associada a uma maior probabilidade (5,63 vezes superior) de ocorrência de malária na população drepanocítica estudada.

Resultado este que pode explicar a baixa frequência dos genótipos com o alelo variante na nossa população, tal como foi já observado e descrito na tabela 4.5. De facto esta correlação com a infeção de malária foi já determinada no estudo de Gomez *et al* onde se aferiu uma associação significativa com a endemicidade da malária com  $p=0,002$ , resultado este semelhante ao obtido no presente estudo (121).

Apesar de ainda não existir uma explicação validada para esta associação com a endemicidade de malária, sabe-se que o SNP ocorre no quarto exão do gene *ICAM-1*, este exão codifica o terceiro domínio extracelular da proteína ICAM-1. O terceiro domínio da proteína apresenta o ligando humano do Mac-1 (ver ponto 1.6.3, Fig.1.11), este que auxilia a transmigração leucocitária e o desencadear da resposta inflamatória do sistema imunitário (122). Ainda que esta variante não altere a proteína, p.Asp282=, pode-se pôr a hipótese de que poderá influenciar negativamente a capacidade de adesão dos leucócitos e o início da resposta inflamatória necessária ao combate ao parasita. Deste modo, ocorrerá uma resposta deficiente, permitindo a propagação do parasita e da infeção.

Concluimos que esta variante apresenta inequivocamente uma relação com a malária e com uma maior probabilidade da sua manifestação, estando associada a um maior número de casos positivos.



## 5 Conclusões e Perspetivas Futuras

O estudo dos modificadores genéticos da malária, da drepanocitose e da interação entre ambas revela-se muito importante, principalmente em populações que habitam em regiões com elevada coexistência destas duas patologias, tal como a população angolana em estudo neste projeto.

Assim, o presente estudo teve como objetivo principal estudar a relação entre variantes genéticas ocorridas nos genes *CD36*, *ICAM-1* e *HBA* e a variabilidade fenotípica da drepanocitose e malária numa população angolana em idade pediátrica. Através da genotipagem destes genes, foram detetadas 15 variantes genéticas e estas estudadas do ponto de vista bioinformático e estatístico para investigar possíveis associações entre os genótipos e os fenótipos (hematológico, bioquímico e clínico).

Com o estudo da co-herança entre a drepanocitose e a  $\alpha$ -talassémia (deleção de 3,7kb) concluiu-se que, como já era esperado para uma população africana, esta deleção apresenta elevada prevalência na nossa população angolana drepanocítica, sendo que 46 (73,0%) crianças apresentam o alelo  $-\alpha^{3,7}$ , em homozigotia ou em heterozigotia. Relativamente a esta deleção  $\alpha$ -talassémica, concluiu-se também que a sua presença, em homozigotia e heterozigotia, se encontra associada significativamente a valores mais baixos de VGM, HGM e reticulócitos e a números superiores de eritrócitos. Os valores de VGM e HGM mais baixos são associados, respetivamente, a fenótipos de microcitose e hipocromia. Já no caso dos resultados obtidos para a percentagem média de reticulócitos e número médio de eritrócitos pôde-se concluir que a coexistência da drepanocitose com a  $\alpha$ -talassémia é benéfica no sentido em que diminui a gravidade da anemia (aumenta o número médio de eritrócitos) e diminui o grau de hemólise (diminui o número médio de reticulócitos), o que se encontra já reportado na literatura. Os nossos resultados validam, portanto, estudos anteriores, efetuados noutras populações drepanocíticas.

Das variantes analisadas no gene *CD36*, duas parecem ter maior impacto como moduladoras. Concluiu-se que o SNP rs3211891 T>C, localizado no quarto intrão do gene, afeta negativamente a concentração média de Hb nestes doentes. Os doentes com genótipos que contêm o alelo *C* apresentam valores de Hb significativamente mais baixos. Por outro lado, a variante *nonsense* rs3211938 T>G, localizada no décimo exão do referido gene, apresenta uma associação negativa com o número de plaquetas. Os

doentes com genótipos que contêm o alelo\_G apresentam um número médio de plaquetas significativamente mais baixo que os doentes com genótipo *wild-type*.

Ambas as variantes se revelaram nefastas para os doentes que possuem genótipos com o alelo variante. Apesar desta observação, no caso do SNP rs3211891 T>C a frequência de genótipos com o alelo variante na população drepanocítica estudada é significativamente superior em relação à população africana e ioruba. Isto poderia querer antever que fosse benéfica para a drepanocitose, uma vez que todas as populações se encontram expostas a malária endémica, porém a concentração de Hb é afetada de forma negativa, ficando assim por explicar as elevadas frequências genótípicas na nossa população estudada. Já para o SNP rs3211938 T>G, a população estudada revelou frequências genótípicas de genótipos contendo o alelo variante significativamente inferiores à população ioruba, resultado este que pode prever que valores inferiores de contagem de plaquetas se podem associar a uma desvantagem para doentes drepanocíticos. Ao mesmo tempo, não foi verificada uma diferença estatisticamente significativa entre as frequências genótípicas da população estudada e da população africana, provavelmente porque esta última corresponde a toda a população africana no geral, enfatizando desta forma a elevada prevalência da mutação rs344 A>T na região africana.

Dada a importância do gene *CD36* na ligação entre a PfEMP1 e o endotélio vascular humano, eram esperadas mais associações significativas entre as variantes ocorridas neste gene e os fenótipos da malária. Estes resultados não se verificaram, muito possivelmente devido ao baixo número de participantes no estudo e à falta de dados clínicos relativos às manifestações da malária.

No caso das variantes genéticas analisadas no gene *ICAM-1*, concluiu-se que a variante *missense* rs5491 A>T, localizada no segundo exão do gene e no primeiro domínio da proteína ICAM-1, vulgarmente conhecida por mutação Kilifi, apresenta uma associação negativa com a drepanocitose. Os doentes com genótipos com o alelo\_T apresentam um número médio de eritrócitos inferior aos restantes, assim como níveis de LDH mais elevados, levantando-se a hipótese deste alelo neste SNP resultar num agravamento da anemia e da taxa de hemólise. No entanto, em termos contrários, apresenta uma associação positiva com a idade da primeira manifestação drepanocítica, para a qual os doentes homozigóticos para o alelo\_T apresentam a primeira manifestação mais tarde que

os restantes. Dito isto, em termos clínicos foi observado um benefício para os doentes que possuem esse genótipo, mas há que salvaguardar que possa haver enviesamento pelo baixo número de amostragem. Quanto à sua relação com os fenótipos de malária não foi verificada qualquer associação, tal como na literatura, continua por esclarecer o seu efeito e explicar a ambiguidade das descobertas realizadas até agora.

Por outro lado, a variante intrónica rs5496 G>A, localizada no quinto intrão do gene *ICAM-1*, revelou uma associação positiva com os fenótipos hematológicos característicos da drepanocitose. A presença do alelo *\_A*, quer em homozigotia quer em heterozigotia, associou-se a uma concentração média de Hb e número de eritrócitos superior e percentagem de reticulócitos inferior, em relação ao genótipo *wild-type*. Dados estes resultados concluímos que este alelo se encontra relacionado com uma diminuição da anemia e da hemólise características da drepanocitose e, portanto, podemos considerá-lo, sem dúvida, um fator genético modificador positivo destes subfenótipos da drepanocitose.

Ainda no gene *ICAM-1*, concluiu-se que a variante *missense* rs5498 A>G, localizada no sexto exão do gene e que afeta a proteína, p.Lys469Glu, apresenta uma relação positiva com a percentagem de reticulócitos. A presença de genótipos homozigóticos e heterozigóticos para o alelo *\_G*, está associada com uma percentagem de reticulócitos superior, o que levantou a hipótese deste alelo se encontrar associado com o aumento da hemólise, sendo assim uma variante nefasta para a drepanocitose. Não foi detetada qualquer associação com os fenótipos de malária, apesar de que caso a associação com o aumento de hemólise se valide, a redução do número de eritrócitos poderá afetar a infeção pelo plasmódio do género *P. falciparum*, sendo os eritrócitos as suas células-alvo e onde se desenvolve e multiplica.

Relativamente ao estudo efetuado sobre o papel das variantes genéticas estudadas sobre a variabilidade fenotípica da malária nestes doentes com drepanocitose, concluiu-se que a variante sinónima rs5494 C>T, localizada no quarto exão do gene *ICAM-1*, se encontra relacionada com um maior número de casos positivos de malária. Os doentes com genótipos que contêm o alelo variante apresentaram uma maior probabilidade de ocorrência de malária, em comparação com os doentes com genótipo *wild-type*, provavelmente devido à destabilização da adesão leucocitária por intermédio do Mac-1 e

à consequente deficiente resposta inflamatória. Não se verificou que tivesse efeito sobre a drepanocitose.

Assim, de entre os genes estudados, concluiu-se que o gene *ICAM-1* apresenta, na nossa população, um importante papel modulador tanto dos fenótipos hematológicos da drepanocitose como da malária, sendo o seu efeito superior quando em comparação com o gene *CD36* ou a com a deleção  $\alpha$ -talassémica.

Em suma, o estudo dos moduladores genéticos e a descoberta de possíveis biomarcadores de alguns subfenótipos da drepanocitose num contexto de malária, pode contribuir para aprimorar os métodos de diagnóstico e de prognóstico, contribuindo para a sua prevenção. Por outro lado, esses marcadores da gravidade da doença são úteis para categorizar os doentes por níveis de risco e assim permitem fornecer os cuidados de saúde mais atempadamente a quem mais precisa. Para além disso, as características individuais de cada doente podem contribuir para uma medicina personalizada e, dessa forma, contribuir para melhorar a qualidade de vida dos doentes e diminuir a morbilidade e até mesmo a mortalidade associada a ambas as patologias.

Como perspetiva futura seria interessante prosseguir com os estudos de associação iniciados neste projeto usando uma amostragem maior de indivíduos, tanto drepanocíticos como não drepanocíticos, para se poder validar com maior robustez os resultados obtidos. Para além disso, um estudo longitudinal mais prolongado e com acompanhamento clínico sem lacunas, principalmente no que toca à malária, seria muito útil para obter um maior número de resultados para esta patologia. Acrescenta-se ainda que os estudos de associação poderão ser mais aprofundados, avaliando-se as possíveis relações epistáticas entre os diversos genes analisados.

Seria ainda muito importante, como referido acima, prosseguir com estudos funcionais sobre as variantes genéticas cuja ação fisiopatológica é ainda desconhecida.

## 6 Referências Bibliográficas

1. Organization WH. World malaria report 2023. 2023. 1–356 p.
2. Los UMDECDE. Evolutionary origin of human and primate malarias: evidence from the circumsporozoite protein gene. *Mol Biol Evol.* 1995 Jul;
3. Saralamba N, Nosten F, Sutherland CJ, Arez AP, Snounou G, White NJ, et al. Genetic dissociation of three antigenic genes in *Plasmodium ovale curtisi* and *Plasmodium ovale wallikeri*. Braga EM, editor. *PLoS One.* 2019 Jun 6;14(6):e0217795.
4. Sutherland CJ, Tanomsing N, Nolder D, Oguike M, Jennison C, Pukrittayakamee S, et al. Two Nonrecombining Sympatric Forms of the Human Malaria Parasite *Plasmodium ovale* Occur Globally. *J Infect Dis.* 2010 May 15;201(10):1544–50.
5. Snounou G, Sharp PM, Culleton R. The two parasite species formerly known as *Plasmodium ovale*. *Trends Parasitol.* 2024 Jan;40(1):21–7.
6. Kariuki SN, Williams TN. Human genetics and malaria resistance. *Hum Genet.* 2020 Jun 1;139(6–7):801–11.
7. Haldane JBS. Disease and evolution. *Ric Sci.* 1949;(19):1–11.
8. Allison AC. Protection afforded by sickle-cell trait against subtertian malarial infection. *Br Med J.* 1954;1(4857):290–4.
9. Roos DS, Crawford MJ, Donald RG, Kissinger JC, Klimczak LJ, Striepen B. Origin, targeting, and function of the apicomplexan plastid. Vol. 2, *Current Opinion in Microbiology.* 1999. p. 426–32.
10. Bañuls AL, Thomas F, Renaud F. Of parasites and men. *Infect Genet Evol.* 2013;20:61–70.
11. Bachmann A, Metwally NG, Allweier J, Cronshagen J, del Pilar Martinez Tauler M, Murk A, et al. CD36—A Host Receptor Necessary for Malaria Parasites to Establish and Maintain Infection. *Microorganisms.* 2022;10(12):1–16.
12. European School of Haematology; Club Du Globule Rouge Et Du Fer. Disorders of erythropoiesis, erythrocytes and iron metabolism. 2009th ed. Beaumont C, Beris P, Beuzard Y, Brugnara C, editors.
13. Sinden RE. The biology of malaria transmission. In: *Advances in Malaria Research.* Hoboken, NJ, USA: John Wiley & Sons, Inc.; 2016. p. 87–124.
14. Leech JH, Barnwell JW, Aikawa M, Miller LH, Howard RJ. *Plasmodium falciparum* malaria: Association of knobs on the surface of infected erythrocytes with a histidine-rich protein and the erythrocyte skeleton. *J Cell Biol.* 1984;98(4):1256–64.

15. Crabb BS, Cowman AF. Plasmodium falciparum virulence determinants unveiled. *Genome Biol.* 2002;3(11):3–6.
16. Sherman IW, Eda S, Winograd E. Cytoadherence and sequestration in Plasmodium falciparum: Defining the ties that bind. *Microbes Infect.* 2003;5(10):897–909.
17. Molina-Franky J, Patarroyo ME, Kalkum M, Patarroyo MA. The Cellular and Molecular Interaction Between Erythrocytes and Plasmodium falciparum Merozoites. *Front Cell Infect Microbiol.* 2022;12(March):1–21.
18. Ebel ER, Kim BY, Mcdew-White M, Egan ES, Anderson TJC, Petrov DA. Antigenic diversity in malaria parasites is maintained on extrachromosomal DNA. *bioRxiv.* 2023;2023.02.02.526885.
19. Akide Ndunge OB, Kilian N, Salman MM. Cerebral Malaria and Neuronal Implications of Plasmodium Falciparum Infection: From Mechanisms to Advanced Models. *Adv Sci.* 2022 Dec 27;9(36):1–23.
20. Flick K, Chen Q. var genes, PfEMP1 and the human host. *Mol Biochem Parasitol.* 2004 Mar 1;134(1):3–9.
21. Baruch DI, Pasloske BL, Singh HB, Bi X, Ma XC, Feldman M, et al. Cloning the P. falciparum Gene Encoding PfEMP1, a Malarial Variant Antigen and Adherence Receptor on the Surface of Parasitized Human Erythrocytes. 1995;92:77–97.
22. Bull PC, Buckee CO, Kyes S, Kortok MM, Thathy V, Guyah B, et al. Plasmodium falciparum antigenic variation. Mapping mosaic var gene sequences onto a network of shared, highly polymorphic sequence blocks. *Mol Microbiol.* 2008;68(6):1519–34.
23. Gullingsrud J, Saveria T, Amos E, Duffy PE, Oleinikov A V. Structure-Function-Immunogenicity Studies of PfEMP1 Domain DBL2 $\beta$ PF11\_0521, a Malaria Parasite Ligand for ICAM-1. *PLoS One.* 2013;8(4):1–8.
24. Cabrera A, Neculai D, Tran V, Lavstsen T, Turner L, Kain KC. Plasmodium falciparum-CD36 structure-function relationships defined by ortholog scanning mutagenesis. *J Infect Dis.* 2019;219(6):945–54.
25. Dimonte S, Bruske EI, Enderes C, Otto TD, Turner L, Kremsner P, et al. Identification of a conserved var gene in different Plasmodium falciparum strains. *Malar J.* 2020;19(1):1–15.
26. Gill J, Singh H, Sharma A. Profiles of global mutations in the human intercellular adhesion molecule-1 (ICAM-1) shed light on population-specific malaria susceptibility. *BMC Genomics.* 2023;24(1):1–10.
27. Smith JE. Erythrocyte Membrane: Structure, Function, and Pathophysiology. *Vet Pathol.* 1987;24(6):471–6.

28. Kuhn V, Diederich L, Keller TCS, Kramer CM, Lückstädt W, Panknin C, et al. Red Blood Cell Function and Dysfunction: Redox Regulation, Nitric Oxide Metabolism, Anemia. *Antioxidants Redox Signal*. 2017;26(13):718–42.
29. M.F. P. I wish I'd made you angry earlier - Essays on science and scientists. Cold Spring Harbor Perspectives in Medicine. 1998;1445.
30. Ahmed MH, Ghatge MS, Safo MK. Hemoglobin: Structure, Function and Allostery. In 2020. p. 345–82.
31. Higgs DR, Engel JD, Stamatoyannopoulos G. Thalassaemia. *Lancet*. 2012;379(9813):373–83.
32. Harewood J, Azevedo AM. Alpha Thalassemia. *StatPearls*. 2023 Sep 4;
33. Kohne E. Hemoglobinopathies. *Dtsch Arztebl Int*. 2011 Aug 8;108(31–32):532–40.
34. Pauling L, Itano HA, Singer SJ, Wells IC. Sickle cell anemia, a molecular disease. *Science* (80- ). 1949 Nov 25;110(2865):543–8.
35. Sundd P, Gladwin MT, Novelli EM. Pathophysiology of Sickle Cell Disease. *Annu Rev Pathol Mech Dis*. 2019;14:263–92.
36. Kato GJ, Piel FB, Reid CD, Gaston MH, Ohene-Frempong K, Krishnamurti L, et al. Sickle cell disease. Vol. 4, *Nature Reviews Disease Primers*. 2018.
37. Kirkham FJ. Sickle cell disease. In: Rosenberg's Molecular and Genetic Basis of Neurological and Psychiatric Disease. Elsevier; 2020. p. 595–609.
38. Petersen JEV, Saelens JW, Freedman E, Turner L, Lavstsen T, Fairhurst RM, et al. Sickle-trait hemoglobin reduces adhesion to both CD36 and EPCR by Plasmodium falciparum-infected erythrocytes. *PLoS Pathog*. 2021;17(6):1–18.
39. Aidoo M, Terlouw DJ, Kolczak MS, McElroy PD, Ter Kuile FO, Kariuki S, et al. Protective effects of the sickle cell gene against malaria morbidity and mortality. *Lancet*. 2002 Apr 13;359(9314):1311–2.
40. Goheen MM, Campino S, Cerami C. The role of the red blood cell in host defence against falciparum malaria: an expanding repertoire of evolutionary alterations. *Br J Haematol*. 2017 Nov 1;179(4):543–56.
41. Piel FB, Steinberg MH, Rees DC. Sickle Cell Disease. Longo DL, editor. <https://doi.org/101056/NEJMra1510865>. 2017 Apr 20;376(16):1561–73.
42. Santos B, Ginete C, Gonçalves E, Delgado M, Miranda A, Faustino P, et al. Characterization of a cohort of Angolan children with sickle cell anemia treated with hydroxyurea. *Blood Cells, Mol Dis*. 2024 Mar;105:102822.

43. Ferraresi M, Panzieri DL, Leoni S, Cappellini MD, Kattamis A, Motta I. Therapeutic perspective for children and young adults living with thalassemia and sickle cell disease. *Eur J Pediatr*. 2023 Mar 31;182(6):2509–19.
44. World Health Organization. Accelerating anaemia reduction. 2023. 1–38 p.
45. Tavares W, Morais J, Martins JF, Scalsky RJ, Stabler TC, Medeiros MM, et al. Malaria in Angola: recent progress, challenges and future opportunities using parasite demography studies. *Malar J*. 2022;21(1):1–13.
46. Associação Angolana de Apoiantes dos Portadores de Falciformação [Internet]. [cited 2023 Feb 20]. Available from: <https://anpaf.org/>
47. Nussbaum R., McInnes R., Willard H. Thompson & Thompson Genetics in Medicine. Elsevier Health Sciences; 2015.
48. McHugh E, Bulloch MS, Batinovic S, Patrick CJ, Sarna DK, Ralph SA. Nonsense-mediated decay machinery in *Plasmodium falciparum* is inefficient and non-essential. Blader IJ, editor. *mSphere*. 2023 Jun 27;
49. Germano I, Santos B, Delgado M, Ginete C, Lopes P, Arez AP, et al. Genetic modulation of anemia severity, hemolysis level, and hospitalization rate in Angolan children with Sickle Cell Anemia. *Mol Biol Rep*. 2022 Sep 12;
50. Fröhlich B, Dasanna AK, Lansche C, Czajor J, Sanchez CP, Cyrklaff M, et al. Functionalized supported membranes for quantifying adhesion of *P. falciparum*-infected erythrocytes. *Biophys J*. 2021;120(16):3315–28.
51. Cojean S, Jafari-Guemouri S, Le Bras J, Durand R. Cytoadherence characteristics to endothelial receptors ICAM-1 and CD36 of *Plasmodium falciparum* populations from severe and uncomplicated malaria cases. *Parasite*. 2008;15(2):163–9.
52. Kalai M, Dridi M, Chaouch L, Moumni I, Ouragini H, Darragi I, et al. The role of rs1984112\_G at CD36 gene in increasing reticulocyte level among sickle cell disease patients. *Hematology*. 2017 Mar 16;22(3):178–82.
53. Coelho A, Dias A, Morais A, Nunes B, Ferreira E, Picanço I, et al. Genetic variation in CD36, HBA, NOS3 and VCAM1 is associated with chronic haemolysis level in sickle cell anaemia: a longitudinal study. *Eur J Haematol*. 2014 Mar;92(3):237–43.
54. Santos B, Delgado M, Ferreira J, Germano I, Miranda A, Arez AP, et al. Co-Inheritance of alpha-thalassemia and sickle cell disease in a cohort of Angolan pediatric patients. *Mol Biol Rep*. 2020;47(7):5397–402.
55. Allen SJ, O'Donnell A, Alexander NDE, Alpers MP, Peto TEA, Clegg JB, et al.  $\alpha^+$ -Thalassemia protects children against disease caused by other infections as well as malaria. *Proc Natl Acad Sci*. 1997 Dec 23;94(26):14736–41.

56. Edelstein SJ. Sick cell disease. *Nature*. 1976;260(5547):98–100.
57. Steinberg MH. Sick cell anemia, the first molecular disease: Overview of molecular etiology, pathophysiology, and therapeutic approaches. *ScientificWorldJournal*. 2008;8:1295–324.
58. Barnwell JW, Asch AS, Nachman RL, Yamaya M, Aikawa M, Ingravallo P. A human 88-kD membrane glycoprotein (CD36) functions in vitro as a receptor for a cytoadherence ligand on *Plasmodium falciparum*-infected erythrocytes. *J Clin Invest*. 1989;84(3):765–72.
59. Ashrafian H, Bogle RG. Invited Review - Molecular mechanisms of cytoadherence in malaria. *J Intensive Care Soc*. 2004;5(3):108–11.
60. Chen Y, Zhang J, Cui W, Silverstein RL. CD36, a signaling receptor and fatty acid transporter that regulates immune cell metabolism and fate. *J Exp Med*. 2022 Jun 6;219(6):1–15.
61. Rać ME, Safranow K, Poncyłjusz W. Molecular Basis of Human CD36 Gene Mutations. *Mol Med*. 2007 May;13(5–6):288.
62. Omi K, Ohashi J, Patarapotikul J, Hananantachai H, Naka I, Looareesuwan S, et al. CD36 polymorphism is associated with protection from cerebral malaria. *Am J Hum Genet*. 2003;72(2):364–74.
63. Lubiana P, Bouws P, Roth LK, Dörpinghaus M, Rehn T, Brehmer J, et al. Adhesion between *P. falciparum* infected erythrocytes and human endothelial receptors follows alternative binding dynamics under flow and febrile conditions. *Sci Rep*. 2020;10(1):1–14.
64. Hsieh FL, Turner L, Bolla JR, Robinson C V., Lavstsen T, Higgins MK. The structural basis for CD36 binding by the malaria parasite. *Nat Commun*. 2016;7(May).
65. Carden MA, Fasano RM, Meier ER. Not all red cells sickle the same: Contributions of the reticulocyte to disease pathology in sickle cell anemia. Vol. 40, *Blood Reviews*. Churchill Livingstone; 2020. p. 100637.
66. Anbarasan C, Bavanilatha M, Latchumanadhas K, Ajit Mulasari S. ICAM-1 molecular mechanism and genome wide SNP's association studies. *Indian Heart J*. 2015;67(3):282–7.
67. Lawson C, Wolf S. ICAM-1 signaling in endothelial cells. *Pharmacol Rep*. 2009 Jan;61(1):22–32.
68. Bui TM, Wiesolek HL, Sumagin R. ICAM-1: A master regulator of cellular responses in inflammation, injury resolution, and tumorigenesis. *J Leukoc Biol*. 2020;108(3):787–99.

69. Nortey LN, Anning AS, Nakotey GK, Ussif AM, Opoku YK, Osei SA, et al. Genetics of cerebral malaria: pathogenesis, biomarkers and emerging therapeutic interventions. *Cell Biosci.* 2022;12(1):1–19.
70. Jensen AR, Adams Y, Hviid L. Cerebral Plasmodium falciparum malaria: The role of PfEMP1 in its pathogenesis and immunity, and PfEMP1-based vaccines to prevent it. *Immunol Rev.* 2020;293(1):230–52.
71. Kucukal E, Man Y, Quinn E, Tewari N, An R, Ilich A, et al. Red blood cell adhesion to ICAM-1 is mediated by fibrinogen and is associated with right-to-left shunts in sickle cell disease. *Blood Adv.* 2020;4(15):3688–98.
72. Plowe C V., Djimde A, Bouare M, Doumbo O, Wellems TE. Pyrimethamine and Proguanil Resistance-Confering Mutations in Plasmodium falciparum Dihydrofolate Reductase: Polymerase Chain Reaction Methods for Surveillance in Africa. *Am J Trop Med Hyg.* 1995 Jun 1;52(6):565–8.
73. Thermo Scientific. NanoDrop One User Guide. Thermo Sci NanoDrop One User Guid. 2016;Revision B(July):18–9.
74. Rahman MT, Uddin MS, Sultana R, Moue A, Setu M. Polymerase Chain Reaction (PCR): A Short Review. *Anwer Khan Mod Med Coll J.* 2013;4(1):30–6.
75. Joshi M, Deshpande JD. Polymerase Chain Reaction: Methods, Principles and Application. *Int J Biomed Res.* 2011;2(1):81–97.
76. Leng Kho S, Heng Chua K, George E, Tan JAMA. A novel gap-PCR with high resolution melting analysis for the detection of  $\alpha$ -thalassaemia Southeast Asian and Filipino  $\beta$  0-thalassaemia deletion. *Sci Rep.* 2015;5(July):1–7.
77. Dodé C, Krishnamoorthy R, Lamb J, Rochette J. Rapid analysis of  $-\alpha$  3.7 thalassaemia and  $\alpha\alpha\alpha$  anti 3.7 triplication by enzymatic amplification analysis. *Br J Haematol.* 1993 Jan;83(1):105–11.
78. Germano I. Anemia das Células Falciformes - influência de fatores genéticos na gravidade da doença em idade pediátrica. Universidade de Lisboa; 2021.
79. Lee PY, Costumbrado J, Hsu CY, Kim YH. Agarose gel electrophoresis for the separation of DNA fragments. *J Vis Exp.* 2012;(62):1–5.
80. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors (DNA polymerase/nucleotide sequences/bacteriophage 4X174). *1977;74(12):5463–7.*
81. Exosap-it U. Using ExoSAP-IT™ Express PCR Product Cleanup to generate high quality BigDye™ Terminator v3.1 Cycle Sequencing Kit Data. 2017;1–17.
82. Fisher Scientific T. BigDye Terminator v3.1 Cycle Sequencing Kit Quick Reference (Pub. no. MAN0015666 Rev A.0). Cat Numbers. 2016;4337454(4337457):1–2.

83. Scientific TF, Biosystems A. DNA Fragment Analysis by Capillary Electrophoresis. In: User Guide. 2014.
84. Tullastr G, Germany E. varSEAK Online - Splice Site Prediction. 2022;1–12.
85. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. *Nat Methods*. 2010;7(4):248–9.
86. Williams TN, Thein SL. Sickle Cell Anemia and Its Phenotypes. *Annu Rev Genomics Hum Genet*. 2018 Aug 31;19:113–47.
87. W.Heil, E.Volker. Reference Ranges for Adults and Children Pre-Analytical Considerations. 9th ed. 2008.
88. Silva M, Vargas S, Coelho A, Ferreira E, Mendonça J, Vieira L, et al. Biomarkers and genetic modulators of cerebral vasculopathy in sub-Saharan ancestry children with sickle cell anemia. *Blood Cells, Mol Dis*. 2020 Jul;83(February):102436.
89. Asad M, Yamaro-Botté Y, Hossain ME, Thakur V, Jain S, Datta G, et al. An essential vesicular-trafficking phospholipase mediates neutral lipid synthesis and contributes to hemozoin formation in *Plasmodium falciparum*. *BMC Biol*. 2021;19(1):1–22.
90. World Health Organization(WHO). Country Disease Outlook Angola. 2023;(August).
91. Steinberg MH, Sebastiani P. Genetic modifiers of sickle cell disease. *Am J Hematol*. 2012;87(8):795–803.
92. Farashi S, Hartevelde CL. Molecular basis of  $\alpha$ -thalassemia. *Blood Cells, Mol Dis*. 2018;70(September 2017):43–53.
93. Olatunya OS, Albuquerque DM, Adekile A, Costa FF. Influence of alpha thalassemia on clinical and laboratory parameters among nigerian children with sickle cell anemia. *J Clin Lab Anal*. 2019;33(2):10–2.
94. Rees DC, Brousse VAM, Brewin JN. Determinants of severity in sickle cell disease. *Blood Rev*. 2022;56(June):100983.
95. Xu X, Zheng X, Zhu F. CD36 gene variants and their clinical relevance: a narrative review. *Ann Blood*. 2021;6(3).
96. De Mendonça VRR, Goncalves MS, Barral-Netto M. The host genetic diversity in malaria infection. *J Trop Med*. 2012;2012.
97. He H, Tang L, Jin Y, Wang Y, Wang H, Ding S, et al. High-Throughput CD36 Phenotyping on Human Platelets Based on Sandwich ELISA and Mutant Gene Analysis. *Transfus Med Hemotherapy*. 2024;51(1):32–40.

98. Pain A, Urban BC, Kai O, Casals-Pascual C, Shafi J, Marsh K, et al. A non-sense mutation in Cd36 gene is associated with protection from severe malaria. *Lancet*. 2001;357(9267):1502–3.
99. Sirisabhabhorn K, Chaijaroenkul W, Na-Bangchang K. Genetic Diversity of Human Host Genes Involved in Immune Response and the Binding of Malaria Parasite in Patients Residing along the Thai-Myanmar border. *Trop Med Infect Dis*. 2021 Sep 24;6(4):174.
100. Fry AE, Ghansa A, Small KS, Palma A, Auburn S, Diakite M, et al. Positive selection of a CD36 nonsense variant in sub-Saharan Africa, but no association with severe malaria phenotypes. *Hum Mol Genet*. 2009;18(14):2683–92.
101. Ellis J, Lange EM, Li J, Dupuis J, Baumert J, Walston JD, et al. Large multiethnic Candidate Gene Study for C-reactive protein levels: identification of a novel association at CD36 in African Americans. *Hum Genet*. 2014 Aug 19;133(8):985–95.
102. Love-Gregory L, Sherva R, Schappe T, Qi JS, McCrea J, Klein S, et al. Common CD36 SNPs reduce protein expression and may contribute to a protective atherogenic profile. *Hum Mol Genet*. 2011;20(1):193–201.
103. KASHIWAGI H. Molecular Basis of CD36 Deficiency. *Japanese J Thromb Hemost*. 1996;7(2):85–92.
104. Farook VS, Puppala S, Schneider J, Fowler SP, Chittoor G, Dyer TD, et al. Metabolic syndrome is linked to chromosome 7q21 and associated with genetic variants in CD36 and GNAT3 in Mexican Americans. *Obesity*. 2012;20(10):2083–92.
105. Elbers CC, Guo Y, Tragante V, van Iperen EPA, Lanktree MB, Castillo BA, et al. Gene-Centric Meta-Analysis of Lipid Traits in African, East Asian and Hispanic Populations. Gong Y, editor. *PLoS One*. 2012 Dec 7;7(12):e50198.
106. Fernández-Rhodes L, Young KL, Lilly AG, Raffield LM, Highland HM, Wojcik GL, et al. Importance of Genetic Studies of Cardiometabolic Disease in Diverse Populations. *Circ Res*. 2020;126(12):1816–40.
107. Love-Gregory L, Sherva R, Sun L, Wasson J, Schappe T, Doria A, et al. Variants in the CD36 gene associate with the metabolic syndrome and high-density lipoprotein cholesterol. *Hum Mol Genet*. 2008 Jun 1;17(11):1695–704.
108. Gao K, Masuda A, Matsuura T, Ohno K. Human branch point consensus sequence is yUnAy. *Nucleic Acids Res*. 2008;36(7):2257–67.
109. Zhang X, Tolzmann CA, Melcher M, Haas BJ, Gardner MJ, Smith JD, et al. Branch point identification and sequence requirements for intron splicing in plasmodium falciparum. *Eukaryot Cell*. 2011;10(11):1422–8.

110. Romasko EJ, Devkota B, Biswas S, Jayaraman V, Rajagopalan R, Dulik MC, et al. Utility and limitations of exome sequencing in the molecular diagnosis of pediatric inherited platelet disorders. *Am J Hematol*. 2018;93(1):8–16.
111. Pace CN, Treviño S, Prabhakaran E, Scholtz JM, Franks F, Wilson K, et al. Protein structure, stability and solubility in water and other solvents. *Philos Trans R Soc B Biol Sci*. 2004;359(1448):1225–35.
112. Sormanni P, Vendruscolo M. Protein solubility predictions using the camsol method in the study of protein homeostasis. *Cold Spring Harb Perspect Biol*. 2019;11(12):1–12.
113. Fernandez-Reyes D, Craig AG, Kyes SA, Peshu N, Snow RW, Berendt AR, et al. A high frequency African coding polymorphism in the N-terminal domain of ICAM-1 predisposing to cerebral malaria in Kenya. *Hum Mol Genet*. 1997;6(8):1357–60.
114. Blankson SO, Dadjé DS, Traikia N, Alao MJ, Ayivi S, Amoussou A, et al. ICAM-1 Kilifi variant is not associated with cerebral and severe malaria pathogenesis in Beninese children. *Malar J*. 2022 Apr 4;21(1):115.
115. Kun JFJ, Klabunde J, Lell B, Luckner D, Alpers M, May J, et al. Association of the ICAM:1(KILIFI) mutation with protection against severe malaria in lambarene, gabon. *Am J Trop Med Hyg*. 1999;61(5):776–9.
116. Ali Al-Barazanchi Z, Abdulateef S, Hassan M. Co-Inheritance of  $\alpha$ -thalassemia gene mutation in patients with sickle cell Disease: Impact on clinical and hematological variables. *Niger J Clin Pract*. 2021;24(6):874.
117. Tang W, Wang Y, Chen Y, Gu H, Chen S, Kang M. Gene and Cancer Risk : a Meta-Analysis. 2015;8(8):11996–2008.
118. Yang M, Fu Z, Zhang Q, Xin Y, Chen Y, Tian Y. Association between the polymorphisms in intercellular adhesion molecule-1 and the risk of coronary atherosclerosis: A case-controlled study. *PLoS One*. 2014;9(10).
119. Silva SL, Fonseca M, Pereira MLM, Silva SP, Barbosa RR, Serra-Caetano A, et al. Monozygotic Twins Concordant for Common Variable Immunodeficiency: Strikingly Similar Clinical and Immune Profile Associated With a Polygenic Burden. *Front Immunol*. 2019;10(November).
120. Coronado LM, Nadovich CT, Spadafora C. Malarial hemozoin: From target to tool. *Biochim Biophys Acta - Gen Subj*. 2014 Jun;1840(6):2032–41.
121. Gomez F, Tomas G, Ko WY, Ranciaro A, Froment A, Ibrahim M, et al. Patterns of nucleotide and haplotype diversity at ICAM-1 across global human populations with varying levels of malaria exposure. *Hum Genet*. 2013;132(9):987–99.

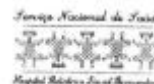
122. Sumagin R, Prizant H, Lomakina E, Waugh RE, Sarelius IH. LFA-1 and Mac-1 Define Characteristically Different Intraluminal Crawling and Emigration Patterns for Monocytes and Neutrophils In Situ. *J Immunol*. 2010;185(11):7057–66.
123. Xu X, Liu Y, Hong X, Chen S, Ma K, Lan X, et al. Variants of CD36 gene and their association with CD36 protein expression in platelets. *Blood Transfus*. 2014;12(4):557–64.
124. Mwanziva C, Mpina M, Balthazary S, Mkali H, Mbugi E, Mosha F, et al. Child hospitalization due to severe malaria is associated with the ICAM-1 Kilifi allele but not adherence patterns of *Plasmodium falciparum* infected red blood cells to ICAM-1. *Acta Trop*. 2010 Oct 1;116(1):45–50.
125. He Q, Lin X, Wang F, Xu J, Ren Z, Chen W, et al. Associations of a polymorphism in the intercellular adhesion molecule-1 (ICAM1) gene and ICAM1 serum levels with migraine in a Chinese Han population. *J Neurol Sci*. 2014 Oct 15;345(1–2):148–53.

## 7 Anexos

### Anexo 1-Documento de consentimento informado para os tutores das crianças participantes



Projeto "Anemia das células falciformes e hemoglobina fetal. Modificadores genéticos numa coorte de crianças Angolanas."



#### Consentimento informado

Estamos a fazer um estudo sobre a anemia das células falciformes em crianças seguidas em consulta no hospital. Este trabalho vai permitir identificar as principais causas genéticas da anemia do seu filho, bem como potenciais marcadores de desenvolvimento da doença e de escolha de terapias farmacológicas mais adaptadas (por exemplo uso de hidroxiureia). Para isso, para além das consultas de rotina que a criança já tem, será recolhida uma amostra de sangue para estudo genético de genes associados à doença da criança e à expressão da hemoglobina fetal (que traz melhorias à condição da criança) e, ainda, que poderão ser modificadores da sua resposta à terapêutica e da interação com outras doenças, como por exemplo, os genes dos agrupamentos génicos das globinas, o gene *BCL11A*, a região intergénica *HBS1L-MYB* e os genes *KLF1* e *SALL2*.

Todas as crianças participantes deste estudo vão também receber tratamento gratuito para a drepanocitose (hidroxiureia) durante 1 ano, não tendo despesas extra com a participação no estudo.

Para participar no estudo é necessário comparecer às consultas de seguimento da drepanocitose (conforme marcação feita pela equipa de saúde) e proceder a colheita de uma amostra de sangue periférico que se destina à realização de estudos genéticos envolvendo o(s) gene(s) e/ou as regiões regulatórias acima mencionado(s) e realização de análises de rotina tais como hemograma e parâmetros bioquímicos.

A participação neste estudo é de carácter voluntário e os riscos físicos, psicológicos ou sociais associados a essa participação são considerados mínimos. Tem o direito de recusar, em qualquer altura do projeto, sem ter nenhuma consequência.

As amostras colhidas no âmbito deste projeto não podem ser utilizadas em atividades com fins lucrativos, pelo que não existe qualquer contrapartida financeira para os indivíduos que as tenham cedido. Assim, não está previsto qualquer pagamento de eventuais despesas derivadas da sua participação.

Para garantir a privacidade das pessoas participantes e a confidencialidade e proteção dos dados, as amostras de produtos biológicos e os eventuais dados demográficos e clínicos correspondentes serão pseudonimizados e tratados exclusivamente por pessoal sujeito a sigilo profissional.

Este estudo envolve várias instituições Nacionais e Internacionais, nomeadamente o Centro de investigação em Saúde de Angola, o Hospital Pediátrico David Bernardino ambas de Angola e 3 instituições Portuguesas, a Escola Superior de Tecnologia da Saúde de Lisboa, o Instituto de Higiene e Medicina Tropical de Lisboa e o Instituto Nacional de Saúde Dr. Ricardo Jorge de Lisboa.

-----  
Cortar pelo picotado - Informação de uma ficha com responsável legal da criança, porção de baixo arquivar no CISA.

Eu (nome do inquirido) \_\_\_\_\_, na qualidade de responsável legal da criança com o nome

(Nome da Criança): \_\_\_\_\_, recebi oralmente as explicações sobre a informação que consta neste termo de

consentimento, compreendi o vosso pedido e aceito participar neste estudo, nomeadamente aceito:

- Responder às questões sobre, problemas de saúde e evolução da doença e autorizo;
- A recolha de sangue (através da picada no braço) para determinar os parâmetros hematológicos da criança, incluindo os níveis de hemoglobina fetal e os estudos genéticos nos âmbitos mencionados.
- Compreendi a informação que me foi dada, tive oportunidade de fazer perguntas e as minhas dúvidas foram esclarecidas.
- Aceito participar de livre vontade e autorizo a apresentação dos resultados obtidos em meio científico, garantindo a confidencialidade.

Data: \_\_ / \_\_ / \_\_ Assinatura ou impressão digital: \_\_\_\_\_

Eu, \_\_\_\_\_ (nome do médico), forneci as explicações referidas neste termo de consentimento e assegurei-me que o participante percebeu a pertinência e implicações deste estudo.

Data: \_\_ / \_\_ / \_\_ Assinatura do médico: \_\_\_\_\_

Investigador responsável: Brígida Santos

Contactos: Hospital Pediátrico David Bernardino; e-mail: santosbrigitaf@yahoo.com.br

## Anexo 1-Códigos União Internacional de Química Pura e Aplicada (IUPAC)

Tabela 7.1-Códigos IUPAC para nucleótidos

<b>IUPAC CÓDIGO NUCLEÓTIDO</b>	<b>BASE</b>
<b>A</b>	Adenina
<b>C</b>	Citosina
<b>G</b>	Guanina
<b>T (OU U)</b>	Timina (ou Uracilo)
<b>R</b>	A ou G
<b>Y</b>	C ou T
<b>S</b>	G ou C
<b>W</b>	A ou T
<b>K</b>	G ou T
<b>M</b>	A ou C
<b>B</b>	C ou G ou T
<b>D</b>	A ou G ou T
<b>H</b>	A ou C ou T
<b>V</b>	A ou C ou G
<b>N</b>	Qualquer base
<b>.OU -</b>	gap

**Tabela 7.2**-Código IUPAC para os diferentes aminoácidos

<b>IUPAC código do aminoácido</b>	<b>Código de 3 letras</b>	<b>Aminoácido</b>
<b>A</b>	Ala	Alanina
<b>C</b>	Cys	Cisteína
<b>D</b>	Asp	Ácido Aspártico
<b>E</b>	Glu	Ácido Glutâmico
<b>F</b>	Phe	Fenilalanina
<b>G</b>	Gly	Glicina
<b>H</b>	His	Histidina
<b>I</b>	Ile	Isoleucina
<b>K</b>	Lys	Lisina
<b>L</b>	Leu	Leucina
<b>M</b>	Met	Metionina
<b>N</b>	Asn	Asparagina
<b>P</b>	Pro	Prolina
<b>Q</b>	Gln	Glutamina
<b>R</b>	Arg	Arginina
<b>S</b>	Ser	Serina
<b>T</b>	Thr	Treonina
<b>V</b>	Val	Valina
<b>W</b>	Trp	Triptofano
<b>Y</b>	Tyr	Tirosina

## Anexo 3-Material Suplementar ao capítulo Materiais e Métodos

Tabela 7.3-Composição dos diferentes reagentes utilizados ao longo do trabalho prático

Tampão David (x10)	Tampão B	Tampão $\alpha$ + $\beta$ ME	BSA (10 mg/mL)
(NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> (2M) – 166mM	KCl <sub>2</sub> (2M) – 500mM	(NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> (1M) – 166mM	
Tris-HCl a 37% - 670mM	Tris-HCl a 37% - 100mM	Tris-HCl a 37% - 670mM	Albumin bovine serum
MgCl <sub>2</sub> (2M) – 67mM	MgCl <sub>2</sub> (1M) – 15mM	MgCl <sub>2</sub> (1M) – 15mM	
Na <sub>2</sub> EDTA (0,25M) – 0,067mM	Gelatina (0,1%)	Na <sub>2</sub> EDTA (0,25M) – 670mM	Lote SLBF1993V
$\beta$ - Mercaptoetanol – 100mM	dH <sub>2</sub> O- 26,73mL	Na <sub>2</sub> $\beta$ - Mercaptoetanol	
	dNTPs (100mM)		

Tabela 7.4-Descrição dos *primers* utilizados na genotipagem das diferentes regiões do gene *CD36*

Gene	Região gênica	Primer	Sequência oligonucleotídica (5' para 3')	Conteúdo GC (%)	T <sub>m</sub> (°C)	Referência
<i>CD36</i>	Região promotora	CD36_33137_Fw	CAT TCT TCT CTT TCT CCT TAA GG	39	52	(49)
		CD36_33137_Rv	CTC AGT TAT CAG GAT CCT ACT C	45	53	
	Exão 5	CD36_Ex5_Fw	CCC CTT CTC GTT AGT TTG CT	50	52	(123)
		CD36_Ex5_Rv	TTT CTT ACA GGC TGC GTT TG	45	50	
	Exão 10	CD36_SNPex10_Fw	GCT GTA TTT GAA TCC GAC GT	45	50	(49)
		CD36_SNPex10_Rv	CTG CCA CCA TTC TTT CTT C	47	51	

**Tabela 7.5-**Descrição dos *primers* utilizados na genotipagem do terceiro intrão do gene *CD36*, para o estudo do STR In3(TG)n

Gene	Variante /Região gênica	Primer	Sequência oligonucleotídica (5'para 3')	Conteúdo GC (%)	Tm (°C)	Referência
<i>CD36</i>	In3(TG)n / Intrão 3	CD36_TGn _Fw	FAM-ATT TAT AAA AGA AGT TGC AC	25	42	(62)
		CD36_TGn _Rv	GTT TAC CTA CAA TTT AAT AA	20	40	
		CD36_IVS3 _FW	TGC AAC TCT ATA TGA TGC G	42	48	Desenhado, PrimerBlast
		CD36_IVS3 _Rv	AGC TAA TTC GAA TAA GCT AGG G	41	51	

**Tabela 7.6-**Descrição dos *primers* utilizados na genotipagem das diferentes regiões do gene *ICAM-1*

Gene	Região gênica	Primer	Sequência oligonucleotídica (5'para 3')	Conteúdo GC (%)	Tm (°C)	Referência
<i>ICAM-1</i>	Exão 2	ICAM1_K29M_ Fw	CCT GAT TTG TAA TGC CTG TCG	48	52	(124)
		ICAM1_K29M_ Rv	CTT GAC CCT ACG AGA GTG	53	51	
	Exão 4 a 6	ICAM1_G241R _Fw	GCG ACT CCC CCA CAA CTT	61	53	(125)
		ICAM1_K469E _Rv	GGT GAG GAT TGC ATT AGG TC	50	52	

**Tabela 7.7-**Descrição dos *primers* utilizados na genotipagem da deleção talassêmica de 3,7kb ocorrida nos genes *HBA2* e *HBA1*

Gene	Variante	Primer	Sequência oligonucleotídica (5'para 3')	Conteúdo GC (%)	Tm (°C)	Referência
<i>HBA</i>	Deleção talassêmica, - $\alpha^{3,7}$	C10- A_5'UTR_Fw	GGG ATG CAC CCA CTG GCA CT	65	58	(77)
		C2- A1_3'UTR_Rv	CTG CTG TCC ACG CCC ATG CC	70	60	

**Anexo 4-Valores de referência dos parâmetros hematológicos para a população portuguesa em idade pediátrica**

<b>Parâmetro</b>	<b>Idade</b>	<b>Valores de referência*</b>
<b>Hb (g/dL)</b>	6 meses-11 anos	11,0-14,0
<b>VGM (fL)</b>	6 meses-11 anos	72,0-86,6
<b>HGM (pg)</b>	>5 anos	26,0-34,0
<b>Eritrócitos (x10<sup>12</sup>/L)</b>	6 meses-11 anos	3,80-5,40
<b>Plaquetas (x10<sup>9</sup>/L)</b>	>2 meses	140-440
<b>Leucócitos (x10<sup>9</sup>/L)</b>	5-11 anos	4,5-13,0
<b>Neutrófilos (x10<sup>9</sup>/L)</b>	5-11 anos	1,5-8,5

\*Segundo a Norma nº 063/2011 da DGS