

UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE BIOLOGIA ANIMAL



Caracterização molecular de Hemoglobinopatias na população portuguesa – um sub-estudo do projeto INSEF

Daniela Filipa Claudino Santos

Mestrado em Biologia Humana e Ambiente

Dissertação orientada por:
Doutora Paula Faustino (paula.faustino@insa.min-saude.pt) – Instituto Nacional de
Saúde Doutor Ricardo Jorge
Prof. Doutora Deodália Dias (dmdias@fc.ul.pt) – Faculdade de Ciências da Universidade
de Lisboa

Agradecimentos

Antes de mais, gostaria de 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 do INSA por permitirem a realização deste trabalho.

À minha orientadora, Doutora Paula Faustino, estou imensamente grata por me ter aceitado no Grupo de Investigação em Hemoglobinopatias, Metabolismo do Ferro e Patologias Associadas, por me ter dado a oportunidade de realizar este trabalho, pela sua disponibilidade e pela orientação neste ano tão atípico, tanto no laboratório como à distância.

Aos restantes membros do grupo, Marisa Silva e Pedro Lopes, agradeço toda a ajuda, acompanhamento, esclarecimentos. Deixo um especial agradecimento à minha companheira de bancada e amiga Isabel Germano. Obrigada por ouvires as minhas preocupações, as dúvidas existenciais e por teres sempre uma palavra amiga para me dar.

Agradeço ainda ao Doutor Carlos Matias Dias, coordenador do Departamento de Epidemiologia do INSA e ao Doutor Baltazar Nunes, coordenador da Unidade de Investigação Epidemiológica do Departamento de Epidemiologia do INSA por também eles permitirem a realização deste trabalho. Deixo também o meu agradecimento à Doutora Marta Barreto, coordenadora do biobanco do INSEF*, pela disponibilização das amostras e à Doutora Irina Kislaya, pela cedência da base de dados contendo informação demográfica e de saúde dos indivíduos estudados.

Também não poderia deixar de agradecer aos membros da Unidade de Tecnologia e Inovação do INSA, tanto o Dr. José Ferrão na execução da metodologia *Multiplex Ligation-Dependent Probe Amplification* (MLPA) como ao Doutor Luís Vieira, Joana Mendonça e Miguel Machado pela ajuda na metodologia *Next-Generation Sequencing* (NGS).

À Professora Doutora Deodália Dias, agradeço por ter aceitado ser minha orientadora na Faculdade de Ciências da Universidade de Lisboa, pelo apoio e inteira disponibilidade que sempre demonstrou ter para com os alunos do ano 2018/2019 do mestrado de Biologia Humana e Ambiente. Estamos gratos pelo seu acompanhamento e carinho. E por falar nisso, obrigada colegas!

Às minhas amigas, Luísa Marques, Mariana Ferreira, Iara Dias, Raquel Silva e Marta Cardoso, não há palavras que exprimam o quão grata estou por ter o vosso carinho, apoio e amizade. Obrigada por todas as energias positivas enviadas, por todas as dúvidas que me esclareceram e por acreditarem em mim, quando eu própria não o conseguia fazer. Por fim queria agradecer à minha colega e amiga Alexandra Oliveira, por ter despendido o seu tempo a ler este trabalho e pela ajuda na sua revisão.

Ao meu namorado, André Lameirinhas, tanto por me secar as lágrimas de ansiedade e nervos, como por partilhar comigo a sua secretária onde fiz grande parte desta dissertação. Obrigada por me motivares e por me ajudares, mesmo não sabendo quase nada desta área. O caminho foi muito mais fácil a teu lado do que se o tivesse percorrido sozinha.

Mas acima de tudo, quero e tenho de agradecer à minha família que foram o meu maior apoio e a minha motivação. Com todo o carinho, dedico esta tese à minha mãe, a minha maior inspiração. Obrigada por acreditares SEMPRE em mim e por me teres ensinado todos os princípios necessários para realizar este trabalho: o sentido de responsabilidade, a vontade de trabalhar e de dar o meu melhor deve-se inteiramente a ti.

*O INSEF, desenvolvido no âmbito do Projeto Pré-definido do Programa Iniciativas em Saúde Pública, foi promovido pelo Instituto Nacional de Saúde Doutor Ricardo Jorge através do Departamento de Epidemiologia e beneficiou de apoio financeiro concedido pela Islândia, Liechtenstein e Noruega, através das EEA Grants.

Resumo

A anemia é um problema de saúde pública a nível mundial que ocorre quando o número de glóbulos vermelhos em circulação ou a sua capacidade de transporte de oxigénio é insuficiente para atender às necessidades do organismo. A anemia pode ter uma origem ambiental, derivada da carência nutricional em ferro, mas também pode ter uma origem genética, onde se enquadram as anemias hereditárias, derivadas de alterações nos genes das cadeias globínicas da hemoglobina. Estas anemias hereditárias são denominadas por hemoglobinopatias, têm transmissão autossómica recessiva e incluem as talassémias e as variantes da hemoglobina. As talassémias, no estado de portador, apresentam-se geralmente com um fenótipo de microcitose e/ou hipocromia e eventualmente anemia.

Os últimos estudos de prevalência de hemoglobinopatias efetuados em Portugal foram realizados há duas décadas e nenhum deles incluiu a análise da população insular. Tendo em conta os movimentos populacionais existentes nos últimos tempos, sobretudo provenientes das regiões Africanas e Sul Americanas para a Europa, é de esperar que o padrão de distribuição destas patologias se tenha vindo a alterar. Deste modo, a presente dissertação tem como principal objetivo contribuir para o conhecimento da realidade atual das hemoglobinopatias em Portugal. Para isso, foram analisados molecularmente 204 indivíduos com fenótipo hematológico de hipocromia e/ou microcitose, residentes em Portugal continental e ilhas há pelo menos 12 meses. Estes indivíduos foram selecionados a partir de 4808 participantes no projeto “Inquérito Nacional de Saúde com Exame Físico” (INSEF), um estudo epidemiológico realizado pelo Instituto Nacional de Saúde Doutor Ricardo Jorge (INSA) em 2015, representativo da população residente em Portugal.

De forma a determinar a base molecular das hemoglobinopatias na população portuguesa, foram analisados os DNAs dos 204 indivíduos acima referidos quanto aos genes *HBA* e *HBB*. Para a pesquisa de deleções/inserções, os genes *HBA* foram estudados através das metodologias de Gap-PCR e *Multiplex Ligation-Dependent Probe Amplification* (MLPA). Identificámos a presença da deleção α -talassémica de 3,7 kb em 53 indivíduos, da deleção de 4,2 kb em um indivíduo e ainda a presença de dois indivíduos com os genes da α -globina triplicados. O gene *HBB* foi estudado através de PCR seguido de *Next-Generation Sequencing* (NGS). Procedeu-se à análise bioinformática dos resultados de NGS, tendo sido identificadas 10 tipos diferentes de variantes genéticas possivelmente patogénicas. Para estas variantes, foram realizados estudos *in silico* através das ferramentas bioinformáticas *PolyPhen-2*, *Sorting Intolerant From Tolerant* (SIFT) e *varSEAK*, de modo a analisar o seu impacto ao nível da estrutura e função das respetivas proteínas e ao nível do *splicing*. As variantes consideradas patogénicas ou possivelmente patogénicas foram confirmadas por sequenciação de Sanger, tendo sido então identificadas sete tipos diferentes de mutações responsáveis por β -talassémia [Cd39 (C>T), IVS-I-6 (T>C), IVS-I-110 (G>A), IVS-I-1 (G>A), Cd15 (G>A), Cd6(-A) e Cd41/42 (-CTTT)] e três variantes de hemoglobina (Hb S, Hb C e Hb D-Portugal).

Neste estudo também se pretendeu determinar a frequência das diferentes hemoglobinopatias detetadas na população portuguesa avaliada no projeto INSEF que apresentava um fenótipo hematológico de hipocromia e/ou microcitose. Obtivemos neste sub-grupo da população uma frequência de 26,6% para a α -talassémia, 10,8% para a β -talassémia e 1,5% para as variantes de hemoglobina. Assim, os nossos resultados revelaram que 38,9% destes casos sintomáticos têm uma origem genética – uma hemoglobinopatia. Este valor, bastante significativo, deverá alertar os clínicos, sobretudo da medicina geral e familiar para esta possibilidade aquando do processo conducente ao diagnóstico de um novo caso. Os nossos resultados permitiram, ainda, determinar uma prevalência de 0,5% de portadores de β -talassémia para a população residente em Portugal continental e ilhas.

Descrevemos ainda as características demográficas e de saúde dos indivíduos analisados. De entre os 204 indivíduos com microcitose e/ou hipocromia, 76 também apresentavam anemia, mas apenas dois deles referiram ter conhecimento desta condição. A grande maioria dos portadores de talassémia detetados percecionava a sua saúde como sendo normal e, eventualmente, desconhece (não declarou ter conhecimento) de que é portador de uma doença genética.

Em conclusão, acreditamos que através deste estudo contribuímos e aprofundámos o conhecimento atual da base molecular das hemoglobinopatias na população residente em Portugal. Acreditamos também que este conhecimento poderá contribuir para que sejam estabelecidas melhores estratégias de deteção e de acompanhamento dos casos de hemoglobinopatias em Portugal, permitindo melhorar a qualidade de vida dos portugueses.

Palavras-chave: Anemia, Hemoglobinopatias, Talassémias, Variantes de Hemoglobina, INSEF.

Abstract

Anemia is a worldwide public health problem that occurs when the number of circulating red blood cells is low or their oxygen carrying capacity is insufficient to meet the body's needs. Its origin could be related with nutritional factors, such as an iron deficient diet or genetic factors, namely hereditary anaemias with an autosomal recessive inheritance. Hereditary anemias occur due to genetic alterations that lead to changes in the hemoglobin molecules and are called hemoglobinopathies, which can be divided in thalassemias or hemoglobin variants. In the carrier state, thalassemia usually presents with a phenotype of microcytosis and/or hypochromia and eventually anemia.

The last studies regarding the prevalence of hemoglobinopathies in Portugal were carried out two decades ago and none of them included an analysis of the insular Portuguese population. In addition, given the population movements that took place in recent years, it is expected that the pattern of distribution of these pathologies has changed. Thus, the main objective of this dissertation is to contribute to the knowledge of the current reality of hemoglobinopathies in Portugal.

With this goal in mind, we analysed molecularly 204 individuals selected from the project “Inquérito Nacional de Saúde com Exame Físico” (INSEF), an epidemiological study carried out in 2015 by Instituto Nacional de Saúde Doutor Ricardo Jorge representative of the resident population in Portugal. These individuals have a hematological phenotype of hypochromia and/or microcytosis and have been resident in mainland Portugal or islands for at least 12 months.

To determine the molecular basis of the hemoglobinopathies present in the sample, we analysed the *HBA* and *HBB* genes. The *HBA* gene was studied for the search for deletions/insertions using Gap-PCR and Multiplex Ligation-Dependent Probe Amplification (MLPA) methodologies. We identified 53 individuals with the α -thalassemic 3.7 kb deletion, one with the 4.2 kb deletion and two with triplicated α -globin genes. The *HBB* gene was studied using PCR, followed by Next-Generation Sequencing (NGS). Through the bioinformatics analysis of the NGS results, we identified 10 possibly pathogenic variants. We carried out *in silico* studies using bioinformatics tools like varSEAK, PolyPhen-2 and Sorting Intolerant From Tolerant (SIFT) to analyse their impact in proteins' splicing, structure and function. Variants considered pathogenic were confirmed by Sanger sequencing. We identified seven different types of mutations responsible for β -thalassemia [Cd39 (C>T), IVS-I-6 (T>C), IVS-I-110 (G>A), IVS-I-1 (G>A), Cd15 (G>A), Cd6(-A) e Cd41/42 (-CTTT)] and three hemoglobin variants (Hb S, Hb C e Hb D-Portugal).

Based on these molecular findings, we discovered a frequency of 26.6% for α -thalassemia, 10.8% for β -thalassemia and 1.5% for hemoglobin variants. Thus, our results revealed that 38.9% of these symptomatic cases have a genetic origin – a hemoglobinopathy. Our results also allowed us to determine a prevalence of 0.5% of β -thalassemia carriers for the Portuguese population.

We also described the demographic and health characteristics of the individuals with hemoglobinopathies. Among the 204 individuals with microcytosis and/or hypochromia, 76 also had anemia, but only two of them reported having knowledge of this condition. Thus, most carriers are unaware of their status as a carrier of a genetic disease.

In conclusion, we believe that with this study we contributed to the current knowledge of the molecular basis of hemoglobinopathies in the population residing in Portugal. This knowledge can contribute to the establishment of better strategies for detecting and monitoring cases of hemoglobinopathies in Portugal, allowing to improve the quality of life of the Portuguese.

Keywords: Anemia, Hemoglobinopathies, Thalassemia, Hemoglobin Variants, INSEF.

Índice

Agradecimentos.....	I
Resumo.....	II
Abstract	IV
Índice.....	VI
Lista de Figuras	VIII
Lista de Tabelas.....	X
Lista de Abreviaturas.....	XII
1. Introdução.....	1
1.1. Estrutura e função da hemoglobina	1
1.2. Organização e estrutura dos genes globínicos humanos	1
1.3. As diferentes hemoglobinas ao longo do desenvolvimento	3
1.4. Patologias relacionadas com défice de hemoglobina – a Anemia.....	4
1.4.1. Anemias hereditárias por alterações nas cadeias globínicas – as Hemoglobinopatias	4
1.4.1.1. Variantes estruturais da hemoglobina	5
1.4.1.2. Talassémias	7
1.4.1.3. A prevalência das talassémias a nível mundial e em Portugal	12
1.4.1.4. Diagnóstico.....	13
1.4.1.5. Tratamento.....	15
2. Objetivos	19
3. Materiais e Métodos	20
3.1. Amostra Populacional	20
3.2. Amostra Biológica.....	21
3.3. Quantificação e verificação da qualidade e pureza do DNA.....	21
3.4. Reação em Cadeia da Polimerase (<i>Polymerase Chain Reaction</i> - PCR).....	21
3.5. Gap-PCR	22
3.6. Amplificação Multiplex de Sondas Dependente de Ligação (<i>Multiplex Ligation-dependent Probe Amplification</i> - MLPA)	23
3.7. Sequenciação de Nova Geração (<i>Next-Generation Sequencing</i> – NGS).....	24
3.7.1. Preparação de Bibliotecas para NGS e Sequenciação	25
3.7.2. Análise Bioinformática e Determinação de Variantes	25
3.8. Sequenciação pelo método de Sanger	27
3.9. Análise bioinformática das variantes encontradas e previsão dos seus efeitos	27
3.9.1. Análise do impacto das variantes ao nível do <i>splicing</i>	28
3.9.2. Análise do impacto das variantes ao nível da proteína.....	29
3.10. Análise Estatística	30
4. Resultados e Discussão	31

4.1. Caracterização da população em estudo.....	31
4.2. Caracterização dos parâmetros hematológicos.....	32
4.3. Diagnóstico de alterações nos genes <i>HBA</i> : α -talassémia	33
4.3.1. Associação da α -talassémia com os parâmetros hematológicos.....	35
4.4. Diagnóstico de alterações no gene <i>HBB</i> : β -talassémia e variantes de hemoglobina.....	37
4.4.1. Hb D-Portugal (<i>HBB</i> :c.364G>C)	38
4.4.2. Cd 41/42 (-CTTT) (<i>HBB</i> :c.126_129delCTTT)	40
4.4.3. Cd 39 (C>T) (<i>HBB</i> :c.118 C>T).....	42
4.4.4. IVS-I-110 (G>A) (<i>HBB</i> :c.93-21G>A)	43
4.4.5. IVS-I-6 (T>C) (<i>HBB</i> :c.92+6T>C).....	46
4.4.6. IVS-I-1 (G>A) (<i>HBB</i> :c.92+1G>A)	50
4.4.7. Cd15 (G>A), (<i>HBB</i> :c.48G>A)	53
4.4.8. Cd 6 (-A), (<i>HBB</i> :c.20delA)	54
4.4.9. Hb S (<i>HBB</i> :c.20A>T).....	55
4.4.10. Hb C (<i>HBB</i> :c.19G>A)	58
4.4.11. Prevalência da β -talassémia e de variantes de hemoglobina	61
4.4.12. Associação da β -talassémia e de variantes de hemoglobina com os parâmetros hematológicos.....	62
4.5. Hemoglobinopatias em Portugal	64
4.5.1. Associação das hemoglobinopatias com as características demográficas e de saúde	65
5. Conclusão.....	73
6. Referências	76
7. Anexos.....	83

Lista de Figuras

Figura 1.1. Estrutura da hemoglobina.	1
Figura 1.2. Agrupamento génico da α -globina e da β -globina humana.	2
Figura 1.3. Representação esquemática da percentagem das várias hemoglobinas durante o desenvolvimento.....	3
Figura 1.4. Tipos de células precursoras eritroides, locais da eritropoiese e níveis de expressão das várias cadeias globínicas nos diferentes estadios de desenvolvimento.....	4
Figura 1.5. Indução do formato de foice nos glóbulos vermelhos pela polimerização da Hb S desoxigenada e fisiopatologia da doença falciforme.....	6
Figura 1.6. Classificação da α -talassémia e representação diagramática das alterações nos genes da α -globina.....	8
Figura 1.7. Eventuais complicações da talassémia <i>major</i>	11
Figura 1.8. Prevalência (%) de portadores de β -talassémia em Portugal continental.....	13
Figura 1.9. Fluxograma para o diagnóstico de portadores de hemoglobinopatias.	15
Figura 3.1. Representação esquemática dos genes <i>HBA1</i> , <i>HBA2</i> e do gene híbrido resultante, caso esteja presente a deleção $-\alpha^{3,7}$	23
Figura 3.2. Sonda de MLPA.....	23
Figura 4.1. Distribuição por idade e por sexo da população selecionada com microcitose e/ou hipocromia.....	31
Figura 4.2. Distribuição da população analisada com fenótipo hematológico por local de nascimento e de residência.....	32
Figura 4.3. Análise do Gap-PCR para o diagnóstico da deleção $-\alpha^{3,7}$	34
Figura 4.4. Análise esquemática de dois resultados de pesquisa de CNVs no braço curto do cromossoma 16 por MLPA.	34
Figura 4.5. Visualização e análise quantitativa e qualitativa do produto do PCR para o diagnóstico de variantes no gene <i>HBB</i>	37
Figura 4.6. Previsão do efeito patogénico da variante p.Glu122Gln na proteína da β -globina através do <i>software PolyPhen-2</i>	39
Figura 4.7. Perfil de conservação evolutiva do aminoácido p.Glu122 na proteína da β -globina.....	39
Figura 4.8. Verificação por sequenciação de Sanger da variante <i>HBB</i> :c.364G>C.....	40
Figura 4.9. Alteração da sequência de nucleótidos no gene <i>HBB</i> , gerada pela variante <i>HBB</i> :c.126_129delCTTT.	41
Figura 4.10. Verificação por sequenciação de Sanger da variante <i>HBB</i> :c.126_129delCTTT.	42
Figura 4.11. Verificação por sequenciação de Sanger da variante <i>HBB</i> : c.118 C>T.....	43
Figura 4.12. Análise do impacto da alteração c.93-21G>A no processo de <i>splicing</i> do RNA codificante da proteína da β -globina através do <i>software varSEAK</i>	44
Figura 4.13. Sequência de aminoácidos correspondente aos exões 1 e 2 do gene da β -globina normal e com a variante c.93-21G>A.	45
Figura 4.14. Verificação por sequenciação de Sanger da variante <i>HBB</i> :c.93-21G>A.	46

Figura 4.15. Análise do impacto da alteração c.92+6T>C no processo de <i>splicing</i> do RNA codificante da proteína da β -globina através do <i>software</i> varSEAK.	47
Figura 4.16. Sequência de aminoácidos correspondente aos exões 1 e 2 do gene da β -globina normal e com a variante c.92+6T>C.	48
Figura 4.17. Verificação por sequenciação de Sanger da variante <i>HBB</i> :c.92+6T>C.	49
Figura 4.18. Análise do impacto da alteração c.92+1G>A no processo de <i>splicing</i> do RNA codificante da proteína da β -globina através do <i>software</i> varSEAK.	50
Figura 4.19. Sequência de aminoácidos correspondente aos exões 1 e 2 do gene da β -globina normal e com a variante c.92+1G>A.	51
Figura 4.20. Verificação por sequenciação de Sanger da variante <i>HBB</i> :c.92+1G>A.	52
Figura 4.21. Verificação por sequenciação de Sanger da variante <i>HBB</i> :c.48 G>A.	53
Figura 4.22. Verificação por sequenciação de Sanger da variante <i>HBB</i> :c.20delA.	54
Figura 4.23. Previsão do efeito patogénico da variante p.Glu7Val na proteína da β -globina através do <i>software</i> PolyPhen-2.	55
Figura 4.24. Perfil de conservação evolutiva do aminoácido p.Glu7 na proteína da β -globina.	56
Figura 4.25. Verificação por sequenciação de Sanger da variante <i>HBB</i> :c.20A>T.	57
Figura 4.26. Previsão do efeito patogénico da variante p.Glu7Lys na proteína da β -globina através do <i>software</i> PolyPhen-2.	58
Figura 4.27. Perfil de conservação evolutiva do aminoácido p.Glu7 na proteína da β -globina.	59
Figura 4.28. Morfologia dos glóbulos vermelhos com Hb C.	59
Figura 4.29. Verificação por sequenciação de Sanger da variante <i>HBB</i> :c.19G>A.	60
Figura 4.30. Distribuição dos portadores de hemoglobinopatias pelos seus países de origem.	67
Figura 4.31. Perceção de saúde dos indivíduos com hemoglobinopatias que possuem médico de família.	71
Figura 7.1. Declaração de Consentimento Informado – Informação ao Participante (Parte I).	83
Figura 7.2. Declaração de Consentimento Informado – Informação ao Participante (Parte II).	84
Figura 7.3. Declaração de Consentimento Informado.	85
Figura 7.4. Marcador de massa molecular usado na eletroforese em gel de agarose.	88
Figura 7.5. Sondas MLPA do agrupamento génico da α -globina.	90
Figura 7.6. Código genético.	91
Figura 7.7. Publicação em <i>Proceedings</i>	95

Lista de Tabelas

Tabela 1.1. Genótipos comuns, classificação e fenótipos da α -talassémia.....	8
Tabela 1.2. Genótipos comuns, classificação e fenótipos da β -talassémia.....	11
Tabela 3.1. Amplicões desenhados para o estudo do gene <i>HBB</i> por NGS.....	25
Tabela 3.2. Sequências <i>RefSeq</i> usadas para a análise comparativa dos resultados obtidos por sequenciação de Sanger.....	27
Tabela 4.1. Estatística descritiva dos parâmetros hematológicos e teste de normalidade Shapiro-Wilk	33
Tabela 4.2. Genótipos detetados no gene <i>HBA</i> e frequências relativas para os 204 indivíduos analisados	35
Tabela 4.3. Associação entre os parâmetros hematológicos estudados e a ausência ou presença de α -talassémia	36
Tabela 4.4. Variantes patogénicas no gene <i>HBB</i> identificadas por NGS	38
Tabela 4.5. Resultados do NGS para a variante c.364G>C no gene <i>HBB</i>	40
Tabela 4.6. Resultados do NGS para a variante c.126_129delCTTT no gene <i>HBB</i>	41
Tabela 4.7. Resultados do NGS para a variante c.118 C>T no gene <i>HBB</i>	43
Tabela 4.8. Resultados do NGS para a variante c.93-21G>A no gene <i>HBB</i>	46
Tabela 4.9. Resultados do NGS para a variante c.92+6T>C no gene <i>HBB</i>	49
Tabela 4.10. Resultados do NGS para a variante c.92+1G>A no gene <i>HBB</i>	52
Tabela 4.11. Resultados do NGS para a variante c.48G>A no gene <i>HBB</i>	53
Tabela 4.12. Resultados do NGS para a variante c.20delA no gene <i>HBB</i>	54
Tabela 4.13. Resultados do NGS para a variante c.20A>T no gene <i>HBB</i>	57
Tabela 4.14. Resultados do NGS para a variante c.19G>A no gene <i>HBB</i>	60
Tabela 4.15. Genótipos detetados no gene <i>HBB</i> e frequências relativas para os 204 indivíduos analisados	61
Tabela 4.16. Associação entre os parâmetros hematológicos estudados e a ausência de talassémia ou presença de β -talassémia	62
Tabela 4.17. Associação entre os parâmetros hematológicos estudados e a ausência ou presença de variantes de hemoglobina.....	64
Tabela 4.18. Características demográficas, de saúde e moleculares dos 204 participantes deste estudo	65
Tabela 4.19. Prevalência da α -talassémia por região de Portugal tendo em conta os seus pesos amostrais	67
Tabela 4.20. Prevalência da β -talassémia por região de Portugal tendo em conta os seus pesos amostrais	68
Tabela 4.21. Diversidade, frequência e distribuição geográfica das mutações no gene <i>HBB</i> detetadas nos 204 participantes	69

Tabela 4.22. Diversidade, frequência e distribuição geográfica das mutações no gene <i>HBA</i> detetadas nos 204 participantes	70
Tabela 7.1. Valores de referência dos parâmetros hematológicos para indivíduos adultos normais	86
Tabela 7.2. Condições do PCR convencional para a amplificação do fragmento de interesse no gene <i>HBB</i>	86
Tabela 7.3. Condições do Gap-PCR para a detecção da deleção $-\alpha^{3,7}$	87
Tabela 7.4. Composição das soluções tampão utilizadas	88
Tabela 7.5. Sondas MLPA do agrupamento génico da α -globina dispostas de acordo com a sua localização cromossómica.....	89
Tabela 7.6. Condições do MLPA para a detecção de CNV no agrupamento génico da α -globina.....	90
Tabela 7.7. Condições para a Sequenciação Automática pelo método de Sanger	91
Tabela 7.8. Código de nucleótidos da União Internacional de Química Pura e Aplicada (IUPAC)	92
Tabela 7.9. Código de aminoácidos da IUPAC.....	92
Tabela 7.10. Variantes detetadas por NGS nos 204 indivíduos analisados neste estudo	93
Tabela 7.11. Sexo, idade e características hematológicas dos indivíduos identificados por NGS como portadores de uma variante no gene <i>HBB</i>	94

Lista de Abreviaturas

A	Adenina
C	Citosina
CHGM	Concentração da Hemoglobina Globular Média
CNV	Variação de número de cópias (<i>Copy Number Variation</i>)
ddNTPs	Didesoxirribonucleótidos trifosfatados
dL	Decilitros
DGS	Direção-Geral da Saúde
DNA	Ácido desoxirribonucleico
dNTPs	Desoxirribonucleótidos trifosfatados
DP	Profundidade de leitura
EDTA	Ácido etilenodiamino tetra-acético
EtBr	Brometo de etídio
FDA	<i>Food and Drug Administration</i>
fL	Fentolitros
G	Guanina
g	Grama
Hb A	Hemoglobina adulta
Hb A₂	Variante normal da hemoglobina adulta
Hb C	Hemoglobina C
Hb F	Hemoglobina fetal
Hb H	Hemoglobina H
Hb S	Hemoglobina S
HBA	Gene que codifica a α -globina
HBA1	Gene que codifica a subunidade 1 da α -globina
HBA2	Gene que codifica a subunidade 2 da α -globina
HBB	Gene que codifica a β -globina
HBD	Gene que codifica a δ -globina
HBE	Gene que codifica a ϵ -globina
HBG1	Gene que codifica a $A\gamma$ -globina
HBG2	Gene que codifica a $G\gamma$ -globina
HBZ	Gene que codifica a ζ -globina
HGM	Hemoglobina Globular Média
HPLC	Cromatografia líquida de alto desempenho

HS	Local hipersensível à DNase I (<i>DNase I Hypersensitive Site</i>)
indel	Inserção/deleção
INSA	Instituto Nacional de Saúde Doutor Ricardo Jorge
INSEF	Inquérito Nacional de Saúde com Exame Físico
IUPAC	União Internacional de Química Pura e Aplicada (<i>International Union of Pure and Applied Chemistry</i>)
IVS	Sequência interveniente (<i>Intervening Sequence</i>)
kb	Quilobases
LCR	<i>Locus Control Region</i>
LVT	Lisboa e Vale do Tejo
MCS	<i>Multispecies Conserved Sequences</i>
MED	Mediterrâneo
mg	Miligrama
mL	Mililitro
MLPA	Amplificação <i>multiplex</i> de sondas dependente de ligação (<i>Multiplex Ligation-dependent Probe Amplification</i>)
mM	Milimolar
mRNA	RNA mensageiro
ng	Nanograma
NGS	Sequenciação de nova geração (<i>Next-Generation Sequencing</i>)
nm	Nanómetro
NMD	<i>Nonsense-Mediated Decay</i>
nt	Nucleótidos
OMS	Organização Mundial de Saúde
pb	Pares de Bases
PNAEQ	Programa Nacional de Avaliação Externa da Qualidade
PCR	Reação em cadeia da polimerase (<i>Polymerase Chain Reaction</i>)
pg	Picograma
pmol	Picomoles
RA	Região Autónoma
RDW	Amplitude de distribuição dos glóbulos vermelhos (<i>Red cell Distribution Width</i>)
RNA	Ácido ribonucleico
SEA	Sudeste Asiático
SIFT	<i>Sorting Intolerant From Tolerant</i>
SNP	Polimorfismo de nucleótido único (<i>Single Nucleotide Polymorphism</i>)
T	Timina

TBE	Tris-Borato-EDTA
U	Uracilo
UTI	Unidade de Tecnologia e Inovação
UTR	Região não traduzida (<i>untranslated region</i>)
UV	Luz Ultravioleta
V	Voltes
VGM	Volume Globular Médio
ZnPP	Protoporfirina de zinco
α	<i>Alpha</i>
β	<i>Beta</i>
γ	<i>Gamma</i>
δ	<i>Delta</i>
ε	<i>Epsilon</i>
ζ	<i>Zeta</i>
θ	<i>Theta</i>
μ	<i>Mu</i>
μg	Micrograma
μL	Microlitro
Ψ	<i>Psi</i>

1. Introdução

1.1. Estrutura e função da hemoglobina

A hemoglobina presente nos glóbulos vermelhos é uma proteína essencial para a vida humana, pois tem como função transportar oxigénio dos pulmões para os tecidos (1).

A estrutura tridimensional da hemoglobina (Figura 1.1) foi descoberta em 1959 por Max Perutz, usando cristalografia de raios X (1). A hemoglobina humana adulta é uma proteína globular com uma estrutura quaternária, constituída por quatro subunidades polipeptídicas - duas cadeias α e duas cadeias β ($\alpha_2\beta_2$) não ligadas covalentemente, compostas, respetivamente, por 141 e 146 resíduos de aminoácidos. Cada subunidade polipeptídica está associada a um grupo heme prostético não proteico, que consiste num ião de ferro contido num anel de porfirina (2). O ião ferro é o local de ligação do oxigénio sendo que no total existem quatro locais de ligação na molécula da hemoglobina (1).

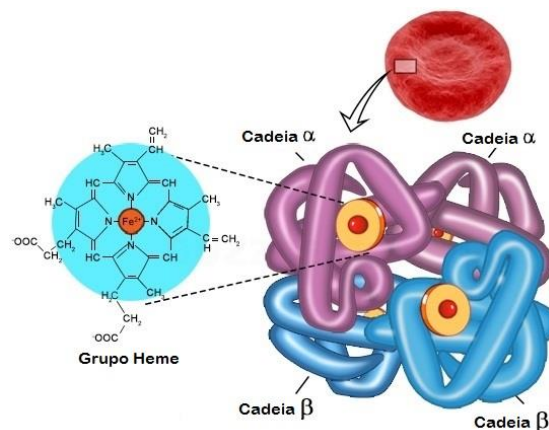


Figura 1.1. Estrutura da hemoglobina. A molécula de hemoglobina é composta por quatro cadeias polipeptídicas (α_1 , α_2 , β_1 e β_2), cada uma delas ligada a um grupo heme composto por um anel de porfirina ligado a um átomo de ferro, onde se liga o oxigénio. São estes complexos ferro-porfirina que conferem à hemoglobina a sua capacidade de transporte de oxigénio no sangue. Adaptado de (<https://www.buzzle.com/images/diagrams/hemoglobin-structure.jpg>).

1.2. Organização e estrutura dos genes globínicos humanos

A síntese da hemoglobina é controlada por dois agrupamentos multigénicos localizados no cromossoma 16 (16p13.3) e no cromossoma 11 (11p15.15), cujos genes estão organizados na direção 5'→3', de acordo com a sua ativação e expressão durante o desenvolvimento humano.

O agrupamento génico no cromossoma 16 que codifica para a cadeia da α -globina localiza-se a 150 quilobases (kb) do telómero e estende-se num segmento de ácido desoxirribonucleico (DNA) de cerca de 28 kb, contendo na seguinte ordem o gene ξ_2 (HBZ), os três pseudogenes $\psi\zeta_1$ (HBZps), $\psi\alpha_2$ (HBA2ps), $\psi\alpha_1$ (HBA1ps), os genes α_2 e α_1 (HBA2 e HBA1) e o pseudogene θ (3). O agrupamento de genes da α -globina é regulado por quatro sequências não codificantes altamente conservadas, denominadas *multispecies conserved sequences* (MCS; MCS-R1 a MCS-R4), localizadas 25-65 kb a montante dos genes da α -globina, que correspondem a locais hipersensíveis à DNase I (*DNase I hypersensitive site* – Hs): HS-48, HS-40, HS-33 e HS-10. O MCS-R2, ou HS-40, demonstrou ser o elemento genómico regulador distal mais importante para a expressão da α -globina pois é o único elemento capaz de conduzir ao elevado nível de expressão da α -globina (4).

O locus da β -globina humana, localizado no cromossoma 11, é composto pelos genes funcionais ϵ (*HBE*), $G\gamma$ (*HBG2*) e $A\gamma$ (*HBG1*), pelo pseudogene $\psi\beta$ (*HBBps*), e por fim pelos genes δ (*HBD*) e β (*HBB*) (5, 6). A montante do agrupamento génico da β -globina está a *locus control region* (LCR), que consiste em cinco locais hipersensíveis à DNase I (HS-1 a HS-5) localizados 6-20 kb a montante dos genes. O HS-3 é o elemento mais importante na abertura da estrutura da cromatina permitindo o acesso dos fatores de transcrição e o HS-2 é essencial para potenciar a síntese das cadeias β -globínicas (3, 7).

A estrutura dos genes globínicos manteve-se altamente conservada ao longo da evolução. Cada gene contém na direção 5'→3' um promotor, o local cap, a 5' *untranslated region* (5' UTR), o codão de iniciação, três exões separados por dois intrões (*intervening sequence* – IVS), o codão de terminação, a 3' *untranslated region* (3' UTR) e a cauda poli A (AATAAAA). O segundo exão codifica resíduos que estão envolvidos na ligação ao grupo heme e na formação do dímero $\alpha\beta$, enquanto o primeiro e o terceiro exões codificam para as outras regiões da cadeia da β -globina não ligadas ao heme (Figura 1.2) (3, 7, 8).

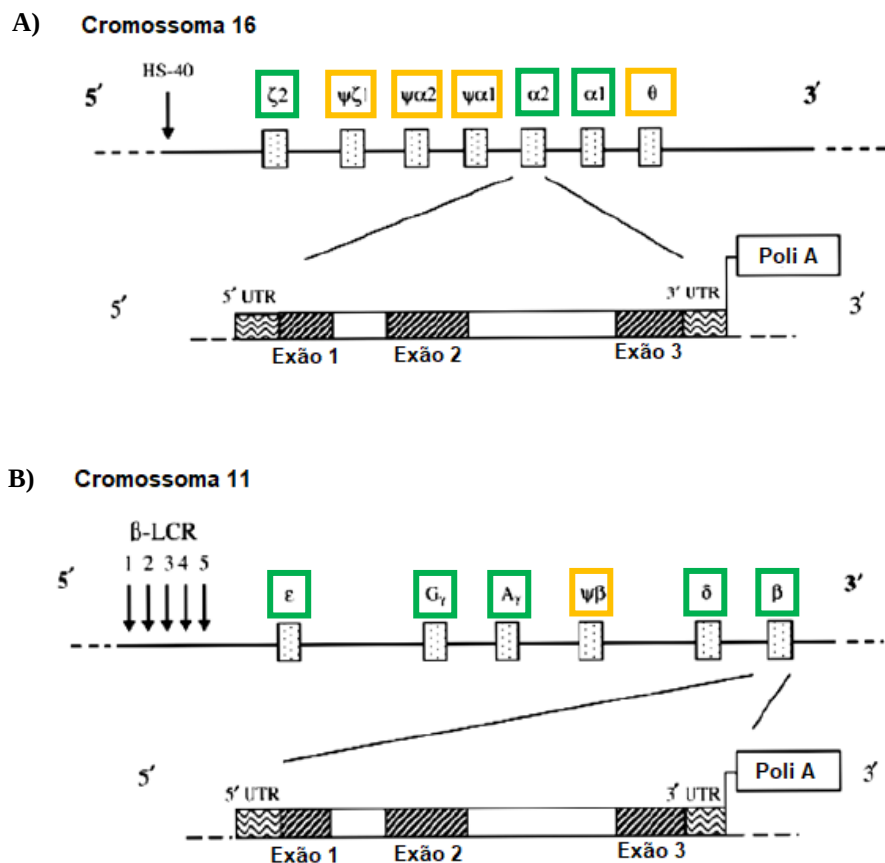


Figura 1.2. Agrupamento génico da α -globina e da β -globina humana. **A)** Representação do *locus* da α -globina humana no cromossoma 16. Dos sete genes presentes, apenas três são funcionais (caixas verdes), nomeadamente o gene $\zeta 2$ (expresso no período embrionário) e os genes $\alpha 1$ e $\alpha 2$ (expressos no período fetal e adulto). Os outros elementos são pseudogenes (caixas amarelas), não expressos - $\psi\zeta 1$, $\psi\alpha 2$, $\psi\alpha 1$ e θ . A seta vertical indica a localização do local hipersensível à DNase I (HS-40), importante para a expressão dos genes. A estrutura do gene $\alpha 2$ é mostrada abaixo, indicando os três exões (caixas rayadas), os dois intrões (caixas abertas) e as regiões UTR (caixas onduladas) comuns a todos os genes da α -globina. **B)** Representação do *locus* da β -globina humana no cromossoma 11 com os cinco genes funcionais ϵ , $G\gamma$ e $A\gamma$, δ e β (caixas verdes) e o pseudogene $\psi\beta$ não expresso (caixa amarela). A montante encontra-se a região de controlo do *locus* da β -globina (β -LCR) composta pelos cinco locais hipersensíveis à DNase I, indicados por setas verticais. A estrutura do gene da β -globina é mostrada abaixo, sendo semelhante à estrutura do gene $\alpha 2$ e à dos outros genes globínicos. Adaptado de (8).

1.3. As diferentes hemoglobinas ao longo do desenvolvimento

Ao longo do desenvolvimento humano é necessária uma expressão coordenada das cadeias de α -globina e β -globina, para originar a concentração certa e equilibrada de hemoglobina nos glóbulos vermelhos (9). A expressão destes genes é regulada pelo estadiu do desenvolvimento, sendo produzidas diferentes moléculas de globina ao longo desse período, as quais dão origem a diferentes hemoglobinas (Figura 1.3).

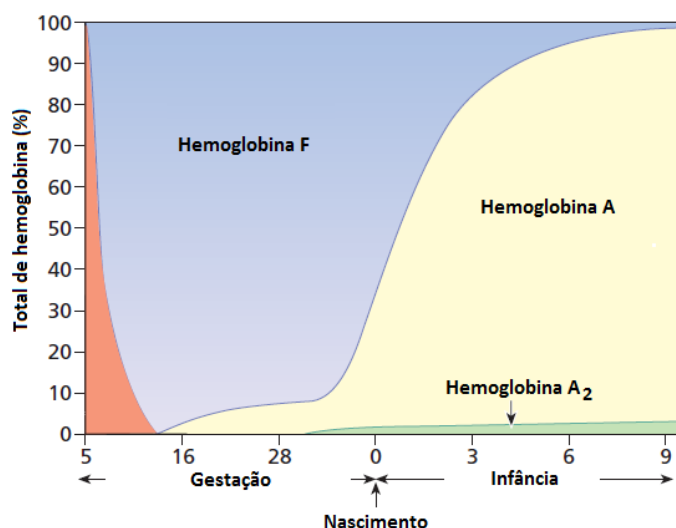


Figura 1.3. Representação esquemática da percentagem das várias hemoglobinas durante o desenvolvimento. Na fase fetal há predomínio de hemoglobina fetal (HbF), marcado a azul. Em indivíduos saudáveis, após o nascimento há um aumento de hemoglobina A (HbA), representado a amarelo, que se torna na principal hemoglobina. Nesta fase do desenvolvimento existe também HbA₂ (a verde) e em muito menor quantidade a HbF, restrita a uma pequena porção de glóbulos vermelhos. Adaptado de (3).

Durante o primeiro mês de gestação são formadas as hemoglobinas embrionárias *Gower 1* ($\zeta 2\epsilon 2$), *Gower 2* ($\alpha 2\epsilon 2$) e hemoglobina de Portland ($\zeta 2\gamma 2$) no saco vitelino. Durante a fase fetal, o local da eritropoiese vai-se alterando gradualmente do fígado, para o baço e de seguida para a medula óssea, com a constante produção da hemoglobina fetal (HbF - $\alpha 2\gamma 2$), a hemoglobina que predomina durante a gestação. Antes do nascimento, começa a ocorrer a transição da expressão da γ -globina para a β -globina, e consequentemente há um decréscimo gradual da expressão de HbF que representa <5% da hemoglobina total, e a sua expressão vai diminuindo até atingir o seu nível de expressão adulto de <1%, a cerca dos 2 anos. Nesta fase, a hemoglobina adulta (HbA - $\alpha 2\beta 2$) torna-se a hemoglobina predominante nos glóbulos vermelhos (cerca de 95% da hemoglobina total) e as mutações que eventualmente existam no gene *HBB* começam-se a manifestar clinicamente. No indivíduo adulto saudável existe ainda uma pequena quantidade de hemoglobina A₂ (HbA₂ - $\alpha 2\delta 2$) bem como de HbF, pois ainda persiste a expressão do gene da γ -globina, embora numa quantidade reduzida (Figura 1.4) (7, 10, 11).

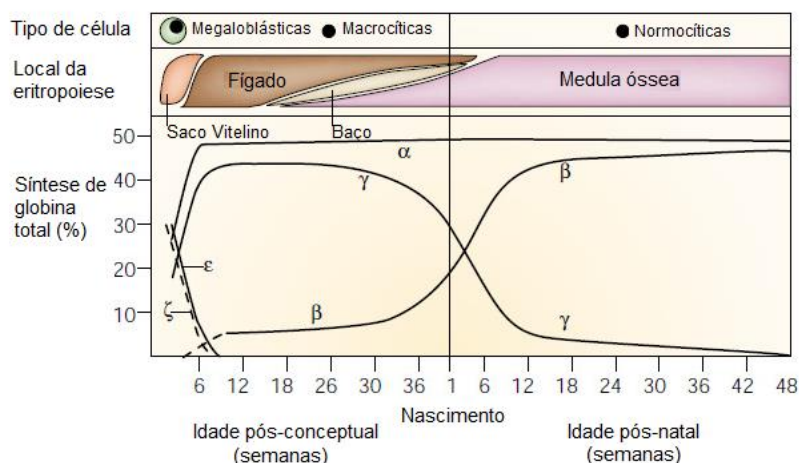


Figura 1.4. Tipos de células precursoras eritroides, locais da eritropoiese e níveis de expressão das várias cadeias globínicas nos diferentes estádios de desenvolvimento. As cadeias de α -globina são expressas ao longo da gestação e na vida adulta, as cadeias γ -globínicas são expressas predominantemente *in útero* e as cadeias β e δ -globínicas são expressas principalmente na fase pós-natal. Esta alteração da expressão gênica é responsável pelas diferentes hemoglobinas presentes no período embrionário, fetal e adulto e também explica a razão pela qual as alterações nos genes da α -globina podem afetar o feto e o adulto, enquanto as alterações na cadeia da β globina são clinicamente manifestadas na fase adulta. Adaptado de (11).

1.4. Patologias relacionadas com défice de hemoglobina – a Anemia

A anemia é um problema de saúde pública, pois é uma das principais causas de morbilidade e mortalidade no mundo inteiro, afetando cerca de um terço da população mundial (12).

A anemia é uma condição na qual existe uma quantidade reduzida de glóbulos vermelhos em circulação ou a sua capacidade de transportar oxigénio é insuficiente para atender às necessidades fisiológicas, que variam com a idade, sexo, altura bem como outros fatores. A Organização Mundial de Saúde (OMS) define anemia quando a concentração de hemoglobina é inferior a 13 g/dL nos homens, 12 g/dL na mulher não grávida e 11 g/dL na mulher grávida (13). As causas da anemia podem ser diversas, embora a mais comum a nível mundial seja a deficiência nutricional em ferro (12). No entanto, a anemia também se pode verificar quando ocorrem alterações hereditárias a nível molecular na hemoglobina, que resultam de mutações nos genes responsáveis pela síntese das cadeias de globina, denominando-se hemoglobinopatias (14). A anemia grave tem como sintomas a fadiga, fraqueza, tontura e sonolência, sendo que as mulheres grávidas e as crianças são as mais vulneráveis (12).

1.4.1. Anemias hereditárias por alterações nas cadeias globínicas – as Hemoglobinopatias

As hemoglobinopatias são um grupo de doenças hereditárias de transmissão autossómica recessiva, que advêm de mutações que afetam os genes responsáveis pela síntese das cadeias globínicas ou mutações que afetem as suas regiões regulatórias (14).

As hemoglobinopatias podem ser classificadas em dois grupos principais: hemoglobinopatias qualitativas ou quantitativas. As hemoglobinopatias qualitativas caracterizam-se pela síntese de hemoglobinas com estruturas anormais devido, geralmente, à substituição de um aminoácido (por exemplo Hb S, Hb C, Hb D, Hb E) e incluem a anemia das células falciformes (drepanocitose). Como consequência são originados glóbulos vermelhos deformados, com capacidade de transporte de oxigénio alterada, instáveis, ou ainda com ganho de outras propriedades bioquímicas anómalas. Por sua vez, nas

hemoglobinopatias quantitativas, ocorre a diminuição ou ausência da síntese de uma ou mais cadeias globínicas (geralmente da β -globina ou da α -globina) e incluem as talassémias (12, 15).

As hemoglobinopatias são as doenças genéticas hereditárias mais comuns no Homem. A frequência do estado de portador foi estimada em 270 milhões de indivíduos, com cerca de 300 a 500 mil nascimentos anuais de crianças com hemoglobinopatias graves. A prevalência das hemoglobinopatias a nível mundial tem vindo a aumentar, o que tem especial importância nos países em desenvolvimento, onde o acesso a serviços de cuidados de saúde é dificultado (2).

1.4.1.1. Variantes estruturais da hemoglobina

Tal como o nome indica, as variantes estruturais de hemoglobina são variantes que alteram a estrutura e as propriedades bioquímicas da molécula da hemoglobina, levando a efeitos fisiológicos que variam na sua gravidade. Esta classe de hemoglobinopatias deriva principalmente de mutações que originam a substituição de um único aminoácido na proteína, resultando num tetrâmero de hemoglobina anormal – daí a designação de "variante". Até à data, foram identificadas mais de 1000 variantes da hemoglobina humana e, embora a sua maioria seja clinicamente silenciosa, algumas conduzem a manifestações clínicas graves, sobretudo quando em homozigotia ou heterozigotia composta. As variantes estruturais de hemoglobina clinicamente significativas podem ser divididas em quatro grupos:

- Variantes com tendência para agregar e sofrer cristalização – este grupo inclui a Hb S e Hb C;
- Variantes com síntese anormal de hemoglobina – inclui, por exemplo, a Hb E;
- Variantes instáveis da hemoglobina – este grupo inclui, por exemplo, a Hb Köln;
- Variantes com transporte anormal de oxigénio – inclui, por exemplo, a Hb Johnstown.

As variantes de hemoglobina mais comuns a nível mundial são a Hb S e Hb C e por isso serão as que vamos dar mais ênfase (16, 17).

➤ Hb S e doença falciforme

A doença falciforme consiste num grupo de doenças resultantes da herança da hemoglobina S (Hb S). A doença falciforme é muitas vezes erroneamente designada por “anemia falciforme”. A anemia falciforme refere-se especificamente à homozigotia para o alelo S no gene da β -globina (β^S) e é a forma mais prevalente de doença falciforme. Contudo existem outros tipos de doença falciforme que incluem hemoglobinopatias nas quais a mutação falciforme no gene da β -globina é herdada em combinação com outra mutação do gene da β -globina (como a HbS/ β -talassémia).

Como referido anteriormente, o grupo heme da hemoglobina contém o ião ferro que é essencial para o transporte de oxigénio, ao passo que a fração globínica protege o grupo heme da oxidação, torna a molécula solúvel e permite a variação na afinidade do oxigénio. Assim, a estrutura da molécula de hemoglobina produz um ambiente interno de radicais hidrofóbicos que protege o ferro da água e, portanto, da oxidação. Os radicais externos são hidrofílicos, tornando a molécula de hemoglobina solúvel, uma condição necessária para que os glóbulos vermelhos consigam acondicionar uma elevada concentração de hemoglobina.

Quando o oxigénio se liga ao ião ferro contido no grupo heme para ser transportado, forma-se o complexo oxi-hemoglobina (ou hemoglobina oxigenada) e os glóbulos vermelhos adquirem uma tonalidade vermelha viva. Quando o oxigénio chega às células do organismo, é libertado do grupo heme e o sangue arterial, com uma tonalidade vermelha vivo, transforma-se em venoso, com uma tonalidade

vermelha arroxeada, característica da desoxi-hemoglobina (ou hemoglobina desoxigenada) (18). Durante esta desoxigenação que segue a passagem dos glóbulos vermelhos na microcirculação, a molécula de hemoglobina sofre uma alteração conformacional. Na Hb S formam-se polímeros alongados que alteram a sua estrutura e a sua flexibilidade, promovendo a desidratação celular. A alteração conformacional que os glóbulos vermelhos sofrem é característica da Hb S, e confere-lhes um formato denominado comumente “em forma de foice” (falciformação). É esta forma que dá o nome à doença (19–23).

Além da falciformação, dá-se um aumento da rigidez, aumento da adesão às células endoteliais e danos na membrana dos glóbulos vermelhos que afetam o seu fluxo através da microvasculatura (Figura 1.5). Consequentemente ocorre vaso-oclusão e, devido ao dano na membrana, o tempo de vida dos glóbulos vermelhos torna-se mais reduzido (15 dias, em vez de 120 dias), resultando numa anemia hemolítica. Devido a estas obliterações vasculares e aos danos em quase todos os órgãos, a doença falciforme é uma das hemoglobinopatias mais nefastas no mundo (2, 17).

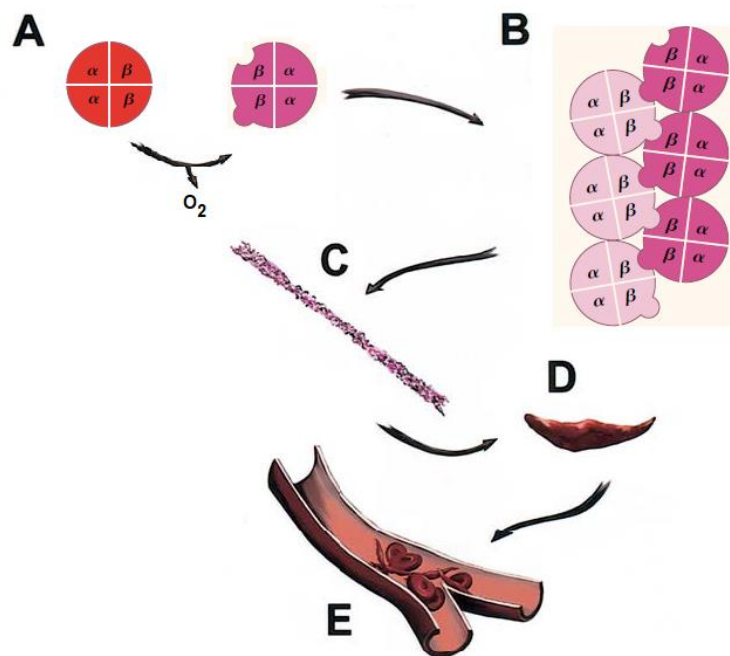


Figura 1.5. Indução do formato de foice nos glóbulos vermelhos pela polimerização da Hb S desoxigenada e fisiopatologia da doença falciforme. A) Nesta imagem é ilustrada uma molécula de hemoglobina S oxigenada, com o tetrâmero $\alpha_2\beta^S_2$ (círculo vermelho). Quando a Hb S é desoxigenada (círculo roxo), ocorre uma mudança conformacional e o local onde se encontra o resíduo de valina hidrofóbica, representado como uma projeção no círculo, pode-se ligar a um local hidrofóbico complementar numa subunidade β de outro tetrâmero de hemoglobina, representado como um recuo no círculo. B) Interação entre três moléculas de desoxi-hemoglobina S numa cadeia (roxo-escuro) e três moléculas de desoxi-hemoglobina S noutra cadeia (roxo-claro), que conduz à formação de um polímero. C) No polímero básico, outras moléculas de desoxi-hemoglobina S podem-se ligar e polimerizar, criando grandes filamentos de polímero. D) Como consequência da polimerização da Hb S, o glóbulo vermelho perde a sua deformabilidade e torna-se rígido, com uma conformação alongada de foice. E) Os glóbulos vermelhos falciformes obstruem pequenos canais vasculares, levando à vaso-oclusão. Adaptado de (22, 23).

O alelo β^S estava originalmente distribuído por toda a África Subsaariana, Médio Oriente, Mediterrâneo e Índia. Esta sua distribuição tão ampla é indicativa da seleção natural dos indivíduos heterozigóticos, devido à proteção relativa que o alelo β^S fornece contra a malária, embora o mecanismo exato dessa proteção ainda não seja totalmente compreendido. O aumento das migrações que se tem vindo a verificar, conduziu também à disseminação do alelo β^S e por isso o número de indivíduos com

doença falciforme é também maior, principalmente em regiões como a América do Norte, Europa Ocidental e Austrália (24).

➤ Hb C e doença de Hb C

A Hb C é a terceira hemoglobinopatia mais prevalente no mundo, com especial incidência nos países africanos e nas zonas endêmicas de malária, tal como acontece com a Hb S (25).

A doença da Hb C, verificada em indivíduos homozigóticos para Hb C, progride de forma semelhante à doença falciforme, mas manifesta-se de um modo menos grave. A Hb C é menos solúvel que a Hb A, levando à formação de cristais no interior dos glóbulos vermelhos. As deformações nas estruturas dos glóbulos vermelhos, derivadas desta formação de cristais, fazem com que as células se tornem rígidas e sejam aprisionadas e destruídas no baço, reduzindo assim o tempo de vida dos glóbulos vermelhos para 30 a 55 dias. Devido a isto, os indivíduos com a doença Hb C podem apresentar sintomas de anemia hemolítica ligeira, em contraste com os indivíduos heterozigóticos, que são geralmente assintomáticos (25).

1.4.1.2. Talassémias

As talassémias são um grupo de condições hereditárias autossómicas recessivas, que afetam aproximadamente 1,7% da população mundial (12). As talassémias consistem em alterações genéticas nas cadeias globínicas que resultam na diminuição da síntese de uma ou mais cadeias e, consequentemente, numa redução da síntese da hemoglobina da qual essas cadeias fazem parte. A condição foi descrita pela primeira vez por Cooley e Lee em 1925 e a designação "talassémia" foi dada por Whipple e Bradford em 1936 (26).

1.4.1.2.1. α -talassémia

A α -talassémia é caracterizada por um défice na produção das cadeias α -globínicas da molécula de hemoglobina. As α -talassémias são causadas geralmente por deleções num ou mais dos quatro genes da α -globina existentes no par de cromossomas 16. Os **indivíduos normais** têm quatro genes α -globínicos e o genótipo normal é designado como $\alpha\alpha/\alpha\alpha$ (27). Os indivíduos que têm uma mutação que afeta apenas um gene α -globínico possuem **α -talassémia silenciosa**, com o genótipo $\alpha\alpha/\alpha-$ (três genes funcionais). Poderão apresentar-se hematologicamente normais ou com ligeira microcitose e/ou hipocromia.

Se dois genes α -globínicos estiverem afetados os indivíduos possuem **traço α -talassémico**, com os genótipos $--/\alpha\alpha$ ou $\alpha-/ \alpha-$ (dois genes funcionais). Os portadores desta condição são clinicamente assintomáticos e o diagnóstico normalmente ocorre por acaso, após análises hematológicas de rotina ou triagem pré-natal, uma vez que apresentam uma anemia ligeira, microcítica e hipocrômica.

Por sua vez, quando só um dos genes α -globínicos está ativo, os indivíduos são portadores da **doença da Hemoglobina H (Hb H)**, com o genótipo $--/\alpha-$ (um gene funcional). Neste último caso, há um grande desequilíbrio entre as cadeias do tipo α -globina e β -globina, pois a expressão do gene da α -globina está reduzida a menos de 30% da quantidade normal. O excesso de cadeias β -globínicas conduz à formação de tetrâmeros β_4 não funcionais, denominados Hb H (6, 28).

Por fim, os indivíduos que possuem os quatro genes afetados e por isso não produzem cadeias α -globínicas têm uma forma muito grave de anemia denominada **Síndrome da Hidrôpsia Fetal por**

Hemoglobina de Bart (*Bart's Hydrops Foetalis Syndrome*), com o genótipo $--/--$. O seu nome deriva do excesso de produção de cadeias γ -globínicas no período fetal, de modo a compensar a ausência de cadeias α -globínicas, com consequente formação de tetrâmeros γ_4 , conhecidos como hemoglobina de Bart (6, 10). Os indivíduos também possuem quantidades variáveis de hemoglobina embrionária de Portland ($\zeta_2\gamma_2$), que é a única hemoglobina funcional e o único transportador de oxigénio que mantém o feto vivo (28). Esta condição é incompatível com a vida humana, pois pode causar a morte ainda no útero ou no período neonatal, contudo, já foram descritos casos cujos fetos que receberam transfusões uterinas ou logo após o nascimento conseguiram sobreviver, embora necessitem de transfusões de sangue para o resto da vida (Figura 1.6 e Tabela 1.1) (6).

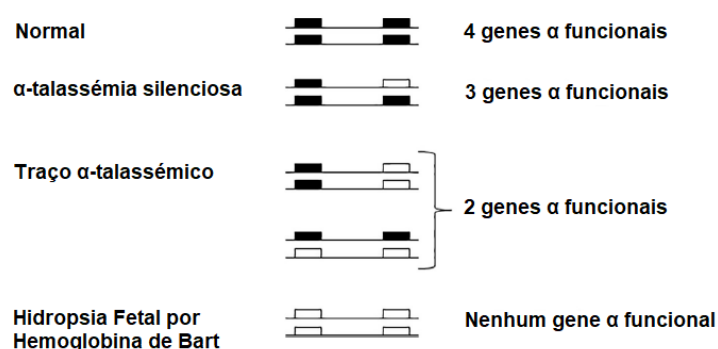


Figura 1.6. Classificação da α-talassemia e representação diagramática das alterações nos genes da α-globina. Cada cromossoma 16 (representado pela linha preta) contém dois genes α (representados pelos retângulos). Os retângulos pretos indicam os genes α funcionais e os retângulos brancos os genes α não funcionais. Adaptado de (6).

Tabela 1.1. Genótipos comuns, classificação e fenótipos da α-talassemia

Genótipos comuns	Número de genes α-globínicos normais	Nome	Fenótipo
$\alpha\alpha/\alpha\alpha$	4	Indivíduo normal	Normal
$\alpha\alpha/\alpha-$	3	α-talassemia silenciosa	Sem alterações hematológicas ou ligeira microcitose e/ou hipocromia
$--/\alpha\alpha$ ou $\alpha-/ \alpha-$	2	Traço α-talassemico	α-talassemia <i>minor</i> : indivíduos com ligeira anemia microcítica e/ou hipocrômica
$--/\alpha-$	1	Doença Hb H	α-talassemia intermédia: indivíduos com anemia microcítica e/ou hipocrômica moderada a grave
$--/--$	0	Hidropsia Fetal por Hemoglobina de Bart	α-talassemia <i>major</i> : hidropsia fetal

Existem outros tipos de α-talassemia mais raras, associadas a deficiências mentais. Uma dessas condições designa-se como síndrome ATR16 (síndrome α-talassemia/atraso mental associada ao cromossoma 16). Esta síndrome rara está associada a deleções muito extensas no cromossoma 16, que removem os genes da α-globina e muitos outros genes dentro e ao redor do agrupamento génico da α-globina. Além da anemia microcítica e hipocrômica acentuada, estes indivíduos apresentam atrasos no desenvolvimento e deficiências mentais ligeiras a moderadas.

Outra síndrome rara associada à α -talassémia é a síndrome ATR-X (síndrome α -talassémia/atraso mental associada ao X), caracterizada por uma deficiência mental mais grave que na síndrome ATRX16. Esta síndrome deriva de mutações pontuais no gene *ATRX*, localizado no cromossoma X e que codifica a proteína ATRX, associada à remodelação da cromatina. A proteína ATRX é considerada um regulador transcricional que afeta a expressão do gene da α -globina.

Além das duas formas raras de α -talassémia acabadas de descrever, existe uma forma adquirida de α -talassémia denominada síndrome de ATMDS (síndrome da α -talassémia mielodisplásica), que envolve mutações adquiridas no gene *ATRX*. Esta síndrome é uma forma adquirida da doença Hb H que ocorre predominantemente em homens idosos com doença hematopoiética clonal pré-maligna, designada síndrome mielodisplásica (6, 28).

As α -talassémias também podem ser diferenciadas entre α^+ ou α^0 . Quando a expressão de um só gene da α -globina está reduzida ou ausente num alelo, estamos perante α^+ -talassémia. A expressão pode estar reduzida devido à deleção de um único gene de α -globina (α^+ -talassémia delecional) ou devido a uma mutação pontual num dos genes duplicados da α -globina (α^+ -talassémia não-delecional). Por outro lado, as α^0 -talassémias caracterizam-se pela ausência total de expressão de ambos os genes da α -globina num alelo, normalmente devido a deleções (aqui inclui-se por exemplo a síndrome da Hidrópsia Fetal por Hemoglobina de Bart).

➤ α^+ -talassémia delecional

A elevada homologia dos genes α -globínicos pode levar a um *crossover* meiótico desigual, que é a base das deleções que causam a α -talassémia. As duas formas mais comuns de α^+ -talassémia delecional são as talassémias $-\alpha^{3.7}$ e $-\alpha^{4.2}$, causadas por recombinações entre os dois genes α -globínicos e que geram respetivamente deleções de 3,7 kb e 4,2 kb. Existem outras α^+ -talassémias delecionais, que normalmente são designadas pelo local da descoberta do primeiro caso, por exemplo, $-\alpha^{\text{Med}}$ e $-\alpha^{\text{SEA}}$, que dizem respeito aos primeiros casos encontrados na região do Mediterrâneo e Sudeste Asiático (6, 10).

➤ α^+ -talassémia não-delecional

As α -talassémias são causadas mais frequentemente por deleções do que mutações pontuais ou pequenas deleções ou inserções em sequências que controlam a expressão génica. No entanto, alguns tipos de α -talassémias não-delecionais podem resultar numa redução mais acentuada da síntese de α -globina do que as deleções, conduzindo a fenótipos mais graves. Estão descritas várias mutações não-delecionais nos genes da α globina que têm impacto na expressão, *splicing*, degradação e estabilidade do ácido ribonucleico mensageiro (mRNA) e que incluem mutações no local de poliadenilação, mutações no codão de terminação ou mutações estruturais que originam variantes de α -globínicas altamente instáveis (28).

1.4.1.2.2. β -talassémia

A β -talassémia é caracterizada pela diminuição da síntese das cadeias de β -globina e pode ser causada por vários tipos de lesões moleculares, tais como mutações pontuais, pequenas deleções limitadas ao gene *HBB*, ou grandes deleções de todo o agrupamento génico da β -globina, dando origem, no estado

de portador, a uma anemia ligeira, microcítica e hipocrômica. Ao contrário da α -talassémia, que deriva sobretudo de deleções, a maioria das mutações β -talassémicas são não-delecionais (5).

Funcionalmente, os alelos da β -talassémia são considerados β^0 quando não há produção de β -globina ou β^+ , quando há produção de alguma β -globina, mas em menor quantidade que o normal. Se a β^0 -talassémia ocorrer no estado de homozigotia ou heterozigotia composta (β^0/β^0), há uma ausência completa de síntese de cadeias β -globínicas e uma falha total na produção de Hb A. Algumas β^+ -talassémias estão associadas a um fenótipo bastante ligeiro, em que há um défice mínimo na produção de cadeias β -globínicas e por isso são designadas como β^{++} -talassémias (5, 26, 29).

As β -talassémias podem também ser classificadas segundo o fenótipo clínico dos indivíduos, que pode ser bastante amplo, indo desde um estado dependente de transfusão até um estado assintomático.

O extremo mais grave do espectro diz respeito à **β -talassémia major**, uma condição que se caracteriza pela ausência completa da produção de β -globina (β^0/β^0). Como na situação normal as cadeias de β -globina substituem as cadeias de γ -globina, produzida durante os primeiros meses de vida, no início da vida as crianças afetadas com talassémia são relativamente normais porque ainda não houve a completa transição da síntese da γ -globina para a β -globina. No entanto, cerca dos seis meses de vida essa transição deveria estar completa, mas neste caso isto não acontece e as formas graves de β -talassémia começam-se a manifestar. Assim, as crianças com esta condição começam a desenvolver uma anemia microcítica grave. A ausência de produção de β -globina leva ao desequilíbrio das cadeias de α/β -globina, havendo um excesso de cadeias α -globínicas livres. Estas cadeias de α -globina livres precipitam nas células precursoras de glóbulos vermelhos, originando corpos de inclusão que por sua vez conduzem a lesões oxidativas da membrana celular, maturação celular anormal e destruição prematura dos glóbulos vermelhos na medula óssea. A consequência deste fenómeno é uma eritropoiese ineficaz. Para manter um nível adequado de hemoglobina e assegurar a sua sobrevivência, os indivíduos são dependentes de transfusões de sangue regulares. No entanto, estas transfusões conduzem a uma sobrecarga em ferro, que se deposita no coração, pâncreas, fígado e outros órgãos danificando os tecidos e levando à insuficiência cardíaca e hepática. Por outro lado, na ausência de transfusões regulares, os indivíduos geralmente morrem antes dos 20 anos. Assim, a terapia ideal consiste nas transfusões regulares de sangue e quelação de ferro, permitindo que muitos portadores desta condição sobrevivam além dos 40 anos sem as várias possíveis complicações que a talassémia *major* pode trazer (Figura 1.7) (26, 29, 30). Este assunto será abordado com mais profundidade no subcapítulo dedicado ao tratamento das hemoglobinopatias.

O outro extremo do espectro das β -talassémias corresponde à **β -talassémia minor** ou traço β -talassémico, em que os indivíduos são heterozigóticos com o genótipo β/β^+ ou β/β^0 . Nesta condição, há uma redução ligeira das cadeias da β -globina dando origem a uma anemia ligeira, microcítica e hipocrômica e os indivíduos são geralmente assintomáticos (26).

A **β -talassémia intermédia** corresponde a uma condição intermédia entre a β -talassémia *major* e a β -talassémia *minor*, ou seja, os indivíduos são sintomáticos, com um quadro clínico ligeiro, mas mais grave do que os estados portadores da β -talassémia *minor*. Os indivíduos também podem precisar de transfusões de sangue ocasionais, mas estas não são essenciais para a sua sobrevivência como ocorre na β -talassémia *major*. As interações génicas que levam ao fenótipo da β -talassémia intermédia são muito heterogéneas: normalmente esta condição ocorre devido à herança de dois alelos β -talassémicos (β^+/ β^+ ou β^+/ β^0), no entanto, a maioria dos casos de β -talassémia intermédia grave pode resultar do aumento relativo da produção de cadeias de α -globina (Tabela 1.2) (7, 29).

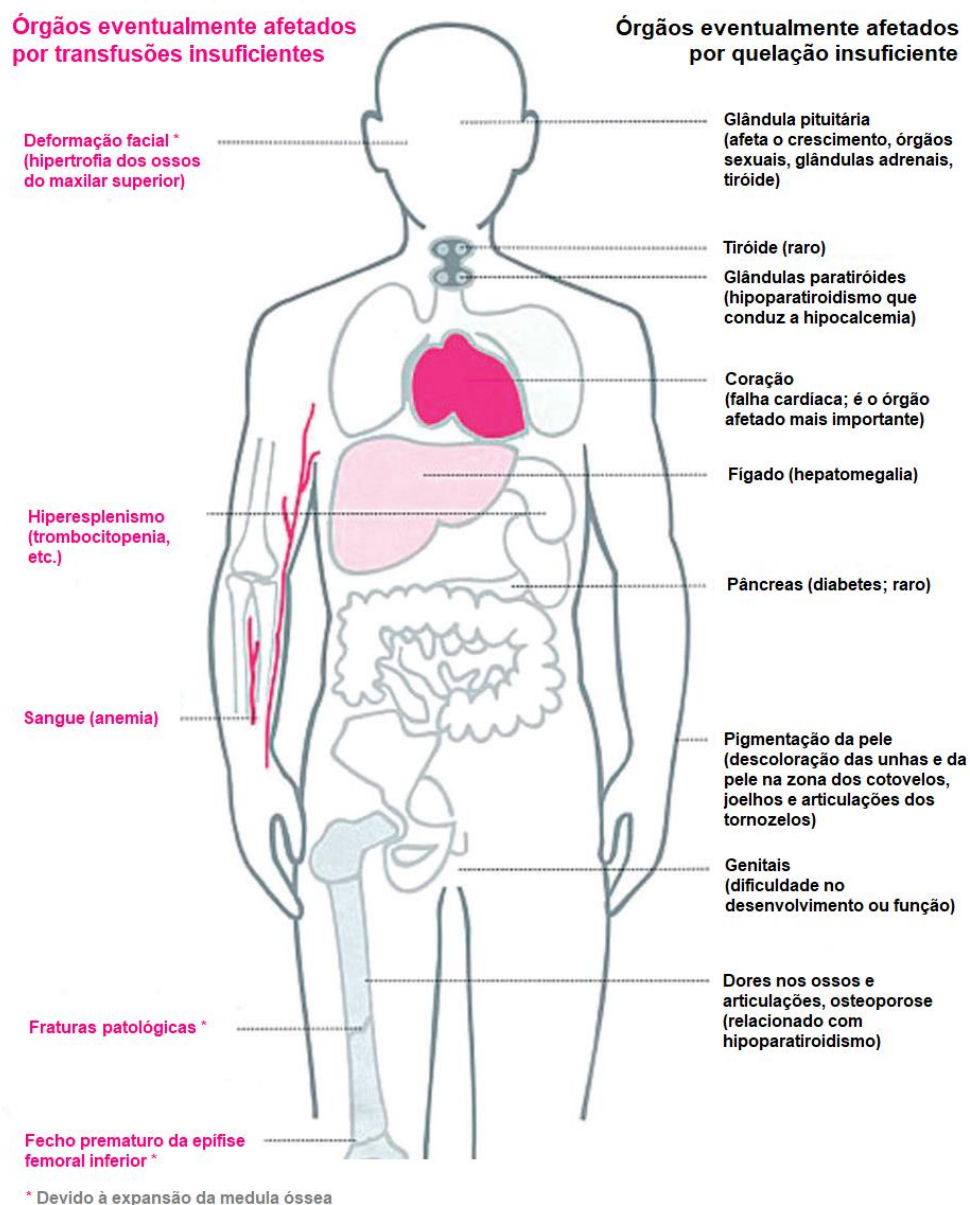


Figura 1.7. Eventuais complicações da talassemia major. Através desta imagem é possível verificar as principais complicações que podem surgir no corpo humano quando a talassemia major não é tratada. Adaptado de (30).

Tabela 1.2. Genótipos comuns, classificação e fenótipos da β -talassemia

Genótipos comuns	Nome	Fenótipo
β/β	Indivíduo normal	Normal
β/β^0 ou β/β^+	Traço β -talassêmico	β -talassemia <i>minor</i> : indivíduos assintomáticos, com microcitose e/ou hipocrômica ligeira e eventualmente anemia
β^0/β^+ ou β^+/ β^+	β -talassemia intermédia	Indivíduos com anemia microcítica e/ou hipocrômica ligeira a moderada, possível hematopoiese extramedular e sobrecarga em ferro
β^0/β^0	β -talassemia <i>major</i>	Indivíduos dependentes de transfusão, com anemia microcítica e/ou hipocrômica grave, hematopoiese extramedular e sobrecarga em ferro

Embora a gravidade clínica da β -talassémia dependa principalmente do grau de deficiência das cadeias de β -globina, o grau de gravidade da condição pode ser atenuado pela co-herança de fatores genéticos que reduzem o excesso de cadeias de α -globina (como a α -talassémia) ou pela co-herança de determinantes genéticos que conduzam à expressão persistente das cadeias γ -globínicas (persistência de Hb F em adultos).

A co-herança da α -talassémia reduz o desequilíbrio das cadeias α/β e deste modo previne o excesso de cadeias α -globínicas que poderiam precipitar nas células e provocar os efeitos deletérios referidos anteriormente. Verifica-se que em várias populações onde a β -talassémia é prevalente, a α -talassémia também ocorre numa elevada frequência e é comum estes indivíduos herdarem ambas as condições.

Em indivíduos β -talassémicos onde há um desequilíbrio das cadeias α/β , pode também ocorrer a co-herança de fatores que conduzam ao aumento das cadeias de γ -globina, que se vão combinar com o excesso de α -globina e formar Hb F. As células que contêm uma quantidade relativamente mais elevada de Hb F conseguem sobreviver, pois estão protegidas do efeito deletério do excesso de α -globina. Assim, todos os indivíduos com β -talassémia apresentam um aumento variável dos níveis de Hb F devido à sobrevivência destas células (5).

1.4.1.3. A prevalência das talassémias a nível mundial e em Portugal

As talassémias são as doenças monogénicas mais difundidas a nível mundial, estimando-se que cerca de 1% a 5% da população mundial seja portadora destas patologias (31). Tanto a α -talassémia como a β -talassémia são predominantes em regiões tropicais e subtropicais onde a malária é, ou foi, endémica. Embora o mecanismo ainda não seja totalmente conhecido, pensa-se que estas patologias tenham sido selecionadas ao longo do tempo por conferirem proteção contra a malária. De facto, as frequências de portadores de β -talassémia naquelas regiões varia entre 2% a 30% e para a α -talassémia a frequência é tão elevada, que se encontra quase fixada na população (de 80% a 90%). Ambas as patologias são bastante frequentes nos países da região do Mediterrâneo, Sudeste Asiático, África Subsariana, Médio Oriente e Índia. Tendo em conta que a sua distribuição se sobrepõe, é bastante comum ocorrer a co-herança de α -talassémia com a β -talassémia, especialmente no Mediterrâneo e no Sudeste Asiático.

Devido à migração populacional, a distribuição das talassémias tem-se vindo a alterar nos últimos 30 anos. Regiões como a Europa e a América do Norte enfrentam agora um grande desafio no diagnóstico e tratamento das hemoglobinopatias, sendo este atualmente um problema global de saúde pública (5, 6, 10).

A epidemiologia das hemoglobinopatias em Portugal só começou a ser sistematicamente estudada a partir de 1986, quando foi instituído o Programa Nacional de Controlo das Hemoglobinopatias. Em 1993, Martins e os seus colaboradores publicaram um estudo de prospeção em Portugal continental, que teve como objetivo estabelecer a prevalência de anemias hereditárias na população portuguesa. Entre as 15208 amostras de sangue analisadas, 0,45% possuíam β -talassémia, apresentado uma distribuição heterogénea no país, que aumenta da região Norte (< 0,1%) em direção à região Sul (> 2% em três distritos) (Figura 1.8). Estas diferenças entre o norte e o sul de Portugal podem estar relacionadas com vários fatores, sobretudo históricos (32).

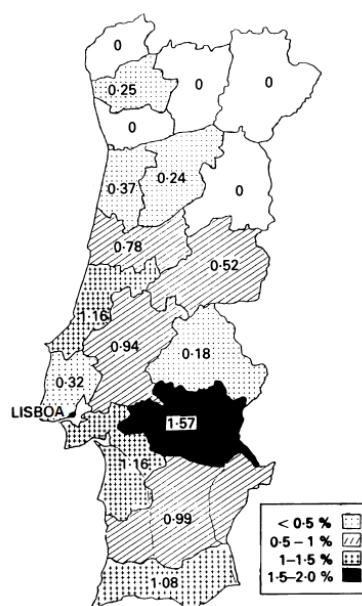


Figura 1.8. Prevalência (%) de portadores de β -talassémia em Portugal continental. É possível perceber pela imagem que distribuição da β -talassémia é heterogênea em Portugal continental, aumentando da região Norte para a região Sul. Adaptado de (32).

Complementarmente, os estudos de genética molecular revelaram que existiam quatro mutações β -talassémicas mais comuns nos países da bacia do Mediterrâneo: Cd39 (C>T), IVS-I-1 (G>A), IVS-I-6 (T>C) e IVS-I-110 (G>A). Estas mesmas quatro mutações eram também as mais frequentemente encontradas em Portugal, possivelmente devido à imigração das populações da área do Mediterrâneo para o sul do nosso país, que oferece condições benéficas para a agricultura, pesca e mineração (32). De acordo com estudos anteriores, na periferia de Lisboa também se encontravam prevalências elevadas de portadores de β -talassémia, resultantes de uma imigração mais recente proveniente dos Países Africanos de Língua Oficial Portuguesa (33).

Relativamente à prevalência da α -talassémia, em 1995 foi publicado um estudo que tinha como objetivo efetuar um rastreio neonatal de α -talassémia numa população residente em Portugal. Os investigadores extraíram o DNA de amostras de sangue recolhidas do cordão umbilical de 100 recém-nascidos e procederam à análise do agrupamento génico da α -globina. Os resultados obtidos demonstraram uma prevalência de 10% de portadores de α -talassémia e ainda um predomínio da deleção de 3,7 kb, sendo que estes resultados se enquadram dentro da prevalência observada na região do Mediterrâneo (34).

1.4.1.4. Diagnóstico

A elevada prevalência das hemoglobinopatias a nível mundial tem levado alguns países a propor programas de rastreio de portadores. Estes programas têm como principal objetivo a identificação de indivíduos portadores de alterações de hemoglobina, de modo a detetar casais em risco de ter um filho com uma hemoglobinopatia grave (como Hidrópsia Fetal por Hemoglobina de Bart ou β -talassémia *major*). Uma vez identificados, os casais são encaminhados para aconselhamento genético de modo a terem um melhor esclarecimento sobre os riscos da doença e para que possam tomar decisões informadas sobre um eventual diagnóstico pré-natal (35, 36). Alternativamente, alguns países ou regiões de elevada prevalência têm implementado o rastreio neonatal para hemoglobinopatias, sobretudo para a

drepanocitose (37). Este permite detetar os doentes logo após o nascimento e oferecer-lhes um regime de terapia preventiva, com vacinação e antibióticos, bem como um acompanhamento médico intensivo, o que permite aliviar-lhes a carga da doença. Em Portugal, encontra-se em estudo, na fase de ensaio piloto, a introdução da pesquisa neonatal de drepanocitose nas patologias pesquisadas no rastreio neonatal de patologias hereditárias conhecido como “teste do pezinho”.

O diagnóstico de hemoglobinopatias, na criança ou no adulto, requer uma avaliação abrangente, combinando a análise hematológica, análise da hemoglobina e análise de DNA (35).

O primeiro passo do diagnóstico consiste na realização de um hemograma completo, onde são discriminados vários parâmetros hematológicos essenciais para o rastreio de hemoglobinopatias, tais como o nível de hemoglobina, a contagem de glóbulos vermelhos e os parâmetros hematimétricos: volume globular médio (VGM), hemoglobina globular média (HGM) e concentração de hemoglobina globular média (CHGM). Cada laboratório deve estabelecer os seus valores de referência para os parâmetros hematológicos, de acordo com a etnia da população, sexo e idade dos indivíduos.

Relativamente ao VGM e HGM, estes parâmetros são essenciais para a identificação de indivíduos portadores de talassémias, pois os seus valores encontram-se reduzidos nestas patologias. Quando os valores de VGM são menores que 80 fL, o valor é indicativo de uma microcitose (glóbulos vermelhos com tamanho inferior ao normal) e quando os valores de HGM são inferiores a 27 pg, indica uma hipocromia (os glóbulos vermelhos possuem uma quantidade de hemoglobina inferior ao normal, o que se reflete numa diminuição da sua coloração) (36).

A redução dos valores de VGM e HGM, além de ser indicadora de talassémias, também podem indicar que um determinado indivíduo possui uma deficiência em ferro. Nestes casos torna-se então necessário avaliar o *status* do ferro, através de testes bioquímicos tais como o doseamento da transferrina (glicoproteína plasmática do sangue que transporta o ferro), ferritina sérica (proteína que acumula o ferro, protegendo a célula dos efeitos tóxicos do metal livre) e protoporfirina de zinco (ZnPP - quando há falta de ferro, é incorporado um ião zinco no grupo heme da hemoglobina, formando ZnPP). A deficiência em ferro é indicada por um nível de transferrina sérica aumentado, nível de ferritina inferior ao normal e o valor de ZnPP superior ao normal (29, 36, 38).

Após este primeiro passo de análise do hemograma, é importante combinar esses dados com a análise da hemoglobina. Os métodos mais usados nessa análise são a cromatografia líquida de alto desempenho (HPLC) e eletroforese capilar.

A HPLC é o método mais utilizado para a deteção e quantificação de Hb A₂, Hb F, e de variantes de hemoglobina. Como o sistema é automatizado, as funções do operador são simples, mas a interpretação dos resultados requer alguma especialização. Embora possua um custo elevado, é uma técnica bastante útil nos programas de rastreio realizados em larga escala. Por seu lado, a eletroforese capilar automatizada é uma técnica de triagem mais recente e complementar à HPLC. Os dados da eletroforese capilar são de mais simples análise, o que torna esta técnica mais adequada para o uso rotineiro. Estas análises ao nível da hemoglobina são especialmente importantes para a determinação quantitativa da Hb A₂, um parâmetro extremamente importante no diagnóstico de portadores de β -talassémia. Assim um quadro de microcitose e hipocromia e valores superiores a 3,5% de Hb A₂ sugere estarmos perante um portador de β -talassémia. Perante um quadro de microcitose e/ou hipocromia e nível de Hb A₂ normal (< 3%) poderemos estar perante um portador de α -talassémia ou um doente com anemia ferropénica.

Como as hemoglobinopatias derivam principalmente de mutações nos genes globínicos, o diagnóstico definitivo deve assentar na análise molecular. Esta análise molecular é especialmente importante no diagnóstico pré-natal, na identificação de casos graves (tais como talassémia *major* e intermédia), em casos de suspeita de portador de β -talassémia com nível *borderline* de Hb A₂ (entre 3,0 e 3,5%), para diagnóstico de α -talassémia e na confirmação da presença de variantes de hemoglobina clinicamente

significativas. A análise molecular é normalmente realizada através de técnicas de análise de DNA baseadas na reação em cadeia da polimerase (*Polymerase Chain Reaction* – PCR), embora alguns laboratórios tenham adotado recentemente técnicas de sequenciação de nova geração (*Next-Generation Sequencing* – NGS). As estratégias utilizadas para o diagnóstico de hemoglobinopatias, que incluem alguns procedimentos acima descritos, estão representadas num fluxograma (Figura 1.9).

O fluxograma está desenhado para detetar os principais tipos de hemoglobinopatias, utilizando cinco parâmetros (percentagem de Hb A₂, percentagem de Hb F, VGM, HGM e variantes de hemoglobina), conduzindo consequentemente a cinco resultados possíveis (indivíduo normal, traço β-talassémico, traço α-talassémico, traço δβ-talassémico ou persistência hereditária de hemoglobina fetal e por último, variante de hemoglobina). É de salientar ainda que os valores de referência indicados no fluxograma são os mais utilizados a nível global, no entanto, como já referido, devem ser definidos de forma independente para cada laboratório (36).

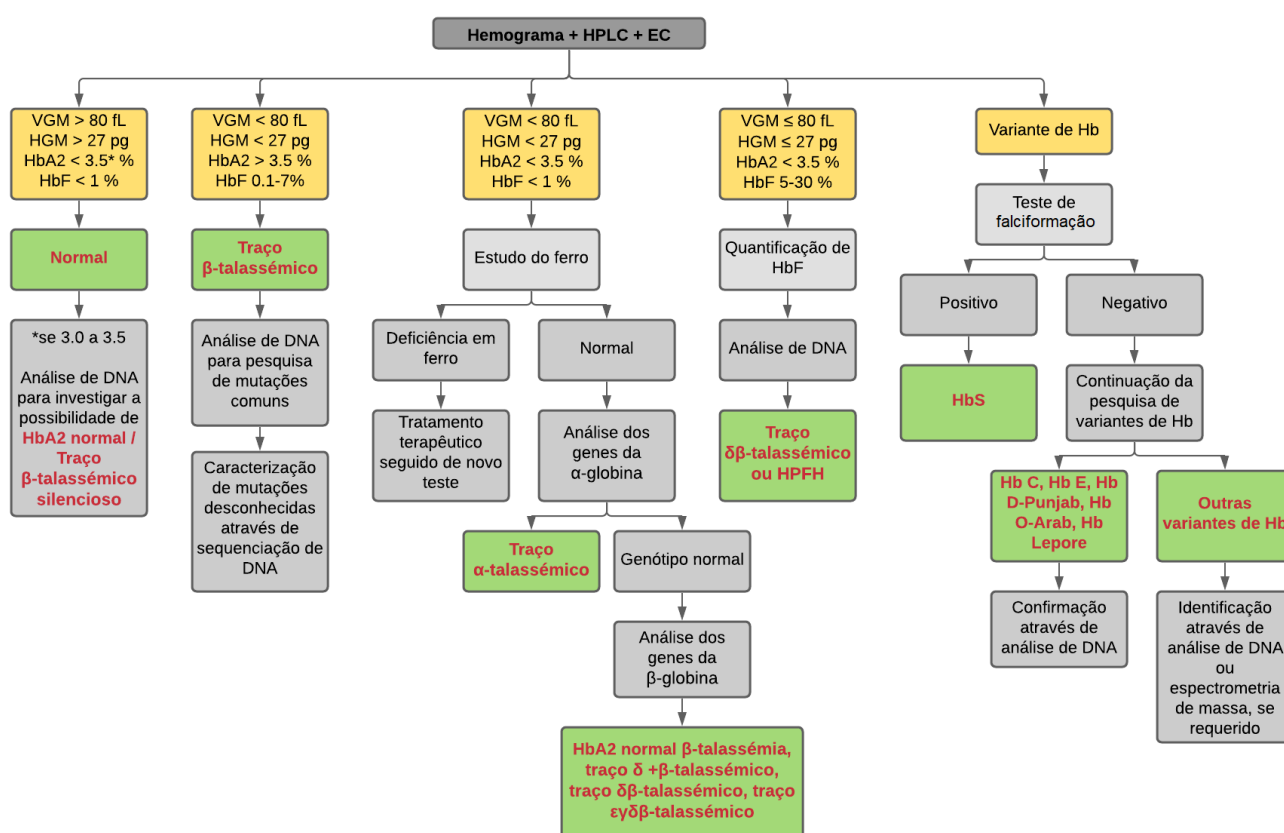


Figura 1.9. Fluxograma para o diagnóstico de portadores de hemoglobinopatias. Neste fluxograma encontram-se representadas as melhores estratégias usadas para diagnosticar os indivíduos portadores de hemoglobinopatias, sendo que as caixas a amarelo contêm os cinco parâmetros a utilizar e as caixas a verde incluem os diagnósticos possíveis. **HPLC** – cromatografia líquida de alto desempenho; **EC** – eletroforese capilar; **VGM** – volume globular médio; **HGM** – hemoglobina globular média; **Hb** – hemoglobina; **HPFH** – persistência hereditária de hemoglobina fetal. Adaptado de (36).

1.4.1.5. Tratamento

As manifestações clínicas das hemoglobinopatias variam bastante, de acordo com a severidade da condição, podendo ir desde hemoglobinopatias ligeiras, que não requerem tratamento médico, até hemoglobinopatias graves que, quando não são tratadas, podem conduzir a uma anemia dependente de transfusões, danos nos órgãos ou mesmo morte. A complexidade da doença implica uma abordagem de

tratamento individualizada para cada paciente. Após o diagnóstico, o paciente deve ser encaminhado a um centro de hematologia para aconselhamento e decisão sobre o tratamento e, se for o caso, para avaliação diagnóstica periódica (17).

1.4.1.5.1. β -talassémia

De uma forma geral, perante a condição mais ligeira - **β -talassémia *minor*** - na maioria das vezes não é necessário recorrer a tratamento médico. Contudo, nalguns casos pode-se observar anemia e por isso deve-se considerar a suplementação com ácido fólico. Nesta condição, os suplementos de ferro são contraindicados, exceto se houver também uma deficiência em ferro.

Nos indivíduos com **β -talassémia *intermédia***, é indicado efetuar terapia de transfusão quando os indivíduos apresentem complicações resultantes de um aumento da eritropoiese ou quando apresentem níveis de hemoglobina abaixo de 8 g/dL. De acordo com severidade da manifestação, deve-se considerar a realização de transfusões intermitentes ou optar por uma terapia de transfusão contínua ao longo da vida, combinada com terapia de quelação do ferro, devido ao alto teor de ferro nas células transfundidas.

A deposição de ferro no fígado, coração e órgãos endócrinos é responsável por uma morbilidade significativa e por isso os níveis de ferro sérico e de outros biomarcadores do *status* do ferro devem ser medidos regularmente. A terapia de quelação de ferro é administrada com o objetivo de reduzir os níveis tóxicos do ferro não ligado à transferrina. Existem três quelantes de ferro principais na prática clínica: desferroxamina (via intravenosa), deferiprone e deferasirox oral (ambos via oral), sendo que o regime de qualquer quelante deve ser ajustado a cada indivíduo. Para obter melhores resultados, a terapia de quelação do ferro deve ser iniciada antes de ocorrerem danos irreversíveis nos órgãos mediados pelo ferro.

Para as formas mais graves - **β -talassémia *major*** - o transplante de células estaminais hematopoiéticas é atualmente o único tratamento curativo. No entanto, o acesso a esta terapia está limitado a sistemas de saúde com infraestruturas mais sofisticadas, à existência de um dador compatível e, além disso, o seu procedimento não está isento de riscos. A mortalidade relacionada com o transplante e os riscos associados ainda representam preocupações consideráveis e por isso é importante ponderar estes riscos *versus* as complicações conhecidas da hemoglobinopatia subjacente. Assim, as principais opções terapêuticas para a maioria dos indivíduos continuam a ser os cuidados de suporte, na forma de transfusões de sangue periódicas para toda a vida combinadas com as terapias de quelação do ferro. Apesar dos avanços científicos das últimas décadas, a disponibilidade e o acesso a transfusões seguras e eficazes continuam a ser um desafio em diversos países (17, 39).

1.4.1.5.2. α -talassémia

Normalmente, os indivíduos com **traço α -talassémico** não precisam de tratamento, embora possam apresentar uma anemia muito ligeira. A causa da anemia deve ser averiguada antes de iniciar a suplementação com ácido fólico e, especialmente, antes de suplementar com ferro, de modo a evitar o risco de desenvolver a sua sobrecarga.

O tratamento da **doença Hb H** depende da gravidade do quadro clínico, sendo que alguns indivíduos podem precisar de terapia de transfusão intermitente ou, em casos de doença mais grave, podem requerer transfusões de sangue regulares, combinadas com terapia de quelação do ferro. Em indivíduos que apresentem esplenomegalia grave, há necessidade de recorrer à esplenectomia, embora seja um método com riscos conhecidos de infeção e sepsia. Por esta razão, as recomendações atuais para esplenectomia

são restritas aos pacientes com incapacidade de receber transfusões e terapias de quelação de ferro. Quando indicado, recomenda-se o tratamento adequado com vacinação, antibióticos e profilaxia para pacientes de alto risco (10, 17).

Relativamente à **Hidrôpsia Fetal por Hemoglobina de Bart**, a maioria das gestações em que o feto é diagnosticado como tendo esta patologia são interrompidas, pois não existem tratamentos eficazes e geralmente é uma condição fatal (existe uma probabilidade muito grande do feto morrer no útero ou logo após o nascimento). Num número reduzido de casos, verificou-se que as transfusões intrauterinas após a deteção pré-natal permitiram a sobrevivência dos bebés, no entanto na maioria os indivíduos apresentavam malformações congénitas urogenitais e nos membros, bem como complicações neurológicas a longo prazo. Mesmo com o tratamento perinatal, esta continua a ser uma condição muito grave e os indivíduos requerem transfusões de sangue ao longo da vida, combinadas com terapias de quelação do ferro. Sempre que possível, é realizado o transplante de células estaminais hematopoéticas. Além do risco fetal, esta patologia está frequentemente associada a complicações sérias na mãe, e por isso, todos estes fatores devem ser tidos em consideração no aconselhamento de casais que se encontrem em risco de ter um filho afetado por Hidrôpsia Fetal por Hemoglobina de Bart (17, 28).

1.4.1.5.3. Doença Falciforme e variantes de hemoglobina instáveis

A estratégia terapêutica da **doença falciforme** baseia-se no suporte profilático através do uso de analgésicos bem como opiáceos no tratamento das crises de dor, vacinação e antibióticos de profilaxia em caso de infeção. A hidroxiureia, através do seu efeito de aumento do nível de Hb F, é o medicamento mais usado para a melhoria da doença. As transfusões de sangue devem ser administradas apenas quando estritamente necessário. Embora estas estratégias possam melhorar os sintomas da doença, o único tratamento curativo é o transplante de células estaminais hematopoéticas, embora não seja indicado para indivíduos mais velhos devido à falta de dadores compatíveis e aos altos riscos que o procedimento implica. A triagem neonatal e o acompanhamento rigoroso, especialmente no início da vida, melhoram significativamente a qualidade de vida dos indivíduos (17, 29).

No que diz respeito às **variantes de hemoglobina instáveis**, o tratamento geralmente consiste no melhoramento da sintomatologia. Raramente se recomendam as transfusões ou esplenectomia, exceto nos casos muito graves. Se a hemólise for crítica, a suplementação com folato pode vir a ser uma opção (29).

1.4.1.5.4. Novas abordagens terapêuticas

Recentemente, tem-se vindo a verificar um aumento do número de estudos relativos às terapias génicas como abordagens alternativas ao transplante de células estaminais hematopoéticas, visto serem potenciais soluções a longo prazo e com menor risco - a terapia génica não requer um dador compatível, visto serem utilizadas as células estaminais hematopoéticas do próprio indivíduo. Os estudos iniciais *in vitro* concentraram-se na adição do gene da β -globina ou γ -globina através do uso de vetores virais, sobreexpressão de ativadores de fatores de transcrição endógenos de γ -globina, e *gene editing* (através da metodologia de CRISPR-Cas9 para o silenciamento de repressores de γ -globina (como BCL11A) ou de mutações β -globínicas. Vários estudos têm como base a γ -globina, pois a reativação da sua expressão pode servir como um substituto da β -globina e pode melhorar o desequilíbrio entre as cadeias α/β -globina (40, 41).

As terapias génicas desenvolvidas por algumas farmacêuticas através de vetores lentivirais estão a mostrar resultados promissores e, em 2019 foi aprovada pela *Food and Drug Administration* (FDA) a primeira terapia génica para o tratamento da β -talassémia, denominada *Zynteglo* e fabricado pela *Bluebird Bio*. No entanto, esta terapia foi aprovada apenas para doentes com β -talassémia a partir dos 12 anos e não abrange casos com ausência total de cadeias β -globínicas. A terapia com *Zynteglo* tem sido utilizada em doentes β -talassémicos que sejam elegíveis para transplante de células estaminais, mas que não possuem um doador compatível. Além disso o custo destas novas terapias, bem como a falta de dados a longo prazo sobre os seus possíveis efeitos secundários são fatores limitantes (41, 42).

Assim, estamos perante o desenvolvimento de uma era de novas terapias, mas são ainda necessários estudos futuros para melhorar os resultados, compreender os efeitos a longo prazo e permitir a sua disponibilidade a nível mundial, visto serem um desafio financeiro, principalmente em países com poucos recursos.

2. Objetivos

Os estudos de prevalência de hemoglobinopatias efetuados em Portugal foram realizados, na sua maioria, nas décadas de 80 e 90 e nenhum deles incluiu a análise das populações residentes nas ilhas portuguesas. Tendo em conta os movimentos populacionais existentes nos últimos tempos, sobretudo provenientes das regiões Africanas e Sul Americanas para a Europa, o padrão de distribuição destas patologias no território português pode, eventualmente, ter-se alterado.

Deste modo, a presente dissertação tem como principal objetivo contribuir para o conhecimento da realidade atual das hemoglobinopatias em Portugal, nomeadamente aquelas associadas a um fenótipo hematológico. Para isso, estabelecemos como **objetivos principais**:

1. Identificar os casos de hemoglobinopatias e determinar a sua base molecular na população avaliada no projeto INSEF (que abrangeu 4808 indivíduos representativos de toda a população residente em Portugal continental e ilhas em 2015) com fenótipo hematológico de hipocromia e/ou microcitose (204 indivíduos).
2. Determinar a frequência das diferentes hemoglobinopatias nos participantes do projeto INSEF e o seu padrão de distribuição geográfica. Determinar a prevalência da β -talassémia na população residente em Portugal.
3. Descrever as características demográficas e de saúde dos indivíduos identificados como portadores de uma hemoglobinopatia, nomeadamente: sexo, nacionalidade, perceção de saúde, existência de acompanhamento por médico de família e autorreporte de anemia.

Este estudo teve ainda como **objetivos secundários**:

- i. Implementar no laboratório uma metodologia de PCR seguido de análise por NGS e validação por sequenciação de Sanger para as hemoglobinopatias associadas ao gene da β -globina. Utilizar as metodologias de Gap-PCR e *Multiplex Ligation-dependent Probe Amplification* (MLPA) para a pesquisa de inserções/deleções no agrupamento génico da α -globina.
- ii. Caracterizar *in silico* o possível efeito patogénico das variantes genéticas detetadas por NGS através do uso de várias ferramentas bioinformáticas.
- iii. Efetuar estudos de associação entre as variantes genéticas detetadas e os vários parâmetros hematológicos correspondentes.

3. Materiais e Métodos

3.1. Amostra Populacional

Este estudo incidiu sobre uma amostra de 4808 indivíduos residentes em Portugal (continental e ilhas) participantes no Inquérito Nacional de Saúde com Exame Físico (INSEF).

O INSEF consistiu num estudo epidemiológico observacional, transversal de base populacional, realizado pelo departamento de epidemiologia do Instituto Nacional de Saúde Doutor Ricardo Jorge (INSA), de modo a ser representativo da população residente em Portugal no ano 2015. O estudo envolveu a avaliação de indivíduos entre os 25 e os 74 anos, residentes há mais de 12 meses em Portugal continental ou nas regiões autónomas, não-institucionalizados e com capacidade para acompanhar a entrevista em língua portuguesa. Os critérios de exclusão para a participação no INSEF foram os seguintes: 1) Ser residente em lares de idosos, hospitais psiquiátricos ou instituições; 2) Residir em unidades habitacionais militares; 3) Ser recluso; 4) Ser incapaz de entender a entrevista em português; 5) Ser incapaz de dar o consentimento informado ou completar o inquérito devido a incapacidade física, doença mental, demência ou outra condição que afete seriamente a capacidade de compreender e dar uma resposta precisa.

Todos os participantes do INSEF e, subsequentemente todos os participantes deste trabalho, assinaram uma declaração de consentimento informado antes do início de qualquer procedimento (Figura 7.1 a 7.3, Anexos).

Todos os procedimentos do INSEF foram conduzidos por profissionais de saúde treinados, nas instalações dos centros de saúde primária. As informações demográficas, socioeconómicas e de saúde dos participantes foram recolhidas através de uma entrevista pessoal assistida por computador.

Para a realização do hemograma completo foram recolhidas amostras de sangue venoso (sem jejum) em tubos EDTA *Vacutainer*®. As amostras de sangue foram pré-processadas nos centros de saúde e armazenadas temporariamente a 4°C. No final de cada dia, as amostras de sangue foram transportadas para o laboratório regional em ambiente refrigerado. O soro e plasma restantes foram distribuídos em alíquotas e armazenados temporariamente a -20°C. Amostras de sangue total foram armazenadas diretamente a -20°C para posterior extração de DNA. Deste modo, a concentração de hemoglobina no sangue foi determinada dentro de 24 horas após a colheita da amostra e o VGM foi medido através de analisadores automáticos presentes nos diferentes laboratórios. Todos os 12 laboratórios envolvidos nas análises hematológicas do INSEF são participantes do Programa Nacional de Avaliação Externa da Qualidade (PNAEQ) e todos os procedimentos foram realizados de acordo com o manual de procedimentos do *European Health Examination Survey* (43).

Uma vez que se pretendiam detetar os portadores de hemoglobinopatia associada a expressão hematológica foi realizada uma análise dos parâmetros hematológicos identificados no hemograma dos participantes e foram selecionados os indivíduos que apresentaram o seguinte fenótipo: anemia (homem, hemoglobina < 13 g/dL; mulher não grávida, hemoglobina < 12 g/dL) e/ou microcitose (VGM <80 fL) e/ou hipocromia (HGM <27pg). Observou-se que de entre os 4808 participantes no projeto INSEF, 204 apresentavam esta condição e estes foram o alvo da nossa investigação molecular. Para tal as amostras de sangue foram anonimizadas e foi-lhes atribuído uma identificação alfanumérica.

Foi construída uma base de dados agregando a informação referente aos dados demográficos (sexo, tipo de zona de residência, região de residência, local de nascimento, acompanhamento por médico de família, última consulta com o médico família, perceção de saúde, gravidez e autorreporte de anemia) e

aos dados hematológicos (níveis de hemoglobina, VGM, HGM, CHGM, hematócrito, *red cell distribution width* - RDW, eritrócitos, leucócitos e plaquetas) dos 204 indivíduos que prosseguiram para a análise molecular.

3.2. Amostra Biológica

A extração do DNA genômico foi realizada pela Dra. Marta Barreto (Departamento de Epidemiologia) na Unidade de Genética Molecular do INSA. Esta extração foi realizada a partir das amostras de sangue periférico dos indivíduos, colhidas em ácido etilenodiamino tetra-acético (*ethylenediaminetetraacetic acid* – EDTA), através do extrator automático *MagNA Pure LC* (Roche®) e permaneceram armazenadas a 4°C.

3.3. Quantificação e verificação da qualidade e pureza do DNA

Para a quantificação e avaliação da pureza das amostras de DNA genômico, utilizou-se o espectrofotômetro *NanoDrop One* (Thermo Fisher Scientific). Determinou-se a concentração de DNA (ng/μL) a 260 nm e, de modo a determinar o grau de pureza do DNA, calculou-se a razão entre as absorvâncias aos comprimentos de onda de 260 nm e 280 nm (A260/280), bem como a 230 nm e 280 nm (A230/280). São considerados “puros” os DNAs que apresentem um valor entre 1,8-2,0 para ambas as razões.

3.4. Reação em Cadeia da Polimerase (*Polymerase Chain Reaction* - PCR)

Em 1984, Kary Mullis desenvolveu uma técnica denominada reação em cadeia da polimerase, mais conhecida como PCR (44). A técnica consiste numa reação que permite a amplificação exponencial de sequências específicas de DNA. Para efetuar esta amplificação é necessário a amostra de DNA molde; uma mistura reacional que contenha a polimerase de DNA, que permite a síntese de novas cadeias de DNA, bem como os oligonucleótidos iniciadores (*primers*) específicos, que consistem em aproximadamente 20 nucleótidos (nt) de cadeia simples, complementares ao fragmento do DNA que pretendemos amplificar; os desoxirribonucleótidos trifosfatados (dNTPs) e uma solução tampão que proporciona um ambiente químico adequado ao funcionamento da polimerase de DNA (45).

A reação é constituída por três etapas principais: 1) Desnaturação – a separação da dupla cadeia em duas cadeias simples por aquecimento, entre 94°C e 98°C, deixando a região alvo disponível; 2) Hibridação – ligação dos *primers* às sequências complementares na cadeia molde de DNA; 3) Extensão – síntese da nova cadeia de DNA, usando uma polimerase de DNA termoestável. Cada uma destas etapas tem uma temperatura específica, sendo que a temperatura de hibridação varia com as características dos *primers*. Num programa de PCR, são realizados entre 25 e 45 ciclos, resultando numa produção exponencial do fragmento desejado (45). Em todos os PCRs realizados foi utilizado um controlo negativo (sem adição de DNA), de forma a controlar possíveis contaminações.

Para cada variante genética em estudo, o DNA genômico foi amplificado num termociclador (*T1 Thermocycler*, *TGradient Thermocycler* e *TProfessional Gradient Thermocycler*; Biometra®) usando *primers* específicos. As condições do PCR estão detalhadas na Tabela 7.2 dos Anexos.

O controlo do PCR foi realizado através da visualização do produto amplificado por eletroforese em gel de agarose. A eletroforese envolve a preparação prévia de um gel cuja concentração de agarose

depende do tamanho dos fragmentos de DNA a serem separados (quanto maior for o tamanho do fragmento, maior terá de ser a porosidade do gel e, consequentemente, menor será a concentração do gel a utilizar). De seguida, cada produto PCR é pipetado num poço do gel, que se encontram numa das extremidades deste. Ao aplicar corrente elétrica ao gel, as moléculas de DNA movem-se através do gel a diferentes velocidades, em função do seu tamanho, o que permite que sejam separadas umas das outras - os fragmentos de menor dimensão migram mais rapidamente do que os fragmentos de maior dimensão, em direção ao eletrodo positivo, visto estarem carregados negativamente (45).

Após a separação dos fragmentos, estes podem ser vistos como “bandas”, cada uma representando um grupo de fragmentos de DNA do mesmo tamanho. O gel contém brometo de etídio (EtBr), que se intercala no DNA e emite fluorescência quando exposto a luz ultravioleta (UV), permitindo deste modo a visualização das bandas.

Ao comparar as bandas do produto PCR com um marcador de peso molecular, que contém vários fragmentos de DNA de tamanhos conhecidos, podemos determinar os tamanhos aproximados dessas mesmas bandas em pares de bases (pb).

O controlo dos resultados obtidos por PCR foi então realizado por eletroforese em gel de agarose. O gel foi preparado com 1% (p/v) de agarose (*Seakem® LE Agarose, Lonza*), em tampão Tris-Borato-EDTA (TBE) 1x, contendo EtBr (*Sigma*). A eletroforese padrão foi realizada numa tina de eletroforese (*PowerPac300* ou *PowerPac Basic, Bio-Rad*) a 70 V por 60 minutos. A revelação do gel foi realizada numa câmara UV (*Uvitec*).

Neste trabalho, o PCR convencional foi realizado para a amplificação do gene *HBB*. Os produtos PCR foram posteriormente estudados por NGS.

3.5. Gap-PCR

O Gap-PCR é uma metodologia variante do PCR e é o método mais comumente utilizado para detetar grandes deleções, através do uso de *primers* específicos que flanqueiam a deleção conhecida (46). Através desta metodologia foi possível detetar deleções no agrupamento génico da α -globina, nomeadamente a deleção $-\alpha^{3,7}$. O Gap-PCR foi realizado segundo as condições descritas na Tabela 7.3 dos Anexos. Nesta variante de PCR, são preparadas duas misturas de PCR (I e II), que diferem apenas no *primer* de sentido reverso.

O diagnóstico da deleção $-\alpha^{3,7}$ é feito através da análise do perfil de bandas. Se o indivíduo apresentar uma banda de 1,9 kb em I e uma banda de 2,1 kb em II, então é homozigótico normal ($\alpha\alpha/\alpha\alpha$); se apresentar uma banda de 1,9 kb em I e simultaneamente uma banda de 1,9 kb e uma de 2,1 kb em II, então é heterozigótico ($\alpha\alpha/-\alpha^{3,7}$); se não apresentar bandas em I e possuir uma banda de 1,9 kb em II, então é homozigótico para a mutação $-\alpha^{3,7}$ ($-\alpha^{3,7}/-\alpha^{3,7}$) (Figura 3.1).

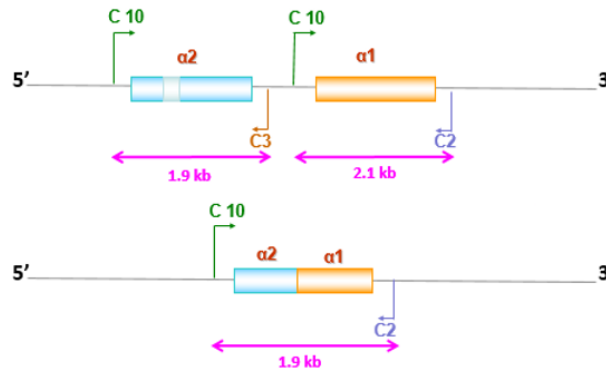


Figura 3.1. Representação esquemática dos genes *HBA1*, *HBA2* e do gene híbrido resultante, caso esteja presente a deleção $\alpha^{3,7}$. Encontra-se representado a azul o gene *HBA2* ($\alpha 2$) e a laranja o gene *HBA1* ($\alpha 1$). O gene híbrido $\alpha 2\alpha 1$ resultante da deleção $\alpha^{3,7}$ está representado pelas duas cores. Encontram-se também esquematizados os locais de ligação dos *primers*, indicados por C10, C2 e C3, bem como o tamanho esperado dos três fragmentos resultantes.

3.6. Amplificação Multiplex de Sondas Dependente de Ligação (*Multiplex Ligation-dependent Probe Amplification - MLPA*)

Embora o Gap-PCR e os outros métodos baseados em PCR tenham simplificado bastante a detecção de deleções α -talassêmicas, estas metodologias não permitem distinguir alelos delecionais desconhecidos. A técnica MLPA, descrita em 2002 por Schouten e os seus colaboradores, permite detetar variações no número de cópias (*copy number variation* – CNV), isto é, segmentos de DNA com mais de 1 kb, que possuem um número de cópias variável em comparação com um genoma de referência, e por isso esta técnica pode ser usada no diagnóstico molecular de várias doenças genéticas cuja patogénese está relacionada com a presença de deleções ou duplicações de genes específicos, como o gene da α -globina (47–49).

O princípio do MLPA baseia-se na amplificação de até 60 sondas (com cerca de 60 nucleótidos), cada uma específica para uma sequência de DNA diferente. Cada sonda de MLPA consiste em 2 oligonucleótidos de DNA de cadeia simples (o oligonucleótido 5' e o 3') que se vão ligar aos respetivos alvos na amostra de DNA. O oligonucleótido 3' contém uma sequência específica (*stuffer sequence*) que serve para definir o tamanho total da sonda de MLPA. Assim, cada sonda vai ter um tamanho único, e por isso os seus produtos de amplificação podem ser separados e identificados posteriormente. Cada sonda possui ainda um par de *primers* que são usados durante o PCR, sendo que este par de *primers* é universal, permitindo a amplificação simultânea e multiplexada por PCR de todas as sondas (Figura 3.2).

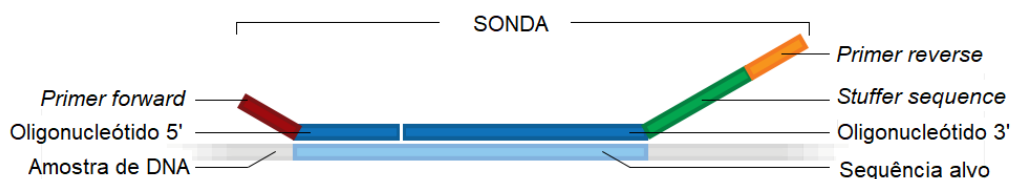


Figura 3.2. Sonda de MLPA. Representação esquemática de uma sonda de MLPA, sendo que cada sonda é constituída por dois oligonucleótidos, uma *stuffer sequence* e uma sequência onde se liga um par de *primers* universal. Adaptado de <https://www.mrcholland.com/technology/mlpa/technique>.

A reação MLPA pode ser dividida em cinco etapas: 1) Desnaturação do DNA e hibridação das sondas, em que o DNA purificado da amostra é desnaturado e incubado durante a noite com as sondas, que vão hibridar com as sequências de DNA alvo; 2) Reação de ligação, na qual ocorre a ligação dos oligonucleótidos 5' e 3' da sonda que hibridou com as sequências alvo; 3) Amplificação por PCR, onde ocorre a amplificação das sondas ligadas por PCR multiplex usando um único par de *primers* universais, marcados com fluorescência, de modo a permitir a visualização dos produtos de amplificação durante a separação de fragmentos; 4) Separação de produtos de amplificação (amplicões com comprimento de 64 a 500 nt) por eletroforese capilar, onde os produtos de PCR são carregados num dispositivo de eletroforese capilar e separados de acordo com o seu tamanho, produzindo um eletroferograma específico da amostra: o padrão de pico da amostra; 5) Análise de dados pelo *software Coffalyser.Net* (MCR-Holland).

O MLPA é uma técnica cujo resultado depende de uma análise comparativa: a altura relativa de cada pico da sonda, em comparação com as alturas relativas do pico da sonda em várias amostras de DNA de referência, reflete o número de cópias da sequência alvo correspondente numa amostra. A inclusão de amostras de referência é, portanto, essencial. A falta de uma ou mais sequências alvo é traduzida como uma diminuição relativa na altura do pico, enquanto um aumento na altura relativa do pico reflete um aumento no número de cópias (50).

Neste estudo, o MLPA foi realizado usando o *kit SALSA® MLPA® probemix P140-C1-0415 HBA* (MCR-Holland), seguindo as instruções do fabricante. Este *kit* foi desenhado para detetar deleções ou duplicações no agrupamento de genes da α -globina humana e das suas regiões regulatórias localizadas no cromossoma 16p13.3, como causa potencial de α -talassémia (51). Numa única reação de MLPA, cada amostra de paciente (≈ 75 ng de DNA) foi testada simultaneamente com um controlo positivo, um controlo negativo e três controlos normais. Os fragmentos amplificados foram separados por eletroforese capilar de acordo com seu tamanho no aparelho *3130xl Genetic Analyzer, ABI PRISM* (Applied Biosystems). Para avaliar a qualidade dos picos, recorreu-se ao *software Microsatellite Analysis v1.0.1* (Applied Biosystems). Os dados quantitativos foram obtidos com o *software Coffalyser.Net* (MCR-Holland) e as áreas dos picos foram utilizadas, após padronização, para determinar o número de cópias das sequências genómicas específicas em cada amostra.

As sequências das sondas, as respetivas localizações e condições de realização do MLPA, encontram-se descritas na Tabela 7.5, Tabela 7.6, e Figura 7.5 dos Anexos.

3.7. Sequenciação de Nova Geração (*Next-Generation Sequencing* – NGS)

A tecnologia de sequenciação de Sanger foi a mais usada pelos laboratórios até à conclusão da primeira sequenciação do genoma humano, em 2004. A partir deste acontecimento, a busca por métodos de sequenciação mais baratos e rápidos aumentou bastante, impulsionando o desenvolvimento de métodos de sequenciação de segunda geração, ou NGS. Estas tecnologias de NGS permitem a sequenciação em paralelo de milhões de fragmentos de DNA de uma única amostra, numa corrida, com um alto rendimento, a um custo muito reduzido e num curto espaço de tempo, tornando-as melhores do que as sequenciações de primeira geração, como a sequenciação de Sanger (52).

A metodologia de NGS inclui quatro etapas essenciais: 1) Preparação de bibliotecas – através da fragmentação aleatória da amostra de DNA, seguida pela ligação de adaptadores a 5' e 3'; 2) Formação de *clusters* – a biblioteca é carregada numa célula de fluxo onde os fragmentos hibridam com os oligonucleótidos que se encontram à superfície da célula e que são complementares à sequência dos adaptadores. Cada fragmento é amplificado por amplificação em ponte; 3) Sequenciação – a tecnologia

de sequenciação da *Illumina* utiliza um método baseado numa terminação reversível que deteta os nucleótidos marcados com fluoróforos à medida que estes vão sendo incorporados na cadeia de DNA que se está a formar; 4) Análise de dados – durante a análise de dados, as sequências novas são alinhadas com o genoma de referência através de *softwares* bioinformáticos. Após o alinhamento, é possível detetar alterações nas sequências recém identificadas, tais como *single nucleotide polymorphisms* (SNPs), pequenas inserções ou deleções (53).

Para cada indivíduo, o gene *HBB* foi previamente amplificado por PCR na sua totalidade (exões, intrões, regiões UTR e regiões promotoras proximais), obtendo-se assim um amplicão de 2106 pb para cada indivíduo (Tabela 3.1).

Tabela 3.1. Amplicões desenhados para o estudo do gene *HBB* por NGS

Gene	Cromossoma	Localização_NC_000011.10	Amplicão_NM_000518.5	Região
<i>HBB</i>	11p15.4	g.5225123 a g.5227229	c.-159 (5'UTR) a c.*474 (3'UTR)	Exão 1 – 3 (2106 pb)

Coordenadas genómicas segundo GRCh38p.12. **NC** – Sequência de referência com base no cromossoma; **NM** – Sequência de referência com base no RNA codificante de proteína (mRNA); **5'UTR** – região a montante não traduzida; **3'UTR** – região a jusante não traduzida.

A purificação dos amplicões foi realizada com o *kit Agencourt AMPure XP PCR Purification* (Beckman Coulter, referência B37419AB de agosto de 2016), de acordo com as instruções do fabricante. Os amplicões foram quantificados utilizando o método fluorométrico, com recurso ao aparelho *Qubit 3.0* (Life Technologies).

3.7.1. Preparação de Bibliotecas para NGS e Sequenciação

Para a preparação das bibliotecas foi utilizado o protocolo *Nextera XT DNA Library Prep Reference Guide* (Illumina, Part#15031942, v05 maio 2019), até ao *pooling* de bibliotecas. No último passo da preparação de bibliotecas, foi incluído um controlo de qualidade interno PhiX (uma sequência conhecida que atua como o controlo de qualidade interno da sequenciação), de acordo com o protocolo *Bead-Based Normalization Method*, do guia *MiSeq System Denature and Dilute Libraries Guide* (Illumina, Doc#15039740, v10, fevereiro de 2019).

As bibliotecas foram sequenciadas no equipamento *MiSeq* (Illumina) usando uma célula de fluxo de 300 ciclos. Durante a sequenciação, o equipamento permite realizar uma análise primária do controlo de qualidade, um parâmetro através do qual se obtém a percentagem de certeza com que a base é atribuída a uma determinada posição. Para o *kit* utilizado, a qualidade mínima definida pelo fabricante é de 30%.

Todo o processo de purificação dos amplicões, preparação das bibliotecas e sequenciação foi realizado no laboratório da Unidade de Tecnologia e Inovação (UTI) do INSA, pela técnica Joana Mendonça sob orientação do Doutor Luís Vieira.

3.7.2. Análise Bioinformática e Determinação de Variantes

A análise dos resultados da sequenciação foi realizada pelo Miguel Machado, bioinformático do laboratório da UTI do INSA, e compreendeu 3 etapas:

- 1) Controlo de qualidade: a qualidade dos dados da sequenciação foi avaliada usando o *software MutiQC* (v1.6.dev0) que fornece uma visão geral da qualidade da sequenciação, ajudando-nos a localizar amostras com falhas ou identificar grupos de amostras com um comportamento irregular (54). O *Illumina InterOp* (v1.0.26-src), foi executado para obter estatísticas detalhadas das corridas de sequenciação, juntamente com a ferramenta *FastQC* (v.0.11.5) (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>) que avalia a qualidade das leituras por amostra;
- 2) Mapeamento das leituras com o genoma de referência: Foi utilizado o genoma de referência humano GRCh38 (versão GRCh38.p12) (55), disponível na base de dados *Ensembl* (<https://www.ensembl.org/index.html>). As leituras limpas (sequências desmultiplexadas e sem sequências de adaptadores) foram mapeadas contra o genoma de referência com o *software Bowtie 2* (v2.3.4, via imagem *Docker* miguelpmachado/bowtie2: 2.3.4.3-02) (56). O alinhamento foi realizado usando o modo *very-sensitive-local*, número máximo de alinhamentos válidos por leitura de 1, um tamanho máximo de inserto para alinhamentos *paired-end* de 1500 e comportamento determinístico. Os ficheiros de mapeamento foram classificados por coordenadas genómicas através do *software SAMtools* (v1.9, via imagem *Docker* miguelpmachado/samtools:1.9) aplicando os módulos *view* e *sort* (57). As etapas seguintes foram executadas usando o ficheiro classificado de cada amostra. As leituras duplicadas foram identificadas usando o *software GATK* (v4.1.5.0, via imagem *Docker* broadinstitute/gatk:4.1.5.0) (58). Foram calculadas diferentes estatísticas usando as ferramentas de estatística e profundidade do *software SAMtools*. Para cada região foi avaliada a profundidade média de cobertura e a percentagem de posições cobertas com o mínimo de 10 leituras;
- 3) Determinação das variantes: Os valores de qualidade das bases das leituras foram corrigidos para erros sistemáticos com o *software GATK* (ferramentas *BaseRecalibrator* e *ApplyRecalibration*) usando variantes de linha germinativa da base de dados *Ensembl Release 96* (depois de remover alelos 0, com códigos de ambiguidade e repetidos para cada variante, bem como entradas sem alelos alternativos). O ficheiro de alinhamento resultante foi utilizado na determinação de variantes. Foram determinadas para cada amostra as variantes, relativamente à sequência referência utilizada, com recurso à ferramenta *GATK HaplotypeCaller* em modo *Genomic Variant Call Format* (GVCF). Foram adicionados os grupos de anotações *StandardAnnotation*, *StandardHCAAnnotation* e *AS_StandardAnnotation*, usando a informação das variantes da linha germinativa da base de dados *Ensembl Release 96*. As amostras foram genotipadas em conjunto com a ferramenta *GATK Genotype GVCFs* para as variantes inicialmente encontradas pelo *HaplotypeCaller*, mas foram consideradas as que possuíam uma confiança de chamada mínima de 20 (numa escala Phred).
Os genótipos das amostras e posições variáveis foram então filtrados usando o *software GATK* (ferramentas *SelectVariants* e *VariantFiltration*) e o *software BCFtools* v1.9 (*annotate*, *filter*, *norm* e *view*) (via imagem do *Docker* miguelpmachado/bcftools: 1.9-01) (59). Em primeiro lugar, as variantes foram filtradas e excluídas quando a profundidade de leitura (DP), também denominada de cobertura foi inferior a 10x. Para aceitar as posições heterozigóticas genotipadas, as leituras de profundidade alélica também foram levadas em consideração para ter em conta o desequilíbrio alélico - foram excluídos genótipos com um equilíbrio alélico inferior a 30% e superior a 70% ($0,3 \leq \text{equilíbrio alélico} \leq 0,7$). De seguida, as variantes foram sequencialmente filtradas se: 1) A variante foi encontrada numa região repetitiva com mais de 8 unidades repetitivas; 2) A variante SNP estava a 2 pb de uma inserção/deleção (indel).

Todas as variantes com previsão de alteração ao nível do *splicing* ou ao nível da estrutura e função da proteína foram confirmadas por sequenciação de Sanger.

3.8. Sequenciação pelo método de Sanger

A sequenciação de DNA é o processo através do qual é possível determinar a sequência de nucleótidos num fragmento de DNA. A sequenciação de Sanger (também conhecido por método de terminação de cadeia), desenvolvida em 1977 por Edward Sanger e pelos seus colaboradores, foi considerado o padrão-ouro para a sequenciação (60). Embora seja um método de sequenciação de primeira geração, ainda é amplamente utilizado para a sequenciação de fragmentos usados na clonagem de DNA ou gerados por PCR.

Nesta técnica, ocorre a síntese de cadeias a partir do fragmento de DNA de interesse. Ao longo da reação, as cadeias são marcadas com fluoróforos de quatro cores diferentes: ddATP – verde; dTTP – vermelho; ddCTP – azul e ddGTP – preto. Os fragmentos de DNA obtidos diferem entre si em apenas um nucleótido, sendo posteriormente separados de acordo com o seu tamanho. O sinal fluorescente que cada fragmento emite quando é iluminado através de um laser permite identificar os fluoróforos através da diferença do comprimento de onda, o que permite estabelecer a sequência de nucleótidos do fragmento de DNA original (61).

Neste estudo, para preparar as amostras para a reação de sequenciação, procedeu-se à purificação dos produtos amplificados por *ExoSAP-IT™ PCR Product Cleanup* (Applied Biosystems), de acordo com as indicações do fabricante (Tabela 7.7, Anexos). As reações de sequenciação foram realizadas usando o kit comercial *ABI Prism BigDye ® Terminator v 1.1 Cycle Sequencing* (Applied Biosystems) e a sequência de nucleótidos do fragmento de DNA original foi determinada por eletroforese capilar no equipamento *3500 Genetic Analyzer* (Applied Biosystems).

Os resultados foram analisados utilizando o software *FinchTV v1.4.0*. (Geospiza) e comparados com as sequências da base de dados *RefSeq* (<https://www.ncbi.nlm.nih.gov/refseq/>) dos genes estudados (Tabela 3.2), tendo em conta a nomenclatura da União Internacional de Química Pura e Aplicada (IUPAC) (Tabelas 7.8 e Tabela 7.9, Anexos) e o código genético (Figura 7.6, Anexos).

A sequenciação automática pelo método de Sanger foi usada para confirmar as variantes patogénicas encontradas no gene *HBB* através de NGS e os resultados foram comparados com a informação presente nas bases de dados *HbVar* (<http://globin.cse.psu.edu/hbvar/menu.html>) e *IthaGenes* (<https://www.ithanet.eu/db/ithagenes>).

Tabela 3.2. Sequências *RefSeq* usadas para a análise comparativa dos resultados obtidos por sequenciação de Sanger

Gene	Cromossoma	NC	NG	NM	NP	Tamanho do transcrito
<i>HBB</i>	11p15.4	NC_000011.10	NG_059281.1	NM_000518.5	NP_000509.1	628 pb

NC – sequência de referência com base no cromossoma; **NG** – sequência de referência com base no gene; **NM** – sequência de referência com base num RNA codificante de proteína (mRNA); **NP** – sequência de referência com base na proteína; **pb** – pares de bases.

3.9. Análise bioinformática das variantes encontradas e previsão dos seus efeitos

Para a análise bioinformática das variantes encontradas utilizaram-se várias bases de dados *online*: *Ensembl* (<https://www.ensembl.org/index.html>); *PubMed* (<https://www.ncbi.nlm.nih.gov/pubmed/>); *ClinVar* (<https://www.ncbi.nlm.nih.gov/clinvar/>); *UniProt* (<https://www.uniprot.org/>); *Genome Browser* (<https://genome-euro.ucsc.edu>) e *Translate Tool* (<https://web.expasy.org/translate/>). Estas

bases de dados permitiram a obtenção da localização genômica dos genes em estudo, informação sobre os seus transcritos, obtenção da sequência de nucleótidos, bem como a sua tradução em sequências de aminoácidos e ainda a previsão *in silico* da importância clínica das variantes encontradas em termos de *splicing* e alterações de proteínas.

Ao realizar essas análises, usamos como referência para o gene *HBB* a sequência *Ensembl* ENSG00000244734, a sequência transcrita ENST00000335295.4 e a sequência de proteína correspondente segundo o identificador *UniProtKB* P68871.

3.9.1. Análise do impacto das variantes ao nível do *splicing*

As regiões codificantes da maioria dos genes eucariotas (exões) encontram-se intercaladas com regiões não codificantes (intrões). Entre a transcrição do mRNA e sua tradução em proteína, ocorre o mecanismo de *splicing*, no qual os intrões são removidos, de forma a produzir um mRNA maduro funcional.

A maioria das variantes genéticas humanas são SNPs, sendo que muitos deles localizam-se em regiões intrônicas, com significado funcional desconhecido. Os SNPs podem afetar o processo de *splicing* e, consequentemente, levar à produção de proteínas truncadas (no caso de *skipping* de exões) ou, pelo contrário, originar proteínas mais extensas (por incorporação no mRNA de intrões ou porções destes). Estes produtos génicos incorretos podem conduzir a graves consequências e a patologia.

De modo a analisar o impacto que as diferentes alterações detetadas no gene *HBB* podem ter ao nível do *splicing*, foi utilizada a ferramenta *varSEAK*® (<https://varseak.bio>). Este *software* tem em consideração as possíveis alterações que podem surgir nos locais de *splicing* canónicos, isto é, no dinucleótido GT para os locais dadores de *splicing* a 5' (5'ss) ou dinucleótido AG nos locais aceitadores de *splicing* a 3' (3'ss).

A previsão geral do impacto que uma determinada variante pode ter no processo de *splicing* é fornecida pela ferramenta *varSEAK* como uma classe de previsão, existindo cinco classes possíveis, cada uma correspondendo a uma cor: **classe 1** (verde-escuro) - sem efeito no *splicing*; **classe 2** (verde-claro) - provavelmente sem efeito no *splicing*; **classe 3** (amarelo) - efeito no *splicing* desconhecido; **classe 4** (laranja) - provável efeito no *splicing*; **classe 5** (vermelho) - efeito no *splicing*.

O *software* gera também um gráfico onde constam características importantes: **sequência referência** - sequência normal, sem a variante; **sequência variante** - sequência com a alteração, os locais de *splicing* autênticos; os novos locais de *splicing* gerados pela variante; **branch point** - uma sequência localizada 18 a 40 nucleótidos a montante da extremidade 3' do intrão e que é caracterizada por conter uma adenina seguida de uma sequência polipirimidínica e de um dinucleótido AG a 3'; **zona de exclusão-AG** (AGEZ) - uma região entre o *branch point* e o 3'ss marcada pela ausência do dinucleótido AG.

Além disso, é fornecida uma tabela com as seguintes informações: **classe de previsão** - como indicado anteriormente pode ir da classe 1 à 5; **score** - probabilidade prevista do local de *splicing* ser funcional (valores positivos) ou não funcional (valores negativos), podendo ir de -100% a +100% e para locais de *splicing* com funcionalidade desconhecida, o *score* é 0%; **Δscore** - a diferença entre o *score* do local de *splicing* na sequência de referência e o *score* do local de *splicing* na sequência variante (62).

3.9.2. Análise do impacto das variantes ao nível da proteína

Como referido anteriormente, a grande maioria das variantes genéticas humanas consistem em SNPs que, quando estão localizados em regiões codificantes podem levar a alterações na sequência de aminoácidos das proteínas. As substituições de aminoácidos (alterações *missense*) estão na génese de várias lesões génicas responsáveis por doenças hereditárias humanas, sendo por isso importante analisar o seu impacto ao nível da estrutura e função das proteínas (63).

➤ **PolyPhen-2**

Para estudar este impacto foi utilizado o *software online PolyPhen-2* (<http://genetics.bwh.harvard.edu/pph2/>). Esta ferramenta prevê o possível impacto das alterações de aminoácidos na estabilidade e função das proteínas humanas, com base na análise da estrutura do aminoácido e na sua conservação entre espécies, permitindo estimar a probabilidade de determinada alteração ser prejudicial.

O *software* usa dois modelos para fazer as previsões: *HumDiv* e *HumVar*. O modelo *HumDiv* consiste numa junção de todos os alelos prejudiciais (presentes na base de dados *UniProtKB*) conhecidos como causadores de doenças mendelianas humanas, juntamente com as diferenças entre as proteínas humanas e as suas homólogas presentes noutros mamíferos relacionados, e consideradas não prejudiciais. O segundo modelo, *HumVar*, agrupa todas as mutações presentes também na base de dados *UniProtKB* e que causam doenças humanas, juntamente com SNPs humanos comuns não prejudiciais. O modelo *HumDiv* é o modelo padrão para avaliar alelos raros e análise de seleção natural, enquanto o modelo *HumVar* é mais adequado para o diagnóstico de doenças mendelianas, que requer a distinção entre mutações com efeitos drásticos das outras variantes humanas existentes, incluindo alelos levemente deletérios e que existem em grande abundância.

Este *software* fornece também um alinhamento múltiplo das sequências de aminoácidos de proteínas homólogas em várias espécies de mamíferos, que é importante de analisar, pois a substituição de aminoácidos pode ser incompatível com o espectro de substituições observado na posição na família de proteínas homólogas. O *PolyPhen-2* identifica estas sequências homólogas através da base de dados *UniRef100* (64).

O *score* indicado pelo programa é o que permite prever o possível impacto de uma substituição na estrutura e função de uma proteína humana e varia de 0,00 a 1,00: de 0,00 a 0,50 as variantes são consideradas benigna e de 0,50 a 1,00 as variantes são consideradas possivelmente prejudiciais.

➤ **SIFT**

Utilizou-se outro *software online*, o *Sorting Intolerant From Tolerant* (SIFT; <https://sift.bii.a-star.edu.sg/>). Esta ferramenta presume que os aminoácidos importantes serão conservados na família de proteínas e, portanto, as alterações de aminoácidos que ocorram em posições bem conservadas serão previsivelmente deletérias para a função da proteína (63). Além de prever o impacto de alterações *missense*, este *software* também pode ser utilizado para prever os efeitos funcionais de inserções/deleções (indels) nas regiões codificantes de um gene que podem causar alterações na grelha de leitura (alterações *frameshift*) ou conduzir à alteração de aminoácidos nas proteínas (65).

Relativamente ao *score*, o *PolyPhen-2* e o *SIFT* usam o mesmo intervalo, de 0,00 a 1,00, mas com a classificação da patogenicidade oposta nos extremos, isto é, no *SIFT* de 0,00 a 0,05 as variantes são consideradas deletérias e de 0,05 a 1,00 as variantes são consideradas benignas (toleradas).

3.10. Análise Estatística

Os dados foram estatisticamente tratados utilizando o *software IBM SPSS Statistics v.26.0*. A significância estatística foi estabelecida para *p-value* <0,05.

Os dados foram testados quando à normalidade da distribuição usando o teste de Shapiro-Wilk e foram usados testes estatísticos não paramétricos quando tal condição não foi satisfeita. As associações do estado portador ou não de hemoglobinopatias com os parâmetros hematológicos foram pesquisadas através do teste paramétrico Teste-t de Amostras Independentes e pelo teste não paramétrico de Mann-Witney.

4. Resultados e Discussão

4.1. Caracterização da população em estudo

Este trabalho trata-se de um sub-estudo transversal de base populacional, baseado no INSEF.

Aos 4911 indivíduos que aceitaram participar no INSEF, foram efetuadas avaliações antropométricas, bioquímicas e hematológicas (incluindo o hemograma) e um questionário autorreporte sobre variáveis demográficas e socioeconómicas, estado de saúde, e sobre os seus comportamentos e utilização de serviços e cuidados de saúde. Foram excluídos alguns indivíduos por motivos de ausência de informação essencial ou de produto biológico, e por isso apenas foram considerados 4808 indivíduos (66).

A partir dos 4808 indivíduos pertencentes à população INSEF, foram selecionados para fazer parte deste sub-estudo os indivíduos que possuíam microcitose (VGM <80 fL) e/ou hipocromia (HGM <27 pg), pois neste perfil hematológico enquadram-se os fenótipos de parte dos portadores de α -talassémia e todos os portadores de β -talassémia. Esta seleção baseada nos parâmetros hematológicos resultou em 208 indivíduos, contudo, devido à falta de amostra biológica, alguns indivíduos não foram passíveis de ser analisados, pelo que a amostra do presente estudo é constituída por 204 indivíduos.

Destes 204 indivíduos, 47 são do sexo masculino e 157 do sexo feminino, com idades compreendidas entre os 25 e os 74 (tal como no INSEF), mediana de 49 anos, uma idade média de 49,3 anos e com desvio-padrão de 11,8 anos (Figura 4.1).

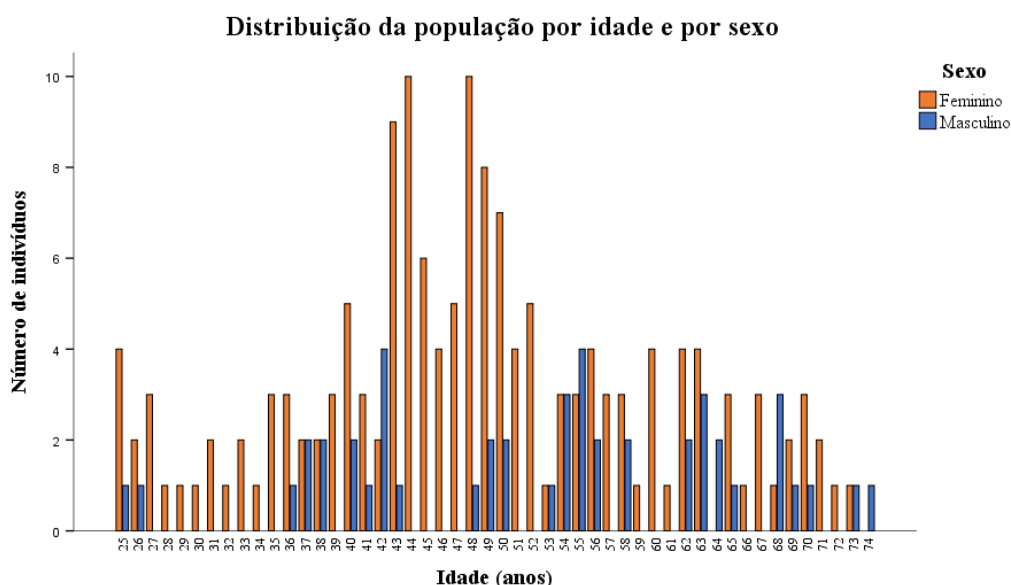


Figura 4.1. Distribuição por idade e por sexo da população selecionada com microcitose e/ou hipocromia. Através deste gráfico é possível verificar que a idade está compreendida entre os 25 e os 74 anos e que o número de indivíduos femininos é superior ao número de indivíduos masculinos em quase todas as idades.

Uma das condições do INSEF era que os indivíduos residissem em Portugal há mais de 12 meses e não, necessariamente, que fossem de origem portuguesa. Assim, embora a maioria dos participantes tenham nascido em Portugal, este estudo abrange indivíduos oriundos de 14 países diferentes, com as seguintes frequências: Portugal (172 indivíduos), Angola (8 indivíduos), Brasil (7 indivíduos), Cabo

Verde (6 indivíduos), França (2 indivíduos), África do Sul (1 indivíduo), Bulgária (1 indivíduo), Canadá (1 indivíduo), China (1 indivíduo), Índia (1 indivíduo), Moçambique (1 indivíduo), Moldávia (1 indivíduo), Paquistão (1 indivíduo) e Síria (1 indivíduo) (Figura 4.2, A).

Relativamente ao local de residência dos participantes, a maioria reside na região do Alentejo e do Algarve, seguindo-se a Região Autónoma dos Açores (RA Açores), Lisboa e Vale do Tejo (LVT), Região Autónoma da Madeira (RA Madeira), Norte e por fim Centro (Figura 4.2, B).

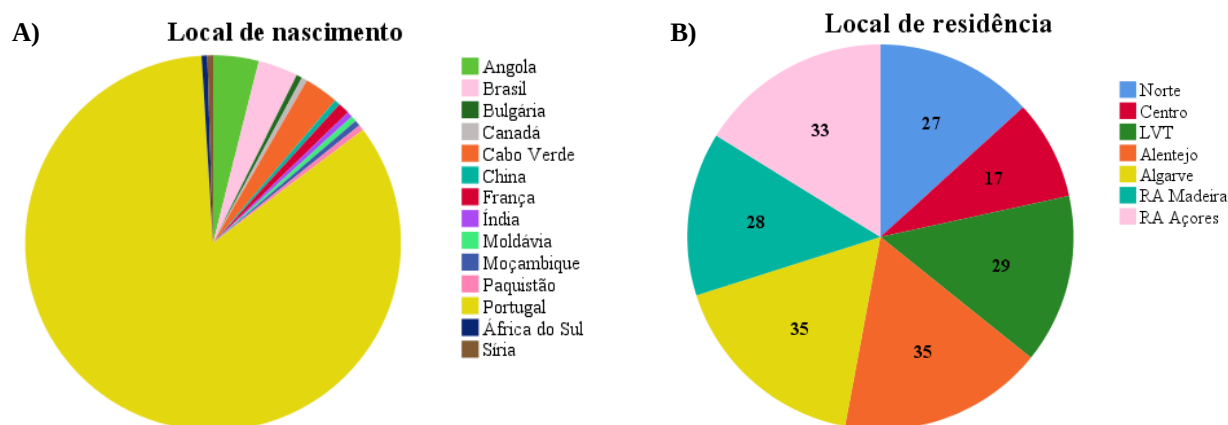


Figura 4.2. Distribuição da população analisada com fenótipo hematológico por local de nascimento e de residência. A) Distribuição da população pelo seu local de nascimento. Através deste gráfico é possível perceber que a maioria dos participantes deste estudo tem origem portuguesa, seguindo-se a origem angolana, brasileira e cabo verdiana. B) Distribuição da população pelo seu local de residência. O gráfico indica o número de indivíduos que residem em cada uma das sete regiões de Portugal. Através deste gráfico é possível perceber que a maioria dos participantes com o referido fenótipo hematológico reside no Alentejo e no Algarve e que a região onde menos indivíduos residem é a região Centro.

4.2. Caracterização dos parâmetros hematológicos

As características hematológicas disponíveis referentes a cada indivíduo neste estudo foram a hemoglobina, VGM, HGM, CHGM, hematócrito (percentagem de volume ocupada pelos glóbulos vermelhos no volume total de sangue), amplitude de distribuição dos glóbulos vermelhos (RDW - é a amplitude de distribuição dos glóbulos vermelhos que avalia a variação de tamanho entre eles), eritrócitos, leucócitos e plaquetas.

Na Tabela 4.1 encontra-se a estatística descritiva com o teste de normalidade Shapiro-Wilk. Os parâmetros que se consideraram como tendo uma distribuição normal ($p\text{-value} \geq 0,05$) foram o hematócrito e os eritrócitos nos homens, e por isso, na avaliação estatística serão usados testes paramétricos. Todos os outros parâmetros (hemoglobina, VGM, HGM, CHGM, RDW, eritrócitos nas mulheres, leucócitos e plaquetas) foram considerados como tendo uma distribuição não normal ($p\text{-value} < 0,05$) e por isso, na avaliação estatística serão usados testes não paramétricos.

Tabela 4.1. Estatística descritiva dos parâmetros hematológicos e teste de normalidade Shapiro-Wilk

Parâmetros	Média	Mediana	Mínimo	Máximo	Desvio padrão	Teste de normalidade Shapiro-Wilk (<i>p-value</i>)
Eritrócitos ($\times 10^{12}/L$)						
Homens	5,67 $\pm$ 0,08	5,71	4,30	6,70	0,57	0,531
Mulheres	4,78 $\pm$ 0,04	4,77	3,75	7,30	0,51	<0,001
Hemoglobina (g/dL)						
Homens	13,9 $\pm$ 0,3	14,2	9,8	16,7	1,8	0,031
Mulheres	11,9 $\pm$ 0,1	12,0	6,5	15,3	1,3	0,002
VGM (fL)	77,7 $\pm$ 0,4	79,5	56,9	86,9	6,4	<0,001
HGM (pg)	24,8 $\pm$ 0,2	25,6	16,2	28,1	2,3	<0,001
CHGM (g/dL)	31,9 $\pm$ 0,1	31,9	28,3	36,1	1,0	<0,001
Hematócrito (%)	38,6 $\pm$ 0,3	38,1	23,0	52,0	4,7	0,060
RDW (%)	15,4 $\pm$ 0,1	15,2	12,1	22,3	1,8	<0,001
Leucócitos ($\times 10^9/L$)	7,44 $\pm$ 0,15	7,17	4,00	16,40	2,13	<0,001
Plaquetas ($\times 10^9/L$)	260 $\pm$ 4	246,5	132	433	63,1	<0,001

VGM – volume globular médio; **HGM** – hemoglobina globular média; **CHGM** – concentração da hemoglobina globular média; **RDW** – *red cell distribution width*. Considera-se a distribuição normal quando *p-value* $\geq$ 0,05 (os valores estão salientados a negrito).

4.3. Diagnóstico de alterações nos genes *HBA*: α -talassémia

O diagnóstico da α -talassémia foi realizado recorrendo a duas técnicas: o Gap-PCR e o MLPA.

Primeiramente, recorrendo ao Gap-PCR, procedeu-se à deteção da deleção de 3,7 kb ($-\alpha^{3,7}$), que é a deleção mais comum no agrupamento génico da α -globina e confirmou-se o perfil eletroforético de cada produto amplificado em gel de agarose (Figura 4.3). Entre os 204 indivíduos analisados foram detetados 52 indivíduos heterozigóticos para a deleção de 3,7 kb e um indivíduo homozigótico (Tabela 4.2).

Para os indivíduos que não apresentaram $-\alpha^{3,7}$, procedeu-se a pesquisa geral de CNVs no agrupamento génico da α -globina através de MLPA. Com este estudo encontramos um indivíduo heterozigótico para a deleção de 4,2 kb ($-\alpha^{4,2}$) e dois indivíduos com genes de α -globina triplicados ($\alpha\alpha\alpha^{\text{anti3,7}}$) (Figura 4.4 e Tabela 4.2). Assim, a α -talassémia foi observada em 54 indivíduos, o que corresponde a uma prevalência de α -talassémia de 26,6% se considerarmos a população com fenótipo de microcitose e/ou hipocromia.

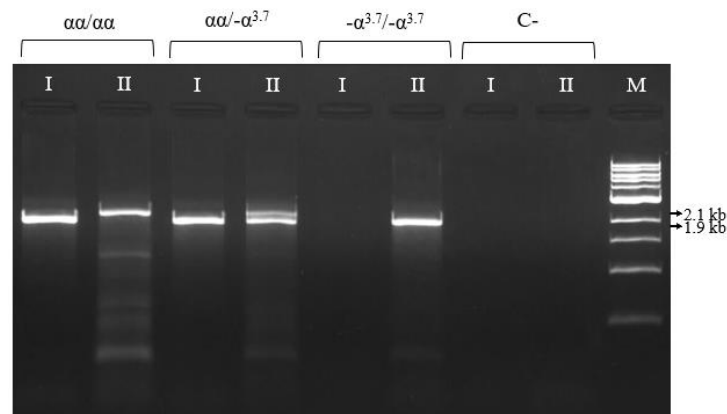


Figura 4.3. Análise do Gap-PCR para o diagnóstico da deleção $-\alpha^{3.7}$. Perfil eletroforético em gel de agarose (1%, p/v) representativo dos três possíveis genótipos que podem surgir aquando da detecção da deleção $-\alpha^{3.7}$. C- – controlo negativo; M – marcador de peso molecular (1 kb DNA Ladder, BioLabs); I – mistura I; II – mistura II. O indivíduo $\alpha\alpha/\alpha\alpha$ é normal, apresentando uma banda de 1,9 kb na mistura I e uma banda de 2,1 kb na mistura II; o indivíduo $\alpha\alpha/-\alpha^{3.7}$ é heterozigótico para a deleção, apresentando três bandas no gel, uma de 1,9 kb na mistura I, uma banda de 1,9 kb na mistura II e outra de 2,1 kb na mistura II; o indivíduo $-\alpha^{3.7}/-\alpha^{3.7}$ é homozigótico para a mutação, apresentando apenas um fragmento de 1,9 kb na mistura II. No controlo negativo não se observam quaisquer bandas quer na mistura I, quer na mistura II. Os tamanhos dos fragmentos encontram-se indicados no lado direito da imagem, em kb.

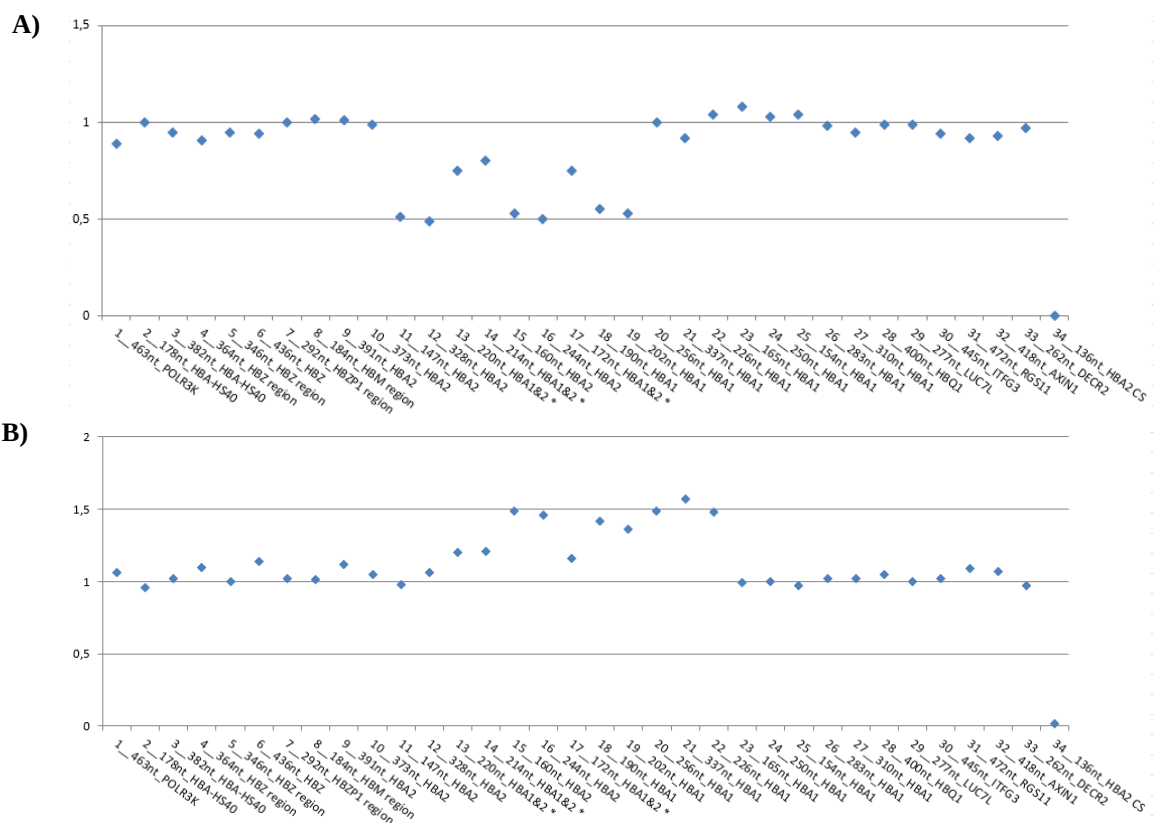


Figura 4.4. Análise esquemática de dois resultados de pesquisa de CNVs no braço curto do cromossoma 16 por MLPA. A) Resultados do MLPA representativos de um indivíduo heterozigótico para a deleção $-\alpha^{4.2}$. B) Resultados do MLPA representativos de um indivíduo com os genes α -globínicos triplicados ($\alpha\alpha\alpha^{\text{anti}3.7}$). Os números das sondas MLPA encontram-se descritos no eixo das abcissas. As proporções das sondas MLPA (eixo das ordenadas) foram determinadas através da comparação da quantificação do seu sinal nos indivíduos em estudo com os controlos normais. Os resultados esperados das proporções da maioria das sondas MLPA são: 1 (sonda normal); 0 (deleção homozigótica); 0,5 (deleção heterozigótica); 1,5 (duplicação heterozigótica) e 2 (triplicação heterozigótica ou duplicação homozigótica). A sonda 34 é amplificada apenas na presença da variante Hb Constant Spring e por isso em indivíduos normais a sua proporção será 0.

Tabela 4.2. Genótipos detetados no gene *HBA* e frequências relativas para os 204 indivíduos analisados

Genótipos	Frequência absoluta	Frequência relativa
$\alpha\alpha/\alpha\alpha$	148	72,6%
$-\alpha^{3,7}/-\alpha^{3,7}$	1	0,5%
$\alpha\alpha/-\alpha^{3,7}$	52	25,5%
$\alpha\alpha/-\alpha^{4,2}$	1	0,5%
$\alpha\alpha/\alpha\alpha^{\text{anti}3,7}$	2	0,9%

Se considerássemos a população de partida (4808 participantes INSEF) este resultado traduzir-se-ia numa prevalência de 1,1%. No entanto, esta extrapolação não deve ser realizada uma vez que a nossa abordagem não considerou os casos de α -talassémia silenciosa, que são hematologicamente indistinguíveis da população normal. Assim, podemos concluir que, na população residente em Portugal, a α -talassémia é uma causa bastante considerável para a existência de um fenótipo hematológico anómalo de microcitose e/ou hipocromia (observada em 26,6% dos casos analisados). Estes resultados vão ao encontro do que é reportado no estudo realizado por Rebordão e Silva numa população portuguesa jovem adulta, clinicamente assintomática. Neste estudo, os autores concluem que a α -talassémia deverá ser o grupo de alterações que apresenta maior frequência, salientando que a α -talassémia é mais frequente do que se pensa e que por isso é uma doença subdiagnosticada. Os autores detetaram na sua população de estudo uma prevalência de 1,25% de portadores de α -talassémia, bastante semelhante à prevalência de 1,1% que também obtivemos se considerarmos o total da nossa amostra de indivíduos (14). Assim, os resultados de ambos os estudos deverão servir para alertar os clínicos para a necessidade de ser pesquisada a α -talassémia como possível causa de um fenótipo de microcitose e/ou hipocromia persistente, acompanhado ou não de anemia. É de salientar que este tipo de fenótipo é muito frequentemente associado pelos clínicos a anemia ferropénica com a consequente administração de medicamentos de reposição de ferro. Contudo, tratando-se de α -talassémia, esse tratamento não irá melhorar a sintomatologia e poderá mesmo ser bastante prejudicial por poder originar patologia de sobrecarga em ferro.

Neste estudo detetámos a presença em heterozigotia de uma triplicação no gene da α -globina em dois casos ($\approx 1\%$ da população analisada). É ainda de salientar que os indivíduos $\alpha\alpha\alpha^{\text{anti}3,7}$ foram considerados como não tendo α -talassémia, pois apresentam um excesso de cadeias de α -globina e não uma redução de cadeias globínicas, que é a característica principal das talassémias.

Os resultados obtidos neste estudo quanto ao tipo de deleções causadoras de α -talassémia presentes na população residente em Portugal vão ao encontro do estudo de 1995, realizado por Peres *et al.* em 100 recém-nascidos consecutivos. Confirmamos assim que a lesão molecular maioritariamente associada a α -talassémia na população residente em Portugal é a deleção de 3,7 kb, seguida da deleção de 4,2 kb (34).

4.3.1. Associação da α -talassémia com os parâmetros hematológicos

Pretendeu-se estudar uma possível associação entre os parâmetros hematológicos disponíveis e a co-herança da α -talassémia. Para isso recorreu-se ao uso do teste estatístico T de duas amostras independentes, para os parâmetros com distribuição normal, para os parâmetros com distribuição não normal, usou-se o teste de Mann-Whitney (Tabela 4.3).

Tabela 4.3. Associação entre os parâmetros hematológicos estudados e a ausência ou presença de α -talassémia

Parâmetros	População com microcitose e/ou hipocromia, mas sem α -talassémia					População com microcitose e/ou hipocromia e com α -talassémia					<i>p-value</i>
	Média	DP	Mediana	Min	Max	Média	DP	Mediana	Min	Max	
Eritrócitos ($\times 10^{12}/L$)											
Homens	5,69	0,630	5,800	4,45	6,70	5,63	0,513	5,705	4,30	6,57	0,719 ^a
Mulheres	4,76	0,547	4,755	3,75	7,30	4,87	0,335	4,825	4,29	5,77	0,074 ^b
Hemoglobina (g/dL)											
Homens	13,3	1,98	13,90	9,8	16,7	14,6	1,28	14,60	11,1	16,5	0,012^b
Mulheres	11,6	1,33	11,75	6,5	15,3	12,7	0,89	12,65	10,9	14,8	<0,001^b
VGM (fL)	76,5	6,77	78,00	56,9	85,9	81,2	3,17	81,50	68,5	86,9	<0,001^b
HGM (pg)	24,4	2,51	25,20	16,2	28,1	25,9	0,99	26,30	22,4	26,9	<0,001^b
CHGM (g/dL)	31,8	1,12	31,90	28,3	36,1	32,0	0,80	31,52	30,3	33,5	0,422 ^b
Hematócrito (%)	37,3	4,35	37,10	23,0	52,0	42,0	4,10	41,90	34,2	51,2	<0,001^a
RDW (%)	15,8	1,77	15,70	12,3	22,3	14,2	1,30	14,05	12,1	19,1	<0,001^b
Leucócitos ($\times 10^9/L$)	7,49	2,221	7,160	4,03	16,40	7,29	1,901	7,140	4,00	12,70	0,823 ^b
Plaquetas ($\times 10^9/L$)	268	65,6	261,0	132	433	238	50,1	227,0	153	415	0,001^b

DP – desvio padrão; Min – mínimo; Max – máximo; VGM – volume globular médio; HGM – hemoglobina globular média; CHGM – concentração da hemoglobina globular média; RDW – *red cell distribution width*. ^a – Teste T; ^b – Teste de Mann-Witney. Consideram-se os valores estatisticamente significativos quando *p-value* < 0,05 (os valores estão salientados a negrito).

Considerando a população estudada com microcitose (VGM < 80 fL) e/ou hipocromia (HGM < 27 pg), não se encontraram diferenças estatisticamente significativas (*p-value* $\geq 0,05$) para os níveis de CHGM, eritrócitos (quer nos homens, quer nas mulheres) e leucócitos entre os indivíduos sem α -talassémia e os portadores de α -talassémia. Contudo, observaram-se diferenças estatisticamente significativas (*p-value* < 0,05) nos níveis de hemoglobina (tanto para os homens, como para as mulheres), hematócrito, VGM, HGM, RDW e plaquetas, entre os indivíduos sem α -talassémia e os portadores de α -talassémia. Verificou-se que a α -talassémia está associada a níveis mais elevados de hemoglobina para homens e mulheres, VGM, HGM e hematócrito e a níveis mais reduzidos de RDW e plaquetas, quando comparados com os indivíduos sem α -talassémia.

Há ainda a realçar que todos os indivíduos que têm α -talassémia apresentam hipocromia (valor mínimo de HGM de 22,4 pg e máximo de 26,9 pg) mas uma proporção considerável deles não apresenta microcitose (valor mínimo do VGM de 68,5 fL e máximo de 86,9 fL). Podemos assim concluir que a HGM é um parâmetro hematimétrico mais fiável a usar do que o VGM para a discriminação dos portadores de α -talassémia entre uma população com microcitose e/ou hipocromia.

Outro parâmetro com utilidade na discriminação dos portadores de α -talassémia é o RDW. No nosso estudo observámos que estes portadores apresentam valores de RDW mais baixos que a restante população com microcitose e/ou hipocromia, mas sem α -talassémia, que será maioritariamente constituída por indivíduos com anemia ferropénica. Esta constatação já tem sido observada por outros autores (67, 68).

Estas pequenas diferenças nas características dos eritrócitos dos indivíduos com α -talassémia têm sido usadas em algoritmos e fórmulas matemáticas para a discriminação destes portadores de entre a população com anemias microcíticas (69). Contudo é de extrema importância lembrar que o diagnóstico correto de α -talassémia tem obrigatoriamente de ser obtido através de estudos moleculares.

4.4. Diagnóstico de alterações no gene *HBB*: β -talassémia e variantes de hemoglobina

A análise da β -globina foi realizada através da amplificação do gene *HBB* por PCR. Os fragmentos de 2106 pb originados por PCR para cada um dos 204 indivíduos deste estudo foram confirmados por eletroforese em gel de agarose (Figura 4.5) e sequenciados por NGS.

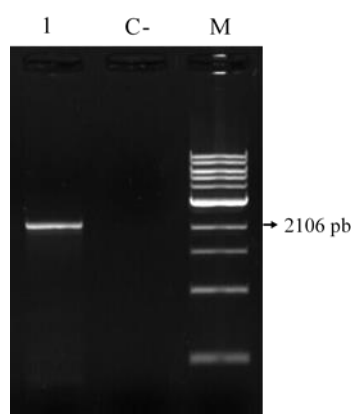


Figura 4.5. Visualização e análise quantitativa e qualitativa do produto do PCR para o diagnóstico de variantes no gene *HBB*. Perfil eletroforético em gel de agarose (1%, p/v) esperado, permitindo a visualização do fragmento de 2106 pb que contém os três exões do gene *HBB*, as duas regiões intrónicas e parte das regiões 5'UTR e 3'UTR. **1** – primeiro poço do gel onde se encontra o fragmento originado por PCR; **C-** – controlo negativo; **M** – marcador de peso molecular (*1kb DNA Ladder*, BioLabs). O tamanho do fragmento encontra-se indicado no lado direito da imagem, em pb.

A análise por NGS revelou um total de 29 variantes genéticas diferentes (Tabela 7.10, Anexos): dez variantes em regiões codificantes, das quais três são *missense*, três são sinónimas, duas são *nonsense* e duas são *frameshift*; 12 variantes intrónicas e sete variantes *downstream*/3'UTR.

Todas as variantes receberam um identificador rs, exceto a variante c.92+25G>T. Além disso, todas as variantes identificadas na base de dados *ClinVar* como sendo patogénicas ou possivelmente patogénicas (ou seja, que causam doença possivelmente devido à sua presença trazer alguma consequência funcional na proteína) foram confirmadas por sequenciação de Sanger em pelo menos um indivíduo identificado como portador da variante. Nesta dissertação, apenas serão abordadas essas 10 alterações consideradas possivelmente patogénicas e, consequentemente, com interesse clínico (Tabela 4.4).

Tabela 4.4. Variantes patogénicas no gene *HBB* identificadas por NGS

Posição (GRCh38.p12)	rs_ID	Nome Comum	Região	Consequência molecular	NM_000518.5	NP_000509.1
g.5225678	rs33946267	Hb D-Portugal	Exão 3	Var <i>Missense</i>	c.364G>C	p.Glu122Gln
g.5226762	rs80356821	Cd41/42 (-CTTT)	Exão 2	Var <i>Frameshift</i>	c.126_129del	p.Phe42LeufsTer19
g.5226774	rs11549407	Cd39 (C>T)	Exão 2	Var <i>Nonsense</i>	c.118C>T	p.Gln40Ter
g.5226820	rs35004220	IVS-I-110 (G>A)	Intrão 1	Var Intrónica	c.93-21G>A	-
g.5226924	rs35724775	IVS-I-6 (T>C)	Intrão 1	Var Intrónica	c.92+6T>C	-
g.5226929	rs33971440	IVS-I-1 (G>A)	Intrão 1	Var Intrónica	c.92+1G>A	-
g.5226974	rs34716011	Cd15 (G>A)	Exão 1	Var <i>Nonsense</i>	c.48G>A	p.Trp16Ter
g.5227001	rs63749819	Cd6 (-A)	Exão 1	Var <i>Frameshift</i>	c.20del	p.Glu7GlyfsTer13
g.5227002	rs334	Hb S	Exão 1	Var <i>Missense</i>	c.20A>T	p.Glu7Val
g.5227003	rs33930165	Hb C	Exão 1	Var <i>Missense</i>	c.19G>A	p.Glu7Lys

rs_ID – Identificação das variantes segundo a referência da base de dados *dbSNP*; **NM** – sequência de referência com base num RNA codificante de proteína (mRNA); **NP** – sequência de referência com base na proteína; **Var** – variante.

4.4.1. Hb D-Portugal (*HBB*:c.364G>C)

Esta alteração, também conhecida como Hb D-Punjab, Hb D-Los Angeles, Hb D-North Carolina, Hb D-Chicago e Hb Oak Ridge é uma das variantes de hemoglobina mais comuns a nível mundial, depois da Hb S, da Hb C e da Hb E. É encontrada principalmente no noroeste da Índia (no estado de Punjab), mas a sua distribuição encontra-se atualmente bastante alargada por todo o mundo (70).

A variante surge de uma mutação pontual na primeira base do codão 121 da β -globina, que tem como consequência a alteração do aminoácido ácido glutâmico (GAA) para glutamina (CAA). O ácido glutâmico está envolvido nos locais de ativação ou ligação da proteína e pode ser substituído por outros aminoácidos polares, em particular o ácido aspártico e a glutamina, que é o que acontece na Hb D-Portugal (71). Assim, é de esperar que esta alteração de aminoácidos não acarrete consigo grandes consequências a nível clínico, embora esteja descrita na base de dados *ClinVar* como havendo conflito de interpretação quanto à sua patogenicidade.

Para esclarecer melhor o impacto desta variante *missense* ao nível da proteína, recorreu-se aos estudos *in silico*, nomeadamente aos softwares *PolyPhen-2* e *SIFT*. O *PolyPhen-2* prevê a variante como sendo uma mutação benigna com um *score* de 0,174 (Figura 4.6) e o mesmo faz o software *SIFT*, que prevê a variante também como benigna, com um *score* de 0,18.

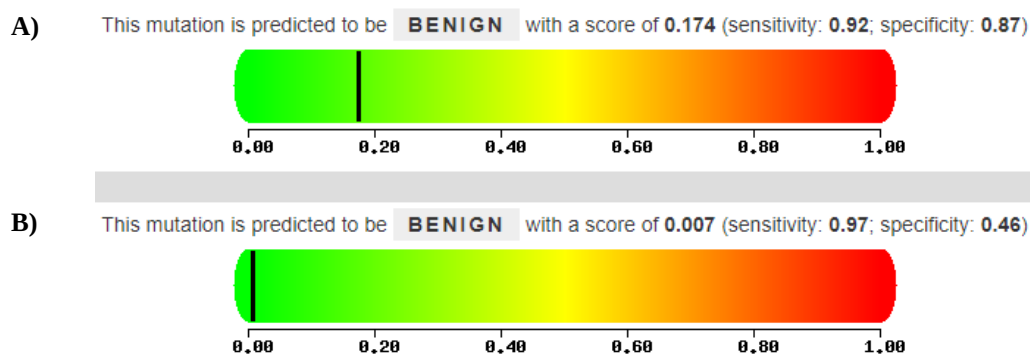


Figura 4.6. Previsão do efeito patogénico da variante p.Glu122Gln na proteína da β -globina através do software PolyPhen-2. A) Dados do modelo *HumDiv* para a alteração, que a prevê como sendo benigna. B) Dados do modelo *HumVar* para a alteração que a prevê como sendo benigna.

A análise do perfil de conservação obtido através do alinhamento múltiplo (Figura 4.7) revelou que o resíduo de ácido glutâmico presente na posição 121 é conservado entre espécies próximas de mamíferos. Para algumas espécies, o perfil mostra que o resíduo de ácido glutâmico pode ser substituído por principalmente por ácido aspártico, devido às suas propriedades polares, tal como já referido anteriormente. Tendo em conta que a glutamina também é um aminoácido polar, embora não esteja ilustrada neste alinhamento múltiplo, faz sentido que, no que diz respeito a sua patogenicidade, esta alteração seja benigna.

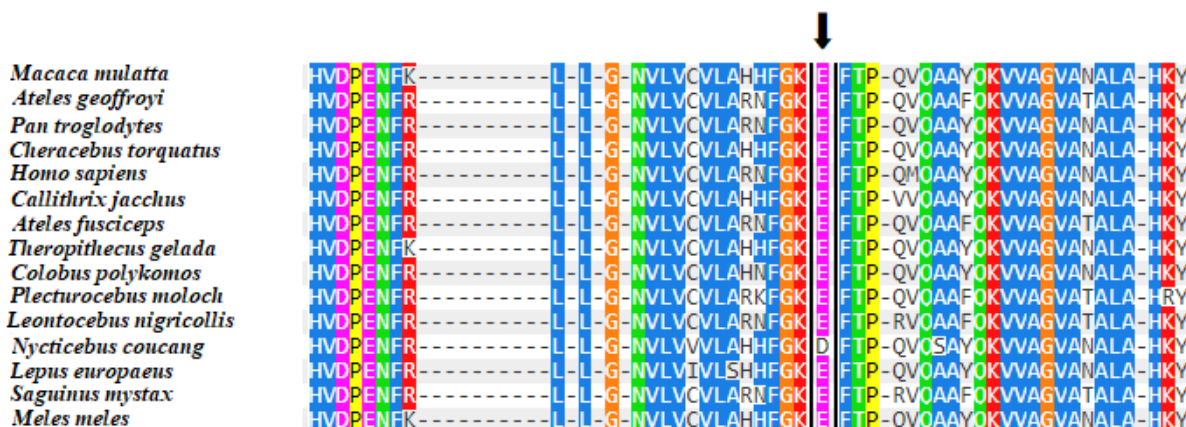


Figura 4.7. Perfil de conservação evolutiva do aminoácido p.Glu122 na proteína da β -globina. Alinhamento múltiplo de sequências de aminoácidos de proteínas homólogas de 15 espécies de mamíferos relacionados, através do software *PolyPhen-2*. O padrão de substituições de aminoácidos está ilustrado dentro da caixa preta e destacado com uma seta, sendo que neste, a letra E corresponde ao ácido glutâmico e a letra D ao ácido aspártico.

Assim, após a análise *in silico* e confirmação com o descrito na literatura, a Hb D-Portugal por si só não parece ter qualquer significado clínico e, em indivíduos heterozigóticos ou homozigóticos geralmente é assintomática ou está associada a uma anemia muito ligeira. Contudo, o conflito de interpretação quanto à patogenicidade indicado na base de dados *ClinVar* pode ser explicado pelo facto desta variante, quando co-herdada com outras variantes como Hb S ou talassémia, conduzir a manifestações clínicas que podem variar de moderadas a graves, semelhantes a um fenótipo de doença falciforme (no caso de co-herança com Hb S) ou anemia leve a grave e esplenomegalia (70, 72).

Esta variante foi detetada por NGS em heterozigotia apenas num indivíduo (Tabela 4.5) e foi confirmada por sequenciação de Sanger (Figura 4.8).

Tabela 4.5. Resultados do NGS para a variante c.364G>C no gene *HBB*

ID	Ref	Alt	Número de participante	Genótipo	Cobertura alélica (Ref;Alt)	Cobertura Total
c.364G>C; p.Glu122Gln	G	C	PTI113	0/1	676;692	1369

ID – Identificação da variante ao nível do gene, seguida da identificação ao nível da proteína; **Ref** – Alelo referência; **Alt** – Alelo alterado; **0/1** – heterozigotia.

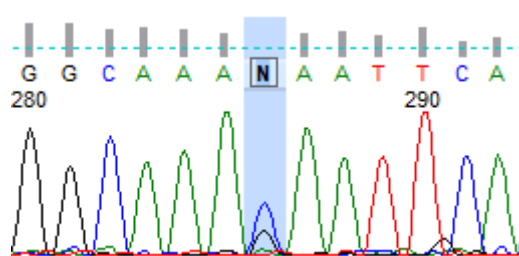


Figura 4.8. Verificação por sequenciação de Sanger da variante *HBB*:c.364G>C. Eletroferograma parcial da sequenciação de Sanger onde se pode verificar na barra sombreada a azul a heterozigotia G/C na posição c.364 da região codificante do gene *HBB* (terceiro exão).

O indivíduo em questão (PTI113) tem 65 anos, é de nacionalidade portuguesa e do sexo masculino. Possui um fenótipo de anemia microcítica e hipocrômica (hemoglobina= 11,1 g/dL; VGM= 72,5 fL e HGM= 22,7 pg – restantes parâmetros hematológicos deste indivíduo na Tabela 7.11 dos Anexos). Além da variante Hb D-Portugal não foram detetadas outras patologias co-herdadas, quer no gene da β -globina, quer no gene da α -globina, e por isso o seu fenótipo de anemia microcítica e hipocrômica não deverá ser resultante da variante Hb D-Portugal por si só em heterozigotia. Deste modo, torna-se necessário avaliar também o *status* do ferro deste indivíduo através de testes bioquímicos (por exemplo o doseamento da ferritina, ferro sérico e da capacidade total de fixação do ferro), pois tendo nós excluído a presença de talassémias, as alterações dos valores dos parâmetros hematológicos neste indivíduo podem ser indicadoras de uma anemia ferropénica.

4.4.2. Cd 41/42 (-CTTT) (*HBB*:c.126_129delCTTT)

A variante *HBB*:c.126_129delCTTT, também conhecida como Cd41/42 (-CTTT) foi descrita pela primeira vez em 1983 por Kimura e os seus colaboradores. É uma variante particularmente prevalente no este asiático, com especial incidência na população chinesa, e geralmente está associada a um fenótipo de β^0 -talassémia, sendo classificada como β^0 (73, 74).

Esta variante *frameshift* surge de uma deleção de quatro nucleótidos envolvendo parcialmente os codões 41 e 42, que conduzem à substituição do aminoácido fenilalanina (TTC) por uma leucina (TTG) e consequentemente é criado um codão de terminação prematuro na posição 19 da nova grelha de leitura (Figura 4.9).

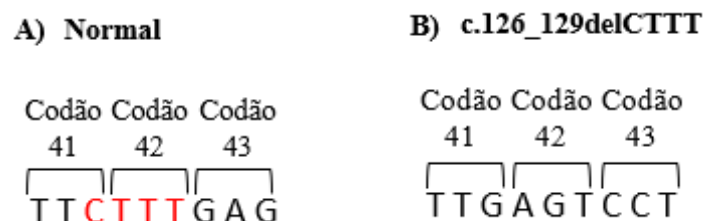


Figura 4.9. Alteração da sequência de nucleótidos no gene *HBB*, gerada pela variante *HBB*:c.126_129delCTTT. A) Sequência de nucleótidos dos codões 41, 42 e 43 do gene *HBB*, num indivíduo normal. A vermelho encontram-se realçados os nucleótidos que serão eliminados se houver a referida deleção. B) Sequência de nucleótidos dos codões 41, 42 e 43 do gene *HBB* dos indivíduos portadores da variante c.126_129delCTTT.

Vários estudos referem que uma característica interessante das variantes que conduzem à formação de codões de terminação prematuros, é que os níveis de mRNA anómalos da β -globina são muito baixos, em comparação com o mRNA normal. É agora claro que estes níveis reduzidos de mRNA anómalo da β -globina são resultado do mecanismo *nonsense-mediated decay* (NMD), uma via que tem como principal função reduzir erros na expressão génica, eliminando os transcritos de mRNA que contêm codões de terminação prematuros (75). Assim, o novo codão de terminação resultante da variante *HBB*:c.126_129delCTTT, vai interromper o normal processo de tradução do mRNA, o que pode conduzir à formação de uma proteína truncada ou de um mRNA sujeito a esse mecanismo NMD. Em qualquer um dos casos, ocorre perda da função normal da proteína β -globina e, por isso, esta variante está descrita na base de dados *ClinVar* como sendo patogénica.

Ainda assim, recorreu-se ao estudo *in silico* para confirmar o impacto desta variante. Por se tratar de uma deleção, não foi possível utilizar o *software PolyPhen-2*, uma vez que este apenas faz análise de variantes *missense* e por isso recorreremos apenas ao SIFT. A variante *HBB*:c.126_129delCTTT foi classificada como sendo prejudicial (*damaging*) com um *score* de 0,858. Esta classificação deve-se ao facto desta deleção afetar bases conservadas de uma sequência de DNA codificante.

Esta variante foi detetada por NGS em heterozigotia apenas num indivíduo (Tabela 4.6) e foi confirmada por sequenciação de Sanger (Figura 4.10).

Tabela 4.6. Resultados do NGS para a variante c.126_129delCTTT no gene *HBB*

ID	Ref	Alt	Número de participante	Genótipo	Cobertura alélica (Ref;Alt)	Cobertura Total
c.126_129del; p.Phe42LeufsTer19	GTTTC	G	PT1196	0/1	863;389	1252

ID – Identificação da variante ao nível do gene, seguida da identificação ao nível da proteína; **Ref** – Alelo referência; **Alt** – Alelo alterado; **0/1** – heterozigotia.

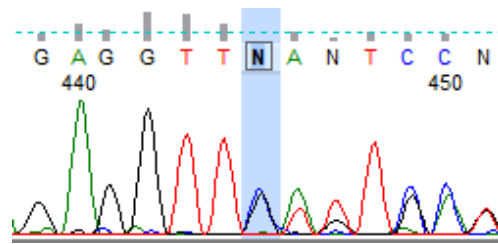


Figura 4.10. Verificação por sequenciação de Sanger da variante *HBB*:c.126_129delCTTT. Eletroferograma parcial da sequenciação de Sanger onde se pode verificar na barra sombreada a azul o início da heterozigotia CTTT/del e é possível observar também a consequente alteração da grelha de leitura.

O indivíduo em questão (PTI196) é uma mulher de 44 anos, com um fenótipo de anemia microcítica e hipocrômica (hemoglobina= 9,0 g/dL; VGM= 60,0 fL e HGM= 17,2 pg – restantes parâmetros hematológicos deste indivíduo na Tabela 7.11, Anexos). Esta participante é de origem chinesa, o que explica o facto de ser portadora desta variante, pois a alteração *HBB*:c.126_129delCTTT é uma das quatro mutações no gene da β -globina mais frequentes na China, descrita com uma frequência relativa de 31,8% nesse local, de acordo com a base de dados *ItaGenes*. Assim, podemos concluir que esta participante é portadora de uma β -talassémia *minor* (β/β^0), responsável pelo seu quadro clínico de anemia microcítica e hipocrômica. Tanto quanto sabemos, esta alteração talassémica ainda não foi reportada na população com origem portuguesa.

4.4.3. Cd 39 (C>T) (*HBB*:c.118 C>T)

A variante *HBB*:c.118C>T, também conhecida por Cd39 (C>T), foi uma das primeiras mutações *nonsense* do gene *HBB* a ser caracterizada e extensivamente estudada e é uma das mutações β -talassémicas mais comuns na população da zona do Mediterrâneo (75).

A Cd39 (C>T) surge de uma mutação pontual no codão 39 do gene da β -globina, responsável pela conversão do aminoácido glutamina (CAG) num codão de terminação prematuro (TAG). Esta alteração, interrompe a tradução do mRNA e por isso prevê-se que resulte numa proteína truncada ou num mRNA sujeito ao mecanismo NMD (76, 77). Esta variante resulta então na perda da função normal da proteína β -globina e por esta razão está classificada na base de dados *ClinVar* como sendo patogénica.

Recorreu-se ao estudo *in silico* para confirmar o impacto desta variante, contudo, por se tratar de uma variante *nonsense*, mais uma vez não foi possível utilizar o *software PolyPhen-2*. O impacto da variante *HBB*:c.118C>T foi avaliado com recurso ao *software* SIFT que, como esperado, a classificou como sendo prejudicial (*damaging*). De facto, os estudos experimentais demonstraram que esta variante *nonsense* está sujeita ao mecanismo NMD que elimina todos os mRNAs que contêm este codão de terminação prematuro, permitindo limitar a síntese de proteínas truncadas e resultando em pouca ou nenhuma proteína detetável *in vitro*, confirmando o resultado do estudo *in silico* (78).

A variante Cd39 (C>T) foi encontrada em oito indivíduos sempre em heterozigotia (Tabela 4.7) e foi confirmada por sequenciação de Sanger num deles (Figura 4.11).

Tabela 4.7. Resultados do NGS para a variante c.118 C>T no gene *HBB*

ID	Ref	Alt	Número de participante	Genótipo	Cobertura alélica (Ref;Alt)	Cobertura Total
c.118 C>T; p.Gln40Ter	C	T	PTI 28	0/1	334;226	560
			PTI 41	0/1	895;480	1375
			PTI 94	0/1	919;504	1423
			PTI 105	0/1	892;526	1418
			PTI 115	0/1	811;453	1264
			PTI 124	0/1	726;395	1121
			PTI 142	0/1	756;424	1180
			PTI 143	0/1	855;421	1276

ID – Identificação da variante ao nível do gene, seguida da identificação ao nível da proteína; **Ref** – Alelo referência; **Alt** – Alelo alterado; **0/1** – heterozigotia.

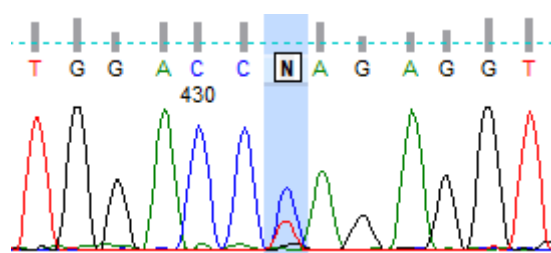


Figura 4.11. Verificação por sequenciação de Sanger da variante *HBB*: c.118 C>T. Eletroferograma parcial da sequenciação de Sanger onde se pode verificar na barra sombreada a azul a heterozigotia C/T.

Esta variante foi a alteração detetada por NGS mais frequente entre os participantes deste estudo e todos os 8 indivíduos identificados como portadores são de origem portuguesa, o que está de acordo com o descrito na literatura, pois a variante Cd39 (C>T) é uma das quatro mutações β -talassémicas mais comuns na população portuguesa (79). Assim, a presença desta variante leva a que os 8 indivíduos portadores da alteração c.118 C>T possuam uma β -talassémia *minor* (β/β^0), que é responsável pelo seu quadro clínico de anemia microcítica e hipocrômica (os parâmetros hematológicos estão todos descritos na Tabela 7.11, Anexos).

4.4.4. IVS-I-110 (G>A) (*HBB*:c.93-21G>A)

A variante *HBB*:c.93-21G>A, também conhecida como IVS-I-110 (G>A) foi a primeira substituição de nucleótidos identificada num gene β -talassémico (80, 81). Esta variante é uma das mutações β -globínicas encontradas com mais frequência em indivíduos β -talassémicos e é um dos alelos da β^+ -talassémia mais comum nos países do Mediterrâneo e do Médio Oriente.

A IVS-I-110 (G>A) é caracterizada pela substituição de uma guanina por uma adenina, 21 nucleótidos a montante do segundo exão do gene *HBB*, ou seja, perto do terminal 3' do primeiro intrão do gene. A análise dos nucleótidos presentes nas fronteiras exão-intrão demonstrou que, nestes locais, existem sequências conservadas, chamadas sequências *consensus*, necessárias para que o processo de *splicing*

ocorra com eficiência e precisão. Na extremidade 5' e na extremidade 3' de cada intrão (e dessas sequências *consensus*), encontram-se também os dinucleótidos GT e AG, respetivamente, que são altamente conservados, ao passo que outras bases presentes na sequência *consensus* podem variar. A extremidade 5' também é denominada de local dador de *splicing*, e a extremidade 3' local aceitador de *splicing* (82).

Por ocorrer numa sequência intrónica, a variante IVS-I-110 (G>A) não altera diretamente a sequência de aminoácidos que codifica a proteína β-globina, contudo cria um novo local aceitador de *splicing* crítico, 19 nucleótidos a montante do local aceitador de *splicing* autêntico. Este local aceitador de *splicing* crítico é gerado diretamente pela substituição G>A, que conduz à formação de um novo dinucleótido AG, conservado em locais aceptadores de *splicing*, como referido anteriormente. Além disso, a sequência que flanqueia este novo dinucleótido assemelha-se muito à sequência *consensus* do local aceitador de *splicing* autêntico e por isso o novo local aceitador de *splicing* possuiu todas as condições para competir com o local autêntico, resultando num *splicing* anormal nalgumas das moléculas precursoras do mRNA que codifica a proteína β-globina (80, 81). Com base nisto, a variante IVS-I-110 (G>A) é considerada patogénica pelo *ClinVar* e está associada à β-talassémia.

Por se tratar de uma alteração numa zona não codificante fez-se o estudo *in silico* do seu impacto ao nível do *splicing*, com recurso ao *software* varSEAK (Figura 4.12).

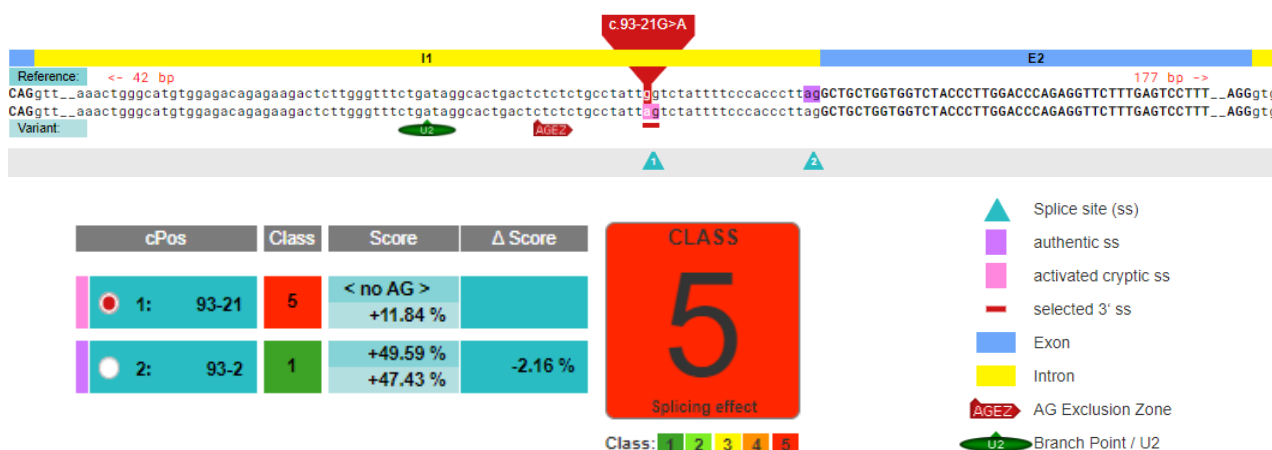


Figura 4.12. Análise do impacto da alteração c.93-21G>A no processo de *splicing* do RNA codificante da proteína β-globina através do *software* varSEAK. I1 – intrão 1; E2 – éxon 2; bp – pares de bases; **Reference** – sequência referência (normal); **Variant** – sequência variante (com a alteração); **c.Pos** – posição na sequência de DNA; **Class** – classe de previsão do efeito no *splicing*; **Score** – probabilidade prevista do local de *splicing* ser funcional ou não; **ΔScore** – variação entre o *score* do local de *splicing* da sequência de referência e da sequência variante.

Os dados obtidos pelo *varSEAK* indicam que o local aceitador de *splicing* autêntico na sequência referência, ou seja, na sequência onde não ocorre a variante, tem um *score* de +49,59%, que reflete a probabilidade deste local de *splicing* ser funcional. Por sua vez, na sequência variante, isto é, na sequência onde ocorreu a alteração *HBB*:c.93-21G>A, o *score* do local de *splicing* autêntico reduz ligeiramente para +47,43%. Ao mesmo tempo que se observa esta pequena redução de *score* no local aceitador de *splicing* autêntico, verifica-se que na sequência variante, o novo local aceitador de *splicing* criado tem *score* de +11,84%, ou seja, adquire uma probabilidade considerável de ser funcional. Este aumento de funcionalidade de 11,84% é previsto pelo *software* como tendo efeito no mecanismo de *splicing*, incluindo-o assim na classe mais alta com efeito ao nível do *splicing* - classe 5. A ativação deste local aceitador de *splicing* crítico, provocada pela variante *HBB*:c.93-21G>A, tem como

consequência a inclusão de parte do primeiro intrão do gene *HBB*, com alteração da grelha de leitura, o que poderá provocar a introdução de um codão de terminação prematuro. Para testar esta hipótese, recorreu-se à ferramenta bioinformática *Translate Tool* da *ExPASy*, de forma a traduzir as sequências de DNA codificante original e mutado nas respetivas sequências de aminoácidos (Figura 4.13).



Figura 4.13. Sequência de aminoácidos correspondente aos exões 1 e 2 do gene da β -globina normal e com a variante c.93-21G>A. A) Sequência parcial de aminoácidos da proteína da β -globina normal. **B)** Sequência de aminoácidos que é gerada pelo uso do local aceitador de *splicing* crítico no mecanismo de remoção do primeiro intrão do gene *HBB*, na presença da variante c.93-21G>A, revelando a introdução de um codão de terminação prematuro. A sequência de aminoácidos correspondente ao exão 1 encontra-se a azul e a sequência correspondente ao exão 2 está indicada a laranja. A sequência que permanece inalterada está realçada a cinzento e corresponde a parte do primeiro exão do gene *HBB*. **Stop** – codão de terminação.

Comparando as duas sequências, verifica-se que a alteração c.93-21G>A leva à introdução de 19 nucleótidos pertencentes ao primeiro intrão do gene *HBB*, o que corresponde a seis codões extra na região 5' do segundo exão. Esta introdução de codões poderia conduzir à produção de uma proteína mais longa e disfuncional, porém o sexto codão extra é um codão de terminação prematuro (TAG). Por isso, apesar de serem incluídos nucleótidos na sequência codificante, a proteína resultante será mais curta, uma vez que o codão de terminação anula a sequência tanto do segundo exão como do terceiro, originando uma proteína anormal.

De facto, encontra-se descrito na literatura que a substituição de G>A na posição 110 do primeiro intrão do gene da β -globina ativa um novo local de *splicing* e, embora o local de *splicing* autêntico possa funcionar normalmente, ambos os locais competem com o processo de *splicing* de RNA, sendo que o novo local de *splicing* é usado preferencialmente 80% a 90% das vezes (83, 84). Em suma, com o estudo *in silico* percebemos que na presença da variante c.93-21G>A existem dois locais que podem ser usados pela maquinaria de *splicing* para a remoção de intrões: o local aceitador de *splicing* autêntico com uma probabilidade de 47,43% de ser funcional e o local aceitador de *splicing* crítico, com 11,84% de probabilidade de ser funcional, o que torna mais provável que a maquinaria de *splicing* utilize o local autêntico que o local crítico para a remoção do intrão. Além disso, nas condições em que o local aceitador de *splicing* crítico é usado, o transcrito anómalo gerado deverá ser direcionado para o mecanismo NMD, sendo deste modo eliminados todos os mRNAs que contêm o codão de terminação prematuro consequente da alteração *HBB*:c.93-21G>A, e permanecendo apenas os transcritos normais. Percebemos assim que, na presença da mutação IVS-I-110 (G>A), continua a haver produção de cadeias β -globínicas normais, embora possivelmente numa quantidade mais reduzida, sendo assim esta mutação responsável pelo fenótipo β^+ .

Neste estudo, a variante IVSI-110 foi detetada por NGS em três indivíduos, sempre em heterozigotia (Tabela 4.8), tendo sido confirmada posteriormente por sequenciação de Sanger (Figura 4.14).

Tabela 4.8. Resultados do NGS para a variante c.93-21G>A no gene *HBB*

ID	Ref	Alt	Número de participante	Genótipo	Cobertura alélica (Ref;Alt)	Cobertura Total
c.93-21G>A	G	A	PTI 43	0/1	431;280	713
			PTI 130	0/1	622;453	1075
			PTI 145	0/1	918;681	1599

ID – Identificação da variante ao nível do gene; **Ref** – Alelo referência; **Alt** – Alelo alterado; **0/1** – heterozigotia.

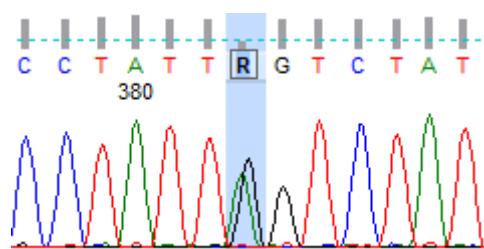


Figura 4.14. Verificação por sequenciação de Sanger da variante *HBB*:c.93-21G>A. Eletroferograma parcial da sequenciação de Sanger onde se pode verificar na barra sombreada a azul a heterozigotia G/A.

A mutação IVS-I-110 (G>A) é também uma das quatro mutações β -talassémicas mais comuns na população portuguesa, e por isso a sua deteção neste estudo era esperada, principalmente em indivíduos de origem portuguesa (79). De facto, os três indivíduos identificados por NGS têm todos origem portuguesa e, embora apenas um deles apresente uma anemia ligeira, os três possuem microcitose e hipocromia (parâmetros hematológicos descritos na Tabela 7.11 dos Anexos), derivadas da presença da variante em heterozigotia.

4.4.5. IVS-I-6 (T>C) (*HBB*:c.92+6T>C)

A variante *HBB*:c.92+6T>C, amplamente conhecida como IVS-I-6 (T>C), é uma causa comum e bem descrita de β -talassémia, com uma elevada frequência entre as populações do Mediterrâneo e do Médio Oriente (85).

A IVS-I-6 (T>C) deriva de uma alteração na sequência de nucleótidos no primeiro intrão do gene *HBB*, nomeadamente a substituição de uma timina por uma citosina. Como já referido, para que o mecanismo de excisão dos intrões ocorra é necessária a presença do dinucleótido GT conservado nos locais dadores de *splicing*, e o grau de conservação das sequências que flanqueiam esse dinucleótido também é importante para o processo de *splicing*. Deste modo, embora a variante *HBB*:c.92+6T>C não altere diretamente a sequência de aminoácidos codificante da proteína β -globina, a substituição de nucleótidos ocorre precisamente na sequência *consensus* que abrange o local dador de *splicing*. A mutação tem como consequência a redução da eficácia através da qual a maquinaria responsável pelo *splicing* consegue discriminar entre o local dador de *splicing* autêntico e os outros locais dadores *splicing* críticos que existem nas proximidades. Deste modo, a variante *HBB*:c.92+6T>C é responsável pela redução da qualidade do local dador de *splicing* autêntico, interferindo com o processamento normal do mRNA,

que resulta numa expressão reduzida das cadeias de β -globina, característica da β^+ -talassémia e a variante é descrita pelo *ClinVar* como patogénica (85, 86).

Para confirmar o impacto da IVS-I-6 (T>C) ao nível do *splicing*, recorreu-se ao estudo *in silico*, através da ferramenta *varSEAK* (Figura 4.15).

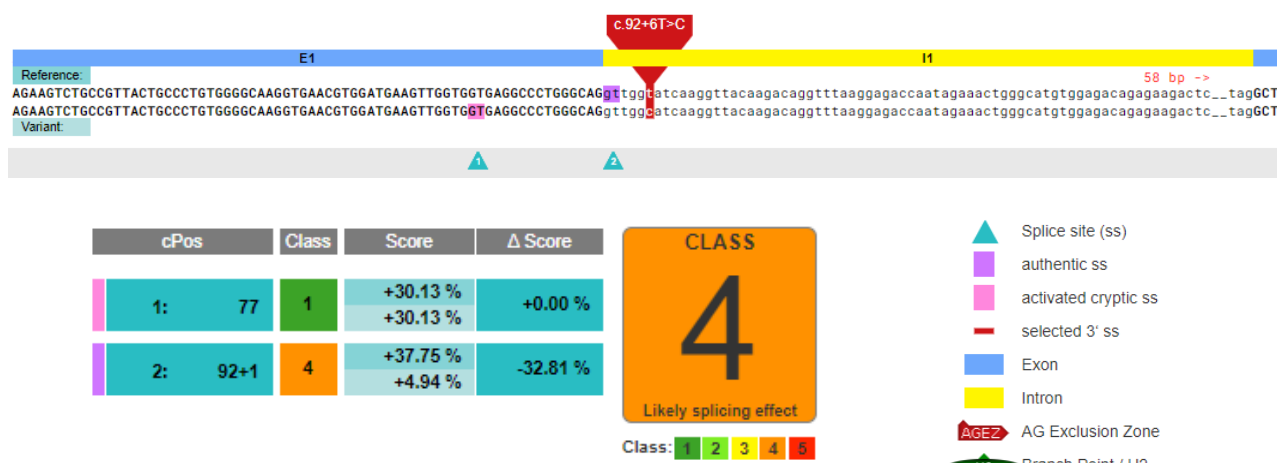


Figura 4.15. Análise do impacto da alteração c.92+6T>C no processo de *splicing* do RNA codificante da proteína da β -globina através do software *varSEAK*. E1 – exão 1; I1 – intrão 1; bp – pares de bases; **Reference** – sequência referência (normal); **Variant** – sequência variante (com a alteração); **c.Pos** – posição na sequência de DNA; **Class** – classe de previsão do efeito no *splicing*; **Score** – probabilidade prevista do local de *splicing* ser funcional ou não; **ΔScore** – variação entre o *score* do local de *splicing* da sequência de referência e da sequência variante.

A análise dos dados obtidos mostra que, tanto na sequência referência, como na sequência variante, o *score* do local crítico de *splicing* (marcado na figura com um triângulo azul contendo o número 1) é de +30,13%, ou seja, a presença da variante *HBB*:c.94+6T>C não altera a probabilidade deste local de *splicing* crítico ser funcional que, por si só, já é elevada. De seguida, ao analisar o local dador de *splicing* autêntico verifica-se que na sequência referência (marcado na figura com um triângulo azul contendo o número 2), este local tem um *score* de +37,75%, um valor ligeiramente superior ao *score* do local crítico de *splicing* (+30,13%), o que nos leva a pensar que ou a maquinaria responsável pelo *splicing* tem uma elevada eficácia, permitindo-lhe discriminar entre estes dois locais, ou podemos colocar a hipótese que, na realidade, em situações normais existe a produção de dois transcritos diferentes, consoante é usado o local dador de *splicing* autêntico ou crítico. Por sua vez, na sequência variante há uma redução drástica do *score* do local dador de *splicing* autêntico de +37,75% para +4,94%, um valor bastante inferior ao *score* do local crítico de *splicing*, de +30,13%. Esta diminuição de funcionalidade do local dador de *splicing* autêntico é por isso interpretada pelo *software* como tendo provavelmente efeito no mecanismo de *splicing*, incluindo-o na classe 4. Além disso, devido a esta redução da funcionalidade do local dador de *splicing* autêntico, é provável que a maquinaria de *splicing* use o local dador de *splicing* crítico existente 16 nucleótidos a montante do local nativo, levando à exclusão da parte 3' do primeiro exão do gene *HBB*, que por sua vez poderá alterar a grelha de leitura.

Utilizou-se a ferramenta bioinformática *Translate Tool* da *ExPASy*, de forma a traduzir as sequências de DNA codificante original e mutado nas respectivas sequências de aminoácidos, para entender o impacto da variante *HBB:c.92+6T>C* (Figura 4.16).

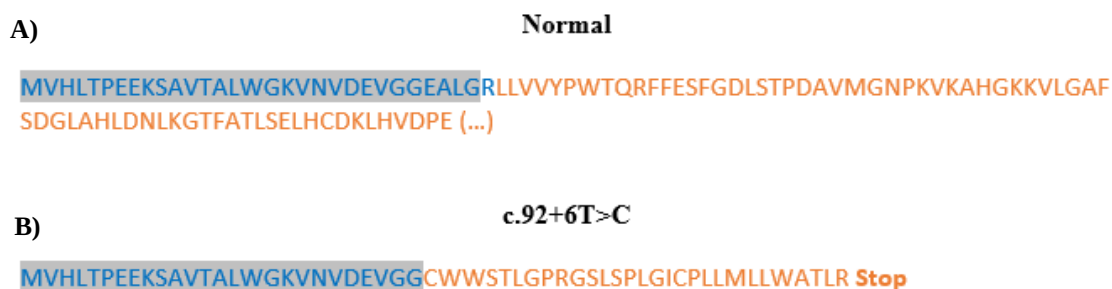


Figura 4.16. Sequência de aminoácidos correspondente aos exões 1 e 2 do gene da β -globina normal e com a variante c.92+6T>C. A) Sequência parcial de aminoácidos da proteína da β -globina normal. B) Sequência de aminoácidos que é gerada pelo uso do local dador de *splicing* crítico no mecanismo de remoção do primeiro intrão do gene *HBB*, na presença da variante c.92+6T>C, revelando a introdução de um codão de terminação prematuro. A sequência de aminoácidos correspondente ao exão 1 encontra-se a azul e a sequência correspondente ao exão 2 está indicada a laranja. A sequência de aminoácidos que permanece inalterada está salientada a cinzento e, como se pode observar, corresponde apenas a parte do primeiro exão do gene *HBB*. **Stop** – codão de terminação.

Como referido, a presença da alteração *HBB:c.92+6T>C* aumenta a probabilidade de ser usado o local dador de *splicing* crítico localizado 16 nucleótidos a montante do local autêntico para a remoção de intrões. Para além disso, a mutação diminui a probabilidade de ser usado o local dador de *splicing* autêntico para a remoção do primeiro intrão. Neste caso, a maquinaria de *splicing* recorre ao local dador de *splicing* crítico localizado 16 nucleótidos a montante do local autêntico, levando assim à exclusão de 16 nucleótidos terminais do primeiro exão, que vai alterar toda a grelha de leitura a jusante e consequentemente toda a sequência de aminoácidos, tendo como resultado a introdução de um codão de terminação prematuro (TGA). A introdução deste codão de terminação prematuro conduz à produção de um transcrito de β -globina aberrante e a proteína resultante será mais curta, uma vez que o codão de terminação anula tanto a sequência do segundo exão como do terceiro.

Apesar da IVS-I-6 (T>C) aumentar a probabilidade de ser produzida uma proteína da β -globina anómala, com o estudo *in silico* percebemos que na sequência referência (sem a variante), os *scores* dos dois locais dadores de *splicing* (autêntico e crítico) são muito semelhantes, o que nos fez levantar a hipótese que, em situações normais, é provável que exista a produção de dois transcritos diferentes, consoante seja reconhecido um ou outro local de *splicing*. Nestas condições normais, o transcrito anómalo deverá ser direcionado para o mecanismo NMD, sendo deste modo eliminados todos os mRNAs que contêm este codão de terminação prematuro e permanecendo apenas os mRNAs com processamento normal. Embora na presença da variante *HBB:c.92+6T>C* ocorra uma produção muito maior desses transcritos aberrantes em comparação com os transcritos normais, após a eliminação dos transcritos anómalos por NMD, permanecem os transcritos de β -globina que sofreram *splicing* normal, embora existam numa quantidade inferior à habitual, ou seja, na presença da variante continua a haver produção de cadeias β -globínicas, mas num número mais reduzido.

As evidências referidas vão ao encontro do que está mencionado na literatura para a IVS-I-6 (T>C), descrita como uma mutação que interfere com o *splicing* do mRNA, o que resulta na expressão reduzida das cadeias de β -globina, característica de uma talassémia do tipo β^+ (85). Os indivíduos homozigóticos para a variante c.92+6T>C diferem amplamente no seu quadro clínico, estando a gravidade deste quadro diretamente relacionada com a quantidade de mRNA de β -globina que sofre ou não um *splicing* correto,

e por isso o fenótipo destes indivíduos pode ir desde uma condição ligeira, independente de transfusão, até uma doença mais grave, dependente de transfusão. Contudo, de um modo geral, a mutação IVS-I-6 é considerada como sendo de gravidade suave a moderada pois permite que ocorra o *splicing* normal numa parte dos transcritos de β -globina. Em situações em homozigotia é bastantes vezes associada a β -talassémia intermédia (86, 87). Neste estudo, a variante foi detetada em heterozigotia por NGS em quatro indivíduos (Tabela 4.9), e foi confirmada por sequenciação de Sanger num deles (Figura 4.17).

Tabela 4.9. Resultados do NGS para a variante c.92+6T>C no gene *HBB*

ID	Ref	Alt	Número de participante	Genótipo	Cobertura alélica (Ref;Alt)	Cobertura Total
c.92+6T>C	T	C	PTI 32	0/1	221;144	365
			PTI 40	0/1	549;444	998
			PTI 154	0/1	551;476	1027
			PTI 157	0/1	660;501	1161

ID – Identificação da variante ao nível do gene; **Ref** – Alelo referência; **Alt** – Alelo alterado; **0/1** – heterozigotia.

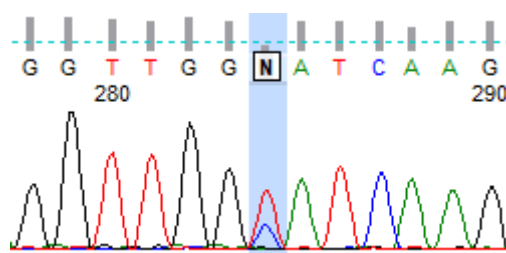


Figura 4.17. Verificação por sequenciação de Sanger da variante *HBB*:c.92+6T>C. Eletroferograma parcial da sequenciação de Sanger onde se pode verificar na barra sombreada a azul a heterozigotia T/C.

A deteção desta variante em quatro indivíduos, torna-a a segunda alteração β -talassémica encontrada com maior frequência neste estudo. Este resultado era já esperado, pois a mutação IVS-I-6 (T>C) é também uma das quatro mutações β -talassémicas mais comuns entre a população portuguesa (79). Os quatro indivíduos portadores da mutação IVS-I-6 (T>C) em heterozigotia têm um quadro hematológico quase normal, apresentando apenas uma leve hipocromia e microcitose, o que vai ao encontro do reportado na literatura, que descreve esta variante como tendo gravidade moderada quando em homozigotia, sendo responsável pela β -talassémia intermédia (87). Esta variante em heterozigotia encontra-se associada a fenótipos hematológicos muito ligeiros e foi descrita como tendo uma elevada prevalência na zona de Coimbra e Aveiro (88). Por isso, é por vezes chamada de β -talassémia portuguesa do tipo β^{++} . De facto, dois dos indivíduos detetados com a alteração IVS-I-6 (T>C) residem na região Centro de Portugal, indo ao encontro do esperado. Os outros dois indivíduos residem na Região Autónoma da Madeira e, tendo em conta que não existem estudos que englobem as populações das regiões autónomas de Portugal, os nossos resultados podem sugerir que esta região também tenha uma elevada prevalência de portadores da alteração IVS-I-6 (T>C). Deste modo, realçamos a importância de os estudos futuros incluírem as populações insulares de Portugal.

O indivíduo PTI 157, além de ter a IVS-I-6 (T>C) em heterozigotia, possui também a deleção de 3,7 kb no gene da α -globina. A combinação de β -talassêmia com α -talassêmia diminui o desequilíbrio de cadeias α/β -globina, resultando, portanto, num fenótipo menos grave. De facto, este indivíduo é o único dos quatro portadores de IVS-I-6 (T>C) que não possui microcitose (VGM= 82,3 fL) e a sua hipocromia é bastante ligeira (HGM= 25,4 pg), sendo também a menos acentuada dos quatro indivíduos. Conclui-se assim que nenhum dos indivíduos possui um quadro clínico grave e que as ligeiras alterações nos parâmetros hematológicos podem ser explicadas pela presença da variante *HBB*:c.92+6T>C.

4.4.6. IVS-I-1 (G>A) (*HBB*:c.92+1G>A)

A variante *HBB*:c.92+1G>A, conhecida de um modo comum por IVS-I-1 (G>A), é também ela uma das mutações β -globínicas mais comuns, estando descrita em muitas populações com frequências diferentes, mas com prevalência elevadas nos países da zona do Mediterrâneo e também na Índia (89).

A IVS-I-1 (G>A) deriva da substituição de uma guanina por uma adenina, no primeiro intrão do gene *HBB*, um nucleótido após o primeiro exão, ou seja, na região 5' do primeiro intrão. Tendo em conta que ocorre numa sequência intrónica, a variante não vai alterar diretamente a sequência de aminoácidos codificante da proteína da β -globina, mas a substituição G>A altera diretamente o dinucleótido conservado no local dador de *splicing* de GT para AT. A variante acaba assim por eliminar este dinucleótido conservado, que é um requisito para que o processo de *splicing* ocorra normalmente e por isso este processo fica completamente impedido de acontecer e nenhum mRNA normal é produzido (86). Por esta razão, a variante *HBB*:c.92 + 1G>A é classificado como patogénica pelo *ClinVar*.

Tendo em conta que a IVS-I-1 (G>A) é uma alteração que ocorre numa zona não codificante, o estudo *in silico* para confirmar o seu impacto ao nível do *splicing*, foi realizado através do *software* varSEAK (Figura 4.18).

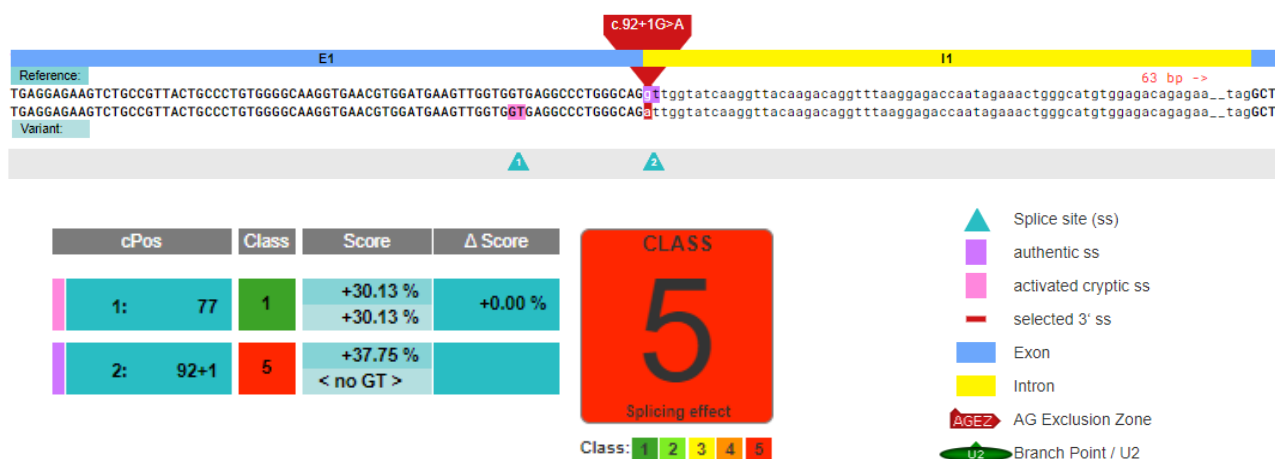


Figura 4.18. Análise do impacto da alteração c.92+1G>A no processo de *splicing* do RNA codificante da proteína da β -globina através do *software* varSEAK. E1 – exão 1; I1 – intrão 1; bp – pares de bases; **Reference** – sequência referência (normal); **Variant** – sequência variante (com a alteração); **c.Pos** – posição na sequência de DNA; **Class** – classe de previsão do efeito no *splicing*; **Score** – probabilidade prevista do local de *splicing* ser funcional ou não; **ΔScore** – variação entre o *score* do local de *splicing* da sequência de referência e da sequência variante.

Os dados obtidos através da ferramenta *varSEAK* para o local dador de *splicing* são semelhantes aos da mutação IVS-I-6 (T>C): tanto na sequência referência como na sequência variante, o *score* do local crítico de *splicing* a montante do nativo (marcado na figura com um triângulo azul contendo o número

1) é de +30,13%, ou seja, a presença da variante c.92+1G>A não altera a probabilidade deste local de *splicing* crítico ser funcional que, por si só, já é elevada. Para o local dador de *splicing* autêntico (marcado na figura com um triângulo azul contendo o número 2) verifica-se que na sequência referência, este local tem um *score* de +37,75%, um valor ligeiramente superior ao *score* do local crítico de *splicing* (+30,13%), que permite à maquinaria responsável pelo *splicing* discriminar entre estes dois locais. Por sua vez, na sequência variante o local dador de *splicing* autêntico deixa de existir (na figura registado como “<no GT>”), o que leva a que a maquinaria de *splicing* tenha de utilizar o local dador de *splicing* crítico, localizado a montante. Ao contrário da IVS-I-6 (T>C), em que apenas havia uma redução do *score* do local dador de *splicing* autêntico que permitia que continuasse a haver produção de cadeias β -globínicas normais, embora numa quantidade mais reduzida, na IVS-I-1 (G>A) o local dador de *splicing* autêntico é completamente abolido e por isso não há nenhuma produção de cadeias β -globínicas normais (mutação β^0) e levando a que o *varSEAK* inclua esta alteração na classe 5, a classe mais alta de variantes com efeito ao nível do *splicing*.

Desta forma, a eliminação do local dador de *splicing* autêntico derivado da variante *HBB*:c.92+1G>A fará com que provavelmente a maquinaria de *splicing* use o local dador de *splicing* crítico existente 16 nucleótidos a montante, levando à exclusão da parte 3' do primeiro exão do gene *HBB*, que consequentemente poderá conduzir à alteração da grelha de leitura. Para testar esta hipótese, recorreu-se à ferramenta bioinformática *Translate Tool* da *ExPASy* (<https://web.expasy.org/translate/>), de forma a traduzir as sequências de DNA codificante original e mutado nas respetivas sequências de aminoácidos (Figura 4.19).

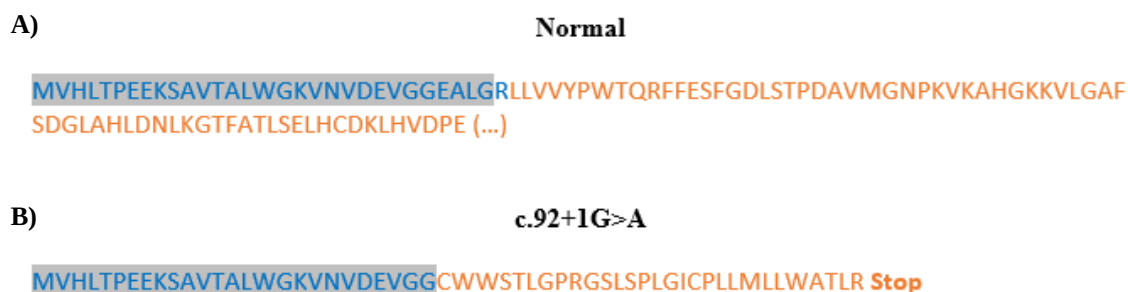


Figura 4.19. Sequência de aminoácidos correspondente aos exões 1 e 2 do gene da β -globina normal e com a variante c.92+1G>A. A) Sequência parcial de aminoácidos da proteína da β -globina normal. **B)** Sequência de aminoácidos que é gerada pelo uso do local dador de *splicing* crítico no mecanismo de remoção do primeiro intrão do gene *HBB*, na presença da variante c.92+1G>A, revelando a introdução de um códon de terminação prematuro. A sequência de aminoácidos correspondente ao exão 1 encontra-se a azul e a sequência correspondente ao exão 2 está indicada a laranja. A sequência de aminoácidos que permanece inalterada está salientada a cinzento e, como se pode observar, corresponde apenas a parte do primeiro exão do gene *HBB*. **Stop** – códon de terminação.

Os resultados do *Translate Tool* para a IVS-I-1 (G>A) são semelhantes aos resultados obtidos para a IVS-I-6 (T>C) uma vez, que para a remoção de intrões, é utilizado em ambas o local dador de *splicing* crítico, localizado a montante do local dador de *splicing* autêntico, tendo como consequência a exclusão de 16 nucleótidos da região 3' do primeiro exão, que vai alterar a grelha de leitura, resultando na introdução de um códon de terminação prematuro (TGA). Como referido, a introdução deste códon de terminação prematuro conduz à produção de um transcrito de β -globina anómalo e a proteína resultante

será mais curta, uma vez que o codão de terminação anula tanto a sequência do segundo exão como do terceiro.

Em suma, através dos estudos *in silico*, percebemos que na presença da variante *HBB*:c.92+1G>A, devido à abolição do dinucleótido conservado GT não há produção de mRNA normal e apenas são originados transcritos anómalos, que deverão ser direcionados para o mecanismo NMD para serem eliminados. Deste modo, e de acordo com a base de dados *HbVar*, quando a variante *HBB*:c.92+1G>A se encontra em heterozigotia há uma redução acentuada da síntese de cadeias β -globínicas, quando se encontra em homozigotia, a síntese de cadeias de β -globina nestes indivíduos é mesmo inexistente, conduzindo a um fenótipo grave de β -talassémia que resulta numa anemia severa e dependente de transfusões. Neste estudo, a variante foi detetada por NGS em três indivíduos, sempre em heterozigotia (Tabela 4.10), e foi posteriormente confirmada por sequenciação de Sanger num deles (Figura 4.20).

Tabela 4.10. Resultados do NGS para a variante c.92+1G>A no gene *HBB*

ID	Ref	Alt	Número de participante	Genótipo	Cobertura alélica (Ref;Alt)	Cobertura Total
c.92+1G>A	G	A	PTI 88	0/1	398;294	692
			PTI 89	0/1	1132;651	1783
			PTI 188	0/1	748;645	1396

ID – Identificação da variante ao nível do gene; **Ref** – Alelo referência; **Alt** – Alelo alterado; **0/1** – heterozigotia.

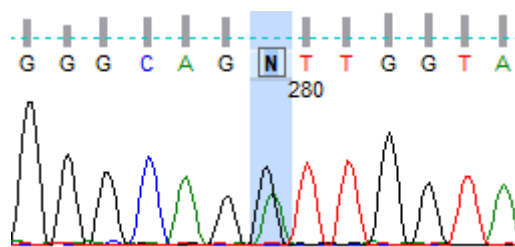


Figura 4.20. Verificação por sequenciação de Sanger da variante *HBB*:c.92+1G>A. Eletroferograma parcial da sequenciação de Sanger onde se pode verificar na barra sombreada a azul a heterozigotia G/A.

A deteção desta mutação no nosso estudo era um resultado esperado, pois a IVS-I-1 (G>A) é outra das quatro mutações β -talassémicas mais comuns na população portuguesa, responsável pelo fenótipo β^0 (79). Os três indivíduos identificados por NGS têm todos origem portuguesa e, embora apenas um deles apresente níveis de hemoglobina *borderline*, os três possuem microcitose e hipocromia (parâmetros hematológicos descritos na Tabela 7.11 dos Anexos), derivadas da presença desta variante em heterozigotia (genótipo β/β^0).

4.4.7. Cd15 (G>A), (*HBB*:c.48G>A)

A variante *HBB*:c.48G>A, também conhecida por Cd15 (G>A), é uma variante β -talassémica com especial prevalência em Portugal, onde possui uma frequência relativa de cerca de 11%, de acordo com as bases de dados *Ithant* e *HbVar*. Esta variante deriva da substituição de uma guanina por uma adenina, no terceiro nucleótido do codão 15 da β -globina, sendo assim responsável pela conversão do aminoácido triptofano (TGG) num codão de terminação prematuro (TGA). Ao introduzir um codão de terminação prematuro, a tradução do mRNA da β -globina é interrompida, o que pode resultar numa proteína truncada ou num mRNA sujeito ao mecanismo NMD. No entanto, o codão de terminação prematuro está localizado no codão 15, o que significa que este se encontra próximo da extremidade 5' do primeiro exão do gene *HBB*, permitindo-lhe inibir o mecanismo NMD, e fazendo com que este transcrito produza uma proteína truncada, que resulta no fenótipo β^0 , descrito para esta variante (90). Com base nestas informações, a variante *HBB*:c.48G>A está classificada na base de dados *ClinVar* como sendo patogénica.

De modo a confirmar a informação indicada no *ClinVar* e por se tratar de uma alteração *nonsense*, recorreu-se ao estudo *in silico* através do *software* SIFT, para verificar o impacto desta variante. Como esperado, o *software* classificou a variante *HBB*:c.48G>A como sendo prejudicial (*damaging*), indo ao encontro da informação descrita no *ClinVar*.

A mutação Cd15 (G>A) foi encontrada em dois indivíduos (Tabela 4.11), e foi confirmada por sequenciação de Sanger num deles (Figura 4.21).

Tabela 4.11. Resultados do NGS para a variante c.48G>A no gene *HBB*

ID	Ref	Alt	Número de participante	Genótipo	Cobertura alélica (Ref;Alt)	Cobertura Total
c.48G>A; p.Trp16Ter	G	A	PTI 30	0/1	229;211	442
			PTI 150	0/1	884;478	1365

ID – Identificação da variante ao nível do gene, seguida da identificação ao nível da proteína; **Ref** – Alelo referência; **Alt** – Alelo alterado; **0/1** – heterozigotia.

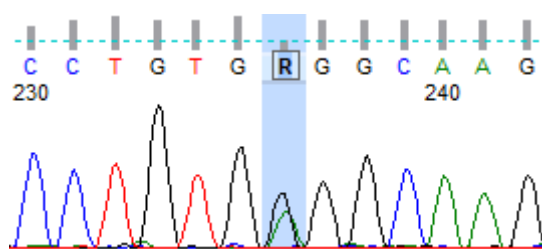


Figura 4.21. Verificação por sequenciação de Sanger da variante *HBB*:c.48 G>A. Eletroferograma parcial da sequenciação de Sanger onde se pode verificar na barra sombreada a azul a heterozigotia G/A.

Embora a variante Cd15 (G>A) não pertença ao grupo conhecido das quatro mutações β -talassémicas principais, descritas na população portuguesa, é a quinta mutação β -talassémica mais comum em Portugal, com especial incidência na região litoral Centro (79, 91). De facto, um dos dois indivíduos

detetados com a variante reside na região Centro e o outro na Região Autónoma da Madeira. A presença da variante *HBB*:c.48 G>A leva a que ambos os indivíduos possuam uma β -talassémia *minor* (β/β^0), responsável pelo seu quadro clínico de anemia microcítica e hipocrômica (parâmetros hematológicos descritos na Tabela 7.11 dos Anexos).

4.4.8. Cd 6 (-A), (*HBB*:c.20delA)

A variante *HBB*:c.20delA, também conhecida como Cd6 (-A) é uma variante com maior frequência relativa na Argélia ($\approx 13\%$) e em Marrocos ($\approx 10\%$), segundo a base de dados *IthaGenes*, estando geralmente associada a um fenótipo de β^0 -talassémia (92). Esta mutação está descrita em Portugal e tem uma prevalência maior na região Norte (cerca de 8%), do que no restante território (79, 93).

Esta variante surge da deleção de um único nucleótido, uma adenosina, no codão seis do gene *HBB*, que conduz à alteração da grelha de leitura e à substituição do aminoácido ácido glutâmico (GAG) por uma glicina (GGG). Consequentemente, é criado um codão de terminação prematuro (TGA) na posição 18 da nova grelha de leitura, prevendo-se que este novo codão de terminação interrompa o processo normal de tradução do mRNA, o que pode conduzir à formação de uma proteína truncada ou de um mRNA sujeito ao mecanismo NMD (94). Assim, a variante Cd6 (-A) leva à perda da função normal da proteína, conduzindo ao fenótipo β^0 e por isso é interpretada pela base de dados *ClinVar* como sendo patogénica.

De novo, para verificar o impacto da variante Cd6 (-A), recorreu-se ao estudo *in silico* através do *software* SIFT, visto tratar-se de uma deleção, e por isso não é possível utilizar o *software* PolyPhen-2. A variante *HBB*:c.20delA foi classificada pelo SIFT como sendo prejudicial (*damaging*) com um *score* de 0,858, confirmando a informação descrita no *ClinVar*.

A variante Cd6 (-A) foi detetada por NGS em heterozigotia apenas num indivíduo (Tabela 4.12) e foi confirmada por sequenciação de Sanger (Figura 4.22).

Tabela 4.12. Resultados do NGS para a variante c.20delA no gene *HBB*

ID	Ref	Alt	Número de participante	Genótipo	Cobertura alélica (Ref;Alt)	Cobertura Total
c.20delA; p.Glu7GlyfsTer13	GA	G	PTI 18	0/1	1040;614	1654

ID – Identificação da variante ao nível do gene, seguida da identificação ao nível da proteína; **Ref** – Alelo referência; **Alt** – Alelo alterado; **0/1** – heterozigotia.

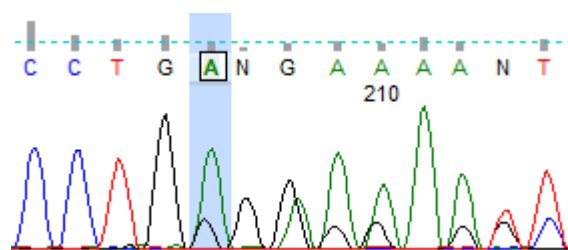


Figura 4.22. Verificação por sequenciação de Sanger da variante *HBB*:c.20delA. Eletroferograma parcial da sequenciação de Sanger onde se pode verificar barra sombreada a azul o início da heterozigotia GA/G- e é possível observar também a consequente alteração da grelha de leitura.

O portador da mutação Cd6 (-A) é um indivíduo do sexo masculino com 54 anos e com anemia microcítica e hipocrômica (hemoglobina= 12,1 g/dL, VGM= 61,8 fL e HGM= 19,8 pg; restantes parâmetros hematológicos na Tabela 7.11 dos Anexos). A variante Cd6 (-A) é a sexta mutação β -talassêmica mais comum em Portugal e por isso era de esperar que a detetássemos neste estudo. Além disso, o indivíduo em causa reside na região Norte de Portugal que, de acordo com a literatura, é a região do país onde a mutação Cd6 (-A) atinge uma maior prevalência (79, 93). Assim, concluímos que este participante possui uma β -talassémia *minor* (β/β^0), responsável pelo seu quadro hematológico.

4.4.9. Hb S (*HBB*:c.20A>T)

A variante *HBB*:c.20A>T, mais conhecida como Hb S, é o genótipo mais prevalente associado à doença falciforme. Como referido anteriormente nesta dissertação, o termo “doença falciforme” por vezes é erradamente igualado a “anemia falciforme” ou mesmo a “traço falciforme”, contudo, é importante definirmos e termos presente as diferenças entre eles. A doença falciforme diz respeito a todo o grupo de doenças que resultam da herança da hemoglobina S, enquanto a anemia falciforme se refere especificamente à homozigotia para o alelo β^S (Hb SS) e, por sua vez, o traço falciforme remete à presença do alelo β^S em heterozigotia (Hb AS). A Hb S estava originalmente distribuída por toda a África Subsaariana, Médio Oriente, Mediterrâneo e Índia, mas tem atualmente uma distribuição global devido em grande parte ao comércio de escravos e, mais recentemente, às migrações populacionais (24).

A variante *missense* *HBB*:c.20A>T surge de uma alteração de nucleótidos (A>T) no sexto codão do gene da β -globina, que tem como consequência a substituição do aminoácido ácido glutâmico (GAG) por uma valina (GTG), na proteína (24). O ácido glutâmico é um aminoácido polar e, por isso, pode ser substituído por outros aminoácidos polares, que não é o caso da valina. A valina é um aminoácido não-polar que se encontra preferencialmente no núcleo hidrofóbico das proteínas e assim, devido a estas diferenças a alteração entre aminoácidos derivada da variante *HBB*:c.20A>T deverá ter impacto na estrutura secundária da proteína (71). Deste modo, e como era de esperar, a variante *HBB*:c.20A>T está descrita na base de dados *ClinVar* como sendo patogénica.

Com o intuito de esclarecer melhor o impacto desta variante ao nível da proteína, recorremos aos estudos *in silico* através dos *softwares* *PolyPhen-2* e *SIFT*. O *PolyPhen-2* prevê a variante como sendo uma mutação benigna (Figura 4.23) com um *score* de 0,001 e o mesmo faz o *software* *SIFT*, que prevê a variante também como benigna, com um *score* de 0,06.

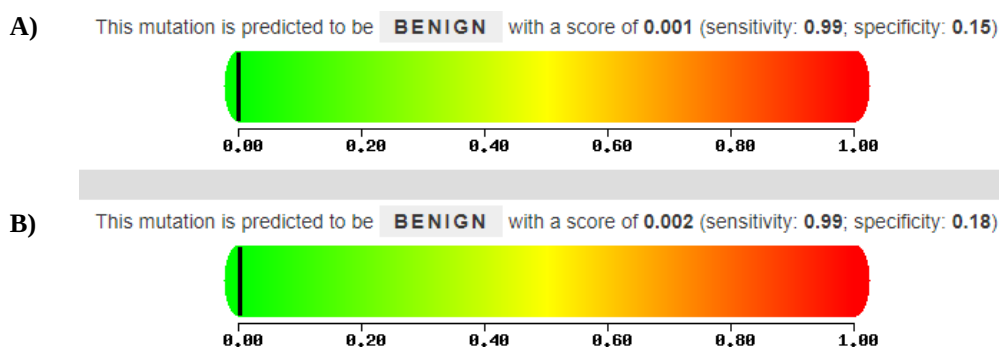


Figura 4.23. Previsão do efeito patogénico da variante p.Glu7Val na proteína da β -globina através do *software* *PolyPhen-2*. **A)** Dados do modelo *HumDiv* para a alteração, que a prevê como sendo benigna. **B)** Dados do modelo *HumVar* para a alteração que a prevê como sendo benigna.

A análise do perfil de conservação obtido através do alinhamento múltiplo (Figura 4.24) revelou que o resíduo de ácido glutâmico presente na sexta posição do gene da β -globina é moderadamente conservado entre espécies próximas de mamíferos. Para algumas espécies, o perfil mostra que o resíduo de ácido glutâmico pode ser substituído por ácido aspártico. Isto deve-se ao facto de o ácido aspártico ser também um aminoácido polar, como referido anteriormente nesta dissertação, ao contrário da valina, com características hidrofóbicas. Tendo em conta este alinhamento múltiplo, faz sentido que, no que diz respeito à sua patogenicidade, esta alteração seja patogénica, embora não seja esse o resultado dos estudos *in silico*.



Figura 4.24. Perfil de conservação evolutiva do aminoácido p.Glu7 na proteína da β -globina. Alinhamento múltiplo de sequências de aminoácidos de proteínas homólogas de 15 espécies de mamíferos relacionados, através do software PolyPhen-2. O padrão de substituições de aminoácidos está ilustrado dentro da caixa preta e destacado com uma seta, sendo que neste, a letra E corresponde ao ácido glutâmico e a letra D ao ácido aspártico.

Quando desoxigenada, a Hb S é extremamente insolúvel e forma polímeros de hemoglobina nos glóbulos vermelhos que, conseqüentemente, deformam a sua estrutura e conferem-lhes a forma falciforme, característica da Hb S. A alteração conformacional que a hemoglobina sofre durante a sua desoxigenação é conferida pela substituição do ácido glutâmico hidrofílico, no sexto codão da β -globina, pela valina hidrofóbica. Esta substituição permite que se estabeleçam interações hidrofóbicas com outros resíduos de valina hidrofóbicos presentes nas cadeias de β -globina de outras moléculas de Hb S desoxigenadas. Devido a estas interações, formam-se polímeros alongados que se vão agrupando e preenchendo o glóbulo vermelho, alterando a sua estrutura e a sua flexibilidade e promovendo a desidratação celular (19). Assim, em desoxigenação, os glóbulos vermelhos destes doentes tomam a forma de foice (os drepanócitos) sendo que uma certa percentagem dessas células fica irreversivelmente deformada.

Sabemos que as causas genéticas da doença falciforme incluem a homozigotia para a variante *HBB:c.20A>T* (Hb SS), geralmente conhecida como anemia falciforme e ainda heterozigotia composta entre a variante *HBB:c.20A>T* e outras variantes estruturais de β -globina (como a Hb C) ou β^+ -talassémias (95). Contudo, neste estudo a variante *HBB:c.20A>T* foi detetada por NGS apenas num indivíduo e em heterozigotia (Tabela 4.13), tendo sido este resultado posteriormente confirmado por sequenciação de Sanger (Figura 4.25).

Tabela 4.13. Resultados do NGS para a variante c.20A>T no gene *HBB*

ID	Ref	Alt	Número de participante	Genótipo	Cobertura alélica (Ref;Alt)	Cobertura Total
c.20A>T; p.Glu7Val	A	T	PTI 47	0/1	328;310	99

ID – Identificação da variante ao nível do gene, seguida da identificação ao nível da proteína; **Ref** – Alelo referência; **Alt** – Alelo alterado; **0/1** – heterozigotia.

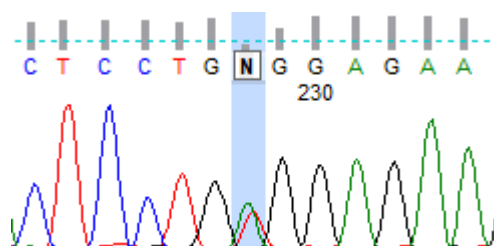


Figura 4.25. Verificação por sequenciação de Sanger da variante *HBB*:c.20A>T. Eletroferograma parcial da sequenciação de Sanger onde se pode verificar na barra sombreada a azul a heterozigotia A/T.

O indivíduo em questão (PTI47) tem 63 anos, é do sexo masculino com origem angolana e embora seja anémico, não possui microcitose e apresenta uma hipocromia bastante ligeira (hemoglobina= 11,1 g/dL, VGM= 83,0 fL, HGM= 25,9 pg e restantes parâmetros hematológicos na Tabela 7.11 dos Anexos).

Tendo em conta que, segundo os resultados do NGS, este indivíduo não apresenta qualquer outra variante estrutural de β -globina nem β -talassémia, a heterozigotia da variante *HBB*:c.20A>T resulta no traço falciforme (Hb AS), em que os glóbulos vermelhos apresentam uma mistura de Hb S e de moléculas normais de Hb A. Sabe-se que a Hb S confere uma proteção relativa contra a malária, embora o mecanismo exato dessa proteção ainda não seja totalmente compreendido. Este benefício permitiu que a Hb S atingisse altas frequências populacionais, nas regiões endémicas de malária, concedendo uma vantagem de sobrevivência aos indivíduos com traço falciforme em regiões como a África Subsaariana, de onde é proveniente o indivíduo portador da variante *HBB*:c.20A>T identificado neste estudo (95). O traço falciforme é uma condição relativamente comum e clinicamente benigna em que os seus portadores são geralmente assintomáticos, sem anormalidades físicas, com esperança média de vida semelhante à da população normal e com parâmetros hematológicos normais (96). Tendo isto em conta, a anemia hipocrômica ligeira detetada neste indivíduo não poderia ser explicada pela presença da Hb S.

Contudo, este indivíduo, além de ser portador do traço falciforme é também heterozigótico para a deleção $-\alpha^{3,7}$. Encontra-se descrito na literatura que aproximadamente 30% dos indivíduos com doença falciforme e de ascendência africana possuem também α -talassémia, geralmente causada pela deleção de 3,7 kb, comum na região africana, tal como verificado neste indivíduo (95). Fisiologicamente, a coexistência da α -talassémia influencia positivamente o fenótipo da doença falciforme pois a diminuição das cadeias de α -globina conduz também à diminuição da concentração da hemoglobina S nos glóbulos vermelhos (refletida na presença de anemia e hipocromia leve no indivíduo em causa), diminuindo a tendência da Hb S para polimerizar. Devido a isto, ocorre menor hemólise que se traduz numa normalização do hematócrito. Os efeitos clínicos da co-herança da α -talassémia são variáveis, mas geralmente benéficos para os pacientes, com risco reduzido de complicações hemolíticas, mas um risco aumentado de complicações vaso-oclusivas, associadas ao aumento do hematócrito e da viscosidade sanguínea (24, 95).

Tendo todas estas informações em conta, a origem africana do indivíduo em questão explica o facto de este ser portador da variante Hb S em co-herança com a α -talassémia, sendo esta coexistência responsável pelo seu quadro hematológico de anemia hipocrômica.

Relativamente à diferença de resultados observada quanto à patogenicidade entre a base de dados *ClinVar* e os estudos *in silico*, esta disparidade pode ser explicada pelas diferenças clínicas acima referidas que se observam quando a variante se encontra em homozigotia ou heterozigotia. Efetivamente, a Hb S é considerada em inúmeros estudos como uma variante patogénica devido a todas as complicações clínicas verificadas na anemia falciforme.

4.4.10. Hb C (*HBB*:c.19G>A)

A Hb C é uma variante estrutural da hemoglobina que deriva de uma mutação no gene da β -globina, localizada no mesmo codão que a Hb S. A Hb C foi descrita pela primeira vez em 1950, por Harvey Itano e James Neel e pensa-se que se tenha originado na África Ocidental, a oeste do Rio Níger, sendo a segunda variante estrutural da hemoglobina mais comum em pessoas de ascendência africana (97–99).

A variante *missense* *HBB*:c.19G>A deriva da substituição de uma guanina por uma adenina na primeira base do sexto codão do gene da β -globina, que leva à alteração do aminoácido ácido glutâmico (GAG) para o aminoácido lisina (AAG), enquanto na Hb S é a valina que substitui o ácido glutâmico, como referido anteriormente. Embora tanto o ácido glutâmico como a lisina sejam aminoácidos polares, a lisina é um aminoácido carregado positivamente, ao contrário do ácido glutâmico, que tem carga negativa e por isso esta substituição poderá ser nefasta. A alteração do ácido glutâmico pela lisina resulta numa molécula de hemoglobina com carga altamente positiva, diminuição da solubilidade e tendência à cristalização, resultantes das mudanças conformacionais que acompanham a mutação Hb C (71, 100). Estas alterações estão de acordo com os resultados de patogenicidade do *ClinVar*, que classifica a Hb C como uma variante patogénica.

Também para esta variante foram realizados estudos *in silico* para obter mais informações sobre o impacto desta variante *missense* *HBB*:c.19G>A ao nível da proteína, utilizando os *softwares* *PolyPhen-2* e *SIFT*. Tanto o *PolyPhen-2* como o *SIFT* preveem a variante como sendo uma mutação benigna, tendo o *PolyPhen-2* atribuído um *score* de 0,005 (Figura 4.26) e o *SIFT* um *score* de 0,31.

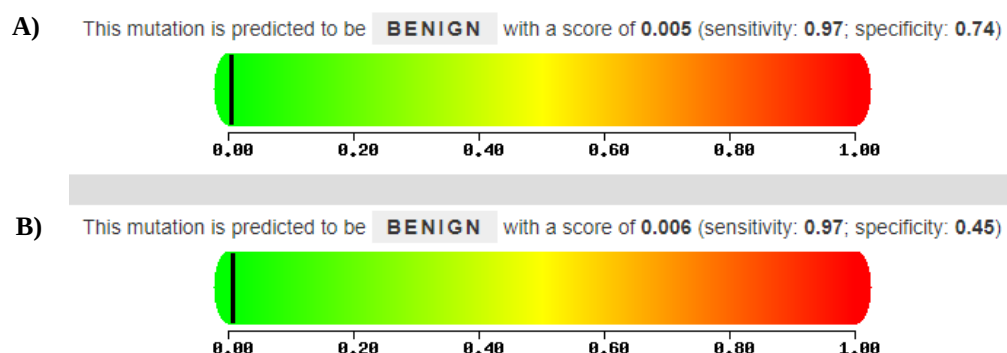


Figura 4.26. Previsão do efeito patogénico da variante p.Glu7Lys na proteína da β -globina através do *software* *PolyPhen-2*. A) Dados do modelo *HumDiv* para a alteração, que a prevê como sendo benigna. B) Dados do modelo *HumVar* para a alteração que a prevê como sendo benigna.

Tendo em conta que o resíduo de ácido glutâmico afetado tanto na Hb C como na Hb S é o mesmo, os resultados do alinhamento múltiplo são semelhantes para ambas (Figura 4.27) e revelam que o resíduo de ácido glutâmico que se encontra na sétima posição do gene da β -globina é moderadamente conservado entre espécies próximas de mamíferos. Tal como já foi referido anteriormente na análise da Hb S, através do alinhamento múltiplo podemos perceber que o resíduo de ácido glutâmico pode ser substituído principalmente por ácido aspártico, pois este também é um aminoácido polar com carga negativa, ao contrário da lisina, com carga positiva.



Figura 4.27. Perfil de conservação evolutiva do aminoácido p.Glu7 na proteína da β -globina. Alinhamento múltiplo de sequências de aminoácidos de proteínas homólogas de 15 espécies de mamíferos relacionados, através do software PolyPhen-2. O padrão de substituições de aminoácidos está ilustrado dentro da caixa preta e destacado com uma seta, sendo que neste, a letra E corresponde ao ácido glutâmico e a letra D ao ácido aspártico.

Como já explicado nesta dissertação, quando o oxigénio se liga ao ião ferro do grupo heme para ser transportado, forma-se o complexo oxi-hemoglobina. A oxi-hemoglobina C é propensa a formar cristais (Figura 4.28), mas ao ocorrer a desoxigenação da hemoglobina estes cristais dissolvem-se e não ocorre obstrução dos capilares. A cristalização resulta em danos na membrana dos glóbulos vermelhos e consequente redução da sua sobrevivência, mas por terem a capacidade de se dissolverem, a formação de cristais não conduz a alterações morfológicas agudas na microcirculação como acontece na Hb S (98, 101).

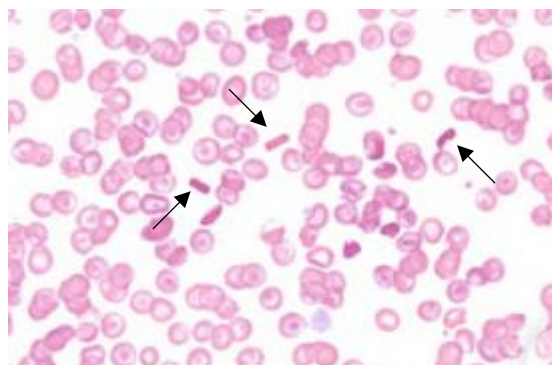


Figura 4.28. Morfologia dos glóbulos vermelhos com Hb C. Imagem de microscopia de um esfregaço periférico que demonstra a morfologia característica dos glóbulos vermelhos induzida pela formação de cristais de oxi-hemoglobina de Hb C indicados pelas setas a preto. Adaptado de (101).

A doença da hemoglobina C descreve o estado homozigótico no qual o indivíduo possui os dois alelos β^C . Esta homozigotia para Hb C conduz a uma anemia hemolítica crónica moderada, mas os indivíduos apresentam uma concentração normal de hemoglobina ou são ligeiramente ou moderadamente anémicos. Por sua vez, o traço Hb C diz respeito à heterozigotia para a Hb C, ou seja, apenas um alelo β^C , sendo que esta condição não conduz a consequências clínicas, e embora os indivíduos sejam considerados assintomáticos, é comum a deteção de microcitose (98).

Como se pode ver pela Tabela 4.14, neste estudo a variante *HBB*:c.19G>A foi detetada em heterozigotia por NGS num indivíduo português de 42 anos, do sexo masculino e que apresentava uma microcitose ligeira (VGM= 75,8 fL e restantes parâmetros hematológicos na Tabela 7.11 dos Anexos), tendo sido a presença desta variante confirmada posteriormente por sequenciação de Sanger (Figura 4.29).

Tabela 4.14. Resultados do NGS para a variante c.19G>A no gene *HBB*

ID	Ref	Alt	Número de participante	Genótipo	Cobertura alélica (Ref;Alt)	Cobertura Total
c.19G>A; p.Glu7Lys	G	A	PTI 184	0/1	555;521	1076

ID – Identificação da variante ao nível do gene, seguida da identificação ao nível da proteína; **Ref** – Alelo referência; **Alt** – Alelo alterado; **0/1** – heterozigotia.

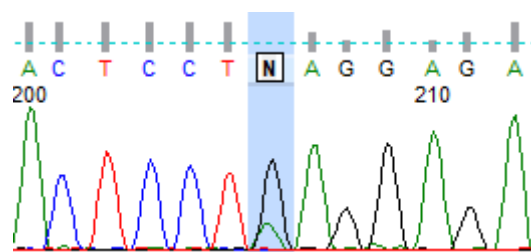


Figura 4.29. Verificação por sequenciação de Sanger da variante *HBB*:c.19G>A. Eletroferograma parcial da sequenciação de Sanger onde se pode verificar na barra sombreada a azul a heterozigotia G/A.

Embora o *ClinVar* descreva a Hb C como sendo patogénica, nos estudos *in silico* e segundo a literatura, no estado heterozigótico a alteração não parece trazer consequências clínicas graves. Contudo a deteção da Hb C torna-se importante para o aconselhamento genético deste indivíduo uma vez que, se o seu cônjuge for portador de Hb C, Hb S, ou de traço talassémico estaremos perante um casal em risco de ter descendência gravemente afetada. A Hb SC pertence ao grupo das doenças falciformes e caracteriza-se pela presença dos alelos β^S e β^C , isto é, nenhum dos alelos da β -globina é normal e, portanto, não existe Hb A. Este estado heterozigótico composto leva a um distúrbio falciforme, mas com um fenótipo menos grave do que a Hb SS, pois o grau de hemólise é menor, e a esperança média de vida também é consideravelmente melhor do que na Hb SS (98).

Podemos concluir então que a microcitose detetada neste indivíduo pode ser explicada pela presença da variante *HBB*:c.19G>A e, embora o indivíduo seja clinicamente estável, seria importante ser encaminhado para o aconselhamento genético. Além disso, sendo o indivíduo português e sendo a Hb

C mais prevalente na região africana, seria interessante explorar o histórico familiar do indivíduo para perceber a origem do alelo β^C .

4.4.11. Prevalência da β -talassémia e de variantes de hemoglobina

A análise por NGS revelou a presença de 10 alterações com interesse clínico na amostra em estudo que foram descritas nesta dissertação, sendo que sete dessas alterações são variantes β -talassêmicas [Cd41/42 (-CTTT), Cd39 (C>T), IVS-I-110 (G>A), IVS-I-6 (T>C), IVS-I-1 (G>A), Cd15 (G>A) e Cd6 (-A)] e três das alterações detetadas são classificadas como variantes de hemoglobina (Hb D-Portugal, Hb S e Hb C). Os genótipos e as suas respectivas frequências encontram-se detalhados na Tabela 4.15.

Tabela 4.15. Genótipos detetados no gene *HBB* e frequências relativas para os 204 indivíduos analisados

Genótipos	Frequência absoluta	Frequência relativa
<i>Genótipos heterozigóticos</i>		
Hb D-Portugal	1	0,5%
Cd41/42 (-CTTT)	1	0,5%
Cd39 (C>T)	8	3,9%
IVS-I-110 (G>A)	3	1,5%
IVS-I-6 (T>C)	4	1,9%
IVS-I-1 (G>A)	3	1,5%
Cd15 (G>A)	2	0,9%
Cd6 (-A)	1	0,5%
Hb S	1	0,5%
Hb C	1	0,5%
<i>Genótipos homozigóticos</i>		
Sem alterações no gene <i>HBB</i>	179	87,8%

Assim, a β -talassémia foi observada, em heterozigotia, em 22 indivíduos dos 204 com microcitose e/ou hipocromia analisados molecularmente, o que corresponde a uma prevalência de β -talassémia de 10,8% nestes indivíduos previamente selecionados pelo seu perfil hematológico. Atendendo à população de origem, os 4808 participantes do estudo INSEF, este resultado traduz-se numa prevalência de 0,5% de indivíduos portadores de β -talassémia. Os resultados obtidos neste estudo são semelhantes ao reportados no estudo realizado por Martins *et al.*, em 1993, cujos resultados revelam uma frequência de 0,45% de β -talassémia em Portugal continental (32).

Relativamente às variantes de hemoglobina, estas foram detetadas em três indivíduos dos 204 analisados, o que se reflete numa prevalência de variantes de hemoglobina de 1,5% nestes indivíduos previamente selecionados pelo seu fenótipo hematológico. Atendendo à população de origem, os 4808 participantes do estudo INSEF, o nosso resultado traduzir-se-ia numa prevalência de 0,1% de portadores de variantes de hemoglobina. No entanto, este nosso estudo não tinha como objetivo a deteção de variantes de hemoglobina na população uma vez que, se assim fosse, não teríamos efetuado uma primeira seleção pelo fenótipo de microcitose e hipocromia que é característica da talassémia e não das

variantes de hemoglobina. Estas variantes de hemoglobina derivam principalmente de mutações *missense* na região codificante dos genes globínicos que não afetam o nível de expressão do gene em causa, mas que originam uma substituição de aminoácidos que poderá alterar a carga superficial da molécula de hemoglobina ou atribuir-lhe propriedades inexistentes na hemoglobina normal (por exemplo a capacidade de falciformizar, no caso da Hb S). Assim, na presença de uma variante de hemoglobina os parâmetros hematimétricos, nomeadamente o VGM e HGM, apresentam-se, geralmente, nos níveis normais. Portanto não é adequado utilizar os nossos resultados para efetuar o cálculo da prevalência de variantes de hemoglobina na população residente em Portugal.

4.4.12. Associação da β -talassémia e de variantes de hemoglobina com os parâmetros hematológicos

Tal como para a α -talassémia, pretendeu-se estudar uma possível associação entre os parâmetros hematológicos, obtidos no estudo INSEF, com a β -talassémia e as variantes de hemoglobina. Para isso utilizou-se o teste estatístico T de duas amostras independentes para os parâmetros com distribuição normal, e o teste de Mann-Whitney para os parâmetros com distribuição não normal (Tabela 4.16).

Tabela 4.16. Associação entre os parâmetros hematológicos estudados e a ausência de talassémia ou presença de β -talassémia

Parâmetros	Indivíduos com microcitose e/ou hipocromia mas sem talassémia					Indivíduos com microcitose e/ou hipocromia e com β -talassémia					<i>p-value</i>
	Média	DP	Mediana	Min	Max	Média	DP	Mediana	Min	Max	
Eritrócitos ($\times 10^{12}/L$)											
Homens	5,46	0,626	5,375	4,45	6,70	6,08	0,409	6,105	5,26	6,60	0,013^a
Mulheres	4,66	0,459	4,695	3,75	7,30	5,64	0,540	5,430	4,80	6,40	<0,001^b
Hemoglobina (g/dL)											
Homens	13,4	2,27	14,00	9,8	16,7	13,2	1,25	13,25	11,4	14,8	0,838 ^b
Mulheres	11,7	1,35	11,85	6,5	15,3	11,3	1,15	11,40	9,0	12,8	0,268 ^b
VGM (fL)	78,3	5,18	78,90	57,2	85,9	66,3	5,95	66,60	56,9	82,3	<0,001^b
HGM (pg)	24,9	2,02	25,60	16,2	28,1	20,8	2,08	20,70	17,2	25,4	<0,001^b
CHGM (g/dL)	31,9	1,07	31,90	28,3	35,5	31,4	1,02	31,45	28,7	32,6	0,036^b
Hematócrito (%)	38,6	4,78	38,20	23,0	52,0	38,7	4,46	37,70	31,4	47,1	0,797 ^a
RDW (%)	15,8	1,89	15,60	12,3	22,3	15,9	0,98	15,95	14,1	18,0	0,377 ^b
Leucócitos ($\times 10^9/L$)	7,55	2,126	7,185	4,03	16,40	6,84	1,914	7,050	4,40	10,80	0,144 ^b
Plaquetas ($\times 10^9/L$)	275	65,2	269,0	154	433	227	51,9	224,5	132	312	0,002^b

DP – desvio padrão; **Min** – mínimo; **Max** – máximo; **VGM** – volume globular médio; **HGM** – hemoglobina globular média; **CHGM** – concentração da hemoglobina globular média; **RDW** – *red cell distribution width*. ^a – Teste T; ^b – Teste de Mann-Witney. Consideram-se os valores estatisticamente significativos quando *p-value* < 0,05 (os valores estão salientados a negrito).

Relativamente à β -talassémia, não se encontraram diferenças estatisticamente significativas (p -value $\geq 0,05$) para os níveis de hemoglobina (quer nos homens, quer nas mulheres), hematócrito, RDW e leucócitos entre os indivíduos sem talassémia e portadores de β -talassémia, mas existem diferenças estatisticamente significativas (p -value $< 0,05$) nos níveis de eritrócitos (quer nos homens, quer nas mulheres), VGM, HGM, CHGM e plaquetas entre os indivíduos sem talassémia e portadores de β -talassémia. Também foi possível constatar que comparativamente aos indivíduos sem talassémia, a β -talassémia está associada a níveis mais elevados de eritrócitos (para homens e mulheres) e a níveis mais reduzidos de VGM, HGM, CHGM e plaquetas.

Os parâmetros hematológicos do traço β -talassémico são muito característicos e estão bem descritos na literatura. Sabe-se através da literatura existente que os níveis de VGM e HGM em indivíduos β -talassémicos encontram-se baixos e por isso os resultados que obtivemos neste estudo para ambos os parâmetros já eram de esperar. Esse conhecimento prévio esteve na origem da seleção dos 204 indivíduos de entre todos os participantes no projeto INSEF. Inclusivamente, a conjugação dos parâmetros hematimétricos diminuídos (VGM < 80 fL e HGM < 27 pg) com níveis aumentados de hemoglobina A_2 ($> 3,5\%$), medida por técnicas bioquímicas fiáveis (como o HPLC) permitem um diagnóstico bioquímico de portador de β -talassémia.

Também está descrito na literatura que os portadores de β -talassémia apresentam níveis de CHGM normais ou reduzidos, traduzindo uma diminuição na concentração da hemoglobina dentro do eritrócito. Este facto poderá explicar a redução dos níveis de CHGM que verificámos nos indivíduos β -talassémicos, comparativamente aos indivíduos sem talassémia (102).

Relativamente à concentração de eritrócitos por μ L de sangue, em indivíduos β -talassémicos esta concentração pode ser normal, mas muitas vezes encontra-se aumentada. Na β -talassémia, a redução da síntese das cadeias β -globínicas e consequente redução da síntese de hemoglobina, desencadeia um processo de compensação pelo organismo aumentando a taxa de eritropoiese. Assim, a quantidade de eritrócitos em indivíduos β -talassémicos é aumentada, traduzindo-se em maiores níveis de eritrócitos no hemograma, embora estes sejam mais pequenos e com menos hemoglobina do que o normal. O aumento de eritrócitos (eritrocitose) é comum aquando da β -talassémia *minor* ou traço β -talassémico e não está associada a um aumento da concentração de hemoglobina em si (102–104).

A contagem de plaquetas pode estar normal ou aumentada. Geralmente, em indivíduos com β -talassémia *major*, quando a doença avança na sua gravidade com esplenomegalia acentuada, a contagem de plaquetas diminui. De acordo com isto, a diminuição dos níveis de plaquetas nos indivíduos β -talassémicos observada por nós só pode ser explicada caso grande parte dos indivíduos se encontre na situação de doença avançada (102).

Relativamente ao estudo de associação entre as variantes de hemoglobina e os parâmetros hematológicos, não se encontraram diferenças estatisticamente significativas ($p \geq 0,05$) para nenhum dos parâmetros hematológicos analisados (Tabela 4.17). Este resultado já era de esperar pois apenas identificámos três indivíduos portadores de uma variante de hemoglobina, pelo que este número não é suficiente para a análise estatística com significância. Para além disso, e como foi referido anteriormente, não é de esperar que os portadores de variantes de hemoglobina apresentem um fenótipo de microcitose e hipocromia.

Tabela 4.17. Associação entre os parâmetros hematológicos estudados e a ausência ou presença de variantes de hemoglobina

Parâmetros	Indivíduos com microcitose e/ou hipocromia mas sem variante de hemoglobina					Indivíduos com microcitose e/ou hipocromia e com variante de hemoglobina					p-value
	Média	DP	Mediana	Min	Max	Média	DP	Mediana	Min	Max	
Eritrócitos (x10¹²/L)											
Homens	5,71	0,539	5,715	4,45	6,70	5,06	0,851	4,900	4,30	5,98	0,058 ^a
Mulheres	4,78	0,510	4,770	3,75	7,30	-	-	-	-	-	-
Hemoglobina (g/dL)											
Homens	13,9	1,71	14,20	9,8	16,7	12,8	3,00	11,10	11,1	16,3	0,459 ^b
Mulheres	11,9	1,31	12,00	6,5	15,3	-	-	-	-	-	-
VGM (fL)	77,7	6,37	79,50	56,9	86,9	77,1	5,37	75,80	72,5	83,0	0,657 ^b
HGM (pg)	24,8	2,31	25,60	16,2	28,1	25,3	2,36	25,90	22,7	27,3	0,654 ^b
CHGM (g/dL)	31,9	1,01	31,90	28,3	35,5	32,9	2,77	31,40	31,2	36,1	0,972 ^b
Hematócrito (%)	38,6	4,73	38,20	23,0	52,0	38,9	5,66	35,70	35,5	45,4	0,923 ^a
RDW (%)	15,4	1,81	15,20	12,1	22,3	16,2	1,27	16,40	14,8	17,3	0,265 ^b
Leucócitos (x10⁹/L)	7,40	2,049	7,160	4,00	16,40	10,21	5,566	9,900	4,80	15,92	0,381 ^b
Plaquetas (x10⁹/L)	260	63,2	246,0	132	433	265	65,6	275,0	195	325	0,851 ^b

DP – desvio padrão; **Min** – mínimo; **Max** – máximo; **VGM** – volume globular médio; **HGM** – hemoglobina globular média; **CHGM** – concentração da hemoglobina globular média; **RDW** – *red cell distribution width*. ^a – Teste T; ^b – Teste de Mann-Witney. Consideram-se os valores estatisticamente significativos quando *p-value* < 0,05. Não foi possível analisar os parâmetros hemoglobina e eritrócitos nas mulheres para o grupo de indivíduos com microcitose e/ou hipocromia e com variante de hemoglobina, pois estes indivíduos são todos do sexo masculino.

4.5. Hemoglobinopatias em Portugal

Através do diagnóstico de α -talassémia, β -talassémia e variantes de hemoglobinas, nos 204 indivíduos com microcitose e/ou hipocromia selecionados a partir dos participantes no projeto INSEF, foram diagnosticados 77 indivíduos com hemoglobinopatias, que corresponde a uma prevalência de hemoglobinopatias de 37,8% nesta sub-população selecionada. Não podemos extrapolar a prevalência geral de hemoglobinopatias para os 4808 participantes do INSEF uma vez que sabemos que, no que diz respeito à α -talassémia silenciosa e à maior parte das variantes de hemoglobina não há associação com um fenótipo hematológico que as permita distinguir do geral da população. No entanto, se considerarmos apenas os portadores de β -talassémia, como todos têm microcitose e hipocromia, todos os casos existentes nos 4808 participantes no estudo INSEF foram detetados e foi diagnosticada a sua base molecular. Assim, consideramos que a prevalência obtida para os portadores de β -talassémia de 0,5% é realista para a população residente em Portugal, continental e ilhas, em 2015.

Também é de salientar que na população com fenótipo hematológico de microcitose e/ou hipocromia se observou que 76 indivíduos (37,3%) tinham como causa desse fenótipo uma talassémia (α ou β). Este

valor tão elevado deverá servir para alertar os clínicos, sobretudo da medicina geral e familiar, da importância de considerarem as doenças hereditárias ao tentarem diagnosticar os indivíduos que os consultam e que apresentam este tipo de sintomatologia.

4.5.1. Associação das hemoglobinopatias com as características demográficas e de saúde

Através dos dados demográficos e de saúde recolhidos no INSEF, foi possível construir uma tabela que fornece um panorama geral dessas várias características demográficas e de saúde da nossa amostra e relacioná-las com o estado portador ou não das diferentes hemoglobinopatias (Tabela 4.18).

Tabela 4.18. Características demográficas, de saúde e moleculares dos 204 participantes deste estudo

	Total	Hemoglobinopatia		α -talassémia		β -talassémia		Variante de hemoglobina	
		Não	Sim	Não	Sim	Não	Sim	Não	Sim
<i>Sexo</i>									
Feminino	157	113	44	124	32	145	12	157	0
Masculino	47	14	33	25	22	37	10	44	3
<i>Tipo de zona de residência</i>									
Rural	44	31	18	40	9	40	9	49	0
Urbano	155	96	59	109	45	142	13	152	3
<i>Região de residência</i>									
Norte	27	19	8	20	7	26	1	27	0
Centro	17	8	9	14	3	11	6	17	0
LVT	29	13	16	13	16	29	0	28	1
Alentejo	35	25	10	29	6	31	4	35	0
Algarve	35	24	11	30	5	30	5	24	1
RA Madeira	28	11	17	14	14	24	4	28	0
RA Açores	33	27	6	29	3	31	2	32	1
<i>Local de nascimento</i>									
África do Sul	1	1	0	1	0	1	0	1	0
Angola	8	6	2	6	2	8	0	7	1
Brasil	7	1	6	1	6	7	0	7	0
Bulgária	1	1	0	1	0	1	0	1	0
Canadá	1	1	0	1	0	1	0	1	0
Cabo Verde	6	1	5	1	5	6	0	6	0
China	1	0	1	1	0	0	1	1	0
França	2	0	2	0	2	2	0	2	0
Índia	1	0	1	0	1	1	0	1	0
Moldávia	1	1	0	1	0	1	0	1	0
Moçambique	1	1	0	1	0	1	0	1	0
Paquistão	1	0	1	0	1	1	0	1	0
Portugal	172	113	59	134	37	151	21	170	2
Síria	1	1	0	1	0	1	0	1	0
<i>Médico de família</i>									
Sim	160	103	57	118	42	145	15	158	2
Não	43	24	19	31	11	36	7	42	1
Não responde	1	0	1	0	1	1	0	1	0
<i>Última consulta com MF</i>									
Menos de 12 meses	135	82	53	95	40	122	13	133	2
12 meses ou mais	50	35	15	43	7	43	7	49	1
Nunca	15	8	7	9	5	13	2	15	0
Não sabe	4	2	2	2	2	4	0	4	0
<i>Perceção de saúde</i>									
Muito boa	7	6	1	6	1	7	0	7	0
Boa	66	40	26	52	13	54	12	65	1
Normal	113	68	45	77	36	104	9	111	2
Má	13	8	5	9	4	12	1	13	0
Muito má	5	5	0	5	0	5	0	5	0

Gravidez									
Sim	1	0	1	0	1	1	0	1	0
Não	82	62	20	67	15	77	5	82	0
Não sabe	2	1	1	0	1	2	0	2	0
Não aplicável	119	64	55	82	37	102	17	116	3
Autorreporte de anemia									
Sim	4	3	1	4	0	3	1	4	0
Não	199	124	75	145	53	178	21	196	3
Não responde	1	0	1	0	1	1	0	1	0

Começando a análise pelo sexo dos indivíduos, embora os participantes deste estudo com microcitose e hipocromia sejam maioritariamente indivíduos do sexo feminino (157 mulheres num total de 204 casos vs 47 homens num total de 204 casos), podemos perceber pelos resultados moleculares que, a percentagem de indivíduos com hemoglobinopatia dentro de cada grupo é semelhante: 1,7% das 2573 mulheres participantes no INSEF têm uma hemoglobinopatia e 1,5% dos 2235 homens participantes no INSEF têm uma hemoglobinopatia. Este facto é explicado pelas hemoglobinopatias serem doenças genéticas autossómicas, ou seja, os genes afetados encontram-se no cromossoma 16 ou 11 e por isso a doença afeta igualmente ambos os sexos. Por outro lado, o facto de haver muito mais mulheres do que homens com o fenótipo de microcitose e/ou hipocromia é, provavelmente, justificado pela presença de anemia ferropénica ou carência nutricional em ferro, patologia que afeta bastante mais as mulheres do que os homens.

Quanto à tipologia da zona de residência, embora os participantes deste estudo residam maioritariamente nas zonas urbanas (155 residentes em áreas urbanas vs 44 residentes em áreas rurais), através do nosso estudo concluímos que a percentagem de indivíduos com hemoglobinopatia dentro de cada grupo é semelhante, tal como acontece para o sexo: 1,7% dos 3428 indivíduos do INSEF que residem em zonas urbanas têm uma hemoglobinopatia e 1,3% dos 1380 indivíduos do INSEF que residem em zonas rurais são portadores de uma hemoglobinopatia. Estes resultados são explicados pelo facto da diversidade genética não ser afetada pelo local de residência.

Relativamente ao local de nascimento dos indivíduos em estudo, é interessante verificar que, embora tenham sido abrangidos poucos indivíduos de outras nacionalidades além da portuguesa (32 indivíduos de nacionalidade estrangeira em 4808 participantes no projeto INSEF, ou seja, 0,67%) muitos desses indivíduos com origem não portuguesa apresentam uma hemoglobinopatia: dois dos seis indivíduos angolanos, seis dos sete indivíduos de nacionalidade brasileira, cinco dos seis indivíduos cabo-verdianos, os dois indivíduos franceses abrangidos, o indivíduo indiano e o indivíduo paquistanês, são portadores de α -talassémia. De facto, a α -talassémia é a hemoglobinopatia estudada onde se verifica um maior número de indivíduos com nacionalidade não portuguesa. Fora isso, apenas detetámos um indivíduo de origem chinesa com β -talassémia e um indivíduo angolano com uma variante de hemoglobina, sendo que os restantes portadores destas duas hemoglobinopatias são todos de origem portuguesa (Figura 4.30). Também nos indivíduos portugueses foi possível constar que a maioria dos indivíduos portadores de uma hemoglobinopatia eram portadores de α -talassémia, seguido dos portadores de β -talassémia e por fim os portadores de variante de hemoglobina. Estes resultados são semelhantes aos obtidos para toda a população em estudo onde se verificou uma maior prevalência de α -talassémia (1,1%), seguida da β -talassémia (0,5%) e das variantes de hemoglobina (0,1%).

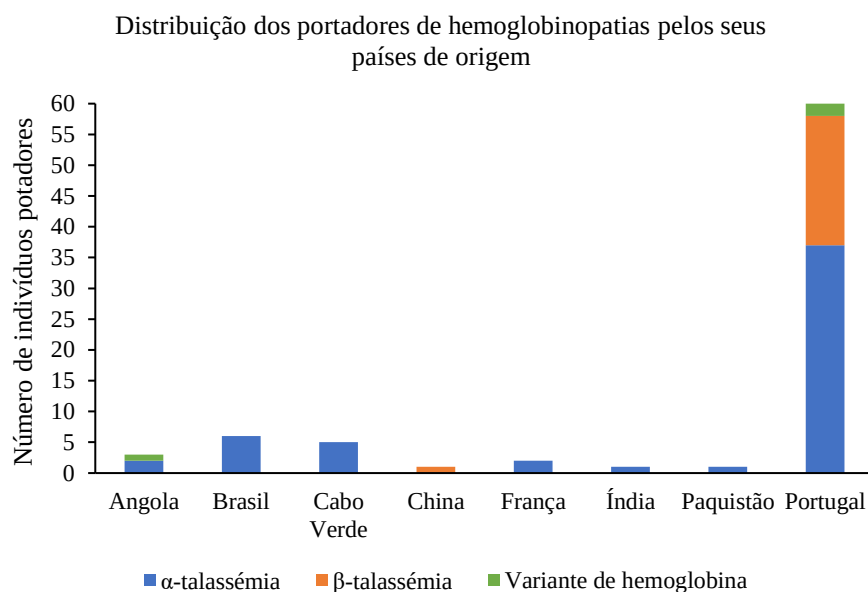


Figura 4.30. Distribuição dos portadores de hemoglobinopatias pelos seus países de origem. Através deste gráfico é possível perceber que a maioria dos indivíduos portadores de hemoglobinopatias são portugueses. Além disso, é possível constatar que existe uma maior prevalência de indivíduos com α -talassémia, seguida dos indivíduos com β -talassémia e por fim, uma menor prevalência de indivíduos com variante de hemoglobina.

Um dos objetivos deste estudo passa por caracterizar o padrão geográfico das hemoglobinopatias na população residente em Portugal em 2015. Para isso, foram realizadas estimativas da prevalência da α -talassémia e da β -talassémia para os 4808 indivíduos, tendo em conta os pesos amostrais das diferentes regiões de Portugal, previamente definidos aquando do INSEF (Tabela 4.19 e Tabela 4.20) (66).

Tabela 4.19. Prevalência da α -talassémia por região de Portugal tendo em conta os seus pesos amostrais

Região	Prevalência	Intervalo de Confiança de 95%
Norte	21,1%	[6,536;50,8]
Centro	5,0%	[1,407;16,35]
Lisboa e Vale do Tejo	64,8%	[41,13;82,86]
Alentejo	2,7%	[0,936;7,548]
Algarve	2,2%	[0,821;5,748]
Região Autónoma da Madeira	3,6%	[1,684;7,437]
Região Autónoma dos Açores	0,6%	[0,152;2,341]

Tabela 4.20. Prevalência da β -talassémia por região de Portugal tendo em conta os seus pesos amostrais

Região	Prevalência	Intervalo de Confiança de 95%
Norte	18,5%	[2,008;71,47]
Centro	52,5%	[21,61;81,53]
Lisboa e Vale do Tejo	0,0%	Não se aplica
Alentejo	10,4%	[4,85;20,85]
Algarve	11,3%	[3,368;31,91]
Região Autónoma da Madeira	5,0%	[1,388;16,58]
Região Autónoma dos Açores	2,3%	[0,4495;11,27]

Os resultados sugerem que a região com maior número de portadores sintomáticos de α -talassémia é a região de Lisboa e Vale do Tejo (64,8%), seguida da região Norte (21,1%). No entanto, dada a elevada amplitude observada nos intervalos de confiança, estes valores poderão não refletir a realidade atual do país.

Por outro lado, o nosso estudo revelou que as regiões com maior prevalência de indivíduos portadores de β -talassémia são a região Centro (52,5%) e a região Norte (18,5%), porém, na região de Lisboa e Vale do Tejo não foram detetados indivíduos com β -talassémia. Estes resultados da prevalência da β -talassémia não estão de acordo com os estudos anteriores que relatavam uma maior prevalência no sul do país, por comparação ao norte. A explicação tanto para esta discrepância, como para a não deteção de indivíduos β -talassémicos em Lisboa e Vale do Tejo possivelmente reside na dimensão da amostra. No estudo anterior de Martins *et al.* foram avaliados do ponto de vista hematológico e bioquímico, 15208 indivíduos do sexo masculino e no nosso estudo apenas foi possível estudar hematologicamente 4808 indivíduos e molecularmente 204 indivíduos sintomáticos (32). De acordo com o *software OpenEpi* (<https://www.openepi.com/SampleSize/SSPropor.htm>), para uma prevalência esperada de 0,5%, a dimensão da amostra deveria consistir em 7164 indivíduos (com uma margem de erro de 0,2%), um número bastante superior aos 4808 indivíduos abrangidos. No entanto, num outro estudo realizado também em 1993 por Tamagnini e os seus colegas, é possível distinguir uma maior prevalência de mutações β -talassémicas na região centro de Portugal, o que vai ao encontro dos nossos resultados. Este estudo teve também uma amostra com uma dimensão mais semelhante à nossa (131 indivíduos), mas os seus resultados não foram ponderados tendo em conta o número de habitantes de cada região (88). Também de acordo com o *software OpenEpi*, para uma prevalência esperada de 10% de α -talassémia, a dimensão da amostra deveria consistir em 208 ou 830 indivíduos (com margens de erro respetivamente de 5% e 2,5%). A dimensão da amostra usada neste estudo para a α -talassémia (204 indivíduos) aproxima-se deste limite inferior, mas, em estudos futuros, dever-se-ia utilizar uma amostra perto do limite superior, bem como ajustar o número de participantes por região, de modo a obterem-se valores de prevalência por região de Portugal mais exatos.

No que diz respeito às variantes de hemoglobina, só foram detetados três indivíduos com variantes de hemoglobina, um na região de Lisboa e Vale do Tejo, um no Algarve e outro na Região Autónoma dos Açores (Tabela 4.18). Como já referido, este estudo não foi desenhado para determinar a prevalência de variantes de hemoglobina. Estas geralmente não se encontram associadas a fenótipo de microcitose e hipocromia. Assim, as variantes de hemoglobina detetadas são justificadas por co-herança com talassémia ou por co-existirem com anemia ferropénica.

No nosso estudo também foi possível perceber que as quatro mutações β -talassémicas mais comuns são a Cd39 (C>T), IVS-I-6 (T>C), IVS-I-110 (G>A) e IVS-I-1 (G>A) (Tabela 4.21). Estes resultados corroboram o que está descrito nos estudos anteriormente publicados sobre as mutações β -talassémicas na população portuguesa. Estas quatro mutações mais comuns da β -talassémia em Portugal são também as mutações mais comuns nos países da bacia do Mediterrâneo. Este facto pode ser explicado pela ocupação ou imigração ao longo dos séculos de populações da área do Mediterrâneo, por exemplo, Romanos, Fenícios e Mouros, para o sul do nosso país, que tem oferecido condições benéficas para a agricultura, pesca e mineração (79).

Tabela 4.21. Diversidade, frequência e distribuição geográfica das mutações no gene *HBB* detetadas nos 204 participantes

	Norte	Centro	LVT	Alentejo	Algarve	RA Madeira	RA Açores	Total
Hb D-Portugal	0	0	0	0	1	0	0	4,0%
Cd41/42 (-CTTT)	0	0	0	0	0	0	1	4,0%
Cd39 (C>T)	0	2	0	2	4	0	0	32,0%
IVS-I-110 (G>A)	0	1	0	0	1	1	0	12,0%
IVS-I-6 (T>C)	0	2	0	0	0	2	0	16,0%
IVS-I-1 (G>A)	0	0	0	2	0	0	1	12,0%
Cd15 (G>A)	0	1	0	0	0	1	0	8,0%
Cd6 (-A)	1	0	0	0	0	0	0	4,0%
Hb S	0	0	1	0	0	0	0	4,0%
Hb C	0	0	0	0	0	0	1	4,0%
Total	1	6	1	4	6	4	3	100,0%

LVT – Lisboa e Vale do Tejo; RA – região autónoma

Através da Tabela 4.21 podemos ainda perceber que mutações β -talassémicas detetadas apresentam alguma heterogeneidade na sua distribuição geográfica. A heterogeneidade identificada neste estudo é semelhante ao relatado por Faustino e os seus colaboradores, na meta-análise realizada em 1999 (79).

Para a mutação β -talassémica mais comum, a Cd39 (C>T), neste estudo obtivemos uma frequência de 36,4%, semelhante aos 33% descritos no estudo de 1999. Além disso, também é possível verificar uma prevalência ligeiramente maior no Algarve, equivalente ao referido no mesmo estudo, que descreve uma maior prevalência da Cd39 (C>T) na região sul. Também para a IVS-I-1 (G>A) é descrito no mesmo estudo uma maior prevalência no sul, comparativamente ao centro e norte de Portugal e, através da Tabela 4.21 podemos verificar que também nos foi possível apurar uma maior frequência da mutação IVS-I-1 (G>A) no sul do país, mais especificamente na região do Alentejo. No entanto, as frequências do nosso estudo e do estudo de Faustino *et al.* são um pouco diferentes, de 13,6% para 22% (79).

Neste trabalho descrevemos uma frequência de 13,6% para a IVS-I-110 (G>A), perto dos 10% indicados no estudo de Faustino *et al.* (79). No nosso estudo, tal como no referido, não foram observadas diferenças significativas entre as regiões de Portugal, indicando que esta mutação apresenta uma distribuição uniforme pelo território.

Relativamente à mutação IVS-I-6 (T>C), detetámos uma frequência de 18,2% e mais uma vez o nosso estudo demonstrou estar de acordo com o estudo de Faustino e os seus colaboradores, que descrevem

uma frequência de 17% para a IVS-I-6 (T>C). Tanto o nosso estudo como o estudo de Faustino *et al.* em 1999 e o estudo de Tamagnini em 1993 descrevem uma maior prevalência desta mutação na região Centro de Portugal (79, 88). Contudo, é importante salientar que também verificámos simultaneamente uma prevalência da IVS-I-6 (T>C) na Região Autónoma da Madeira. Até onde temos conhecimento, este estudo é o primeiro que inclui as regiões autónomas de Portugal na sua análise.

No que diz respeito à mutação Cd15 (G>A) obtivemos uma frequência de 9,1%, um valor próximo à frequência de 6%, descrita no estudo de Faustino *et al.* Também semelhante a este estudo de 1999, obtivemos uma maior prevalência na região Centro, o que corrobora outro estudo realizado em 1993 por Tamagnini *et al.* que verificaram uma maior prevalência de indivíduos portadores de Cd15 (G>A) na região que se estende de Aveiro a Lisboa (79, 88).

Por fim, para a mutação β -talassémica Cd6 (-A) obtivemos uma frequência de 4,6%. Esta mutação foi identificada por nós apenas na região Norte, estando mais uma vez de acordo com o descrito no estudo de Faustino *et al.* que, embora tenham uma frequência ligeiramente superior (8%), também só descreveram a alteração na região Norte de Portugal (79). A meta-análise de 1999 realizada por Faustino *et al.* inclui os dados de um estudo também de 1999 realizado por Cabeda *et al.* onde é demonstrado que a mutação Cd6 (-A) é uma alteração β -talassémica frequente na região Norte em contrapartida com o Centro e Sul, onde é quase ausente (79, 93).

No que concerne às mutações no agrupamento génico da α -globina, poucos estudos foram realizados em Portugal e nenhum se debruçou na análise do padrão geográfico destas alterações. Tendo em conta que identificámos um indivíduo homozigótico para a $-\alpha^{3,7}$, um indivíduo heterozigótico para a $-\alpha^{4,2}$, dois heterozigóticos para a $\alpha\alpha^{\text{anti } 3,7}$ e 52 indivíduos heterozigóticos para a $-\alpha^{3,7}$ é possível fazer a distinção da distribuição geográfica destas alterações pelo território português.

Relativamente a estes, através da Tabela 4.22 podemos concluir que existe um maior número de indivíduos portadores da deleção de 3,7 kb na região de Lisboa e Vale do Tejo comparativamente às restantes regiões de Portugal continental. Também detetámos um elevado número de indivíduos portadores da deleção de 3,7 kb na Região Autónoma da Madeira. No entanto, tendo em conta os pesos amostrais das diferentes regiões de Portugal, a região Norte é a segunda região com maior prevalência de α -talassémia, como referido anteriormente. De facto, seria importante caracterizar melhor a população α -talassémica e conhecer a sua distribuição geográfica de modo a ser possível adotar melhores estratégias de diagnóstico.

Tabela 4.22. Diversidade, frequência e distribuição geográfica das mutações no gene *HBA* detetadas nos 204 participantes

	Norte	Centro	LVT	Alentejo	Algarve	RA Madeira	RA Açores	Total
$\alpha\alpha/-\alpha^{3,7}$	7	2	16	5	5	14	3	92,9%
$-\alpha^{3,7}/-\alpha^{3,7}$	0	0	0	1	0	0	0	1,8%
$\alpha\alpha/-\alpha^{4,2}$	0	1	0	0	0	0	0	1,8%
$\alpha\alpha/\alpha\alpha^{\text{anti } 3,7}$	0	1	0	0	0	0	1	3,6%
Total	7	4	16	6	5	14	4	100,0%

LVT – Lisboa e Vale do Tejo; **RA** – região autónoma

No que diz respeito aos dados de saúde dos participantes fornecidos através do INSEF, alguns pontos são interessantes de explorar. Um dos pontos importantes é o conhecimento se os indivíduos portadores de hemoglobinopatias têm ou não médico de família. Pela Tabela 4.19 podemos constatar que 57 dos 77 indivíduos com hemoglobinopatias possuem médico de família. A importância do médico de família prende-se com a sua intervenção na deteção destes indivíduos portadores de hemoglobinopatias pois, desde 2004 existe uma circular normativa promulgada pela Direção-Geral da Saúde (DGS), a Circular Normativa da DGS nº 18/DSMIA de 2004, onde são definidas as condutas e regras para a prevenção das formas graves de hemoglobinopatias. Assim, é de esperar que estes 57 indivíduos com médico de família tenham conhecimento através deste do seu estado de portador de uma hemoglobinopatia, embora a grande maioria tenha uma perceção da sua saúde como sendo normal e eventualmente desconhece que é portador de doença genética (Figura 4.31). Mais especificamente, no que diz respeito ao indivíduo que nunca teve consulta com o médico de família (Tabela 4.19) esta é uma condição grave pois o sujeito em causa poderá não estar ciente do seu estado portador de hemoglobinopatia e poderá não ter em conta que têm 25% de probabilidade de ter um filho doente caso o seu cônjuge também seja portador de uma hemoglobinopatia.

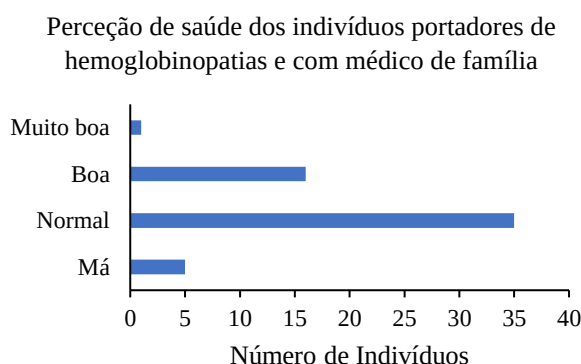


Figura 4.31. Perceção de saúde dos indivíduos com hemoglobinopatias que possuem médico de família. Os dados deste gráfico são referentes aos dados obtidos pelo INSEF. Através do gráfico podemos constatar que a maioria dos indivíduos portadores de hemoglobinopatias e com médico de família percecionam a sua saúde como sendo normal e apenas uma minoria perceciona-a como sendo má. Fica por esclarecer se esta maioria que perceciona a saúde como sendo boa tem ou não conhecimento do seu estado de portador de hemoglobinopatia, indicado pelo seu médico de família.

Com os dados de saúde do INSEF também é possível verificar que existe uma participante grávida e com α -talassémia, mais especificamente a deleção de 3,7 kb em heterozigotia. Esta participante tem médico de família e teve a sua última consulta com este há menos de 12 meses por isso também é de esperar que esteja alerta para o seu estado de portadora de α -talassémia.

Quanto ao autorreporte de anemia, dos 204 indivíduos deste estudo, apenas 4 têm o conhecimento que são anémicos. Este reduzido número de pessoas com conhecimento da sua anemia torna-se ainda mais preocupante se forem acompanhados por um médico de família. Na realidade, dos 160 indivíduos que indicaram possuir médico de família, 76 possuem anemia, mas apenas dois deles têm conhecimento da sua condição. A Circular Normativa da DGS nº 18/DSMIA de 2004 recomenda a deteção e informação precoce, preferencialmente pré-concepcional, de adultos portadores de hemoglobinopatias cujo primeira etapa passa pela solicitação do hemograma. Tendo em conta que os 74 indivíduos que desconhecem que possuem anemia são todos adultos, poderá estar a ocorrer alguma falha no Sistema Nacional de Saúde

(SNS), pois os respetivos médicos de família não estão a seguir a Circular Normativa da DGS, caso contrário, os indivíduos anémicos deveriam ser conhecedores da sua condição.

5. Conclusão

Tendo como principal objetivo contribuir para o conhecimento da realidade atual das hemoglobinopatias sintomáticas, sobretudo das talassémias, em Portugal continental e ilhas, este trabalho incluiu várias etapas.

Na primeira fase deste trabalho, conseguimos esclarecer a base molecular das hemoglobinopatias presentes na população residente em Portugal avaliada no projeto INSEF, com fenótipo hematológico de hipocromia e/ou microcitose e cumprimos assim o primeiro objetivo a que nos propusemos. No total encontramos dois tipos de deleções responsáveis por α -talassémia (deleção de 3,7 kb e deleção de 4,2 kb), uma triplicação dos genes *HBA* ($\alpha\alpha^{\text{anti } 3,7}$), sete tipos de mutações diferentes responsáveis por β -talassémia [Cd39 (C>T), IVS-I-6 (T>C), IVS-I-110 (G>A), IVS-I-1 (G>A), Cd15 (G>A), Cd6(-A) e Cd41/42 (-CTTT)] e três variantes de hemoglobina (Hb S, Hb C e Hb D-Portugal). Através deste estudo também concluímos que a deleção α -talassémica de 3,7 kb é a mais prevalente em Portugal e que as quatro mutações β -talassémicas mais comuns são a Cd39 (C>T), IVS-I-110 (G>A), IVS-I-6 (T>C) e IVS-I-1 (G>A), tal como descrito anteriormente para a população portuguesa. Além disso, nesta dissertação também conseguimos caracterizar *in silico* o efeito patogénico das variantes no gene *HBB* detetadas por NGS, cumprindo assim o segundo objetivo.

Neste trabalho também nos propusemos determinar a frequência das diferentes hemoglobinopatias detetadas na população residente em Portugal avaliada no projeto INSEF, com fenótipo hematológico de hipocromia e/ou microcitose. Este objetivo também foi alcançado, tendo sido obtido uma frequência de α -talassémia de 26,6%, β -talassémia de 10,8% e variantes de hemoglobina de 1,5%. Contudo é importante referir que, tendo em conta que partimos de uma amostra da população com fenótipo hematológico de hipocromia e/ou microcitose, a extrapolação destas prevalências para a população portuguesa deve ter em atenção o facto de nem todos os portadores de α -talassémia e de hemoglobinopatias qualitativas possuírem hipocromia e microcitose. O estado de portador de α -talassémia é, por vezes, hematologicamente assintomático e, por isso, não distinguível da população normal. Assim, para concluirmos sobre a prevalência da α -talassémia na população portuguesa em geral, um possível estudo futuro passaria por realizar a análise molecular dos genes *HBA* dos restantes participantes do INSEF, de forma a incluir os indivíduos portadores de α -talassémia assintomáticos. Relativamente às hemoglobinopatias qualitativas, apenas verificámos por NGS a presença de três indivíduos com variantes de hemoglobina, pelo que seria também interessante analisar os restantes participantes do INSEF quanto a este tipo de hemoglobinopatias. No entanto, o estudo molecular da restante população do INSEF tornar-se-ia muito dispendioso. Normalmente a análise a executar para detetar variantes de hemoglobina seria, por exemplo, a eletroforese de hemoglobinas ou a análise por HPLC. Ambas teriam de ser executadas em amostras de sangue fresco, o que não é compatível com as amostras atuais do projeto INSEF que foram armazenadas sob congelação. Como de entre as variantes de hemoglobina a Hb S é a que tem maior prevalência na nossa população seria de todo o interesse realizar um estudo especificamente para detetar esta variante e determinar se a sua prevalência se alterou na população (32). Contudo, tal como referido, esse estudo necessitaria uma análise bioquímica realizada em sangue fresco, o que requereria outro rastreio à população.

O objetivo de determinar a prevalência da β -talassémia na população residente em Portugal em 2015 foi plenamente alcançado. Uma vez que todos os portadores de β -talassémia apresentam um fenótipo de microcitose e hipocromia, isto implica que ao termos selecionado dos 4808 participantes do projeto INSEF aqueles que apresentavam este fenótipo e seguidamente ao analisarmos o gene *HBB* em todos eles, conseguimos detetar cerca de 100% dos casos existentes. Eventualmente, se a mutação talassémica se situar na região regulatória distal do gene (no LCR) poderemos não ter detetado esse caso, mas tais

mutações são muitíssimo raras. Assim, os 22 indivíduos detetados como sendo portadores de β -talassémia nos 4808 do projeto INSEF, refletem efetivamente uma prevalência de 0,5% na população residente em Portugal, continental e ilhas, em 2015. Este valor é muito semelhante ao anteriormente determinado para a população de Portugal continental e sugere que esta patologia não aumentou muito na nossa população (32). No entanto, a amostragem de 4808 indivíduos situa-se abaixo do desejado para que a margem de erro seja pequena, pelo que será benéfico que num próximo estudo epidemiológico a realizar na nossa população, o número de participantes seja significativamente maior.

Um outro ponto importante de lembrar é que a hipocromia e/ou microcitose, além de serem indicadores de talassémia, são também indicadores de deficiência em ferro. Como dos 204 indivíduos selecionados do INSEF com hipocromia e/ou microcitose, foram diagnosticados 22 indivíduos com β -talassémia, 54 com α -talassémia e 1 com uma variante de hemoglobina que justificam o fenótipo, seria interessante, no futuro, avaliar o *status* do ferro, através de testes bioquímicos para os restantes 126 indivíduos que ficaram com o seu fenótipo de hipocromia e/ou microcitose por esclarecer. Contudo, uma conclusão bastante importante que este estudo permitiu obter é o facto de 38,9% destes fenótipos serem justificados por uma doença genética – uma hemoglobinopatia. Este valor bastante elevado deverá alertar os clínicos, sobretudo os médicos de clínica geral e familiar para este tipo de causa.

Relativamente aos estudos de associação entre as variantes genéticas detetadas e os vários parâmetros hematológicos da população avaliada no projeto INSEF, apenas devemos retirar conclusões relativamente aos portadores de β -talassémia, devido aos fatores anteriormente referidos. Assim, quanto à β -talassémia podemos concluir que os indivíduos portadores desta patologia apresentam níveis mais reduzidos de VGM, HGM e CHGM e níveis mais aumentados de eritrócitos e plaquetas, comparativamente aos outros indivíduos que também têm um fenótipo de microcitose e hipocromia, mas que não têm β -talassémia.

O último objetivo a que nos propusemos consistia em descrever as características demográficas e de saúde dos indivíduos. Nesta etapa, o passo de maior importância consistiu na análise do padrão de distribuição geográfica dos indivíduos portadores de hemoglobinopatias. Analisando este grupo de indivíduos, os resultados obtidos por nós indicam uma maior prevalência de indivíduos portadores de α -talassémia na região de Lisboa e Vale do Tejo (64,8%) e uma maior prevalência de indivíduos portadores de β -talassémia na região Centro (52,5%). No entanto, dada a elevada amplitude dos intervalos de confiança, estes valores estão afetados por um grande nível de incerteza e poderão não refletir a realidade do país. Da mesma forma, a prevalência de portadores de β -talassémia contrasta com os estudos anteriores, que indicavam uma maior prevalência no sul do país pelo que, como perspetiva futura, sugerimos a realização de um novo estudo com uma amostragem total de cerca de 7164 indivíduos, cujo número de participantes por região deverá ser ajustado face ao número de habitantes.

O conhecimento da distribuição geográfica das hemoglobinopatias numa determinada população, bem como da ancestralidade dos indivíduos são pontos importantes para que os laboratórios possam estabelecer uma melhor estratégia de diagnóstico e de caracterização molecular destas patologias. A deteção precoce, o correto diagnóstico dos portadores destas doenças genéticas e a identificação dos casais em risco é de extrema importância, de modo a encaminhá-los para o aconselhamento genético, evitando desta forma graves problemas de saúde dos descendentes e também problemas sociais e económicos para a sociedade.

Em conclusão, embora ainda possam ser realizados estudos complementares principalmente ao nível da α -talassémia, hemoglobinopatias quantitativas e anemia ferropénica, acreditamos que, através deste estudo, contribuímos para o conhecimento atual da base molecular das hemoglobinopatias na população portuguesa. Este conhecimento poder-se-á refletir no estabelecimento de melhores estratégias de

diagnóstico e análise molecular, permitindo também um melhor encaminhamento para aconselhamento genético, bem como uma melhor qualidade de vida dos portugueses.

6. Referências

1. A. J. Marengo-Rowe, Structure-function relations of human hemoglobins. *Proc. (Bayl. Univ. Med. Cent)*. **19**, 239–245 (2006).
2. S. A. Sarnaik, Thalassemia and Related Hemoglobinopathies. *Indian J. Pediatr.* **72**, 319–324 (2005).
3. B. J. Bain, in *Haemoglobinopathy Diagnosis*, B. J. Bain, Ed. (Blackwell Publishing Ltd, 2^a Ed., 2006; <https://doi.org/10.1002/9780470988787.ch1>), Wiley Online Books, pp. 1–25.
4. D. R. Higgs, The Molecular Basis of α -Thalassemia. *Cold Spring Harb. Perspect. Med.* **3** (2013), doi:10.1101/cshperspect.a011718.
5. S. L. Thein, Molecular basis of β thalassemia and potential therapeutic targets. *Blood Cells, Mol. Dis.* **70**, 54–65 (2018).
6. S. Farashi, C. L. Hartevel, Molecular basis of α -thalassemia. *Blood Cells, Mol. Dis.* **70**, 43–53 (2018).
7. S. L. Thein, Genetic insights into the clinical diversity of β thalassaemia. *Br. J. Haematol.* **124**, 264–274 (2004).
8. F. Samarah, M. Athanasiou-Metaxa, N. Vavatsi-Christaki, Genetic modulators of sickle cell disease. *Aristotle Univ. Med. J.* **34**, 25–35 (2007).
9. B. E. Clark, S. L. Thein, Molecular diagnosis of haemoglobin disorders. *Clin. Lab. Haematol.* **26**, 159–176 (2004).
10. A. T. Taher, D. J. Weatherall, M. D. Cappellini, Thalassemia. *Lancet.* **391**, 155–167 (2018).
11. M. M. Moosa, Analysis of β -globin gene mutations in a β -thalassaemic thalassaemic patient and predicting their relative role in disease prognosis (2008), doi:10.13140/2.1.4692.5768.
12. C. M. Chaparro, P. S. Suchdev, Anemia epidemiology, pathophysiology, and etiology in low- and middle-income countries. *Ann. N. Y. Acad. Sci.* **1450**, 15–31 (2019).
13. World Health Organization, Haemoglobin concentrations for the diagnosis of anaemia and assessment of severity (2011), pp. 1–6.
14. M. Rebordão, M. Silva, Estudo de microcitose numa população jovem adulta assintomática. *Rev. Port. Clínica Geral.* **27**, 508–516 (2011).
15. Association of Public Health Laboratories, Centers for Disease Control and Prevention, Hemoglobinopathies: Current Practices for Screening , Confirmation and Follow-up. *Assoc. Public Heal. Lab.* (2015), (available at https://www.cdc.gov/ncbddd/sicklecell/documents/nbs_hemoglobinopathy-testing_122015.pdf).
16. C. S. Thom, C. F. Dickson, D. A. Gell, M. J. Weiss, Hemoglobin variants: Biochemical Properties and Clinical Correlates. *Cold Spring Harb. Perspect. Med.* **3** (2013), doi:10.1101/cshperspect.a011858.
17. E. Kohne, Hemoglobinopathies: clinical manifestations, diagnosis, and treatment. *Dtsch. Arztebl. Int.* **108**, 532–540 (2011).
18. B. J. Bain, in *Haemoglobinopathy Diagnosis*, B. J. Bain, Ed. (Blackwell Publishing Ltd, 2^a Ed., 2006; <https://doi.org/10.1002/9781119579977.ch4>), Wiley Online Books, pp. 139–189.
19. M. H. Odièvre, E. Verger, A. C. Silva-Pinto, J. Elion, Pathophysiological insights in sickle cell

- disease. *Indian J. Med. Res.* **134**, 532–537 (2011).
20. D. C. Rees, T. N. Williams, M. T. Gladwin, Sick cell disease. *Lancet*. **376**, 2018–2031 (2010).
 21. A. Chatterjee, A. Agrawal, D. Adapa, S. Tk, Sick Cell Anaemia- A Synopsis of the Inherited Ailment. *Arch. Med.* **10**, 1–10 (2018).
 22. G. J. Lonergan, D. B. Cline, S. L. Abbondanzo, Sick cell anemia. *RadioGraphics*. **21**, 971–994 (2001).
 23. H. F. Bunn, Pathogenesis and treatment of sick cell disease. *N. Engl. J. Med.* **337**, 762–769 (1997).
 24. M. E. Houwing, P. J. de Pagter, E. J. van Beers, B. J. Biemond, E. Rettenbacher, A. W. Rijneveld, E. M. Schols, J. N. J. Philipsen, R. Y. J. Tamminga, K. F. van Draat, E. Nur, M. H. Cnossen, Sick cell disease: Clinical presentation and management of a global health challenge. *Blood Rev.* **37** (2019), doi:10.1016/j.blre.2019.05.004.
 25. K. L. Schmidt, T. R. Randolph, Development of a Microscopic Method to Identify Hemoglobin C Conditions for Use in Developing Countries. *Am. Soc. Clin. Lab. Sci.* (2019), doi:10.29074/ascls.119.001412.
 26. B. J. Bain, in *Haemoglobinopathy Diagnosis*, B. J. Bain, Ed. (Blackwell Publishing Ltd, 2^a Ed., 2006; <https://doi.org/10.1002/9781119579977.ch3>), Wiley Online Books, pp. 63–138.
 27. M. Grosso, R. Sessa, S. Puzone, M. R. Storino, P. Izzo, in *Anemia*, D. Silverberg, Ed. (IntechOpen, Rijeka, 1^a Ed., 2012; <https://www.intechopen.com/books/anemia/molecular-basis-of-thalassemia>), pp. 341–360.
 28. C. L. Harteveld, D. R. Higgs, α -thalassaemia. *Orphanet J. Rare Dis.* **5** (2010), doi:10.1186/1750-1172-5-13.
 29. J. Old, in *Principles and Practice of Medical Genetics*, D. L. Rimoim, R. e Pyeritz, B. R. Korf, Eds. (Academic Press, Oxford, 6^a Ed., 2013), pp. 1–44.
 30. A. Machin, Thalassaemia. *InnovAiT*. **7**, 558–565 (2014).
 31. D. Rund, E. Rachmilewitz, Beta-thalassemia. *N. Engl. J. Med.* **353**, 1135–1146 (2005).
 32. M. C. Martins, G. Olim, J. Melo, H. A. Magalhães, M. O. Rodrigues, Hereditary anaemias in Portugal: epidemiology, public health significance, and control. *J. Med. Genet.* **30**, 235–239 (1993).
 33. F. Inez, M. Sequeira, P. Santos, R. Santos, E. Nunes, A. Cavaco, L. Carvalho, A. Calado, M. S. J. Tavares, L. Portugal, J. Rodrigues, O. Mendes, M. G. Sousa, E. Esteves, M. C. Almeida, M. F. Breia, M. L. Fialho, Contribuição do rastreio de portadores para a prevenção da beta-talassémia e da drepanocitose na população portuguesa: um estudo multicêntrico. *Arq. do Inst. Nac. Saúde.* **19**, 27–31 (1993).
 34. M. J. Peres, M. H. Carreiro, M. C. Machado, T. Seixas, I. Picanço, L. Batalha, J. Lavinha, M. C. Martins, Rastreio neonatal de hemoglobinopatias numa população residente em Portugal. *Acta Med. Port.* **9**, 135–139 (1996).
 35. V. Viprakasit, S. Ekwattanakit, Clinical Classification, Screening and Diagnosis for Thalassemia. *Hematol. Oncol. Clin. North Am.* **32**, 193–211 (2018).
 36. J. Old, Prevention and Diagnosis of Haemoglobinopathies: A Short Guide for Health Professionals and Laboratory Scientists (2016), pp. 1–79.
 37. S. Lobitz, P. Telfer, E. Cela, B. Allaf, M. Angastiniotis, C. Backman Johansson, C. Badens, C. Bento, M. J. Bouva, D. Canatan, M. Charlton, C. Coppinger, Y. Daniel, M. de Montalembert, P.

- Ducoroy, E. Dulin, R. Fingerhut, C. Frömmel, M. García-Morin, B. Gulbis, U. Holtkamp, B. Inusa, J. James, M. Kleanthous, J. Klein, J. B. Kunz, L. Langabeer, C. Lapoumèroulie, A. Marcao, J. L. Marín Soria, C. McMahon, K. Ohene-Frempong, J.-M. Périni, F. B. Piel, G. Russo, L. Sainati, M. Schmugge, A. Streetly, L. Tshilolo, C. Turner, D. Venturelli, L. Vilarinho, R. Yahyaoui, J. Elion, R. Colombatti, Newborn screening for sickle cell disease in Europe: recommendations from a Pan-European Consensus Conference. *Br. J. Haematol.* **183**, 648–660 (2018).
38. R. F. Labbé, H. J. Vreman, D. K. Stevenson, Zinc Protoporphyrin: A Metabolite with a Mission. *Clin. Chem.* **45**, 2060–2072 (1999).
39. E. Khandros, J. L. Kwiatkowski, Beta Thalassemia: Monitoring and New Treatment Approaches. *Hematol. Oncol. Clin. North Am.* **33**, 339–353 (2019).
40. A. El-Beshlawy, M. El-Ghamrawy, Recent trends in treatment of thalassemia. *Blood Cells, Mol. Dis.* **76**, 53–58 (2019).
41. M. Shamoun, M. Callaghan, in *Benign Hematologic Disorders in Children: A Clinical Guide*, D. M. Kamat, M. Frei-Jones, Eds. (Springer, 1^a Ed., 2020), pp. 91–98.
42. C. Harrison, First gene therapy for β -thalassemia approved. *Nat. Biotechnol.* **37**, 1102–1103 (2019).
43. H. Tolonen, *EHES Manual-Part B-Fieldwork Procedures* (National Institute for Health and Welfare, 1^a Ed., 2013).
44. K. Mullis, F. Faloona, S. Scharf, R. Saiki, G. Horn, H. Erlich, Specific Enzymatic Amplification of DNA In Vitro: The Polymerase Chain Reaction. *Cold Spring Harb. Symp. Quant. Biol.* **51**, 263–273 (1986).
45. M. Rahman, M. S. Uddin, R. Sultana, A. Moue, M. Setu, Polymerase Chain Reaction (PCR): A Short Review. *Anwer Khan Mod. Med. Coll. J.* **4**, 30–36 (2013).
46. S. L. Kho, K. H. Chua, E. George, J. A. M. A. Tan, A novel gap-PCR with high resolution melting analysis for the detection of α -thalassaemia Southeast Asian and Filipino β 0-thalassaemia deletion. *Sci. Rep.* **5** (2015), doi:10.1038/srep13937.
47. X. Jia, R. Huang, Z. Lei, L. Yao, L. Wang, Y. Li, D. Meng, Y. Zhou, J. Liu, X. Zhang, Detection of a novel large deletion causing α -thalassemia in South China. *Exp. Mol. Pathol.* **95**, 68–73 (2013).
48. L. Stuppia, I. Antonucci, G. Palka, V. Gatta, Use of the MLPA Assay in the Molecular Diagnosis of Gene Copy Number Alterations in Human Genetic Diseases. *Int. J. Mol. Sci.* **13**, 3245–3276 (2012).
49. J. P. Schouten, C. J. McElgunn, R. Waaijer, D. Zwiijnenburg, F. Diepvens, G. Pals, Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification. *Nucleic Acids Res.* **30**, 1–13 (2002).
50. MRC Holland, MLPA (Multiplex Ligation-dependent Probe Amplification) General Protocol for the detection and quantification of DNA sequences. **7** (2019).
51. MRC Holland, Product Description SALSA® MLPA® Probemix: P140-C1 HBA. **C1-04** (2019).
52. A. Grada, K. Weinbrecht, Next-generation sequencing: methodology and application. *J. Invest. Dermatol.* **133** (2013), doi:10.1038/jid.2013.248.
53. Illumina Inc., An introduction to Next-Generation Sequencing Technology (2017).
54. P. Ewels, M. Magnusson, S. Lundin, M. Käller, MultiQC: summarize analysis results for

- multiple tools and samples in a single report. *Bioinformatics*. **32**, 3047–3048 (2016).
55. V. Schneider, T. Graves-Lindsay, K. Howe, N. Bouk, H.-C. Chen, P. Kitts, T. Murphy, K. Pruitt, F. Thibaud-Nissen, D. Albracht, R. Fulton, M. Kremitzki, V. Magrini, C. Markovic, S. McGrath, K. M. Steinberg, K. Auger, W. Chow, J. Collins, G. Harden, T. Hubbard, S. Pelan, J. T. Simpson, G. Threadgold, J. Torrance, J. M. Wood, L. Clarke, S. Koren, M. Boitano, P. Peluso, H. Li, C.-S. Chin, A. M. Phillippy, R. Durbin, R. K. Wilson, P. Flicek, E. E. Eichler, D. M. Church, Evaluation of GRCh38 and de novo haploid genome assemblies demonstrates the enduring quality of the reference assembly. *Genome Res.* **27**, 849–864 (2017).
 56. B. Langmead, C. Wilks, V. Antonescu, R. Charles, Scaling read aligners to hundreds of threads on general-purpose processors. *Bioinformatics*. **35**, 421–432 (2019).
 57. H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. **25**, 2078–2079 (2009).
 58. G. Van der Auwera, M. Carneiro, C. Hartl, R. Poplin, G. Del Angel, A. Levy-Moonshine, T. Jordan, K. Shakir, D. Roazen, J. Thibault, E. Banks, K. V. Garimella, D. Altshuler, S. Gabriel, M. A. DePristo, From FastQ data to High-Confidence Variant Calls: The Genome Analysis Toolkit Best Practices Pipeline. *Curr. Protoc. Bioinforma.* **43**, 1–33 (2013).
 59. H. Li, A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics*. **27**, 2987–2993 (2011).
 60. F. Sanger, S. Nicklen, A. R. Coulson, DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. U. S. A.* **74**, 5463–5467 (1977).
 61. M. Kchouk, J.-F. Gibrat, M. Elloumi, Generations of Sequencing Technologies: From First to Next Generation. *Biol. Med.* **9** (2017), doi:10.4172/0974-8369.1000395.
 62. JSI Medical Systems, varSEAK Online: Splice Site Prediction (2021).
 63. P. C. Ng, S. Henikoff, SIFT: Predicting amino acid changes that affect protein function. *Nucleic Acids Res.* **31**, 3812–3814 (2003).
 64. I. Adzhubei, D. M. Jordan, S. R. Sunyaev, Predicting functional effect of human missense mutations using PolyPhen-2. *Curr. Protoc. Hum. Genet.* **76**, 1–41 (2013).
 65. J. Hu, P. C. Ng, SIFT Indel: Predictions for the Functional Effects of Amino Acid Insertions/Deletions in Proteins. *PLoS One*. **8** (2013), doi:10.1371/journal.pone.0077940.
 66. M. Barreto, V. Gaio, I. Kislalya, Li. Antunes, A. P. Rodrigues, A. C. Silva, P. Vargas, T. Prokopenko, A. J. Santos, S. Namorado, A. P. Gil, C. A. Alves, E. Castilho, E. Cordeiro, A. Dínis, B. Nunes, C. M. Dias, 1º Inquérito Nacional de Saúde com Exame Físico (INSEF 2015): Estado de Saúde (2016), pp. 1–64.
 67. J. F. Matos, L. M. S. Dusse, K. B. G. Borges, R. L. V de Castro, W. Coura-Vital, M. G. Carvalho, A new index to discriminate between iron deficiency anemia and thalassemia trait. *Rev. Bras. Hematol. Hemoter.* **38**, 214–219 (2016).
 68. J. Eivazi-Ziaei, S. Dastgiri, S. Pourebrahim, R. Soltanpour, Usefulness of red blood cell flags in diagnosing and differentiating thalassemia trait from iron-deficiency anemia. *Hematology*. **13**, 253–256 (2008).
 69. J. J. M. L. Hoffmann, E. Urrechaga, Verification of 20 Mathematical Formulas for Discriminating Between Iron Deficiency Anemia and Thalassemia Trait in Microcytic Anemia. *Lab. Med.* **51**, 628–634 (2020).
 70. L. S. Torres, J. V. Okumura, D. G. H. Silva, C. R. Bonini-Domingos, Hemoglobin D-Punjab:

- origin, distribution and laboratory diagnosis. *Rev. Bras. Hematol. Hemoter.* **37**, 120–126 (2015).
71. M. J. Betts, R. B. Russell, in *Bioinformatics for Geneticists*, M. J. Betts, R. B. Russell, Eds. (John Wiley & Sons Ltd, 1^a Ed., 2003; <https://doi.org/10.1002/0470867302.ch14>), *Wiley Online Books*, pp. 289–316.
 72. Z. Rahimi, R. Akramipour, R. L. Nagel, A. S. Ahmadi, A. Merat, F. Bahrehmand, The β -globin gene haplotypes associated with Hb D-Los Angeles [β 121(GH4)Glu $\rightarrow$ Gln] in Western Iran. *Hemoglobin.* **30**, 39–44 (2006).
 73. A. Kimura, E. Matsunaga, Y. Takihara, T. Nakamura, Y. Takagi, S. Lin, H. Lee, Structural Analysis of a β -Thalassemia Gene Found in Taiwan. *J. Biol. Chem.* **258**, 2748–2749 (1983).
 74. Y. Zeng, S. Huang, The studies of hemoglobinopathies and thalassemia in China-the experiences in Shanghai Institute of Medical Genetics. *Clin. Chim. Acta.* **313**, 107–111 (2001).
 75. S. L. Thein, The molecular basis of β -thalassemia. *Cold Spring Harb. Perspect. Med.* **3**, 1–24 (2013).
 76. W. R. Gomes, R. A. Santos, J. G. Cominal, C. F. F. Tavares, Frequencies of CD39, IVS1-1, IVS1-6 and IVS1-110 mutations in beta-thalassemia carriers and their influence on hematimetric indices. *J. Bras. Patol. e Med. Lab.* **53**, 362–367 (2017).
 77. R. Trecartin, S. Liebhaber, J. Chang, K. Lee, Y. W. Kan, M. Furbetta, A. Angius, A. Cao, β^0 Thalassemia in Sardinia Is Caused by a Nonsense Mutation. *J. Clin. Invest.* **68**, 1012–1017 (1981).
 78. G. Neu-Yilik, B. Amthor, N. H. Gehring, S. Bahri, H. Paidassi, M. W. Hentze, A. E. Kulozik, Mechanism of escape from nonsense-mediated mRNA decay of human β -globin transcripts with nonsense mutations in the first exon. *RNA.* **17**, 843–854 (2011).
 79. P. Faustino, P. Pacheco, P. Loureiro, P. Nogueira, J. Lavinha, The geographic pattern of β -thalassaemia mutations in the portuguese population. *Br. J. Haematol.* **107**, 903–904 (1999).
 80. D. Westaway, R. Williamson, An intron nucleotide sequence variant in a cloned β^+ -thalassaemia globin gene. *Nucleic Acids Res.* **9**, 1777–1788 (1981).
 81. R. Spritz, P. Jagadeeswaran, P. Choudary, P. A. Biro, J. Elder, J. DeRiel, J. Manley, M. Geftter, B. Forget, S. Weissman, Base substitution in an intervening sequence of a β^+ -thalassemic human globin gene. *Proc. Natl. Acad. Sci. U. S. A.* **78**, 2455–2459 (1981).
 82. G. Deidda, A. Novelletto, M. Hafez, Y. Al-Tonbary, L. Felicetti, L. Terrenato, B. Colombo, A new β -thalassemia mutation produced by a single nucleotide substitution in the conserved dinucleotide sequence of the IVS-I consensus acceptor site (AG $\rightarrow$ AA). *Hemoglobin.* **14**, 431–440 (1990).
 83. J. Vadolas, M. Nefedov, H. Warden, S. Mansooriderakshan, L. Voullaire, D. Jamsai, R. Williamson, P. A. Ioannou, Humanized β -Thalassemia Mouse Model Containing the Common IVSI-110 Splicing Mutation. *J. Biol. Chem.* **281**, 7399–7405 (2006).
 84. S. M. Derakhshan, M. S. Khaniani, Restoration of correct splicing in IVSI-110 mutation of β -globin gene with antisense oligonucleotides: implications and applications in functional assay development. *Iran. J. Basic Med. Sci.* **20**, 700–707 (2017).
 85. J. Wayne, B. Eng, M. Patterson, P. Wasi, D. Chui, W. Francombe, G. Sher, N. Olivieri, Severity of β -Thalassemia Due to Genotypes Involving the IVS-I-6 (T $\rightarrow$ C) Mutation. *Am. J. Hematol.* **50**, 15–19 (1995).
 86. R. Treisman, S. Orkin, T. Maniatis, Specific transcription and RNA splicing defects in five cloned β -thalassaemia genes. *Nature.* **302**, 591–596 (1983).

87. M. A. El-Latif, D. Filon, D. Rund, A. Oppenheim, M. Kanaan, The β^{+} -IVS-I-6 (T->C) mutation accounts for half of the thalassemia chromosomes in the Palestinian populations of the mountain regions. *Hemoglobin*. **26**, 33–40 (2002).
88. G. P. Tamagnini, P. Gonçalves, M. L. Ribeiro, J. Kaeda, F. Kutlar, E. Baysal, T. H. Huisman, β -thalassemia mutations in the portuguese; high frequencies of two alleles in restricted populations. *Hemoglobin*. **17**, 31–40 (1993).
89. P. Lahiry, S. Al-Attar, R. Hegele, Understanding Beta-Thalassemia with Focus on the Indian Subcontinent and the Middle East. *Open Hematol. J.* **2**, 5–13 (2008).
90. L. Romão, Â. Inácio, S. Santos, M. Ávila, P. Faustino, P. Pacheco, J. Lavinha, Nonsense mutations in the human β -globin gene lead to unexpected levels of cytoplasmic mRNA accumulation. *Blood*. **96**, 2895–2901 (2000).
91. M. L. Ribeiro, E. Baysal, F. Kutlar, G. P. Tamagnini, P. Gonçalves, D. Lopes, T. H. Huisman, A novel β° -thalassaemia mutation (codon 15, TGG->TGA) is prevalent in a population of central Portugal. *Br. J. Haematol.* **80**, 567–568 (1992).
92. W. Bouyarmene, J. Uwingabiye, A. Biaz, A. Rachid, Y. Mechali, A. Dami, S. Bouhsain, Z. Ouzzif, S. El Machtani Idrissi, Unexpected discovery of hemoglobinopathy C/ β° thalassemia. *Clin. Case Reports*. **6**, 2117–2120 (2018).
93. J. M. Cabeda, C. Correia, A. Estevinho, C. Simões, M. L. Amorim, L. Pinho, B. Justiça, Unexpected pattern of β -globin mutations in β -thalassaemia patients from northern Portugal. *Br. J. Haematol.* **105**, 68–74 (1999).
94. J. Chang, A. Alberti, Y. Kan, A β -thalassemia lesion abolishes the same Mst II site as the sickle mutation. *Nucleic Acids Res.* **11**, 7789–7794 (1983).
95. T. N. Williams, S. L. Thein, Sickle Cell Anemia and Its Phenotypes. *Annu. Rev. Genomics Hum. Genet.* **19**, 113–147 (2018).
96. M. Murao, M. H. C. Ferraz, Traço falciforme - heterozigose para hemoglobina S. *Rev. Bras. Hematol. Hemoter.* **29**, 223–225 (2007).
97. H. Itano, J. Neel, A new inherited abnormality of human hemoglobin. *Proc. Natl. Acad. Sci. U. S. A.* **36**, 613–617 (1950).
98. B. J. Bain, in *Haemoglobinopathy Diagnosis*, B. J. Bain, Ed. (Blackwell Publishing Ltd, 2^a Ed., 2006; <https://doi.org/10.1002/9780470988787.ch5>), Wiley Online Books, pp. 190–233.
99. B. J. Wild, B. J. Bain, in *Dacie and Lewis Practical Haematology: Twelfth Edition*, E. Lewis, B. J. Bain, I. Bates, M. Laffan, Eds. (Elsevier Limited, 12^a Ed., 2017), pp. 282–311.
100. A. R. Feeling-Taylor, S.-T. Yau, D. N. Petsev, R. L. Nagel, R. E. Hirsch, P. G. Vekilov, Crystallization mechanisms of hemoglobin C in the R state. *Biophys. J.* **87**, 2621–2629 (2004).
101. A. Perumbeti, in *Pathobiology of Human Disease: A Dynamic Encyclopedia of Disease Mechanisms*, L. Mcmanus, R. Mitchell, Eds. (Academic Press, 1^a Ed., 2014), pp. 1506–1531.
102. B. J. Bain, in *Blood Cells*, B. J. Bain, Ed. (John Wiley & Sons Ltd, 5^a Ed., 2015; <https://doi.org/10.1002/9781118817322.ch8>), Wiley Online Books, pp. 295–415.
103. Y. M. Yousafzai, S. Khan, F. Raziq, β -thalassaemia trait: haematological parameters. *J. Ayub Med. Coll.* **22**, 84–86 (2010).
104. J. Gao, S. A. Monaghan, in *Hematopathology: A Volume in the Series: Foundations in Diagnostic Pathology*, E. D. Hsi, J. Goldblum, Eds. (Elsevier Inc., Philadelphia, 3^a Ed., 2018; <https://www.sciencedirect.com/science/article/pii/B978032347913400001X>), pp. 3–56.

105. Proceedings of the 24th Annual Meeting of the Portuguese Society of Human Genetics (SPGH - Sociedade Portuguesa de Genética Humana). *Medicine (Baltimore)*. **100**, e23585 (2021).

7. Anexos



GOVERNO DE
PORTUGAL
MINISTÉRIO DA SAÚDE

Instituto Nacional de Saúde Doutor Ricardo Jorge, I.P.
Av. Padre Cruz, 1649-016 Lisboa Portugal
tel. +351 217 519 200 fax. +351 217 526 400
http://www.insa.pt info@insa.min-saude.pt



Instituto Nacional de Saúde
Doutor Ricardo Jorge



INSEF
Inquérito Nacional de Saúde
com Exame Físico - 2013/2016

Declaração de Consentimento Informado - Informação ao participante

Por favor, leia com atenção a seguinte informação. Se considera que algo não está suficientemente claro, não hesite em solicitar mais informações. Se concordar com a proposta que lhe é feita para participar neste estudo, assine este documento.

Título do estudo
Inquérito Nacional de Saúde com Exame Físico (INSEF)

Objetivo do estudo
O Inquérito Nacional de Saúde com Exame Físico (INSEF) tem como objetivo contribuir para o conhecimento da saúde da população portuguesa, dos fatores que determinam esse estado de saúde, e a utilização que os portugueses fazem dos cuidados de saúde, incluindo a utilização de medicamentos e dos serviços de saúde. A informação recolhida permitirá conhecer o estado de saúde da população portuguesa e assim, poder contribuir para a melhoria dos cuidados de saúde prestados, promovendo uma melhor qualidade de vida das atuais e futuras gerações.

Participação no INSEF
Este estudo envolve 4200 indivíduos com idades entre os 25 e os 74 anos, residentes em Portugal Continental e nas Regiões Autónomas, que foram aleatoriamente selecionados a partir da lista de utentes do Serviço Nacional de Saúde (SNS). No caso de aceitar participar neste estudo e assinar este consentimento informado, estará a concordar em:

- Permitir que um profissional de saúde efetue um exame físico que consiste na medição da sua altura, peso, perímetro da cintura e da anca e pressão arterial;
- Permitir a colheita de uma amostra de sangue venoso de aproximadamente 20 ml, a qual será utilizada para realizar análises clínicas (hemograma, hemoglobina glicosilada, triglicéridos, colesterol total, HDL e LDL);
- Responder a um questionário sobre o seu estado de saúde, determinantes de saúde e utilização de medicamentos e dos serviços de saúde;
- Permitir que os seus dados, previamente anonimizados, sejam analisados e armazenados em instalações seguras no Instituto Nacional de Saúde Doutor Ricardo Jorge (INSA), I.P., para que se consigam cumprir os objetivos do presente estudo, sem que seja possível identificar diretamente cada participante.

A sua participação neste estudo é gratuita e voluntária, podendo interromper a sua colaboração assim que achar conveniente, sem necessidade de dar qualquer justificação e não sofrendo por isso, qualquer consequência ao nível da prestação dos cuidados de saúde. Deverá manifestar por escrito a sua vontade em interromper a participação no estudo, sendo que nesta situação as informações recolhidas e as suas amostras serão imediatamente destruídas.

Figura 7.1. Declaração de Consentimento Informado – Informação ao Participante (Parte I). Documento informativo sobre o Inquérito Nacional de Saúde com Exame Físico (INSEF).

Confidencialidade e anonimização

Este estudo foi previamente autorizado pela Comissão Nacional de Proteção de Dados, pela Comissão de Ética do INSA e pelas Comissões de Ética das Administrações Regionais de Saúde. Todas as informações que prestar são confidenciais e será sempre salvaguardada a sua privacidade e anonimato. Apenas os investigadores devidamente autorizados terão acesso à informação e às amostras deste estudo, sendo esse acesso sujeito à supervisão e aprovação prévia do responsável do INSEF e da Comissão de Ética do INSA.

Criação de um Biobanco no INSA

O Biobanco do INSA consiste numa coleção de material biológico com informação associada que é armazenada num sistema organizado e monitorizado, de acesso restrito, obedecendo aos mais elevados padrões de qualidade ética e técnica, aceites a nível internacional. É um recurso fundamental para apoiar a investigação em saúde, de forma sustentada, possibilitando a partilha de material biológico e informação pela comunidade científica nacional e internacional. Todos os investigadores que pretenderem fazer uso das amostras armazenadas no Biobanco do INSA terão de submeter os seus projetos a uma Comissão de Ética e a uma Comissão Científica que avaliará a pertinência dos mesmos. Assim, caso concorde, a sua amostra de sangue, incluindo os seus derivados (soro, plasma e DNA) e a informação associada serão armazenadas para fins de investigação científica, passando a fazer parte do Biobanco do INSA. As amostras serão preservadas em condições apropriadas e as informações associadas serão introduzidas numa base de dados, passando a sua identificação pessoal a estar codificada.

Benefícios para os participantes

Caso concorde, os resultados das análises clínicas serão enviados ao seu médico assistente para que ele possa incluir esta informação no acompanhamento que faz da sua saúde. Para além disto, a sua participação proporcionará a aquisição de conhecimentos que poderão vir a beneficiá-lo a si, à sua família e a toda a população portuguesa.

Riscos previsíveis

Os riscos associados à sua participação neste estudo são mínimos ou inexistentes. Apenas a colheita da amostra de sangue poderá envolver riscos e desconforto para o participante, inerentes ao procedimento em si, e não mais do que o desconforto normal nestas situações. No entanto, a colheita de sangue será efetuada por um profissional de saúde experiente.

Detalhes de contacto para mais informações

Se tiver qualquer dúvida sobre este estudo poderá, em qualquer momento, contactar-nos por telefone (217526487), e-mail (insef@insa.min-saude.pt) ou correio (INSEF, Departamento de Epidemiologia, Instituto Nacional de Saúde Doutor Ricardo Jorge Av. Padre Cruz, 1649-016 Lisboa) dirigindo-se ao Coordenador Geral do INSEF, Doutor Carlos Matias Dias. Poderá também obter informações detalhadas em www.insef.pt.

Figura 7.2. Declaração de Consentimento Informado – Informação ao Participante (Parte II). Documento informativo sobre o Inquérito Nacional de Saúde com Exame Físico (INSEF).

Declaração de Consentimento Informado

Declaro que li, tive oportunidade de considerar e compreendi toda a informação previamente fornecida sobre o **Inquérito Nacional de Saúde com Exame Físico**. Fui informado dos riscos e benefícios da minha participação neste estudo e tive a possibilidade de colocar questões e esclarecer as minhas dúvidas. Compreendo que a minha participação é voluntária, podendo retirar-me do estudo a qualquer momento, sem qualquer repercussão em relação aos cuidados médicos ou de saúde que posso receber, bastando para isso escrever para o Departamento de Epidemiologia do Instituto Nacional de Saúde Dr. Ricardo Jorge, Avenida Padre Cruz, 1649-016 Lisboa.

No âmbito deste estudo, declaro que aceito participar, voluntariamente, e concordo especificamente com os seguintes pontos:

Autorizo ser contactado novamente pelo INSA, IP, para participar num estudo de seguimento.	Sim ☐ Não ☐
Autorizo a conservação das minhas amostras no Biobanco do INSA, I.P., e da informação associada de modo a que possam ser utilizadas para estudos futuros devidamente autorizados.	Sim ☐ Não ☐
Autorizo a ligação futura dos dados, agora recolhidos, a outros registos nacionais, de forma anonimizada, para fins de investigação científica devidamente autorizada.	Sim ☐ Não ☐
Autorizo que os meus resultados das análises de rotina sejam enviados para o meu médico assistente.	Sim ☐ Não ☐

Concordo em participar neste estudo de livre vontade e receberei uma cópia deste consentimento informado após a sua assinatura.

Nome legível: _____

Data: ____/____/____ Assinatura: _____

Confirmação do Investigador

Confirmo que descrevi o Inquérito Nacional de Saúde com Exame Físico ao participante, utilizando uma linguagem compreensível e apropriada, tendo respondido a todas as questões colocadas. Informei adequadamente o participante sobre os objetivos do estudo e sobre os seus possíveis benefícios e riscos, considerando que o participante compreendeu a minha explicação.

Nome legível: _____

Data: ____/____/____ Assinatura: _____

Por favor, assine as duas cópias deste formulário e entregue esta ao participante.

Figura 7.3. Declaração de Consentimento Informado. Documento assinado por todos os participantes do Inquérito Nacional de Saúde com Exame Físico (INSEF).

Tabela 7.1. Valores de referência dos parâmetros hematológicos para indivíduos adultos normais

Sexo	Eritrócitos (x10 ¹² /L)	Hemoglobina (g/dL)	VGM (fL)	HGM (pg)	CHGM (g/dL)	Hematócrito (%)	RDW (%)	Leucócitos (x10 ⁹ /L)	Plaquetas (x10 ⁹ /L)
Masculino	4,3-5,7	13,3-16,7	82-98	27-33	32-36	39-50	9,9-15,5	3,7-9,5	140-440
Feminino	3,9-5,0	11,8-14,8	82-98	27-33	32-36	39-50	9,9-15,5	3,9-11,1	140-440

VGM – volume globular médio; **HGM** – hemoglobina globular média; **CHGM** – concentração da hemoglobina globular média; **RDW** – *red cell distribution width*. Adaptado de (102).

Tabela 7.2. Condições do PCR convencional para a amplificação do fragmento de interesse no gene *HBB*

Gene	Primers				Mistura de Reação			Condições do PCR		Tamanho do fragmento (pb)
	ID	Sequência	Tamanho (nt)	Tm (°C)	Reagente	Concentração	Volume (µL)	Temperatura (°C)	Tempo	
<i>HBB</i>	L5'β (Fw)	5'-TAAGCCAGTGCCAGAAGAG-3'	19	56,8	Tampão David	10x	2,5			2106
					BSA	10 mg/mL	0,35	94	5 min	
					dNTPs	100 mM	0,5	94	45 seg	
					Primer L5'β (Fw)	25 pmol/µL	0,5	63	45 seg	
	14 (Rv)	5'-GAGTCAAGGCTGAGAGATGCAGGA-3'	24	64,4	Primer 14 (Rv)	25 pmol/µL	0,5	72	90 seg	
					H ₂ O	-	19,75	72	10 min	
					AmpliTaq	5 U/ µL	0,2	4	10 min	
					DNA	≈ 70 ng	0,7			

ID – Identificação; **Tm** – Temperatura de fusão; **nt** – nucleótidos; **pb** – pares de base; **Fw** – *forward* (sentido direto); **Rv** – *reverse* (sentido reverso); **BSA** – Albumina do soro de bovino; **dNTPs** – Desoxirribonucleótidos Trifosfatados; **min** – minutos; **seg** – segundos.

Tabela 7.3. Condições do Gap-PCR para a detecção da deleção – α ^{3,7}

Gene	Primers				Mistura de Reação			Condições do PCR		Tamanho do fragmento (pb)														
	ID	Sequência	Tamanho (nt)	Tm (°C)	Reagente	Concentração	Volume (µL)	Temperatura (°C)	Tempo															
HBA2 e HBA1	C10 (Fw)	5' - GGGATGCACCCACTGGCACT - 3'	20	63,5	Mistura I	Tampão α	10x	2,5	94	5 min	x34	1900												
						DMSO	5%	2,5					94	60 seg										
						BSA	10 mg/mL	0,4							63	60 seg								
						MgCl ₂	0,1 M	0,5									72	90 seg						
						dNTPs	100 mM	0,5											72	10 min				
						Primer C10 (Fw)	25 pmol/µL	0,5													4	10 min		
						Primer C3 (Rv)	25 pmol/µL	0,5															2100	
						H ₂ O	-	16,6																
	C3 (Rv)	5' - CTCCATTGTTGGCACATTCCGGG - 3'	23	64,2		Ampli ^{Taq}	5 U/ µL	0,3																
						DNA	≈ 70 ng	0,7																
				Mistura II	Tampão α	10x	2,5	2100																
					DMSO	10%	2,5																	
					BSA	10 mg/mL	0,4																	
					MgCl ₂	0,1 M	0,5																	
					dNTPs	100 mM	0,5																	
					Primer C10 (Fw)	25 pmol/µL	0,5																	
					Primer C2 (Rv)	25 pmol/µL	0,5																	
	C2 (Rv)	5' - CTGCTGTCCACGCCCATGCC - 3'	20		65,5		H ₂ O		-	16,6														
							Ampli ^{Taq}		5 U/ µL	0,3														
							DNA		≈ 70 ng	0,7														

ID – Identificação; **Tm** – Temperatura de fusão; **nt** – nucleótidos; **pb** – pares de base; **Fw** – *forward* (sentido direto); **Rv** – *reverse* (sentido reverso); **DMSO** – Dimetilsulfóxido; **BSA** – Albumina do soro de bovino; **MgCl₂** – Cloreto de magnésio; **dNTPs** – Desoxirribonucleótidos Trifosfatados.

Tabela 7.4. Composição das soluções tampão utilizadas

Tampão David (10x)		Tampão α (10x)		Tampão de deposição da eletroforese	
(NH ₄) ₂ SO ₄	166 mM	(NH ₄) ₂ SO ₄	166 mM	H ₂ O	8,75 μ L
Tris-HCl (pH = 8,8)	670 mM	Tris-HCl (pH = 8,8)	670 mM	Azul de Bromofenol	0,22 g
MgCl ₂	67 mM	MgCl ₂	15 mM	Glicerol	18,75 μ L
EDTA	0,067 mM	EDTA	0,67 mM	EDTA	75 μ L
β -Mercaptoetanol	100 mM	β -Mercaptoetanol	100 mM	NaOH	5 M

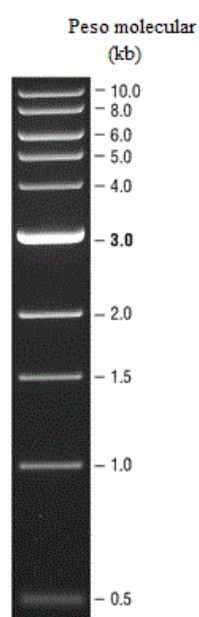


Figura 7.4. Marcador de massa molecular usado na eletroforese em gel de agarose. Marcador 1 kb DNA Ladder (BioLabs) corrido num gel de agarose a 0,8%. O peso molecular de cada banda está indicado em kb.

Tabela 7.5. Sondas MLPA do agrupamento génico da α -globina dispostas de acordo com a sua localização cromossómica

Tamanho (nt)	Sonda SALSA MLPA	Gene / Exão	Sequência Parcial (24 nt adjacentes ao local de ligação)	Distância até à sonda seguinte
-	-	-	Do telómero à sonda POLR3K	37,2 kb
463	19236-L25316	<i>POLR3K</i>	ATATGCTCTGTG-TAAAGTCCTATT	66,4 kb
178	04799-L04797	HS-40	TGCCCCAAGCCA-AGGGTGGAGGCA	0,2 Kb
382	04800-L04175	HS-40	GGTACTGCTGAT-TACAACCTCTGG	30,0 kb
364	04926-L23886	9,2 kb a montante de <i>HBZ</i>	GGCTGGGGCTCA-AACCAAGGCCCA	5,7 kb
346	04622-L04001	3,5 kb a montante de <i>HBZ</i>	CGCAGTGCTAGA-AGGGAGTTCCTG	3,3 kb
436	17214-SP0457-L20489	0,2 kb a montante de <i>HBZ</i>	TTCCTCTCCTGT-37 nt abrangendo oligo-AGTCTAGGAGAG	6,9 kb
292	04624-L04004	Entre <i>HBZ</i> e <i>HBZP1</i>	GTGGAGTAGGCT-TTGTGGGGAAGT	7,8 kb
184	04637-L04018	Entre <i>HBAP2</i> e <i>HBAP1</i>	AGTGGCCACAAT-TTGGCAGACAGA	2,5 kb
391	18097-L22521	3,0 kb a montante de <i>HBA2</i>	AGGAAGGGGTGA-GAATGAGAGAAA	0,5 kb
373	18090-L08415	2,5 kb a montante de <i>HBA2</i>	ATGTCCAGAAGA-AAAGCGGTGACA	1,6 kb
147	18098-L22522	0,9 kb a montante de <i>HBA2</i>	GATAAACAAACT-TGGCTCTGGGTA	0,2 kb
328	18092-L22516	0,6 kb a montante de <i>HBA2</i>	CCGGGAAGGAAC-AAACACCAGGAC	0,6 kb
220	18099-L22524	<i>HBA1</i> + <i>HBA2</i> , exão 1	AAGAGTGCCGGG-CCGCGAGCGCGC	0,1 kb
214	18881-L06288	<i>HBA1</i> + <i>HBA2</i> , exão 1	CGCGCCGACCTT-ACCCCAGGCGGC	0,4 kb
160	08498-L08422	<i>HBA2</i> , intrão 2	GCGCCTTCCTCT-CAGGGCAGAGGA	0,1 kb
244	04633-L23748	<i>HBA2</i> , intrão 2	GGGCCTGGGCCG-CACTGACCCTCT	0,1 kb
172	15857-L21812	<i>HBA1</i> + <i>HBA2</i> , exão 3	GCAGGGGTGAAC-TCGGCGGGGAGG	0,1 kb
136	S0585-SP0043-L09493	Mutação <i>Hb Constant Spring</i>	CCAAATACCGTC-28 nt abrangendo oligo-TGCCCCGCTGGGC	0,2 kb
190	18096-L22520	2,5 kb a montante de <i>HBA1</i> e 0,4 kb a jusante de <i>HBA2</i>	TTCTCTCATTCC-CACCCCTTCCTG	0,5 kb
202	18880-L24428	2,0 kb a montante de <i>HBA1</i>	GTCAGTGCTTTC-CTTCTGGACATG	0,6 kb
256	08494-L08417	1,5 kb a montante de <i>HBA1</i>	TTCTCTGCCCAA-GGCAGCTTACCC	0,6 kb
337	14855-L23604	0,9 kb a montante de <i>HBA1</i>	CTGGAGCATTCA-ACCTCCTCTGGG	0,3 kb
226	18093-L22517	0,6 kb a montante de <i>HBA1</i>	TCCTGGTGTTTA-TTCCTTCCCGGT	0,6 kb
220	18099-L22524	<i>HBA1</i> + <i>HBA2</i> , exão 1	AAGAGTGCCGGG-CCGCGAGCGCGC	0,1 kb
214	18881-L06288	<i>HBA1</i> + <i>HBA2</i> , exão 1	CGCGCCGACCTT-ACCCCAGGCGGC	0,4 kb
165	08498-L21607	<i>HBA1</i> , intrão 2	GCGCCTTCCTCG-CAGGGCAGAGGA	0,1 kb
250	04633-L23600	<i>HBA2</i> , intrão 2	GGCCCTCGGCC-CACTGACCCTCT	0,1 kb
172	15857-L21812	<i>HBA1</i> + <i>HBA2</i> , exão 3	GCAGGGGTGAAC-TCGGCGGGGAGG	0,3 kb
154	08499-L23594	0,1 kb a jusante de <i>HBA1</i>	TGGGACACACAT-GGCTAGAACCTC	0,3 kb
283	04638-L23602	0,4 kb a jusante de <i>HBA1</i>	AAGTCCCCTCC-AGCATGGCTGCA	1,9 kb
310	04639-L04020	2,3 kb a jusante de <i>HBA1</i>	GTTCAGTCCCT-GAAGAAACACCT	1,4 kb
400	19233-L25313	<i>HBQ1</i> , exão 3	TGCTCTCTCGAG-GTCAGGACGCGA	25,1 kb
277	15859-L21960	<i>LUC7L</i>	ATGTTCCAATGA-AACCAGTGGCAC	33,5 kb
445	17227-L20554	<i>ITFG3</i>	GCTGTGATACTT-TTGCCTTTGTCA	31,9 kb
472	18102-L20488	<i>RGS11</i>	GAACCTGAGGTT-CCGTGGAATATT	16,3 kb
418	17212-L13393	<i>AXIN1</i>	GATCATCGGCAA-AGTGGAGAAGGT	119,4 kb
262	17613-L23601	<i>DECR2</i>	CAGAATCGACAT-TCTCATTAAGT	-

nt – nucleótidos; kb – quilo bases. Adaptado de (51).

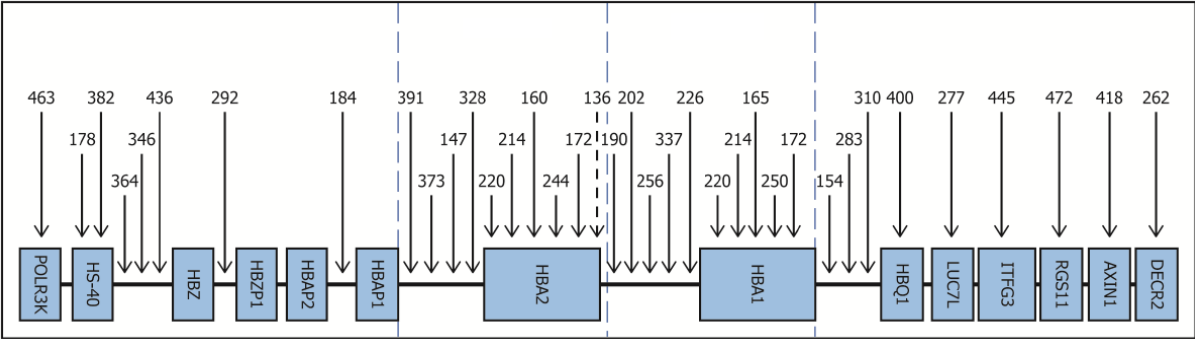


Figura 7.5. Sondas MLPA do agrupamento gênico da α -globina. Representação esquemática da localização das sondas MLPA. As caixas a azul representam os genes do agrupamento gênico da α -globina. Os números em cima das setas indicam o tamanho do fragmento amplificado (em nt) das respectivas sondas. As duas regiões entre as linhas azuis tracejadas são regiões que apresentam bastantes semelhanças quanto às sondas. Adaptado de (51).

Tabela 7.6. Condições do MLPA para a detecção de CNV no agrupamento gênico da α -globina

Mistura de Reação				Condições da Reação			
Mistura	Reagente	Volume (μL)	Volume por tubo (μL)	Passo	Temperatura (°C)	Tempo	
M0	RNase (0,5 mg/mL) DNA (≈ 70 ng) H ₂ O	1,0	5	Digestão com RNase	37	30 min	
		1,0		Desnaturação do DNA	98	10 min	
		3,0			25	Pausa	
MI	SALSA probe mix Tampão MLPA	1,5	3	Hibridação das sondas	95	1 min	
		1,5			60	16 h	
MII	H ₂ O Ligase-65 buffer A Ligase-65 buffer B Ligase-65	25,0	32	Ligação das sondas	54	Pausa	
		3,0			54	15 min	
		3,0			98	5 min	
		1,0			20	Pausa	
MIII	H ₂ O Primer mix SALSA polimerase	7,5	10	Amplificação das sondas por PCR	95	30 seg	x35
		2,0			60	30 seg	
		0,5			72	60 seg	
					72	20 min	
					15	Pausa	

Todos os reagentes de MI, MII e MII pertencem ao kit SALSA® MLPA® Probemix P140-C1 HBA. Adaptado de (51).

Tabela 7.7. Condições para a Sequenciação Automática pelo método de Sanger

Mistura de Reação			Condições da Reação	
Reagente	Concentração	Volume (μL)	Temperatura (°C)	Tempo
Tampão Big Dye	5x	1,75	96	4 min
BigDye	2x	0,50	96	10 seg
Primer Fw ou Rv	2 pmol/μL	1,00	55	5 seg
H ₂ O	-	5,25	60	4 min
Produto PCR Purificado	-	1,50	60	8 min
			4	10 min

Primer Fw – oligonucleótido iniciador de sentido direto; **Primer Rv** – oligonucleótido iniciador de sentido reverso

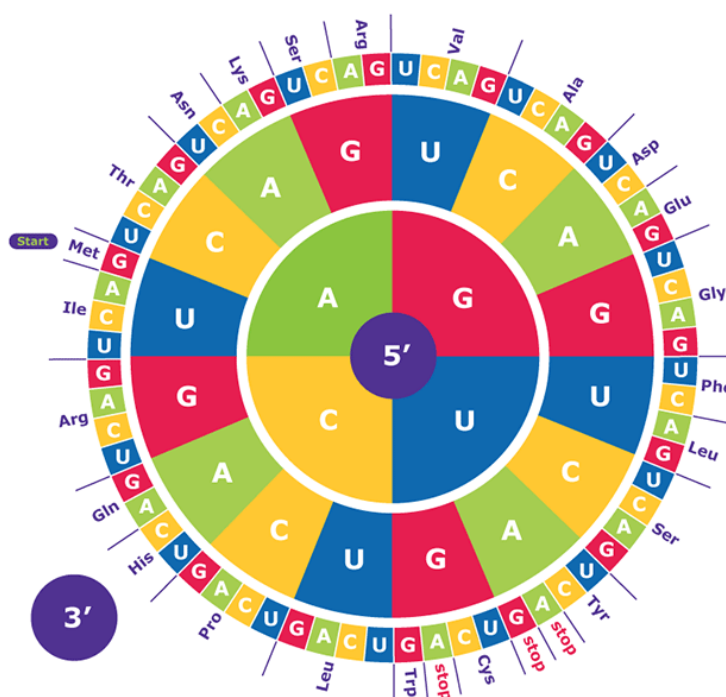


Figura 7.6. Código genético. Relação entre a sequência de nucleótidos no DNA e os respectivos aminoácidos na proteína: um conjunto de 3 nucleótidos (codão) corresponde a um aminoácido, um codão de iniciação (*start*) ou um codão de terminação (*stop*). Retirado de (<https://www.sigmaaldrich.com/technical-documents/articles/biology/amino-acid-codon-wheel.html>).

Tabela 7.8. Código de nucleótidos da União Internacional de Química Pura e Aplicada (IUPAC)

Símbolo	Nucleótido	Descrição
G	G	Guanina
A	A	Adenina
T	T	Timina
C	C	Citosina
R	G ou A	Purina
Y	T ou C	Pirimidina
M	A ou C	Amina
K	G ou T	Cetona
S	G ou C	Interação forte
W	A ou T	Interação fraca
H	A ou C ou T	Não G (H vem depois do G no alfabeto)
B	G ou T ou C	Não A (B vem depois do A no alfabeto)
V	G ou C ou A	Não T/não U (V vem depois do U no alfabeto)
D	G ou A ou T	Não C (D vem depois do C no alfabeto)
N	G ou A ou T ou C	Qualquer nucleótido
. ou -	Nenhum	Lacuna (gap)

Tabela 7.9. Código de aminoácidos da IUPAC

Símbolo	Codão	Aminoácido
A	Ala	Alanina
C	Cys	Cisteína
D	Asp	Ácido Aspártico
E	Glu	Ácido glutâmico
F	Phe	Fenilalanina
G	Gly	Glicina
H	His	Histidina
I	Ile	Isoleucina
K	Lys	Lisina
L	Leu	Leucina
M	Met	Metionina
N	Asn	Asparagina
P	Pro	Prolina
Q	Gln	Glutamina
R	Arg	Arginina
S	Ser	Serina
T	Thr	Treonina
V	Val	Valina
W	Trp	Triptofano
Y	Tyr	Tirosina

Tabela 7.10. Variantes detetadas por NGS nos 204 indivíduos analisados neste estudo

Posição (GRCh38.p12)	rs_ID	Nome Comum	Região	Consequência molecular	NM_000518.5	NP_000509.1
g.5225245	rs112988270	-	3'ÚTR	Variante Intrónica	c.*132+221G>A	-
g.5225282	rs7110263	-	3'ÚTR	Variante Intrónica	c.*132+184A>C	-
g.5225284	rs12574989	-	3'ÚTR	Variante Intrónica	c.*314G>A	-
g.5225302	rs570662275	-	3'ÚTR	Variante Intrónica	c.*296T>C	-
g.5225365	rs12788013	-	3'ÚTR	Variante Intrónica	c.*233G>C	-
g.5225367	rs567259408	-	3'ÚTR	Variante Intrónica	c.*132+99T>C	-
g.5225416	rs113969885	-	3'ÚTR	Variante Intrónica	c.*182G>A	-
g.5225610	rs36020563	-	Exão 3	Variante Sinónima	c.432C>T	p.His144=
g.5225678	rs33946267	Hb D-Portugal	Exão 3	Variante <i>Missense</i>	c.364G>C	p.Glu122Gln
g.5225754	rs771736852	-	Intrão 2	Variante Intrónica	c.316-28G>A	-
g.5225911	rs1609812	-	Intrão 2	Variante Intrónica	c.316-185C>T	-
g.5225964	rs558554234	-	Intrão 2	Variante Intrónica	c.316-238C>T	-
g.5226099	rs78815705	IVS II-478 (C>A)	Intrão 2	Variante Intrónica	c.316-373C>A	-
g.5226496	rs7946748	-	Intrão 2	Variante Intrónica	c.315+81C>T	-
g.5226503	rs7480526	IVS-II-74 (T>G)	Intrão 2	Variante Intrónica	c.315+74T>G	-
g.5226551	rs368604295	-	Intrão 2	Variante Intrónica	c.315+26T>G	-
g.5226561	rs10768683	IVS-II-16 (G>C)	Intrão 2	Variante Intrónica	c.315+16G>C	-
g.5226646	rs145669504	-	Exão 2	Variante Sinónima	c.246C>T	p.Leu82=
g.5226762	rs80356821	Cd41/42 (-CTTT)	Exão 2	Variante <i>Frameshift</i>	c.126_129del	p.Phe42LeufsTer19
g.5226774	rs11549407	Cd39 (C>T)	Exão 2	Variante <i>Nonsense</i>	c.118C>T	p.Gln40Ter
g.5226820	rs35004220	IVS-I-110 (G>A)	Intrão 1	Variante Intrónica	c.93-21G>A	-
g.5226905	-	-	Intrão 1	Variante Intrónica	c.92+25G>T	-
g.5226924	rs35724775	IVS-I-6 (T>C)	Intrão 1	Variante Intrónica	c.92+6T>C	-
g.5226929	rs33971440	IVS-I-1 (G>A)	Intrão 1	Variante Intrónica	c.92+1G>A	-
g.5226974	rs34716011	Cd15 (G>A)	Exão 1	Variante <i>Nonsense</i>	c.48G>A	p.Trp16Ter
g.5227001	rs63749819	Cd6 (-A)	Exão 1	Variante <i>Frameshift</i>	c.20del	p.Glu7GlyfsTer13
g.5227002	rs334	Hb S	Exão 1	Variante <i>Missense</i>	c.20A>T	p.Glu7Val
g.5227003	rs33930165	Hb C	Exão 1	Variante <i>Missense</i>	c.19G>A	p.Glu7Lys
g.5227013	rs713040	-	Exão 1	Variante Sinónima	c.9T>C	p.His3=

rs_ID – Identificação dos SNPs segundo a referência da base de dados *dbSNP*; **NM** – sequência de referência com base num RNA codificante de proteína (mRNA); **NP** – sequência de referência com base na proteína.

Tabela 7.11. Sexo, idade e características hematológicas dos indivíduos identificados por NGS como portadores de uma variante no gene *HBB*

Nº de participante	Sexo	Idade (anos)	Eritrócitos (x10 ¹² /L)	Hemoglobina (g/dL)	VGM (fL)	HGM (pg)	CHGM (g/dL)	Hematócrito (%)	RDW (%)	Leucócitos (x10 ⁹ /L)	Plaquetas (x10 ⁹ /L)
PTI 18	Homem	54	6,10	12,1	61,8	19,8	32,1	37,7	16,0	7,84	158
PTI 28	Homem	63	5,26	12,1	70,9	23,0	32,5	37,3	16,0	4,70	222
PTI 30	Mulher	69	6,06	11,1	56,9	18,3	32,1	34,5	16,8	5,60	198
PTI 32	Homem	70	6,11	14,8	75,4	24,2	32,1	46,1	14,5	8,40	132
PTI 40	Mulher	35	5,37	12,7	72,3	23,6	32,6	38,8	15,4	7,00	158
PTI 41	Homem	26	6,09	11,9	60,3	19,6	32,4	36,8	15,9	7,10	312
PTI 43	Mulher	50	6,23	12,8	64,1	20,5	32,1	39,9	16,0	5,40	235
PTI 47	Homem	63	4,30	11,1	83,0	25,9	31,2	35,7	16,4	9,90	275
PTI 88	Mulher	55	6,07	11,6	60,7	19,1	31,4	36,8	15,2	5,79	246
PTI 89	Homem	50	6,36	13,9	67,2	21,8	32,5	42,8	16,5	7,50	255
PTI 94	Mulher	58	5,46	11,4	66,3	20,9	31,5	36,2	16,5	5,46	227
PTI 105	Mulher	38	6,29	11,4	60,0	18,1	30,1	37,7	15,6	7,25	237
PTI 113	Homem	65	4,90	11,1	72,5	22,7	31,4	35,5	17,3	4,80	325
PTI 115	Mulher	56	6,40	12,8	65,0	20,1	31,0	41,5	15,4	10,80	292
PTI 124	Mulher	39	5,20	10,2	63,2	19,6	31,0	33,0	16,5	9,60	282
PTI 130	Mulher	56	5,40	11,5	68,3	21,4	31,4	36,7	15,8	4,90	195
PTI 142	Mulher	54	4,80	10,2	67,4	21,1	31,3	32,5	17,5	4,60	270
PTI 143	Homem	55	6,60	12,6	61,8	19,1	31,0	40,6	18,0	8,30	218
PTI 145	Homem	68	6,60	14,0	68,7	21,3	31,0	45,3	16,2	7,90	187
PTI 150	Homem	36	5,80	11,4	67,1	19,6	29,2	38,9	14,7	4,40	185
PTI 154	Homem	48	6,12	14,2	72,9	23,2	31,8	44,7	15,2	4,60	163
PTI 157	Homem	68	5,72	14,5	82,3	25,4	30,8	47,1	14,1	7,10	220
PTI 184	Homem	42	5,98	16,3	75,8	27,3	36,1	45,4	14,8	15,92	195
PTI 188	Mulher	27	5,13	10,8	66,9	21,1	31,5	34,3	15,2	10,57	301
PTI 196	Mulher	44	5,22	9,0	60,0	17,2	28,7	31,4	17,5	5,68	301

VGM – volume globular médio; **HGM** – hemoglobina globular média; **CHGM** – concentração da hemoglobina globular média; **RDW** – *red cell distribution width*.

Mutational Spectrum and Geographic Distribution of Alpha-thalassemia in an Adult Microcytic and/or Hypochromic Population Living in Portugal: Results from the First National Health Examination Survey (INSEF 2015)

Daniela Santos¹, Irina Kislaya^{2,3}, Pedro Lopes¹, Carlos Matias-Dias^{2,3}, Marta Barreto^{2,3}, Paula Faustino^{1,4}

1. Departamento de Genética Humana, Instituto Nacional de Saúde Doutor Ricardo Jorge, Lisboa

2. Departamento de Epidemiologia, Instituto Nacional de Saúde Doutor Ricardo Jorge, Lisboa

3. Centro de Investigação em Saúde Pública, Escola Nacional de Saúde Pública, Universidade NOVA de Lisboa, Lisboa

4. Instituto de Saúde Ambiental, Faculdade de Medicina, Universidade de Lisboa, Lisboa

Alpha-thalassemia (α -thal) is one of the most common monogenic disorders in the world. Its clinical severity varies from almost asymptomatic, mild microcytic hypochromic anemia, to a lethal hemolytic condition, depending on the number of affected α -globin genes (1 to 4). The disease is most commonly originated by deletions on 16p13.3. The aim of this study was to identify the molecular basis, geographic distribution and prevalence of mild forms of α -thal in Portugal.

This is a cross-sectional population-based study, based on the first Portuguese National Health Examination Survey (INSEF), which included individuals living in Portugal for more than 12 months, aged between 25 and 74 years old. For this INSEF sub-study, we analysed 4812 participants from whom a Complete Blood Count was performed and selected the 204 participants presenting red blood cell microcytosis (Mean Corpuscular Volume, MCV <80fL) and/or hypochromia (Mean Corpuscular Hemoglobin, MCH <27pg). DNA from these samples was used to search for deletions in *HBA* cluster by Gap-PCR and Multiplex Ligation-dependent Probe Amplification.

We found 52 individuals heterozygous for the $-\alpha^{3.7\text{kb}}$ deletion, one homozygous for this deletion and one heterozygous for the $-\alpha^{4.2\text{kb}}$ deletion. Two cases presented triplicated α -globin genes ($\alpha\alpha\alpha^{\text{anti } 3.7\text{kb}}$). Thus, α -thal was observed in 54 individuals (26.5 %) of the analysed population. Carriers of the $-\alpha^{3.7\text{kb}}$ deletion have hypochromic red blood cells (MCH mean 26.0 ± 0.9 pg) but normal or borderline volume (MCV mean 81.4 ± 2.7 fL). The geographic distribution of affected participants showed two regions with highest prevalence of α -thal: LVT and RA Madeira.

Although the mild forms of α -thal themselves are of no clinical significance, their major importance is the modifying effect that they have on various severe forms of β -thalassemia and sickle cell disease. Furthermore, α -thal trait can be confused with iron deficiency anemia as the hematological parameters are quite similar. Therefore, iron status should be properly assessed to distinguish between the two conditions and α -thal confirmation at DNA level is necessary for a definitive diagnosis.

Figura 7.7. Publicação em *Proceedings*. D. Santos, I. Kislaya, P. Lopes, C. Matias-Dias, M. Barreto, P. Faustino, Mutational Spectrum and Geographic Distribution of Alpha-thalassemia in an Adult Microcytic and/or Hypochromic Population Living in Portugal: Results from the First National Health Examination Survey (INSEF 2015). *Proceedings of the 24 th Annual Meeting of the Portuguese Society of Human Genetics. Medicine. 100, e23585_P33 (2021).* doi.org/10.1097/MD.00000000000023585 (105).