

UNIVERSIDADE DE LISBOA
Faculdade de Ciências
Departamento de Biologia Animal



**Modelação do Risco Trombótico da População
Portuguesa Portadora de FV Leiden por duas
Variantes Alélicas do gene do Fibrinogénio Gama**

Cátia Marlene Marques Correia Sousa Pinto

Mestrado em Biologia Humana e Ambiente

2010

UNIVERSIDADE DE LISBOA
Faculdade de Ciências
Departamento de Biologia Animal



**Modelação do Risco Trombótico da População
Portuguesa Portadora de FV Leiden por duas
Variantes Alélicas do gene do Fibrinogénio Gama**

Cátia Marlene Marques Correia Sousa Pinto

Dissertação orientada por:

Doutor Dezső David
Instituto Nacional de Saúde Dr. Ricardo Jorge

Professor Doutor Octávio S. Paulo
Faculdade de Ciências da Universidade de Lisboa

Mestrado em Biologia Humana e Ambiente

2010

Aos meus pais,
Com amor

Agradecimentos

Sei que os meus estudos vão aprofundar cada vez mais os meus conhecimentos, pois na ciência, assim como na vida, nunca se sabe tudo, nunca se aprende tudo. A conclusão desta tese é apenas mais um passo. Foram muitos os momentos de entusiasmo e alegria. Foram alguns os momentos de desânimo perante as dificuldades que surgiam! Mas nos bons e nos maus momentos, nunca me senti só! Sei que este trabalho só foi possível devido à existência dessas pessoas que estiveram ao meu lado e a quem eu gostaria de agradecer:

Ao Doutor David, pela amabilidade, paciência e sabedoria com que sempre me demonstrou qual o melhor caminho para atingir os meus objectivos.

Ao professor Octávio Paulo, por ter aceite ser o meu orientador e por todas as indicações e ensinamentos, que tornaram este trabalho muito mais fácil de realizar.

Às coordenadores deste mestrado em Biologia Humana e Ambiente, professora Deodália Dias e professora Ana Crespo, por demonstrarem sempre um sorriso amigo.

À Doutora Teresa Gago, do Hospital de Santa Cruz, por toda a disponibilidade demonstrada em fornecer as amostras e as indicações clínicas imprescindíveis para este estudo.

À Isabel, à Célia, à Barbara e à Cristina pela ajuda prestada na realização deste trabalho.

À Catarina, que mais do que uma colega foi uma verdadeira amiga, por me ajudar sempre, em especial nos primeiros passos, quando eu apenas sabia gatinhar.

À Inês e ao Carlos, companheiros de aventuras científicas no laboratório, pela boa disposição e mão amiga sempre que era necessário, tanto para desabafar como para carregar um gel de electroforese.

A todos os meus amigos de curso, que foram a minha alegria durante estes 5 anos... não posso mencionar o nome de todos, mas estão todos no meu coração.

Ao André, por estar sempre do meu lado e me ensinar que o verdadeiro poder está dentro de mim, que eu posso alcançar o Universo... e por ser o meu sorriso. Bom Dia.

Aos meus pais por serem a minha força e me suportarem, quando o mundo parecia desabar em cima de mim; por serem a minha luz, quando andava cega de ansiedade e stress.

Resumo

Actualmente, a trombose venosa (TV), é uma das causas mais comuns de morbidade e mortalidade a nível mundial. A TV é uma doença poligénica e multifactorial, causada por alterações em factores de risco *major*, como o Factor V Leiden, proteínas pró-coagulantes, ou outros factores de risco *minor*, como o Fibrinogénio. Apesar da mutação FV Leiden representar a base molecular de TV em cerca de 8% dos casos, é possível verificar a existência de heterogenicidade no tipo e na idade da primeira manifestação clínica. Este facto sugere que existe uma modulação por um conjunto de factores hereditários adicionais e adquiridos.

O objectivo deste estudo é avaliar o efeito modulador de potenciais factores hereditários e ambientais no risco trombótico de indivíduos portadores de FV Leiden. Os polimorfismos FGG 10034 C>T (rs2066865) e FGG 9340 T>C (rs1049636) do gene *FGG* são bons candidatos para esta função moduladora.

Embora não tenham sido verificadas diferenças estatisticamente significativas entre as frequências alélicas e genotípicas, os dados preliminares adquiridos neste estudo apontam para uma modulação da idade de ocorrência do acidente trombótico por estes polimorfismos. O polimorfismo FGG 10034 C>T actua como potenciador do risco trombótico, diminuindo a idade do primeiro acidente (*pvalue*=0,06). Por sua vez, o polimorfismo FGG 9340 T>C apresenta um efeito protector, aumentando-a (*pvalue*=0,07).

Ao considerar os haplótipos constituídos pelas combinações alélicas dos dois polimorfismos, a modulação é mais evidente: o haplótipo protector/neutro (PTH2/PTH1) aumenta a idade do primeiro acidente em cerca de 4 anos, enquanto que o haplótipo potenciador/neutro (PTH3/PTH1) a diminui em aproximadamente 5 anos, quando comparados com o haplótipo neutro.

A continuação e aprofundamento deste estudo, com uma amostra de maiores dimensões, poderão conduzir à identificação de factores moduladores do risco trombótico, permitindo uma melhor prevenção do mesmo.

PALAVRAS-CHAVE:

Trombose Venosa, FV Leiden, Fibrinogénio, Modulação, Polimorfismo

Abstract

Nowadays, Venous Thrombosis (VT) is one of the world-wide most common causes of morbidity and mortality . VT is a multi-factorial and poligenic disease, caused by alterations in major risk factors, as Factor V Leiden, pro-coagulants proteins or in other minor risk factors as Fibrinogen. In spite of mutation FV Leiden represents the molecular base of VT in nearly 8% of the cases, is possible to verify the existence of heterogeneity in the type and age of the first clinical manifestation. This suggests that there is a modulation by a set of additional hereditary and acquired factors.

The goal of this study is to evaluate the modulating effect of hereditary and environmental potential factors in the thrombotic risk of FV Leiden carrying individuals. The FGG 10034 C>T (rs2066865) and FGG 9340 T>C (rs1049636) polymorphisms are suitable candidates for this modulation function.

Although there haven't been verified significant statistical differences between allelic and genotypic frequencies, the preliminary acquired data in this study suggests a modulation of the thrombotic occurrence age by these polymorphisms. FGG 10034 C>T polymorphism behaves as an enhancer to the thrombotic risk, decreasing the age of the first thrombotic event (pvalue=0,06). On the other hand, FGG 9340 T>C polymorphism has a protector effect, increasing it (pvalue=0,07).

When considering the haplotypes established by the allelic combination of the two polymorphisms, the modulation is more visible: the protector/neutral haplotype (PTH2/PTH1) increases the age of the first accident in almost 4 years, whereas the enhancer/neutral polymorphism (PTH3/PTH1) decreases this same age in approximately 5 years, when compared with the neutral haplotype.

The extension and deepening of this study, with a larger population sample, will lead to the identification of thrombotic risk modulating factors, allowing a better prevention for the patients.

KEYWORDS:

Venous thrombosis, FV Leiden, Fibrinogen, Modulation, Polymorphism

Conteúdo

Lista de Figuras	ix
Lista de Tabelas	x
1 Introdução	1
1.1 Trombose Venosa (TV)	1
1.2 Factores de Risco de TV	2
1.3 Polimorfismos genéticos e o risco de TV	3
1.3.1 Factor V e Factor V Leiden	4
1.3.2 Protrombina	5
1.3.3 PAI-1	5
1.3.4 MTHFR	5
1.4 Fibrinogénio, o Factor I da coagulação sanguínea	6
1.4.1 Variantes de <i>splicing</i> do fibrinogénio – Fibrinogénio γA e γ'	9
1.4.2 Polimorfismo FGG 10034 C>T	11
1.4.3 Polimorfismo FGG 9340 T>C	12
1.4.4 Haplótipos	12
1.5 Objectivos	13
2 Metodologia	14
2.1 População em Estudo	14
2.2 Análise Molecular	16
2.2.1 Extracção de DNA genómico em larga escala	16
2.2.2 Quantificação de DNA	16
2.2.3 Amplificação de DNA (<i>Polimerase Chain Reaction</i> - PCR)	17
2.2.4 SSCP – <i>Single Strand Conformation Polymorphism</i>	18
2.2.5 Purificação do produto de PCR	19
2.2.6 Sequenciação Automática	20
2.2.7 TaqMan [®] Genotyping Assay	20
2.2.8 Outros métodos utilizados	22
2.3 Análise Estatística	25
2.4 Haplótipos	28

3	Resultados e Discussão	30
3.1	Caracterização dos grupos de estudo e controlo	30
3.1.1	Grupo de estudo	30
3.1.2	Grupo controlo	32
3.2	Determinação da presença do Polimorfismo FGG 10034 C>T	32
3.2.1	SSCP	32
3.2.2	Sequenciação	33
3.3	Determinação da presença do Polimorfismo FGG 9340 T>C	34
3.3.1	SSCP	34
3.3.2	dHPLC	34
3.3.3	ARMS	34
3.3.4	Sequenciação	35
3.3.5	HRM	36
3.3.6	<i>TaqMan Genotyping Assay</i>	37
3.4	Prevalência dos polimorfismos do FGG	39
3.4.1	Equilíbrio de Hardy-Weinberg	39
3.5	Efeito dos polimorfismos do FGG no risco de TV	40
3.5.1	<i>Efeito dos variantes na idade do primeiro acidente trombótico</i>	42
3.6	Haplótipos	43
4	Conclusão	46
	Bibliografia	54
A	Anexos	55
A.1	Anexo A – Procedimento para extracção de DNA utilizando QIAamp Midi	55
A.2	Anexo B – Tabela de Primers	57
A.3	Anexo C – SSCP	58
A.4	Anexo D – Colunas de purificação	62
A.5	Anexo E – Tabela Geral das Famílias dos Indivíduos	63

Lista de Figuras

1.1	Interacção dos factores de risco TV	2
1.2	Fibrinogénio e Fibrina	7
1.3	Processamento do RNAm do FGG	10
2.1	Marcador Molecular	17
2.2	<i>TaqMan</i> [®] <i>Genotyping Assay</i>	21
2.3	Cálculo de <i>odds ratio</i> (OR)	27
2.4	Haplótipos	28
3.1	Intervalo de idade de ocorrência do Primeiro acidente trombótico	31
3.2	Gel de Electroforese	32
3.3	Gel de poliacrilamida de SSCP	33
3.4	Electroferograma da Sequenciação do polimorfismo FGG 10034 C>T	33
3.5	ARMS ou PCR alelo-específico	35
3.6	Electroferograma da Sequenciação do polimorfismo FGG 9340 T>C	35
3.7	HRM	36
3.8	Comportamento dos diferentes componentes fluorescentes	38
3.9	Discriminação alélica do polimorfismo FGG 9340 T>C	38
3.10	Variação da idade com a presença de haplótipos	45
3.11	Variação da idade, haplótipos e factores secundários	45
A.1	Tabela de Primers	57

Lista de Tabelas

2.1	Haplótipos constituídos com os polimorfismos analisados nestes estudo	28
3.1	Caracterização do grupo de estudo	31
3.2	Frequências Genotípicas e Alélicas dos Polimorfismos FGG 10034 C>T e FGG 9340 T>C no grupo controlo e no subgrupo TEV	39
3.3	Análise Alélica e Genotípica dos polimorfismos FGG 10034 C>T e FGG 9340 T>C	41
3.4	Comparação entre Haplótipos	44

Capítulo 1

Introdução

1.1 Trombose Venosa (TV)

As doenças cardiovasculares mais comuns são o enfarte do miocárdio, o acidente vascular cerebral e a trombose venosa (TV). A TV apresenta uma incidência anual de cerca de 1 a 2 por cada 1000 indivíduos, com uma mortalidade estimada em cerca de 30% e elevada morbidade [1].

A coagulação sanguínea, pela qual um vaso danificado cessa de perder sangue, tem fascinado leigos e estudiosos em todas as culturas desde o início dos tempos. Contudo, nenhuma explicação real das reacções que levam à formação dos coágulos sanguíneos foi formulada antes da evolução da química na área das proteínas, a bioquímica [2].

A trombose é uma doença associada a distúrbios da coagulação sanguínea e pode ser classificada em termos gerais como o desenvolvimento de um coágulo no sistema venoso (baixa pressão e baixo fluxo) ou no sistema arterial (alta pressão e fluxo elevado). Certas distinções podem ser estabelecidas entre trombose venosa e arterial: a composição dos trombos (rico em fibrina no venoso e rico em plaquetas no arterial) e a presença de danos na parede dos vasos (ateroma), que está ausente na trombose venosa. No entanto as diferenças não são absolutas e há mecanismos adjacentes que são comuns [3, 4].

Quando o coágulo sanguíneo é desenvolvido no interior de um vaso venoso o resultado é uma TV. Usualmente afecta as veias dos membros inferiores, com consequente reacção inflamatória do vaso e podendo determinar a obstrução venosa, total ou parcial, impedindo a circulação de sangue e, como tal, dos nutrientes e oxigénio necessários aos tecidos que são irrigados por esse vaso sanguíneo [4]. Os sintomas são variáveis, havendo sinais clássicos, como o aumento da temperatura local, edema, dor e rigidez da musculatura. No entanto, cerca de 50% dos casos de TV podem passar despercebidos. A apresentação clínica depende, em grande medida, do local e da extensão da formação do coágulo. Existem várias manifestações

clínicas que podem ser incluídas dentro da categoria de TV, nomeadamente a trombose venosa profunda (TVP), tromboflebite superficial (TFS) e embolia pulmonar (EP). Na TVP, a veia afectada é uma veia profunda, enquanto que na TFS a veia afectada é superficial. Quando o coágulo de sangue proveniente de uma veia provoca o bloqueio da artéria pulmonar, dá-se uma EP. Outras manifestações clínicas podem ser também associadas à trombofilia, entre elas o aborto [5].

1.2 Factores de Risco de TV

A perturbação da hemostase, o equilíbrio do sistema de coagulação sanguínea, é de fundamental importância na patogénese da trombose. O processo de coagulação e hemostase compreende os sistemas fibrinolíticos (que desfazem os coágulos de sangue) com activadores, inibidores, zimogénios e co-factores. A função e a regulação destes sistemas compostos têm sido extensivamente estudadas. Contudo, compreender a sua integração *in vivo* representa ainda um grande desafio [6].

Recentes progressos na tentativa de entender a fisiologia da TV estabeleceram que esta apresenta uma natureza multi-causal, na qual a interacção entre factores genéticos e factores adquiridos, tem o papel mecanístico mais importante (figura 1.1).

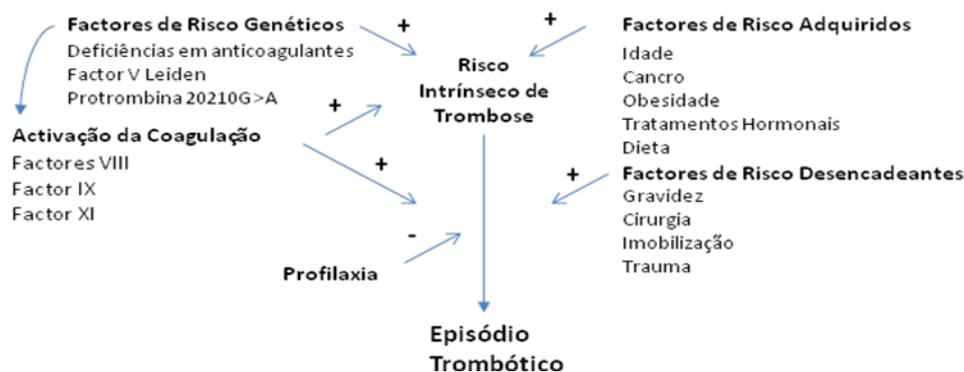


Figura 1.1: Quadro conceptual da interacção dos factores de risco para a ocorrência de TV. O risco intrínseco de trombose é definido pela combinação dos factores de risco genéticos e adquiridos, incluindo a modulação da activação da coagulação por factores conhecidos e desconhecidos. O risco intrínseco é alterado pela ocorrência de factores de risco desencadeantes, que podem ser equilibrados com o uso de medicamentos profiláticos. [Adaptado de [7]]

Em condições fisiológicas, existe um delicado balanço entre os sistemas pró-coagulante (que promovem a formação de coágulos) e anti-coagulante (que previnem a formação de coágulos), de maneira a garantir a hemostase nos locais onde ocorrem danos vasculares. Contudo, vários factores genéticos ou ambientais podem provocar o desequilíbrio deste sistema, resultando numa tendência para episódios de hemorragia ou eventos trombóticos [6]. A hemostase é determinada pelos níveis (funcionais) de todos os factores de coagulação (pró e anti-coagulantes) e factores fibrinolíticos.

O aumento dos níveis de factores pró-coagulantes, a diminuição dos níveis de factores fibrinolíticos e inibidores naturais da coagulação, fortalecem a probabilidade de ocorrência de um evento trombótico [3].

As influências ambientais, transitórias ou de longa duração, podem desempenhar um papel importante na hemostase, perturbando e influenciando o risco de trombose em ambos os sistemas, venoso e arterial [4]. O termo “ambiental” é usado no seu sentido mais amplo, para indicar mudanças induzidas por influências tão diversas como gravidez, administração de contraceptivos orais (hormonas), cirurgia, dieta, tabagismo. As perturbações na hemostase podem ser ainda determinadas geneticamente e, quando é o caso, a sua influência parece ser maior devido à sua presença ao longo da vida [8, 9].

Na ausência de deficiências genéticas, a trombose ocorre na população idosa, sobretudo no contexto de influências ambientais como cirurgia, imobilização e obesidade. Em contrapartida, na população jovem é mais comum que deficiências genéticas na cascata de coagulação sanguínea estejam relacionadas com a diminuição da idade em que começam a ocorrer os episódios trombóticos. O estímulo ambiental necessário é menor e os coágulos formam-se em locais pouco comuns. Embora alguns factores ambientais tenham uma influência inquestionavelmente importante, as modificações genéticas têm um papel crucial, e por vezes podem conduzir a TVs aparentemente espontâneas [5, 10].

A interacção entre dois ou mais factores pode atenuar os sintomas clínicos, mas, por outro lado pode ocorrer uma acção sinérgica, tendo como resultado um risco maior do que aquele que seria de esperar através da combinação dos efeitos. Porém, existe uma fraca relação entre os sintomas clínicos e as alterações moleculares conhecidas actualmente. Isto pode indicar a presença de outros factores genéticos ou adquiridos que sejam capazes de modular a expressão clínica da desordem hemostática [6].

1.3 Polimorfismos genéticos e o risco de TV

As variações do genoma humano podem afectar os portadores, quer aumentando o risco de desenvolvimento de doenças, quer alterando a resposta dos indivíduos a medicamentos, como vacinas e outros agentes. Estas podem ocorrer ao nível da sequência e ao nível da estrutura do DNA. As variações que ocorrem num único par de bases e cuja frequência na população é superior a 1% são designadas SNPs (*Single Nucleotide Polymorphism* – polimorfismo num único nucleótido) e representam o tipo de alteração genética mais comum. Mais de 10 milhões de SNPs foram já catalogados em diversas bases de dados (por exemplo, Ensembl [11]). O conhecimento de SNPs ligados a doenças é o conceito essencial para a concretização da medicina

personalizada. Através de estudos de comparação entre indivíduos com patologia e indivíduos saudáveis é possível determinar se uma alteração genética está ou não a influenciar uma dada patologia [12].

1.3.1 Factor V e Factor V Leiden

Algumas variações genéticas presentes nos factores participantes na cascata de coagulação sanguínea podem em grande medida influenciar o risco de ocorrência de TV.

O Factor V (FV) é uma glicoproteína com 300kD, que circula no plasma como um molécula de cadeia simples, com uma semi-vida de cerca de 12 horas. É codificada pelo gene *factor 5*, localizado no cromossoma 1 (1q24.2) que contém 25 exões. [13].

O FV funciona como um co-factor que permite ao Factor X (FX) activar a trombina. A trombina, por sua vez, cliva o fibrinogénio em fibrina, que polimeriza espontaneamente para formar densos aglomerados, que são o principal componente dos coágulos. A proteína C activada (APC - *Activated Protein C*) é um anti-coagulante natural que actua limitando a extensão da coagulação, através da clivagem e degradação do FV. Se o FV nunca for degradado, vai proporcionar-se uma maior tendência de formação de coágulos sanguíneos. [14].

A mutação FV Leiden (rs6025) foi identificada como um dos factores de risco hereditários de TV mais prevalentes na população caucasiana. Esta mutação é caracterizada pela substituição de uma base azotada G por A no nucleótido 1691 do FV, levando à substituição do aminoácido arginina por ácido glutâmico na posição 506 da proteína (Arg506Gln) [15].

O FV Leiden resulta numa variação do factor V que é activado pela trombina, mas não é tão prontamente inactivado pela APC, fenómeno conhecido como resistência à APC (APCR - *Activated Protein C resistance*). A inactivação ocorre com mais dificuldade, pois a substituição Arg506Gln altera o local de clivagem da APC. Quando o FV permanece activo existe uma sobreprodução de trombina, o que conduz ao excesso de produção de fibrina e conseqüentemente um excesso de coagulação. A coagulação em demasia está restrita às veias (pois estas apresentam um menor fluxo sanguíneo do que as artérias), podendo originar uma TV [16, 17].

A descoberta da mutação FV Leiden resultou numa visão mais clara da natureza molecular do episódio trombótico e da importância de alguns factores genéticos e adquiridos na ocorrência de uma trombose [18].

Existe uma elevada frequência do FV Leiden na população caucasiana (aproximadamente 5%). A elevada prevalência associada a um risco elevado (cerca de 8 vezes maior), tem possibilitado o estudo das suas interacções com outros factores genéticos de risco relativamente menos comuns (como diversas deficiências em proteínas funcionais, como a proteína C, proteína S e antitrombina III) e também com

factores ambientais (como contraceptivos orais) [4, 19].

1.3.2 Protrombina

Além do FV Leiden, foram identificados outros factores de risco *major*, como a Protrombina. A Protrombina é um zimogénio (precursor inactivo de uma proteína, que precisa ser clivado num ponto específico para dar origem à proteína funcional) com 72kD, também denominada de Factor II (FII). A protrombina transforma-se em trombina (através do complexo formado pelo Factor X activado (FXa) e o FV activado (FVa)), a qual catalisa a reacção de transformação do fibrinogénio em fibrina, como explicado anteriormente. Deste modo é um elemento fundamental da cascata de coagulação [16]. A substituição na região 3' não traduzida da protrombina, no nucleótido 20210 de um G para A (PT20210A – rs1799963), dá origem a uma mutação comum que aumenta o risco de TV. Este variante alélico está associado a níveis elevados de protrombina, parecendo ser este o facto responsável pelo aumento do risco de trombose. O mecanismo pelo qual os níveis de protrombina são alterados ainda permanece por esclarecer, contudo é sugerido que existe um aumento da eficiência da poliadenilação do RNAm. Este variante alélico tem um efeito mais pronunciado quando combinado com outros factores de risco genéticos [20].

Além dos variantes genotípicos cientificamente conhecidos por terem um papel relevante no risco de TV, existem ainda outros factores de risco, os chamados factores de risco *minor*. São exemplos, o PAI-1, o MTHFR e o Fibrinogénio.

1.3.3 PAI-1

O inibidor do activador do plasminogénio, (PAI-1 – *Plasminogen Activator Inhibitor*), é um membro da família das *Serpinas*, um grupo de proteínas identificadas como inibidores de proteases. A sua principal função é inibir o activador do plasminogénio tecidual (*tissue Plasminogen Activator* – tPA), um agente anti-coagulante. O PAI-1 existe na circulação em concentração superior ao tPA, permitindo a lise de coágulos de sangue sem permitir que haja uma hemorragia sistémica. Concentrações elevadas de PAI-1 são consistentemente associadas com o risco de acidente vascular, estando presentes em placas arterioscleróticas [3].

1.3.4 MTHFR

O aumento dos níveis de homocisteína também parece influenciar o risco de TV. O MTHFR é um enzima envolvido no metabolismo da homocisteína, convertendo o folato na sua principal forma de circulação: 5-metiltetrahydrofolato [21]. A variação alélica 677C>T (rs1801133) resulta numa substituição de uma alanina por uma valina, tornando o enzima MTHFR termolábil, o que resulta num aumento dos

níveis de homocisteína. Esta variação é comum, 10-12% da população caucasiana apresenta esta variante em homozigotia, com a actividade do enzima reduzida a 50% a 37°C. O MTHFR 1298A>C, que ocorre em heterozigotia em aproximadamente 9-20% da maioria dos grupos étnicos, foi a segundo variante descrito neste enzima [22]. Esta variante por si só não parece estar associada à hiper-homocisteinémia, todavia a combinação das duas variantes, MTHFR 677 e MTHFR 1298, parece resultar numa diminuição da actividade do enzima e no aumento dos níveis de homocisteína. Vários estudos têm sido publicados neste sentido, no entanto os resultados apresentados são contraditórios [23, 13]. Apesar disto, está estabelecido que a hiper-homocisteinémia é um factor de risco para a TV, podendo aumentá-lo de 2 a 4 vezes [24, 25].

1.4 Fibrinogénio, o Factor I da coagulação sanguínea

De todos os componentes do sistema de coagulação sanguínea, o fibrinogénio é aquele que tem sido mais consistentemente associado a distúrbios vasculares oclusivos. O fibrinogénio é um factor fundamental da cascata de coagulação sanguínea da fase aguda, participando em duas vias distintas: primariamente na parte final da via comum da cascata de coagulação, sendo o precursor da fibrina (produto final da cascata de coagulação sanguínea), secundariamente através da sua ligação aos receptores de membrana das plaquetas, formando uma rede que providencia estabilidade aos trombos recém-formados. O fibrinogénio participa também no desenvolvimento de aterosclerose, promovendo a adesão das plaquetas e dos glóbulos brancos à superfície endotelial e a proliferação e migração das células musculares [26, 27].

Várias vias biológicas ligam o fibrinogénio plasmático ao desenvolvimento de patologias trombóticas (venosas e arteriais) e desordens a nível cardiovascular e cerebrovascular. Estas vias incluem a formação de coágulos de fibrina, agregação plaquetária, fibrinólise, inflamação, aterogénese e viscosidade do plasma. [1, 28].

Evidências do *Heart Scottish Health Study* indicam que os níveis de fibrinogénio aumentam com diversos marcadores de risco cardiovascular, incluindo hipertensão, diabetes, tabagismo e doença vascular periférica. Assim, o fibrinogénio fornece uma ligação entre o desenvolvimento da patologia trombótica e os factores ambientais e clínicos, como por exemplo a influência do tabagismo nos níveis de fibrinogénio [3].

Em termos estruturais, o fibrinogénio é uma glicoproteína com peso molecular de 340kD que consiste em duas moléculas simétricas, cada uma contendo três cadeias polipeptídicas: $A\alpha$, $B\beta$ e γ . Estas três cadeias são codificadas por três genes parálogos, *fibrinogénio alfa (FGA)*, *fibrinogénio beta (FGB)* e o *fibrinogénio gama (FGG)* agrupados num *cluster* situado numa região de aproximadamente 50kb no cromossoma 4 (4q31.3). É primariamente sintetizada nos hepatócitos e circula no plasma

a uma concentração aproximada de $9\mu\text{M}$ (3g/L) [11].

O gene *FGA* contém 6 exões, está localizado no centro do *cluster* do fibrinogénio e produz uma cadeia polipeptídica que contém 625 resíduos de aminoácidos.

O gene *FGB* está localizado 13kb a jusante do gene *FGA*. Contém 8 exões que são transcritos na direcção oposta aos genes da cadeia α e da cadeia γ , codificando um polipéptido com 461 resíduos de aminoácidos.

O terceiro e último gene do *cluster* do fibrinogénio, o gene *FGG*, possui 10 exões e está localizado 10kb a montante do gene *FGA*, codificando uma cadeia polipeptídica com 411 resíduos de aminoácidos.

Pensa-se que o gene *FGB* esteja envolvido na determinação da quantidade plasmática de fibrinogénio, e que os genes *FGA* e *FGG* desempenhem um papel na regulação da estrutura do coágulo de fibrina [29, 30].

A Fibrina é o componente maioritário dos coágulos sanguíneos e é formada a partir do fibrinogénio. A molécula, de 45nm, consiste em três domínios estruturais principais: o domínio E, que inclui os fibrinopéptidos A e B e os aminoácidos terminais de todas as 6 cadeias polipeptídicas e dois domínios D distais, conectados à região E por dois segmentos de hélice- α super enrolados, que contêm as terminações carboxilo das cadeias B- β e A- α . Na figura 1.2 estão representadas as moléculas de fibrinogénio e fibrina [1, 30].

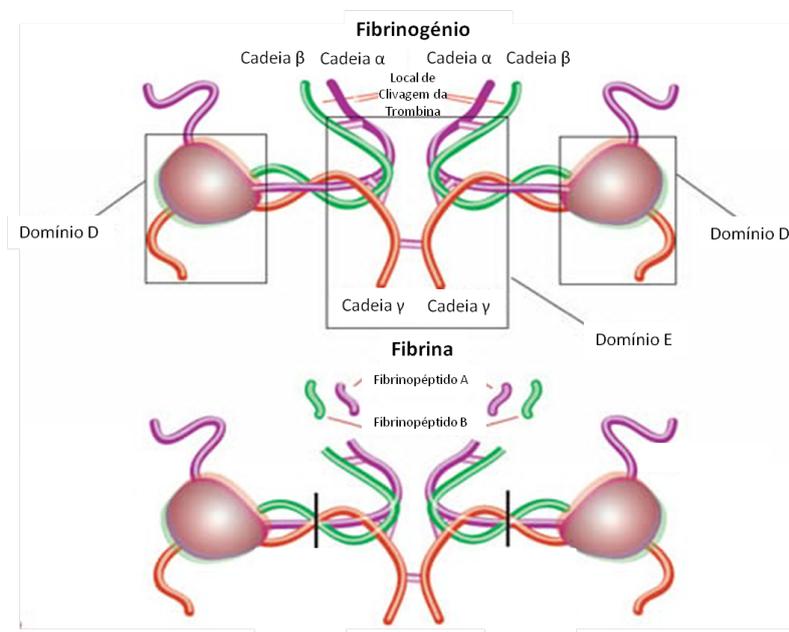


Figura 1.2: Fibrinogénio (em cima) e Fibrina (em baixo). Nesta figura estão identificadas as cadeias α , β e γ e os domínios estruturais (D e E). Está ainda evidenciado o local onde a trombina realiza a clivagem do fibrinogénio, libertando o fibrinopéptido A e B, convertendo-o em Fibrina [adaptado de [31]].

Em situações de dano do tecido vascular e inflamações, a cascata de coagulação é activada. A protrombina é clivada dando origem à trombina. O fibrinogénio é

convertido em monómeros de fibrina através da remoção dos fibrinopéptidos A e B na extremidade amínica das cadeias $A\alpha$ e $B\beta$ pela trombina, como demonstrado na figura 1.2. A libertação do fibrinopéptido A é mais rápida e expõe um sítio de polimerização no domínio E da fibrina. Este sítio tem complementaridade para o domínio D de uma molécula de fibrina adjacente, formando uma protofibrilha. A clivagem do fibrinopéptido B é mais lenta e expõe também um sítio de ligação na região E. O Factor XIII (FXIII) liga-se posteriormente ao fibrinogénio estabelecendo ligações cruzadas, o que torna a fibrina mais forte e resistente à degradação pelo sistema fibrinolítico [32, 33].

A estrutura do coágulo de fibrina parece ser clinicamente relevante, pois coágulos mais rígidos e fortemente empacotados (*tightly packed*) com fibras menos porosas são menos susceptíveis à acção dos agentes trombolíticos. Esta estrutura é determinada por fenótipos quantitativos (como por exemplo, a concentração de fibrinogénio) e por fenótipos qualitativos (por exemplo, as variantes da cadeia do fibrinogénio γ) [34].

Mutações que ocorram nos genes do fibrinogénio podem, por um lado causar deficiência de fibrinogénio (levando a situações de afibrinogenemia ou hipofibrinogenemia), por outro lado causar situações em que a concentração de fibrinogénio aumenta, ou a sua estrutura é modificada podendo tornar-se mais resistente à fibrinólise (potenciando a formação de coágulos que viajam na corrente sanguínea, provocando trombozes). As mutações são causadas por diversos mecanismos e actuam ao nível do DNA, RNA, afectando por exemplo o *splicing* do RNA ou a sua estabilidade; ou ainda ao nível da proteína, afectando a síntese da proteína, a sua montagem (*assembly*) ou a sua secreção [35].

Foram já descritos vários polimorfismos nos genes do fibrinogénio. No que diz respeito às doenças cardiovasculares, os estudos de associação focam mais extensivamente a cadeia β , pois esta parece regular o passo limitante na síntese do fibrinogénio. A maioria dos polimorfismos localiza-se em regiões não traduzidas dos genes. Entre estes inclui-se um polimorfismo da região 5' promotora do gene *FGB*. Esta alteração foi descrita como estando associada ao aumento dos níveis de fibrinogénio [36, 37]

Para além das variantes não codificantes, existem alguns polimorfismos em regiões codificantes do *cluster* do fibrinogénio, que introduzem uma substituição de aminoácidos na proteína madura. Um destes polimorfismos ocorre no gene *FGA* e conduz à substituição de uma treonina por uma alanina no resíduo 312 (Thr312Ala) da proteína. O Thr312Ala está localizado na zona da molécula de fibrinogénio que interage com o FXIII, levando ao aumento de ligações cruzadas (*cross-linking*) e da rigidez do coágulo. Estudos clínicos referem que esta substituição de aminoácidos predispõe para a ocorrência de embolias em sistemas venosos e arteriais [1].

Em vários estudos prospectivos e meta-análises, como o *Framingham study* [38], *Northwick Park Heart study* [39] e *Fibrinogen Studies Collaboration* [40], os níveis plasmáticos de fibrinogénio têm sido associados ao risco de desenvolvimento de doenças coronárias arteriais (DAC) e enfarte do miocárdio (EM). Contudo ainda não é claro se o aumento dos níveis de fibrinogénio são a causa do desenvolvimento da doença, ou apenas um efeito secundário, podendo ser um marcador de inflamação [32, 26]. Estudos efectuados em ratos transgénicos com hiperfibrinogenemia têm sugerido que o fibrinogénio é um verdadeiro modificador das doenças vasculares, aumentando a deposição de fibrina em certos órgãos e regulando a taxa de conversão da fibrina (fibrin *turnover*) [1].

No entanto, os mecanismos que regulam os níveis de fibrinogénio ainda não estão totalmente esclarecidos, e é possível que a regulação se dê através de um complexo sistema de interacções entre os genes e o ambiente [41].

O fibrinogénio circulante e os coágulos de fibrina apresentam uma estrutura bastante heterogénia que pode resultar de proteólise, *splicing* alternativo ou modificações pós-transcricionais ou pós-traducionais nas cadeias α , β e γ do fibrinogénio. A heterogenicidade pode ser avaliada a partir de dois testes: teste funcional do fibrinogénio, que depende de taxa de coagulação na presença de trombina (teste de Clauss), e um teste com anticorpos contra o fibrinogénio que mede o fibrinogénio total em circulação (independentemente de haver coagulação) [36]. Estudos populacionais que comparam estes dois métodos de ensaio observam uma forte associação da incidência de doenças cardiovasculares com o imunoensaio, mas não tanto com o ensaio funcional. Estes resultados são potencialmente importantes, não só para a avaliação de risco, mas para esclarecimento do mecanismo subjacente à associação entre a concentração de fibrinogénio e o aumento de risco cardiovascular, incluindo os efeitos sobre a formação do coágulo de fibrina, viscosidade do sangue ou agregação plaquetária [42, 43].

1.4.1 Variantes de *splicing* do fibrinogénio – Fibrinogénio γA e γ'

O processamento de RNA é um processo que está conservado entre diferentes espécies, o que sugere uma importância fisiológica [34]. Envolve reacções de *capping*, *splicing* e poli-adenilação, em que vários componentes, como pequenos RNAs nucleares (snRNAs – *small nuclear RNAs*) e complexos de proteínas, interagem de uma forma específica e eficiente [44].

Nos genes *FGA* e *FGG* pode ocorrer *splicing* alternativo. A cadeia $A\alpha$ predominante na circulação contém 610 resíduos de aminoácidos e corresponde aos exões de 1 a 5. A forma alternativa da cadeia $A\alpha$ contém 846 resíduos de aminoácidos, sendo uma isoforma prolongada, que remove o intrão 5 e os últimos 15 codões do exão 5,

mas conserva o exão 6. Esta forma alternativa representa cerca de 1 a 2% da cadeia α circulante [35].

A forma mais abundante da cadeia γ , γA , consiste em 411 resíduos de aminoácidos, enquanto que a cadeia variante γ' contém 427 resíduos de aminoácidos e representa cerca de 7 a 15% da cadeia γ circulante. A cadeia γA é formada pela tradução de todos os 10 exões. A variante γ' é formada após *splicing* alternativo, ocorrendo a poli-adenilação no intrão 9 e a tradução dos exões de 1 a 9, como demonstrado na figura 1.3. Como o intrão 9 não é removido, a grelha de leitura é alterada (ORF - *open reading frame*), havendo a substituição dos 4 codões do exão 10 por 20 codões alternativos. Como resultado, a cadeia γA e γ' diferem na sua sequência carboxilica terminal. A cadeia γA termina com uma sequência única que é codificada pelo exão 10, que está envolvida na ligação das plaquetas. Os 20 aminoácidos codificados pelo intrão 9 formam um sítio de ligação para o FXIII e um sítio de ligação à trombina [45, 46].

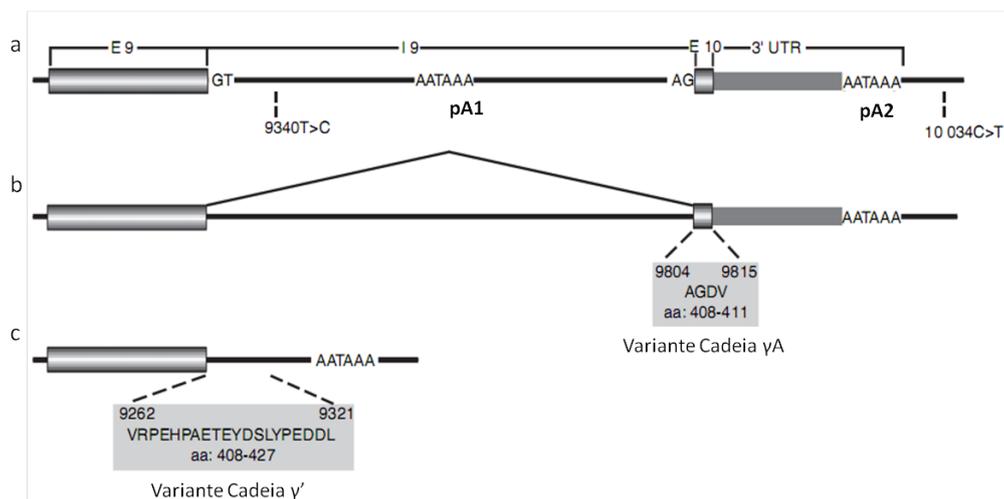


Figura 1.3: Processamento alternativo do RNAm do gene *FGG*. (a) Exão 9, Intrão 9, Exão 10 e a região 3' não traduzida (3'UTR) do gene *FGG*. Estão evidenciados os polimorfismos FGG 9340 T>C e FGG 10034 C>T, assim como os locais de poliadenilação pA1 e pA2. (b) A cadeia γA , que contém 411 resíduos de aminoácidos, surge quando todos os intrões do pré-RNAm do *FGG* são removidos e a poliadenilação ocorre a jusante do exão 10 (E10), no local pA2. (c) A cadeia γ' contém 427 resíduos de aminoácidos porque o intrão 9 (I9) é parcialmente retido e a poliadenilação ocorre no local de pA1, a jusante do exão 9 (E9). [Adaptado de [34]].

Através de cromatografia de troca iônica é possível separar as duas fracções distintas de fibrinogénio γ : a fracção principal, que contém a forma mais abundante γA (cerca de 84%), é um homodímero $\gamma A/\gamma A$ e a fracção menor, que contém a cadeia menos abundante é a variante cadeia γ' , podendo ser um heterodímero ($\gamma A/\gamma'$) (aproximadamente 15%), ou um homodímero (γ'/γ') que constitui cerca de 1% do fibrinogénio plasmático [34].

Apesar da estrutura ser semelhante, os coágulos de fibrina formados a partir das cadeias γ'/γ' são cerca de três vezes mais condensados e apresentam uma taxa

de fibrinólise 10 vezes mais lenta do que os coágulos de fibrina formados a partir das cadeias $\gamma A/\gamma A$. Provavelmente o Fibrinogénio γ' contém sítios de ligação para a trombina e para a subunidade B do FXIII, permitindo um maior número de ligações cruzadas (*cross linking*) [30]. Devido a este aumento de ligações cruzadas, os coágulos originados a partir de fibrinogénio $\gamma A/\gamma'$ têm uma estrutura de fibrina mais resistente à lise, mais estável e mais condensada, comparativamente aos coágulos formados a partir de fibrinogénio $\gamma A/\gamma A$.

Todavia, esta terminação alternativa da cadeia γ' define, não só um sítio de ligação para o FXIII, mas também um sitio de ligação para a trombina. Este local de ligação funciona como um inibidor da trombina e, ainda mais importante, contribui para a actividade da antitrombina. A antitrombina funciona como inibidor da coagulação, neutralizando vários enzimas como a trombina e o FX activado durante a formação da fibrina. Deste modo evita-se que a trombina clive o Fibrinopéptido A. A geração de Fibrinopéptido A é mais lenta com fibrinogénio $\gamma'/\gamma A$ do que com fibrinogénio $\gamma A/\gamma A$ – quer dizer que existe mais regulação negativa (*downregulation*) [1, 47].

1.4.2 Polimorfismo FGG 10034 C>T

O polimorfismo FGG 10034 C>T (rs2066865) está localizado a jusante da região 3' UTR do gene *FGG*, numa região rica em GT (DSE – *Downstream sequence element*), que compreende um sítio de ligação para um factor de estímulo para a clivagem, o CstF (*cleavage stimulation factor*). O CstF é um complexo com múltiplas subunidades necessário para a clivagem e poliadenilação eficientes dos pré-RNA, que se liga a uma região DSE rica em GU e estabiliza a ligação ao sítio do factor específico para a clivagem e poliadenilação para o sinal pA. A região DSE regula o uso do sinal de clivagem pA2 (2º sítio de sinal de poliadenilação - AATAAA) do FGG, e como tal, a formação da variante do fibrinogénio γ' , como indicado na Figura 1.3. Esta região tem 78% de homologia com o local de ligação do CstF [48, 49].

Quando existe um T no nucleótido 10034 do FGG, a região DSE é fortalecida (mais rica em GT, apresentando 83% de homologia com o local de ligação de CstF) relativamente à existência de um nucleótido C. Este fortalecimento é importante, na medida em que a afinidade desta região DSE vai ser aumentada na competição pelos sítios de poliadenilação. No FGG, o sinal pA2 é usado para a formação do RNA da cadeia γA , no qual o intrão 9 foi removido, enquanto que o sinal pA1 está localizado no intrão 9, nos nucleótidos 9558-9563, e é usado para a formação do RNA da cadeia γ' [50, 51]. Assim, quando esta região está fortalecida, resulta numa maior frequência do uso do local pA2 de poliadenilação, havendo mais transcriptos γA [30].

É aparentemente por esta razão que o polimorfismo FGG 10034 C>T está associado a uma maior susceptibilidade para patologia trombótica [34].

1.4.3 Polimorfismo FGG 9340 T>C

O polimorfismo FGG 9340 T>C (rs1049633) está localizado no intrão 9, a montante do sítio de poliadenilação 1 (pA1), como representado na figura 1.3. Ao contrário do polimorfismo FGG 10034 C>T, este parece estar relacionado com o aumento da concentração do fibrinogénio γ' . Este aumento poderá reduzir o risco de TV (conferindo protecção) parecendo, no entanto, aumentar o risco de EM. Esta conclusão não é de estranhar considerando que os dois processos são bastante diferentes, principalmente no que diz respeito à formação de placas arterioscleróticas [34].

A concentração do fibrinogénio γ' plasmático aumenta com o alelo FGG 9340C. É sugerido que este polimorfismo influencie a estabilidade do RNAm e/ou interfira com a maquinaria de *splicing*. É de notar que diversas variações alélicas que influenciam a transcrição dos genes (pela alteração da estabilidade do RNAm, da eficiência do processamento de RNA ou ainda causando modificações epigenéticas) têm sido detectadas tanto dentro como fora das regiões promotoras dos genes [34].

1.4.4 Haplótipos

As variações ou polimorfismos do fibrinogénio constituem os haplótipos existentes e podem afectar os níveis do fibrinogénio, a formação das estruturas de rede de fibrina ou a sensibilidade dos coágulos de fibrina ao sistema fibrinolítico [52, 53].

Os haplótipos (combinação de alelos) mais comuns de cada um dos três genes do fibrinogénio já foram descritos em SeattleSNPs (*Variation Discovery Resource* [54]). Os SNPs e os padrões específicos de haplótipo diferem de acordo com a população e com a análise fenotípica. Através da informação obtida até ao momento, pensa-se que os haplótipos enumerados representem todas as variantes frequentes, genotípicas e proteicas, que existem na população caucasiana [36].

Apesar de não ter sido encontrado nenhum estudo que reporte a presença de ambos os polimorfismos FGG 10034 C>T e FGG 9340 T>C num único haplótipo, estes são incluídos em dois haplótipos diferentes: um haplótipo que parece potenciar o risco de TV (FGG-H2), onde é incluído o polimorfismo FGG 10034T e o haplótipo FGG-H3, que parece constituir um haplótipo protector, incluindo o polimorfismo FGG 9340C [55, 29, 56].

O haplótipo FGG-H2 é comum na população europeia e está associado à diminuição do fibrinogénio γ' , assim como à diminuição da taxa fibrinogénio γ' /fibrinogénio total. Desta forma o haplótipo FGG-H2 sugere uma associação com o aumento do risco trombótico, podendo estas propriedades ser devidas à presença do polimorfismo FGG 10034T [30, 57].

Por outro lado, o haplótipo FGG-H3 está associado ao aumento das concentrações plasmáticas de fibrinogénio γ' , e aparenta estar pouco representado nos pa-

cientes com TV, o que aponta para a ideia de que é um haplótipo protector [55, 26].

1.5 Objectivos

A Trombose Venosa é uma doença poligénica e multifactorial. Apesar do Factor V Leiden ser um factor com elevada prevalência em indivíduos com trombose venosa, a manifestação clínica e a idade em que ocorre o primeiro acidente trombótico são muito variáveis. Esta variação apenas pode ser explicada pela múltipla interacção entre factores de risco hereditários e adquiridos, ou ambientais, que exercem uma modulação.

Deste modo o objectivo do presente estudo foi avaliar a modulação do risco trombótico, pelos polimorfismos FGG 10034 T>C e FGG 9340 T>C, em famílias portadoras de um factor de risco *major*: o FV Leiden. Para atingir o objectivo principal, foi necessário o cálculo de frequências alélicas e genotípicas. Este resultado também é interessante do ponto de vista da caracterização da população portuguesa relativamente aos polimorfismos em estudo.

Capítulo 2

Metodologia

2.1 População em Estudo

Para a avaliação do papel dos polimorfismos do Fibrinogénio Gama, FGG 10034 C>T (rs2066865) e FGG 9340 T>C (rs1049636), na modulação do risco trombótico em portadores de um factor de risco *major*, o FV Leiden, foi delineado um estudo para analisar dois grupos de indivíduos: o grupo de estudo e o grupo controlo.

O **grupo de estudo** é constituído por 110 famílias com pelo menos um portador do FV Leiden, sendo este o critério de inclusão no estudo, num total de 280 indivíduos. Neste grupo foram incluídos pacientes com Trombose Venosa Profunda (TVP), Tromboflebite Superficial (TFS), Embolia pulmonar (EP), Doença Arterial Coronária (DAC), Acidente Vascular Cerebral (AVC), Enfarte Agudo do Miocárdio (EAM), Aborto e indivíduos saudáveis. Os pacientes com historial trombótico, foram classificados consoante o tipo de trombose diagnosticada pelos clínicos, arterial ou venoso e o local onde ocorreu. Relativamente à trombose venosa (TV) foi determinado se esta era superficial ou profunda. Outra informação fundamental foi a idade de cada acidente trombótico (em especial do primeiro).

O grupo de estudo foi posteriormente subdividido em dois: um subgrupo apenas com indivíduos com patologia trombótica e um subgrupo constituído pelos familiares saudáveis. Para o cálculo de frequências alélicas e genotípicas foram considerados apenas os indivíduos não relacionados. Deste modo, foi constituído um subgrupo, onde apenas se incluíram os indivíduos não relacionados com Tromboembolismo Venoso – **subgrupo TEV** (TVP, TFS, EP, TV noutras localizações e Aborto). O critério de selecção escolhido para a eliminação dos indivíduos relacionados foi manter o indivíduo com idade do primeiro acidente trombótico mais baixa. De forma a avaliar o efeito dos factores secundários, o subgrupo TEV foi dividido em indivíduos com e sem factores secundários. Os factores ambientais considerados para este trabalho foram: o uso de contraceptivos orais, terapias hormonais de substituição, cirurgias, traumas, condições médicas específicas (como por exemplo cancro,

diabetes, hipertensão e obesidade) e imobilidade (como por exemplo viagens prolongadas ou internamento hospitalar). Estes factores são essenciais para determinar se o acidente trombótico foi secundário – sempre que teve causas secundárias, como os factores ambientais; ou se foi esporádico – quando não tem factores secundários que potenciem o acidente trombótico. Os factores secundários ou ambientais foram considerados quando apresentados num intervalo de tempo correspondente aos três meses precedentes ao episódio trombótico.

No grupo de indivíduos saudáveis foram incluídos os indivíduos não relacionados que não apresentavam acidente trombótico com mais de 42 anos de idade – **subgrupo saudáveis + 42** (esta idade corresponde a mais 5 anos do que a média de idades da ocorrência do primeiro acidente trombótico na população em estudo, 37 anos) e portadores do FV Leiden. Neste caso, o critério de selecção para que os indivíduos deste grupo não fossem relacionados entre si, foi manter o indivíduo com idade superior, pois atingiu uma idade mais avançada sem apresentar patologia trombótica. A constituição deste subgrupo foi pensada para uma posterior comparação com o **subgrupo TEV - 32**, portadores de FV Leiden que tiveram o primeiro acidente trombótico com idade inferior em 5 anos relativamente à média da idade do primeiro acidente trombótico (32 anos).

Os pacientes analisados neste estudo são maioritariamente provenientes da zona de Lisboa, (Hospital de Santa Cruz, Lisboa) e do Alentejo (Hospital do Litoral Alentejano). Este é um estudo retrospectivo, em que as amostras foram recolhidas entre 1997 e 2010.

Foi elaborado um questionário padronizado para caracterizar os pacientes, relativamente ao seu historial médico, e outros parâmetros de interesse entre os quais se encontram a origem geográfica e a história familiar, que foi considerada positiva sempre que um familiar directo do paciente teve um episódio trombótico (arterial ou venoso). Toda a informação relativamente aos pacientes e familiares foi revista de forma a completar os dados, através dos registos fornecidos pelos médicos.

O **grupo controlo** é constituído por 110 indivíduos dadores de sangue, escolhidos de uma forma aleatória. Estes indivíduos apresentam idades compreendidas entre os 20 e os 64 anos.

Todos os participantes deram o seu consentimento informado, e o estudo foi aprovado pelos comités de ética médica dos hospitais e entidades participantes, de acordo com a Declaração de Helsínquia redigida pela Associação Médica Mundial.

Tanto nos indivíduos da população em estudo como da população controlo, os polimorfismos do FV Leiden, Protrombina 20210 G>A, PAI-1 5G/4G, MTHFR 677C >T, MTHFR 1298 A>C, FV haplótipo HR2 haviam já sido determinados, em

estudos anteriores ou como diagnóstico dos pacientes.

2.2 Análise Molecular

2.2.1 Extração de DNA genómico em larga escala

O DNA foi extraído a partir de sangue periférico, através de um kit da *QIAGEN*, o *QIAamp*[®] *DNA Blood Mini/Midi Kit* [58].

Este kit destina-se à extração e purificação do DNA total de amostras de sangue. O DNA liga-se à membrana existente na coluna, enquanto que todos os outros componentes do sangue são efectivamente eliminados através de duas etapas de centrifugação. O DNA é posteriormente eluído com um tampão salino, que ajuda a conservar a integridade do DNA durante o tempo em que é armazenado. O protocolo de extração de DNA através do uso deste kit está descrito no anexo A.1 [59].

Resumidamente, as células presentes no sangue são sujeitas a uma lise, através do tampão AL (fornecido pelo kit), agitação vigorosa e temperatura de 70°C durante 10 minutos. Foram introduzidas algumas modificações no protocolo para um aumento de eficiência da lise celular, adicionando-se um passo de lise mecânica. Posteriormente são removidos os vestígios de proteínas, restos celulares e possíveis contaminantes através da centrifugação com dois tampões de lavagem (AW1 e AW2). O DNA, que fica retido na coluna durante os passos de lavagens, é eluído com o tampão AE [59].

2.2.2 Quantificação de DNA

A quantificação exacta de DNA presente em cada amostra foi efectuada no NanoDrop[®], usando o tampão em que a amostra estava diluída para calibrar o zero do espectrofotómetro [60].

Apesar de fornecer uma quantificação exacta, não é possível conferir a qualidade do DNA através do NanoDrop[®], apenas quando as amostras de DNA são visualizadas num gel de electroforese.

Deste modo, a quantificação de DNA foi também realizada através de quantificação visual, correndo as amostras de DNA num gel de agarose. A quantificação foi feita através da comparação da intensidade das bandas das amostras com a intensidade das bandas dos fragmentos de DNA do fago Lambda/Hind III em duas concentrações distintas (400ng e 200ng). Os fragmentos do DNA do fago Lambda/Hind III têm tamanhos exactamente determinados e para cada fragmento a percentagem de DNA correspondente é conhecida. Pela comparação da intensidade entre as bandas, é possível estimar a concentração de DNA presente nas amostras desconhecidas [61].

A terceira alternativa corresponde a uma quantificação visual, em tudo idêntica

à anterior excepto no marcador molecular utilizado. A comparação das intensidades das bandas de DNA corridas no gel de Agarose, a determinada concentração, foi feita com um marcador específico, o *HyperLadder II* da Bioline [62]. A figura 2.1 apresenta ambos os marcadores moleculares.

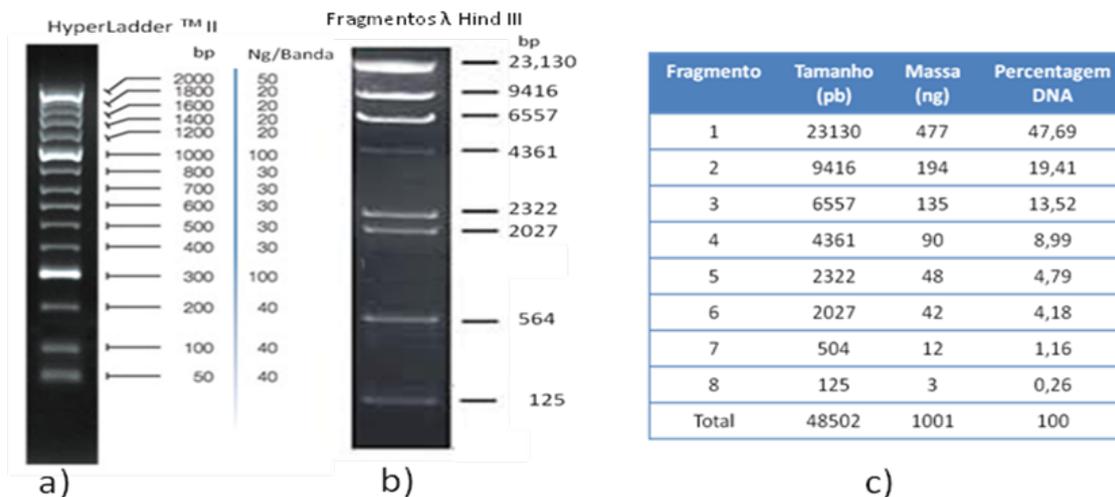


Figura 2.1: (a) Marcador molecular *HyperLadder II*, num gel de agarose de 1,5%. É descrito o tamanho em pares de base (pb) de cada banda no gel, sendo também efectuada a relação entre a quantidade de DNA em ng de cada banda assinalada no gel. (b) Marcador molecular DNA do fago Lambda/Hind III Digerido, num gel de agarose de 1%. É descrito o tamanho em pb de cada banda no gel. (c) Tabela de conversão para o marcador DNA do fago Lambda/Hind III [62]

2.2.3 Amplificação de DNA (*Polimerase Chain Reaction - PCR*)

O fragmento do FGG, do exão 10, contendo o polimorfismo FGG 10034 C>T (rs2066865), com 330pb, foi amplificado por PCR utilizando os *primers forward* FGG E10-1 e *reverse* FGG E10-R (informações adicionais sobre os *primers* na tabela apresentada no anexo A.2). Cada reacção de amplificação continha, por amostra, 50 μ L de *Pré-mix*, 0,5 μ L do *primer forward* e do *primer reverse* (300ng/ μ L) e 0,3 μ L de Taq Polimerase. Esta *mix* é distribuída por cada poço da placa de PCR e 1 μ L de DNA amostra é adicionado posteriormente.

A *pré-mix* é o conjunto do tampão, dos dNTPs e da água, necessários à reacção de PCR: 9mL de água bidestilada tratada com DPEC (inibidor de RNases), 1mL de tampão de PCR (10x, 15mM MgCl₂) e por fim 20 μ L de cada um dos dNTPs (dATP, dCTP, dGTP, dTTP), perfazendo um volume final de aproximadamente 10mL. Estes 10mL foram divididos em aliquotas de 1mL (para minimizar a probabilidade de contaminação) em *eppendorfs* e congelados a -20°C.

O PCR foi realizado num termociclador (Biometra[®]) com as seguintes condições: desnaturação inicial de 94°C durante 3 minutos, seguida de 35 ciclos com desnaturação a 94°C durante 45 segundos, emparelhamento dos *primers* (*annealing*) a

56°C durante 30 segundos, e extensão do fragmento a 72°C durante 30 segundos. A extensão final é realizada a 72°C durante 3min.

A amplificação foi confirmada através de um gel de agarose com concentração de 2% (2g de Agarose em 100mL de tampão TBE), numa tina de electroforese, sujeita a uma voltagem de 70V, durante 1 hora. A revelação foi feita através da adição de 2 μ L de Brometo de Etídio (EtBr) ao gel e a visualização efectuada num transiluminador, com Ultra-Violetas. Um marcador molecular com fragmentos de DNA de tamanho conhecido era carregado no gel de electroforese: através da comparação do marcador molecular com o tamanho das bandas amplificadas era determinado o comprimento dos fragmentos amplificados por PCR.

2.2.4 SSCP – *Single Strand Conformation Polymorphism*

O SSCP (*Single Strand Conformation Polymorphism* ou Polimorfismo de Conformação de Cadeia Simples) é uma metodologia que conduz à identificação de alterações de sequência de fragmentos de DNA amplificados por PCR.

O princípio do funcionamento do SSCP é muito simples: a molécula de DNA de cadeia simples forma estruturas tridimensionais complexas, as quais são estabilizadas por ligações intramoleculares fracas determinadas pela sequência de DNA.

A mobilidade electroforética que as estruturas tridimensionais apresentam em géis não desnaturantes (como os géis de poliacrilamida) depende do comprimento das cadeias e da sua conformação. Qualquer modificação da sequência do DNA pode provocar uma alteração conformacional da molécula, que poderá ser distinguida por apresentar uma diferente mobilidade electroforética (*shift*). Através do elevado poder de resolução da electroforese realizada em gel de poliacrilamida, conseguem ser distinguidas a maioria das mudanças conformacionais, inclusive aquelas que são causadas por uma diferença tão subtil como uma substituição numa única base (SNP) num fragmento com centenas de bases.

O SSCP é aplicável a fragmentos de DNA amplificados com tamanhos entre os 80 e os 600pb, com a vantagem que pode ser aplicado universalmente, sem ser necessário conhecer previamente a sequência para o fragmento ser analisado. É um método relativamente simples, barato e fidedigno: um mesmo padrão num gel de SSCP corresponde sempre a uma mesma sequência de DNA. Qualquer variação num padrão de migração corresponde a uma mudança no fragmento de DNA.

As condições do SSCP podem variar de acordo com os fragmentos analisados, em particular a composição do gel (no tampão utilizado, TBE, TPE e MDE, e na sua concentração) e nas condições da electroforese (temperatura, voltagem e tempo de corrida).

Para aplicação de SSCP, o fragmento de DNA foi amplificado por PCR, o produto de PCR foi desnaturado através do calor (95°C durante 5min) e da formamida

(presente na solução de deposição desnaturante), ficando o DNA em cadeia simples. Algumas amostras não eram desnaturadas antes da aplicação no gel, com o objectivo de visualizar o comportamento da dupla cadeia no gel de poliacrilamida. O marcador molecular, também não era desnaturado. No caso destas amostras foi utilizada uma solução de deposição não desnaturante.

O gel que permitia identificar a alteração da sequência que determinava o polimorfismo FGG 10034 C>T (a substituição de um C para um T) era o gel 6% TBE 1x. As condições da electroforese para este fragmento com comprimento de 300pb, eram corridas de aproximadamente 9h e 30min com intervalo de 2h e 30min entre cada aplicação, a 40W e dentro de uma câmara fria a 4°C. Todo o protocolo para a elaboração do SSCP está descrito em pormenor no anexo A.3.

As amostras foram aplicadas num gel de poliacrilamida em condições não desnaturantes. Como o aumento da temperatura pode afectar a corrida de electroforese e limitar a obtenção de resultados reprodutíveis, a electroforese era realizada numa câmara fria. Após a electroforese, o padrão de bandas do gel era revelado através da incorporação de nitrato de prata nas cadeias de DNA. No final, o gel era transferido para um papel de filtro e seco sob vácuo.

Sempre que as amostras apresentavam um padrão de mobilidade diferente nos fragmentos analisados, esta migração diferencial correspondia a uma variação da sequência. A identificação da alteração molecular presente nessa amostra foi realizada através de sequenciação directa dos fragmentos. A partir do momento em que um padrão era associado a uma alteração, não havia necessidade de repetir a sequenciação: o mesmo padrão correspondia sempre à mesma alteração.

2.2.5 Purificação do produto de PCR

As amostras foram purificadas usando o kit da Microcon[®] com um sistema de filtros e através de centrifugação (*Centrifugal Filter Devices*).

Este kit purifica os fragmentos amplificados por PCR através da passagem por colunas que retêm os fragmentos de DNA com interesse, fazendo a eluição de todos os produtos e componentes que não são de interesse. Estas colunas possuem filtros selectivos que permitem a passagem dos componentes da reacção que sejam inferiores a 125pb para fragmentos em cadeia dupla e 300pb para fragmentos em cadeia simples. Tudo o que seja superior a estes valores fica retido no filtro da coluna e tudo o que seja inferior, como sais, *primers* e outras impurezas que estejam presentes no produto do PCR e que possam interferir com os processos seguintes são eluídas. As lavagens foram efectuadas com água destilada autoclavada. No final das lavagens o filtro é invertido, e pela centrifugação o DNA é eluído da membrana da coluna. O protocolo utilizado para a purificação de DNA está descrito no anexo A.4.

2.2.6 Sequenciação Automática

A sequenciação directa foi realizada num sequenciador automático *ABI Prism*[®], segundo o protocolo *ABI Prism*[®] *BigDye Terminator Cycle Sequencing Ready Reaction Kits Original* (*Applied Biosystems*[®]).

A 8,5µL do produto de PCR já purificado foi adicionado 1µL do *primer* desejado, *forward* ou *reverse* (10ng/µL) e 0,5µL de *BigDye Terminator Reaction v1.1*. Os *primers* utilizados para a reacção de sequenciação foram os mesmos *primers* utilizados para a reacção de PCR. A reacção de amplificação foi efectuada num termociclador (Biometra) com um passo inicial de desnaturação de 30 segundos a 96°C, seguido de 25 ciclos de desnaturação a 96°C durante 10 segundos. A temperatura utilizada para o emparelhamento é de menos 2°C do que o utilizado para o emparelhamento na reacção de PCR por 5 segundos e extensão a 60°C durante 4 min.

2.2.7 TaqMan[®] Genotyping Assay

O Polimorfismo FGG 9340 T>C (rs1049636) foi analisado através do *kit TaqMan*[®] *Genotyping Assay* [63]. O princípio deste *kit* assenta na utilização de sondas oligonucleotídicas marcadas com fluorocromos que são específicas para cada alelo. A sonda emparelha com o produto amplificado de PCR sendo depois clivada pela actividade 5'-3'exonucleotídica da enzima Taq, libertando a marca fluorescente da extremidade 5'da sonda.

O kit possui dois componentes: *TaqMan*[®] *Genotyping Master Mix* e *TaqMan*[®] *Genotyping Assay Mix*. O primeiro componente é constituído pela DNA polimerase *AmpliTaq Gold*[®], Ultra Pura (UP), dNTPs, pelos componentes do tampão optimizados para a reacção e pelo marcador de referência *ROX* (referência interna pela qual o sinal da fluorescência do repórter é normalizada durante a análise de dados). A normalização é necessária para corrigir flutuações de fluorescência devido a diferenças nas concentrações ou volume das diferentes amostras. O segundo componente contém dois *primers* para amplificar a sequência de interesse e duas sondas *TaqMan*[®] *MGB* para detectar os alelos. O marcador fluorescente ligado à extremidade 5'da sonda que identifica o alelo A do polimorfismo FGG 9340 T>C é o *VIC*[®], e o marcador da sonda que identifica o alelo G é o *6FAM*[™]. Cada sonda possui um *quencher* na extremidade 3'.

Durante o PCR, cada sonda emparelha especificamente com a sua sequência complementar, se presente, no fragmento amplificado. Quando a sonda está intacta, a proximidade do *quencher* ao marcador fluorescente suprime a fluorescência do repórter. A DNA polimerase *Amplitaq Gold*, *UP* apenas cliva as sondas que estão hibridadas na sequência alvo. A clivagem separa o repórter do *quencher*, aumen-

tando a fluorescência do repórter, como demonstrado na figura 2.2. Deste modo, apenas ocorre um aumento de fluorescência se a sequência alvo amplificada é complementar da sonda: o sinal de fluorescência gerado na amplificação do PCR indica qual dos alelos está presente na amostra.

A presença das duas sondas em cada reacção permite fazer a discriminação alélica das possíveis variantes deste SNP numa sequência de DNA alvo. Este teste detecta a presença ou ausência de cada um dos alelos baseado na mudança de fluorescência associada a cada uma das sondas, como está exemplificado na figura 2.2.

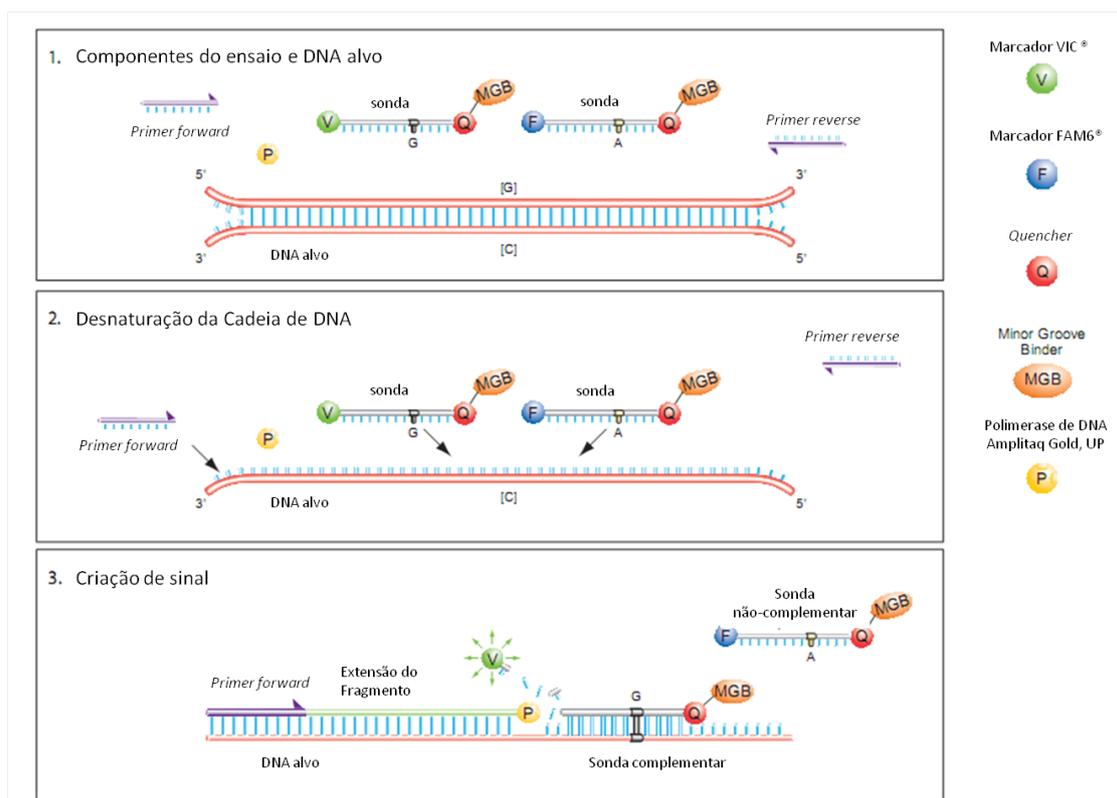


Figura 2.2: *TaqMan*[®] *Genotyping Assay*. A sonda fluorescente complementar à cadeia alvo de DNA emparelha. Durante a extensão a polimerase de DNA AmpliTaq Gold, UP cliva o marcador (porque possui a característica de revisão de provas e detecta um “erro”). Quando o marcador fluorescente fica distante do *Quencher* emite fluorescência que é detectada [adaptado [63]].

Para cada amostra foram utilizados 9 μL de *TaqMan*[®] *Genotyping Master Mix*, 0,25 μL *TaqMan*[®] *Genotyping Assay Mix* (40x) e 4,5 μL de água destilada. O volume de amostra de DNA foi de 11,25 μL com uma concentração de DNA de 1 a 10ng. Eram sempre incluídos em cada ensaio controlos negativos, substituindo o volume de amostra de DNA por água destilada.

A análise através deste *kit* divide-se em 4 passos: 1º é feita uma leitura da fluorescência da placa antes da reacção de PCR; 2º é realizada a amplificação através de PCR; 3º é feita uma leitura da fluorescência da placa depois da reacção de PCR; 4º é feita a discriminação alélica baseada na diferença das fluorescências medidas

antes e depois da realização do PCR.

Tanto o PCR como a leitura de fluorescência terminal para a discriminação alélica foram efectuadas num termociclador com leitura óptica ABI PRISM[®] 7000. Os passos que constituíam o PCR eram uma desnaturação inicial a 95°C durante 10 minutos, seguido de 40 ciclos com desnaturação a 95°C durante 15 segundos e emparelhamento/extensão a 60°C durante 1 minuto.

2.2.8 Outros métodos utilizados

Para a análise do polimorfismo FGG 9340 T>C foram efectuadas inúmeras tentativas, utilizando diversos métodos, até alcançar um método que conseguisse identificar o polimorfismo de uma forma clara e reprodutível.

dHPLC – Denaturing High-Pressure Liquid Chromatography

A técnica de dHPLC (*Denaturing High-Pressure Liquid Chromatography* ou Cromatografia Líquida desnaturante de Alta Pressão) é baseada na formação de heteroduplexes e homoduplexes de DNA, e na sua posterior separação através de cromatografia líquida de fase reversa.

Usando a reacção de PCR, o fragmento de interesse, que inclui o polimorfismo FGG 9340 T>C foi amplificado através do uso dos *primers* FGG E9-1 e FGG E9-R. Durante o dHPLC o fragmento é desnaturado sendo depois a sua renaturação feita gradualmente. Este passo permite a formação de heteroduplexes (cadeias de DNA que diferem num par de bases e portanto possuem um *mismatch* na sua sequência) e homoduplexes (cadeias de DNA idênticas em que o emparelhamento é perfeito). Assim, os heteroduplexes formam-se quando a amostra apresenta heterozigotia para o *locus* em estudo e os homoduplexes formam-se quando a amostra for homozigótica. Os heteroduplexes apresentam uma temperatura de desnaturação diferente dos homoduplexes e não serão retidos na coluna de cromatografia durante o mesmo tempo. As diferenças de tempo de retenção da coluna de cromatografia introduzem um padrão particular que permite distinguir amostras homozigóticas de heterozigóticas. Contudo, esta técnica não permite distinguir os dois tipos de homozigóticos, pois apenas aparece um pico que correspondente ao tempo de retenção dos homoduplexes. A temperatura a que será efectuada a renaturação dos fragmentos é calculada com base na sequência de bases do fragmento. Para o fragmento em estudo esta temperatura foi de 56°C, e representa a temperatura em que as cadeias começam a desnaturar, mas ainda não se encontram totalmente separadas.

Optimização de PCR

Através da análise do padrão de desnaturação do fragmento amplificado pela reacção de PCR utilizando os *primers* FGG E9-1 e FGG E9-R, obtido na análise por dH-PLC, foi possível identificar a existência de uma região onde a temperatura de desnaturação era bastante mais elevada do que no resto do fragmento. Como esta discrepância de temperaturas de desnaturação poderiam estar a influenciar negativamente a análise do polimorfismo, desenhou-se um novo *primer forward*, (FGG E9-2) que eliminava essa região (cerca de 50pb).

A reacção de PCR para amplificação do fragmento que continha o polimorfismo FGG 9340 T>C, com os *primers* FGG E9-2 e FGG E9-R, foi otimizada de modo a eliminar fragmentos inespecíficos e a melhorar a eficiência da amplificação. Para a optimização, foram efectuados gradientes de temperatura de emparelhamento, modificação do tempo de emparelhamento e extensão, uso de DMSO (Dimetilsulfóxido) a 5% e a 10% e uso de diferentes enzimas Taq polimerase (marcas diferentes, com diferentes eficiências). Foi também utilizado um *kit* de optimização de PCR, *Stratagene's Opti-Prime PCR optimization kit*. Este *kit* consiste numa matriz de tampões e aditivos de PCR que aumentam a qualidade das reacções de amplificação. Os tampões diferem entre si na concentração de MgCl₂, Cloreto de Potássio (KCl) e no pH.

A reacção foi optimizada com a utilização de um tampão do *kit* referido que continha 100mM Tris-HCl (pH 8.3), 15mM de MgCl₂ e 250mM de KCl.

ARMS – *Amplified Refractory Mutation System*

O ARMS (*Amplified Refractory Mutation System* ou PCR alelo-específico) permite detectar SNPs e pequenas deleções, tendo como princípio o uso de *primers* com sequências específicas que emparelham no local onde se encontra a alteração. Como são específicos, apenas emparelham quando a sequência do *primer* é complementar da sequência do DNA, e apenas nestas condições é permitida a amplificação do mesmo. Assim, este é um PCR refractário: se não ocorre emparelhamento não poderá haver amplificação. A presença ou ausência de produto de PCR determina a presença ou ausência do alelo em estudo [64].

É necessário realizar sempre duas reacções de PCR para cada amostra. Estas duas reacções diferem apenas no *primer* alelo-específico, sendo todos os outros componentes iguais.

Os *primers* para ARMS foram desenhados sabendo a sequência de DNA que continha o polimorfismo em questão. Os *primers* tinham 25 bases e diferiam entre si de acordo com os alelos do SNP em estudo: *primer* FGG-9340T, contém um T na extremidade 3' que emparelha com o alelo A e o *primer* FGG-9340C, contém um C que emparelha com o alelo G. O *primer reverse* utilizado era o FGG E9-R.

Para a realização de ARMS é necessário que na reacção existam *primers* controlo, que avaliem a reacção de PCR. Deste modo, quando há amplificação do fragmento de DNA utilizando *primers* controlo, mas não há amplificação do fragmento com os *primers* de ARMS, pode afirmar-se que a ausência de fragmento amplificado se deve à ausência do alelo em estudo, e não porque as condições de PCR não permitiam a amplificação de qualquer fragmento. Os *primers* escolhidos para fazer o controlo do ARMS eram os usados para amplificar o exão 25 do FV da cascata de coagulação desenhados para o estudo do mesmo a decorrer no laboratório.

HRM - *High Resolution Melting*

A técnica de HRM (*High Resolution Melting* ou Desnaturação de Alta Resolução) assenta na previsibilidade da temperatura a que duas cadeias de DNA de fragmentos amplificados começam a separar-se (esta temperatura depende da sequência de bases de DNA). Quando se comparam duas amostras com a mesma sequência de DNA a forma da curva de desnaturação há-de ser exactamente a mesma. Contudo se uma das amostras tiver uma mutação na região de DNA que foi amplificada, isto irá alterar a temperatura a que as cadeias de DNA começam a separar-se. Esta diferença de temperatura reflecte-se em curvas de desnaturação distintas. A diferença poderá ser pequena, apenas uma fracção de um grau, mas o instrumento onde é efectuado o HRM e que efectua as leituras de fluorescência das sondas que hibridam com o produto de PCR (*LightCycler*[®] 480) tem a capacidade de monitorizar este processo em "alta resolução", sendo possível documentar essas pequenas mudanças com precisão e identificar a presença de uma mutação [65].

Na análise de um SNP existem três possibilidades, que irão dar origem a três curvas de desnaturação distintas: nenhum dos alelos está alterado (homozigótico *wild-type*), um dos alelos está alterado (heterozigótico), ambos os alelos estão alterados (homozigótico mutado). Com a técnica de HRM é possível distinguir entre estas três possibilidades [55].

O HRM foi efectuado utilizando um *kit LightCycler*[®] 480 *High Resolution Melting Master* da Roche[®], que contém *Master Mix* composta por *FastSart Taq* DNA polimerase, dNTPs e um marcador fluorescente próprio para HRM. No kit estava também disponível MgCl₂ para ajustar a sua concentração na reacção.

Para a reacção de HRM foram utilizados 10μL de Master Mix, 1μL de cada um dos *primers* FGG E9-2 e FGG E9-R (0,2μM cada), 1,6μL de MgCl₂ (que equivale a concentração final na reacção de 2mM) e 30ng de DNA por reacção. O volume final era ajustado com água, também fornecida pelo *kit*. É muito importante que as amostras tenham a mesma concentração de DNA genómico, para que as curvas de desnaturação não variem muito entre si. Deste modo a concentração exacta de DNA foi determinada usando *NanoDrop*[®], e todas as amostras foram ajustadas

para a mesma concentração, diluindo as amostras mais concentradas. Foram sempre introduzidos controlos negativos, substituindo o volume de amostra de DNA por água.

O programa de HRM é constituído por 4 passos: o 1º passo é uma Pré-incubação a 95°C durante 1 minuto, para activação da Taq DNA polimerase e desnaturação das cadeias de DNA; o 2º passo é uma reacção de amplificação do fragmento com 45 ciclos de desnaturação a 95°C durante 10 segundos, seguido de emparelhamento a 55°C durante 15 segundos e extensão a 72°C durante 15 segundos. O 3º passo é o HRM em si, desnaturação a 95°C, e posteriormente a temperatura decresce até aos 40°C (numa cadência de 1 grau por minuto). Este gradiente permite a formação de heteroduplexes. Depois a temperatura sobe novamente até aos 85°C, temperatura em que todos os fragmentos estão desnaturados novamente, e as cadeias de DNA estão completamente separadas. O 4º passo é o arrefecimento do sistema.

O 3º passo é monitorizado em tempo-real, através do uso de um marcador fluorescente, sendo efectuada a leitura da fluorescência presente na reacção. Os marcadores usados em HRM são marcadores intercalantes que têm uma propriedade única: ligam-se especificamente à dupla cadeia de DNA e apenas quando estão ligados existe fluorescência. Na ausência de dupla cadeia de DNA os marcadores não estão ligados e apenas existe uma pequena fluorescência basal. Assim, no início da análise de HRM existe uma elevada fluorescência nas amostras, porque existem biliões de cópias dos fragmentos de DNA amplificados, com o marcador intercalado entre as duas cadeias. Quando as amostras começam a ser desnaturadas e as duas cadeias de DNA começam a separar-se a fluorescência é reduzida. A fluorescência é medida durante este processo e no final os dados são organizados num gráfico, conhecido como curva de desnaturação, que mostra os níveis de fluorescência em função da temperatura.

2.3 Análise Estatística

Em ambos os grupos (subgrupo TEV e grupo controlo) foi calculado o Equilíbrio de Hardy-Weinberg (EHW) para cada um dos polimorfismos através do teste Qui-Quadrado.

O EHW prevê que, dentro de determinadas condições, as frequências alélicas permanecem constantes com o passar das gerações. As condições (ou pressupostos) necessárias para o EHW são que a população considerada seja ideal: uma população infinita (no sentido de eliminar a deriva genética), com reprodução sexuada, os indivíduos acasalem de uma forma aleatória (não há selecção sexual ou desvio de aleatoriedade por dispersão geográfica), diplóide, com número de fêmeas e machos igual, todos os casais são férteis e têm o mesmo número de prole, e não sofre de

selecção natural, mutações, nem migração (não ocorre fluxo genético). Quando algum destes pressupostos do EHW não é estabelecido na população em estudo, ocorre um desvio das expectativas.

No caso de um único locus com dois alelos **A** e **a**, com frequências alélicas p e q , respectivamente, o princípio de Hardy-Weinberg prediz que a frequência genotípica para o homozigótico **AA** será p^2 , para o heterozigótico **Aa** será $2pq$ e para o homozigótico **aa** será q^2 .

Para saber se existe um desvio do EHW, foi utilizado o teste Qui-Quadrado de Pearson, usando as frequências genotípicas observadas (obtidas através dos dados observados na população) e as frequências esperadas que são calculadas através do princípio de Hardy-Weinberg. A hipótese nula é que a população está em EHW e a hipótese alternativa é que a população não está nessas condições. Os graus de liberdade para o teste Qui-Quadrado são calculados através da fórmula: Graus de liberdade = (número de alelos*(Número de alelos-1))/2 [66].

Para testar a semelhança entre variáveis categóricas discretas (que resultam da contagem de "objectos" que pertencem a uma de várias categorias), por exemplo o número de indivíduos com determinado polimorfismo na população em estudo e na população controlo, é utilizado o teste Qui-Quadrado de Pearson. Para a utilização deste teste é necessário que os dois grupos sejam independentes e que as observações sejam frequências ou contagens. A hipótese H_0 é que não existe diferença entre as frequências (contagens) dos grupos e a hipótese alternativa é que existe diferença. O resultado observado deve ser comparado com que seria obtido se não houvesse diferença entre os grupos. Os dados foram organizados em Excel. Os valores esperados para cada célula foram calculados através da multiplicação do total da coluna pelo total da linha e dividindo este resultado pela soma total. A fórmula usada para calcular o Qui-Quadrado é a seguinte:

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

O valor do teste Qui-Quadrado, obtido a partir da soma de todos os Qui-Quadrado, deve ser comparado com o valor de Qui-Quadrado tabelado. Quanto maior o valor calculado, maior a diferença entre as duas populações em estudo. A decisão de rejeitar a hipótese nula está associada a uma probabilidade de erro – o nível de significância do teste, representa a máxima probabilidade de erro que se tem ao rejeitar uma hipótese. O *pvalue* associado ao teste do Qui-Quadrado foi calculado no Excel através da fórmula: *Test Qui*.

O teste *t-Student* foi utilizado quando os dados seguem uma distribuição normal e se pretende observar se duas populações analisadas na mesma unidade estatística têm um valor médio diferente. Neste caso, o intuito é saber se os polimorfismos em

questão aumentam ou diminuem a idade média do primeiro acidente trombótico. Assim, as idades do primeiro acidente trombótico são separadas consoante a presença de cada um dos genótipos. Como o teste é utilizado em tamanhos amostrais diferentes e variância diferente, sabemos que valor do teste t e os graus de liberdade (g.l.) se calculam através das seguintes fórmulas:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

$$\text{g.l.} = \frac{(s_1^2/n_1 + s_2^2/n_2)^2}{(s_1^2/n_1)^2/(n_1 - 1) + (s_2^2/n_2)^2/(n_2 - 1)}$$

A hipótese nula e a hipótese alternativa exprimem convicções sobre a realidade, convicções decorrentes das condições do problema em análise. Neste caso é possível afirmar que a idade média do primeiro acidente trombótico será mais elevada no caso de um polimorfismo protector estar a influenciar a ocorrência do mesmo; por outro lado, a idade será mais baixa no caso de um polimorfismo potenciador. A hipótese nula (H_0) exprime que não há qualquer relação entre o polimorfismo e a idade do primeiro acidente trombótico e a hipótese alternativa (H_A) exprime a opinião de que a idade é superior ou inferior, considerando polimorfismos protectores ou potenciadores, respectivamente. O objectivo é provar que a evidência factual rejeita a hipótese nula. Rejeitar a hipótese nula é uma decisão forte, quer dizer que os dados continham informação suficiente para isso. Manter a hipótese nula é em geral uma decisão fraca, no sentido em que pode decorrer apenas de a evidência factual ser insuficiente para rejeitar aquela hipótese [66].

Para comparar o número de indivíduos com patologia trombótica que apresentam cada um dos genótipos dos polimorfismos em estudo, foi calculado o *odds ratio* (OR). O cálculo de OR é um método de comparação entre a ocorrência de um genótipo/alelo entre casos e controlos, permitindo obter uma estimativa do risco relativo. O OR é definido como a razão entre a probabilidade de ocorrência de determinado genótipo/alelo nos casos e nos controlos, e é obtido através da equação demonstrada na figura 2.3.

	Casos	Controlos
Genótipo comum	A	B
Variante alélica	C	D

$$\text{OR} = \frac{A/C}{B/D}$$

Figura 2.3: Cálculo de odds ratio num estudo de casos e controlos (estimativa do risco relativo). Tabela exemplificativa da equação realizada para o cálculo de OR. A, B, C e D são o número de indivíduos com as características apresentadas na tabela.

Se a frequência de determinado genótipo/alelo do polimorfismo é maior entre os casos, o OR será superior a 1, indicando um risco aumentado de patologia trombótica. Quanto mais forte for a associação entre o polimorfismo e a patologia trombótica, maior será o resultado do OR. Inversamente, se a frequência de determinado genótipo/alelo for mais baixa entre os casos, o OR será menor que 1, indicando a presença de um factor protector em relação à patologia.

2.4 Haplótipos

Com a informação relativa aos polimorfismos em estudo foram determinados os haplótipos. Como existem dois polimorfismos bi-alelicos, foram determinados 4 haplótipos, que constituem todas as hipóteses de associação (Tabela 2.1):

Tabela 2.1: Haplótipos constituídos com os polimorfismos analisados nestes estudo

Haplótipo	FGG 10034 C>T	FGG 9340 T>C
PTH1	C	T
PTH2	C	C
PTH3	T	T
PTH4	T	C

Para a constituição destes haplótipos foram consideradas as combinações genotípicas dos polimorfismos FGG 10034 C>T e FGG 9340 T>C, como exemplificado no esquema da figura 2.4.

Considerando as informações disponíveis na bibliografia, que descrevem o efeito do polimorfismo FGG 10034TT como potenciador do acidente trombótico e o efeito do polimorfismo FGG 9340CC como protector, espera-se que o haplótipo PTH1 tenha um efeito neutro, o haplótipo PTH2 um efeito protector, o haplótipo PTH3 um efeito potenciador e o haplótipo PTH4 terá também um efeito neutro, como o haplótipo PTH1.

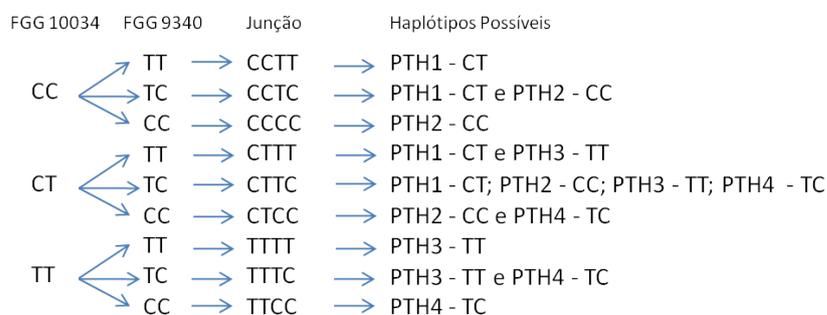


Figura 2.4: Esquema elucidativo da determinação dos haplótipos a partir das combinações genotípicas de cada um dos polimorfismos

Foram calculadas as frequências dos haplótipos para o grupo controle e subgrupos de estudo (TEV, grupo saudável + 42 anos e TEV - 32 anos), e foram comparadas entre si, usando os testes estatísticos mencionados anteriormente. Foram calculadas as idades médias do primeiro acidente trombótico para cada uma das combinações genotípicas dos dois polimorfismos. Sempre que o número era consideravelmente grande, era aplicado o teste estatístico *T-student*, para avaliar se as diferenças entre a média de idade do primeiro acidente trombótico eram significativas.

Capítulo 3

Resultados e Discussão

3.1 Caracterização dos grupos de estudo e controlo

3.1.1 Grupo de estudo

O grupo de estudo inclui famílias com pelo menos um indivíduo portador da mutação FV Leiden e seus familiares, num total de 110 famílias com 280 indivíduos. Contudo, dado que a informação clínica não se encontrava completa, foram excluídas 9 famílias, com 31 indivíduos. Deste modo, o grupo de estudo passou a ser constituído por 101 famílias com 249 indivíduos. Destes, 173 eram portadores da mutação FV Leiden em heterozigotia (R/Q), 7 em homozigotia (Q/Q), e 63 não eram portadores (R/R); em 6 casos não foi efectuada a determinação. Dos indivíduos portadores da mutação FV Leiden, 81 não apresentavam patologia trombótica até ao momento.

Nos 99 indivíduos com patologia trombótica, a primeira manifestação clínica foi em 42 casos TVP, em 23 TFS, em 2 EP, em 8 TV, em 2 CAD, em 6 Aborto, em 4 AVC e em 2 EAM. Num caso o indivíduo apresentava hipertensão pulmonar, noutro o indivíduo teve uma TFS juntamente com EP, em 7 dos casos ocorreu TVP juntamente com EP e houve um caso de parto prematuro no 3º trimestre da gravidez.

Foram definidos dois tipos principais de acidente trombótico: arterial e venoso. No acidente trombótico venoso, foram considerados TV, TVP, EP, TFS e Aborto. Como os indivíduos em que a primeira manifestação clínica foi trombose arterial não faziam parte do objectivo do estudo e eram em número insuficiente para constituírem um grupo com significado estatístico (apenas 10 indivíduos) foram excluídos. A média das idades do conjunto de indivíduos que participaram no estudo é de 43 anos (n=137), variando entre os 7 e os 84 anos. A idade do primeiro acidente trombótico varia entre os 17 e os 77 anos de idade, caso se trate de um acidente espontâneo, e os 13 e os 70 anos de idade caso se trate de um acidente secundário. A tabela 3.1 mostra, as características das famílias que constituem o grupo de estudo. No anexo ?? pode observar-se uma tabela com as informações detalhadas correspondentes a cada uma das famílias consideradas para este estudo.

Tabela 3.1: Caracterização do grupo de estudo

Características dos indivíduos em estudo	
Pacientes com acidente trombótico	45,0% (n=112)
Indivíduos saudáveis	55,0% (n=137)
Idade média 1º Acidente trombótico	37 anos
Espontâneo	40 anos (17-77) (n=42) 40,8%
Secundário	35 anos (13-70) (n=61) 59,2%
Tipo do 1º Acidente Trombótico	
Arterial (AVC, CAD, EAM)	8,9% (n=10)
Venoso	
TVP	41,1% (n=46)
TFS	24,1% (n=27)
Aborto	6,3% (n=7)
Outros (TV, EP entre outros)	19,6% (n=22)
Factores Secundários:	
Fumador	3,6% (n=4)
Uso de Anti-contraceptivos Orais	25,5% (n=28)
Gravidez	13,6% (n=15)
Imobilização	9,1% (n=10)
Cirurgia	11,8% (n=13)

O gráfico 3.1 ilustra o intervalo de idades em que ocorre o primeiro acidente trombótico, subdividido em indivíduos com e sem factores secundários.

É possível identificar uma maior ocorrência do primeiro acidente trombótico no intervalo correspondente aos 20-29 anos de idade, sendo mais notório este acontecimento nos indivíduos que apresentam um factor de risco secundário. A partir de idades superiores a 50 anos o acidente trombótico do tipo espontâneo é muito mais frequente do que o acidente trombótico do tipo secundário.

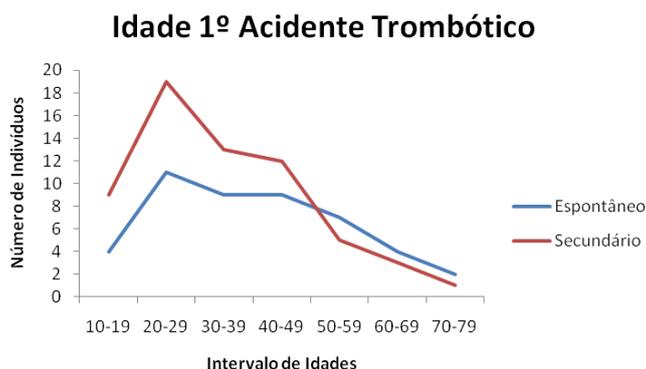


Figura 3.1: Intervalo de idades em que ocorre o primeiro acidente trombótico.

3.1.2 Grupo controlo

O grupo controlo é constituído por indivíduos saudáveis, dadores de sangue. A mutação FV Leiden apenas estava presente em 2 indivíduos do grupo controlo (1,9%), enquanto que a variante da Protrombina 20210GA estava presente em heterozigotia em 5 (4,6%).

3.2 Determinação da presença do Polimorfismo FGG 10034 C>T

O Polimorfismo FGG 10034 C>T foi determinado através de SSCP. Os diferentes padrões foram sequenciados para confirmar qual a alteração molecular que estava subjacente aos diferentes padrões de SSCP.

3.2.1 SSCP

Para determinar a presença do polimorfismo FGG 10034 C>T nas amostras em estudo, primeiro foi realizado um PCR que originava um fragmento de aproximadamente 330pb. Na figura 3.2 é apresentada uma fotografia representativa dos resultados obtidos por PCR.

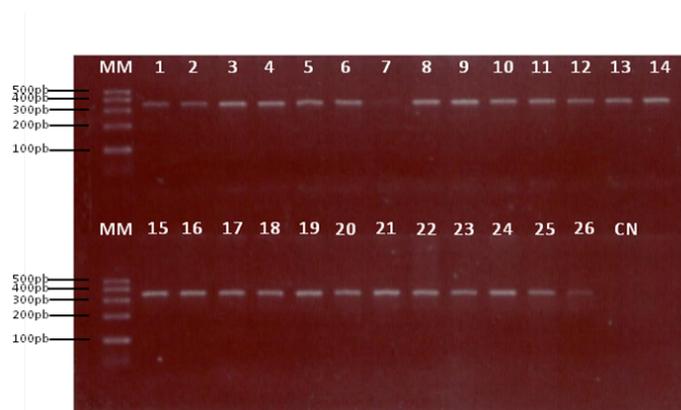


Figura 3.2: Fotografia do Gel de Electroforese (2% Agarose) exemplificativo dos resultados obtidos pela amplificação por PCR do fragmento de aproximadamente 330pb, produzido através do uso dos *primers* FGG E10-1 e FGG E10-R. MM, Marcador Molecular. CN, Controlo Negativo.

Posteriormente à amplificação da amostra de interesse por PCR, a discriminação da presença de alterações genéticas no fragmento amplificado era feito através de SSCP. As amostras apresentavam padrões de mobilidade diferentes, indicando a existência de três variações a nível da sequência de DNA. A figura 3.3 corresponde a uma imagem representativa dos padrões de migração do fragmento.

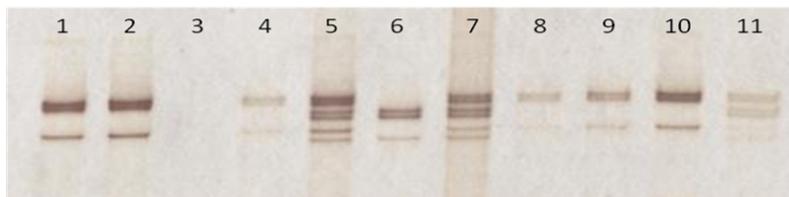


Figura 3.3: Fotografia do gel de poliácridamida de SSCP, representativa dos padrões observados. É possível observar os padrões de migração distintos (melhor definidos nas linhas 5, 6 e 10). Nas linhas 5, 7 e 11 é possível distinguir um padrão em heterozigotia (C/T), nas linhas 1, 2, 4, 8, 9, 10 um padrão distinto que corresponde ao genótipo mais comum, homozigótico (C/C). Na linha 6 é possível identificar o 3º padrão distinto, que corresponde ao genótipo menos frequente, homozigótico (T/T).

3.2.2 Sequenciação

Através da sequenciação directa dos fragmentos que apresentavam padrões de migração diferentes em SSCP, foi possível identificar e confirmar a alteração molecular que estava subjacente a cada padrão. Os três genótipos correspondentes ao polimorfismo FGG 10034 C>T, foram identificados como sendo os responsáveis pela alteração do padrão de migração observado nos gels de SSCP. Na figura 3.4 é possível observar o resultado da reacção de sequenciação correspondente aos 3 genótipos do polimorfismo FGG 10034.

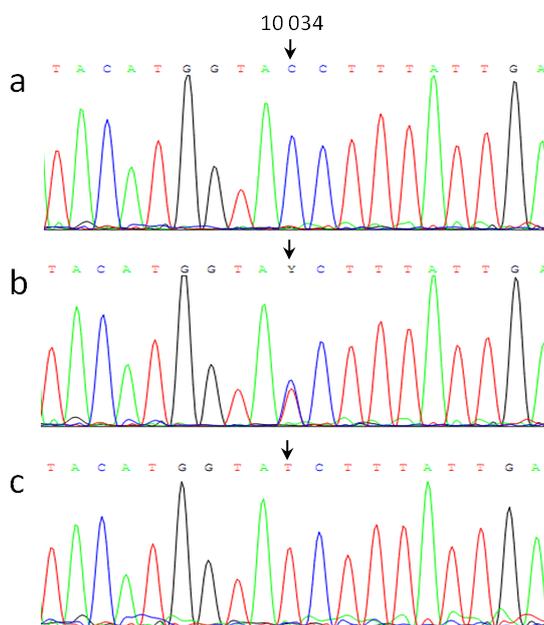


Figura 3.4: Electroferograma da reacção de sequenciação, identificando o polimorfismo FGG 10034 C>T. O produto de PCR é sequenciado e a mutação detectada. A transição C>T, evidenciada pela seta, é assim identificada e associada ao padrão de SSCP. (a) Sequência com o alelo C na posição 10034 (c) Sequência com o alelo T. (b) Sequência com ambos os alelos.

3.3 Determinação da presença do Polimorfismo FGG 9340 T>C

3.3.1 SSCP

Para a determinação da presença do polimorfismo FGG 9340 T>C nas amostras das duas populações em estudo, primariamente foi utilizada SSCP, utilizando os fragmentos da reacção de amplificação com os *primers* FGG E9-1 e FGG E9-R, que originava fragmentos de aproximadamente 230pb. Contudo, não foi possível observar diferenças de migração nas várias amostras aplicadas, não sendo apresentados padrões distintos. A transição de bases observada neste polimorfismo não se revelou suficiente para provocar uma alteração na conformação da molécula de DNA que fosse visível num gel de poliacrilamida, através de uma migração diferencial.

A primeira tentativa de identificação do polimorfismo não se revelou eficiente, portanto foi efectuada uma nova tentativa, ainda com recurso a SSCP, com um *primer forward* diferente (FGG E9-2), que originava um fragmento mais pequeno, com cerca de 180pb. A finalidade desta tentativa era eliminar um domínio da molécula que pudesse estar a influenciar a sua conformação: através da sua eliminação a molécula apresentaria uma conformação distinta, dependente do polimorfismo em estudo, alterando o padrão de migração. Contudo os resultados foram infrutíferos, continuando a não haver alterações na migração das várias amostras num gel de SSCP.

3.3.2 dHPLC

Posteriormente para a análise deste mesmo polimorfismo, foi utilizado dHPLC. Os resultados de dHPLC também não se apresentaram conclusivos, sendo impossível identificar com clareza o polimorfismo.

3.3.3 ARMS

O ARMS é um método simples e rápido, sendo apenas necessário realizar um PCR com *primers* específicos para cada alelo em estudo, sem gasto adicional de reagentes ou enzimas. Esta técnica tem como limitação o facto de não detectar nenhuma alteração para além daquela que é pretendida através da análise – os *primers* são especificamente escolhidos para identificar determinada alteração e mesmo que outras alterações estejam presentes no fragmento amplificado, essas não são identificadas.

Com ARMS, foi possível identificar todas as alterações no polimorfismo em questão. Todavia, os resultados não eram reprodutíveis. Um exemplo dos resultados obtidos por ARMS está representado na figura 3.5 a) e b).

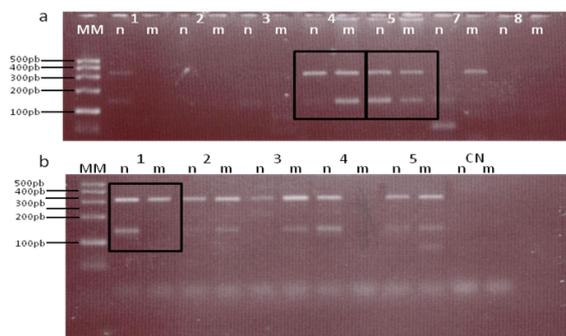


Figura 3.5: (a) e (b) ARMS para o polimorfismo FG9340 T>C. As bandas de 350pb, são os fragmentos controle co-amplificados, correspondentes ao exão 25 do FV. Para cada amostra são utilizados os *primers* específicos para o alelo normal (n) (FG9340T) e para o alelo mutado (m) (FG9340C). MM, marcador molecular. CN, controle negativo. A amostra 4 da figura 3a é exemplo do alelo mutado (C) em homozigotia. A amostra 5 da figura 3a é exemplo de heterozigotia. A amostra 1 da figura 3b é exemplo do alelo normal (T) em homozigotia.

3.3.4 Sequenciação

O produto de PCR, utilizando o *primer reverse* FG9 E9-R foi sequenciado e o polimorfismo anteriormente identificado através de ARMS foi detectado. A Sequenciação automática confirmou os três genótipos correspondentes ao polimorfismo FG9340 T>C (Figura 3.6).

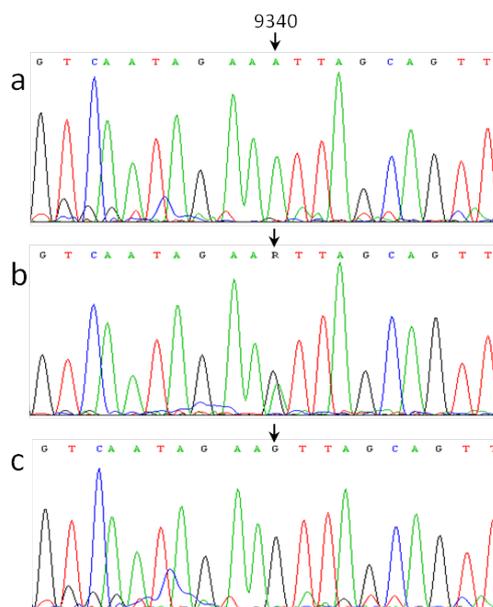


Figura 3.6: Electroferograma da sequenciação, identificando o polimorfismo FG9340 T>C. Como é o *primer reverse* o utilizado, a sequencia apresentada é da cadeia não codificante. Deste modo a transição T>C, é representada pela sua complementar, A>G evidenciada pela seta. (a) Sequencia com A na posição 9340 (c) Sequencia apresentando G. (b) Sequencia em heterozigotia.

3.3.5 HRM

Visto que as metodologias mais tradicionais se mostraram incapazes de demonstrar resultados reprodutíveis, foi necessária a aplicação de técnicas mais recentes e mais específicas. O HRM é um método que não necessita de nenhum processamento ou adição de reagentes depois da realização do PCR: a partir do momento em que são obtidos os fragmentos de DNA amplificados, o processo de desnaturação é monitorizado através de um marcador fluorescente que não inibe a reação de PCR.

Quando combinado com PCR em tempo real (RT-PCR – *Real Time PCR*) apresenta-se como uma solução simples, semi-automática sem custos elevados para a análise de SNPs e pequenas deleções/inserções. Os resultados obtidos por HRM são de fácil análise, pois são apresentados na forma de gráficos de curvas de desnaturação. Estes gráficos colocam em evidência o diferente comportamento de cada amostra de acordo com o seu genótipo, através de perfis de desnaturação distintos, como se pode observar na figura 3.7a, permitindo uma fácil identificação de cada uma das variantes do polimorfismo FGG 9340 T>C. Para evidenciar a interpretação dos resultados, é possível efectuar um *difference plot*. Para a realização deste, escolhe-se uma amostra e considera-se que a fluorescência dessa amostra é 0. Depois, à fluorescência de cada amostra é subtraída a fluorescência da amostra que se considerou 0. As amostras que tiveram fluorescência superior vão ser positivas, e as amostras que tiverem fluorescência inferior vão ser negativas. Quando esta informação é ilustrada sob a forma de gráfico os diferentes perfis obtidos tornam-se mais evidentes.

Exemplo de um *difference plot* está ilustrado na figura 3.7b, onde se definiu que o homozigótico CC seria o 0 de fluorescência.

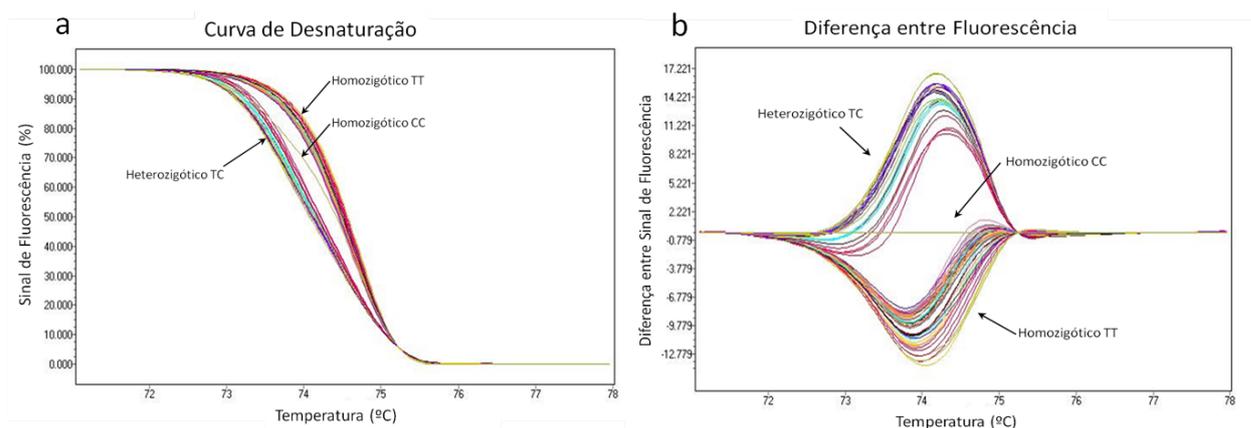


Figura 3.7: Para cada amostra é medida a fluorescência, em tempo real, durante o processo de desnaturação. Em ambas as figuras é realizada a correspondência entre cada perfil de desnaturação e o genótipo correspondente. (a) Curva de Desnaturação Normalizada. (b) Diferença entre fluorescências ou *difference plot*. Em evidência a disparidade entre cada curva, acentuando a distinção entre os diversos genótipos.

Muitos métodos de detecção de variações da cadeia de DNA entre indivíduos são utilizados, como por exemplo SSCP, dHPLC, ARMS, entre outros. Todos estes métodos requerem a separação das amostras num gel ou numa matriz, muitos são manuais e requerem muito trabalho, enquanto que outros são complexos e requerem material apropriado. As vantagens do HRM advêm do facto de ser simples, não precisando nem de separação dos fragmentos nem de processamento das amostras. Como são necessários menos passos de manuseamento, o risco de contaminação da amostra é reduzido. Comparativamente com outras técnicas, nomeadamente dHPLC, esta técnica apresenta uma maior sensibilidade, especificidade, praticabilidade e custo acessível. A reprodutibilidade e eficiência da detecção foram testados através de duplicados de amostras, em experiências diferentes, que eram coerentemente determinados.

3.3.6 *TaqMan Genotyping Assay*

Apesar do HRM se ter demonstrado eficiente, não foi possível analisar todas as amostras em questão através do mesmo, tendo sido utilizado um *kit TaqMan Genotyping Assay*.

Durante o PCR, cada sonda emparelha especificamente com a sua sequência complementar, se presente, no fragmento amplificado. O aumento de fluorescência apenas ocorre se a sequência alvo amplificada é complementar da sonda: o sinal de fluorescência gerado na amplificação do PCR indica qual dos alelos está presente na amostra. Na figura 3.8 é possível observar como as fluorescências de cada um dos marcadores é alterada durante a reacção de PCR.

Depois da reacção de PCR, realiza-se a leitura de fluorescência na placa.

Para a análise final dos resultados é possível subtrair a fluorescência obtida depois da amplificação (*Post-read*) da fluorescência obtida antes de amplificação (*Pre-Read*), permitindo eliminar o “ruído de fundo”, e flutuações devido a diferenças de concentração, como demonstrado na figura 3.9 Toda a informação é organizada num gráfico, onde se obtém a fluorescência correspondente a cada um dos marcadores dos diferentes alelos. Quando o indivíduo apresenta heterozigotia para o alelo em estudo, ambos os marcadores vão apresentar fluorescência.

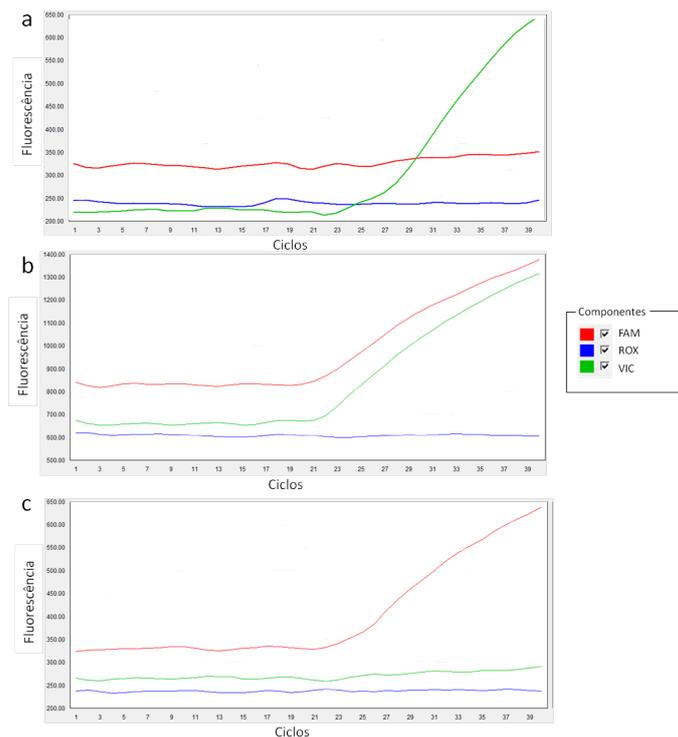


Figura 3.8: Comportamento dos diferentes componentes fluorescentes durante a reacção de PCR. Nesta figura são demonstrados os três componentes utilizados no *kit*. FAM é o fluorocromo que identifica o alelo C e o VIC, fluorocromo que identifica o alelo T do polimorfismo FGG 9340 T>C. O ROX é o componente normalizador. (a) Verifica-se um aumento do componente VIC, o que indica que o alelo T se apresenta em homozigotia. (b) Existe um aumento de ambos os componentes, o que indica heterozigotia. (c) O aumento da fluorescência emitida pelo componente FAM, indica que a amostra é homozigótica para o alelo C.

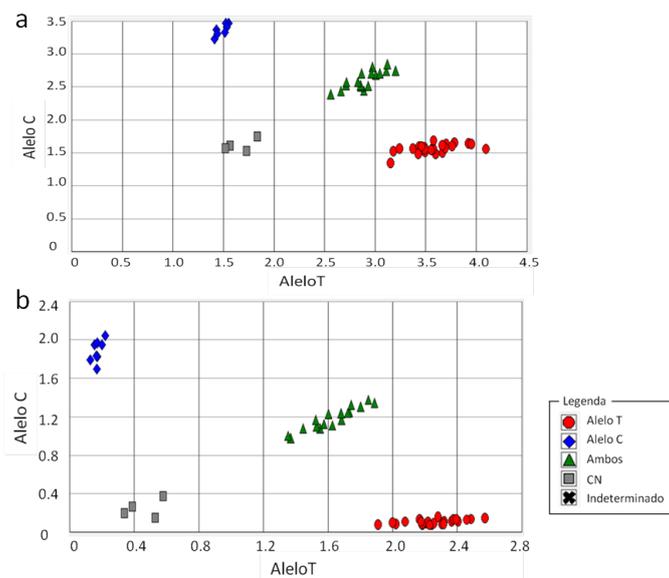


Figura 3.9: Discriminação alélica do polimorfismo FGG 9340 T>C. Em (a) apresenta-se o gráfico relativo à leitura de fluorescências apenas considerando *Post-read*, enquanto que (b) introduz a diferença das fluorescências obtidas em *Pré-read* com as de *Post-read*.

3.4 Prevalência dos polimorfismos do FGG

Depois de determinada a presença dos diversos genótipos presentes em cada amostra, para cada um dos polimorfismos, foi efectuada uma análise estatística, de modo a relacionar os dados. Primeiramente foram calculadas as frequências alélicas e genotípicas, fazendo um levantamento dos genótipos em cada um dos indivíduos não relacionados, tanto no grupo de estudo (subgrupo TEV) como no grupo controlo. Os resultados são apresentados na tabela 3.2.

Tabela 3.2: Frequências Genotípicas e Alélicas dos Polimorfismos FGG 10034 C>T e FGG 9340 T>C no grupo controlo e no subgrupo TEV

	Frequência Genotípica			Frequência Alélica	
FGG 10034 C>T	C/C	C/T	T/T	C	T
Grupo Controlo	59,36% (n=64)	38,96% (n=41)	2,78% (n=3)	78,24% (n=169)	21,76% (n=47)
Subgrupo TEV	61,45% (n=51)	38,55% (n=32)	0% (n=0)	80,72% (n=134)	19,28% (n=32)
FGG 9340 T>C	T/T	T/C	C/C	T	C
Grupo Controlo	48,15% (n=52)	50,00% (n=54)	1,85% (n=2)	73,15% (n=158)	26,85% (n=58)
Subgrupo TEV	53,01% (n=44)	37,35% (n=31)	9,64% (n=8)	71,69% (n=119)	28,31% (n=47)

Em negrito apresenta-se o alelo predisponente. O n representa o número de indivíduos.

3.4.1 Equilíbrio de Hardy-Weinberg

Para o cálculo do Equilíbrio de Hardy-Weinberg (EHW) foram utilizadas as frequências genotípicas observadas, anteriormente calculadas, e as frequências esperadas, calculadas através do princípio de Hardy-Weinberg. O teste de Qui-Quadrado foi utilizado para verificar se os resultados esperados e observados apresentavam semelhanças entre si, isto é, se a população se encontra em EHW. A hipótese nula formulada foi que as populações se encontravam em EHW e seria rejeitada, com 99% de confiança, sempre que o valor obtido pelo teste do Qui-Quadrado fosse superior a 9,2, o que corresponde a um *pvalue* de 0,001.

Subgrupo de estudo TEV

No grupo de estudo TEV, para o polimorfismo FGG 10034 C>T o valor de Qui-Quadrado obtido foi de 4,73 (*pvalue*=0,09). Para o polimorfismo FGG 9340 T>C, o valor de Qui-Quadrado obtido foi de 0,53 (*pvalue*=0,767). Estes resultados indicam que a hipótese nula não pode ser rejeitada, encontrando-se esta população em EHW para ambos os *locus*.

Grupo controlo

No grupo controlo, para o polimorfismo FGG 10034 C>T o valor de Qui-Quadrado obtido foi de 1,427 ($pvalue=0,489$). Este resultado indica que esta população se encontra em EHW para este *locus*. Para o polimorfismo FGG 9340 T>C, o valor de Qui-Quadrado obtido foi de 8,037 ($pvalue=0,018$). Para o intervalo de confiança considerado, de 99%, é possível afirmar que a população se encontra em EHW. Contudo, se fosse considerado um intervalo de confiança inferior, de 95%, já seria possível rejeitar a hipótese nula (rejeitada a partir de valores superiores a 5,99, $pvalue=0,05$).

Quando algum dos pressupostos do equilíbrio de Hardy-Weinberg não é estabelecido na população, ocorre um desvio das expectativas, como o que poderá estar a ocorrer na população controlo para o polimorfismo FGG 9340 T>C. Este desvio pode ser causado por uma falha em qualquer um dos pressupostos, contudo o mais frequente é efeito Wahlund. Este é causado por uma estratificação da população: é efectuada a análise de duas subpopulações com características distintas em conjunto. A relação de HW entre frequências alélicas e genotípicas, embora se mantenham correctas dentro de cada subpopulação separadamente, não se mantém se as subpopulações forem consideradas em conjunto como uma única população. O efeito Wahlund dificulta a interpretação das frequências alélicas e genotípicas em populações naturais, porque elas frequentemente se apresentam subdivididas e o grau de subdivisão raramente é conhecido. São também frequentes o acasalamento não aleatório, selecção natural ou deriva genética. O grupo controlo utilizado neste estudo, trata-se de uma amostra cega, constituída a partir de dadores de sangue, desconhecendo-se as características do mesmo, excepto ser constituído por indivíduos saudáveis. Deste modo, é impossível estabelecer se os pressupostos do EHW estão a ser respeitados ou não, o que pode explicar deste grupo não estar perfeitamente dentro do que seria expectável.

3.5 Efeito dos polimorfismos do FGG no risco de TV

Para verificar se existiam diferenças significativas entre o grupo controlo e o subgrupo TVE, e entre os indivíduos do subgrupo TEV que tiveram acidente com menos de 32 anos e os indivíduos saudáveis com mais de 42 anos, foi efectuada o teste de Qui-Quadrado como teste de semelhança, para as frequências alélicas e genotípicas dos polimorfismos FGG 10034 C>T e FGG 9340 T>C.

Como se pode verificar na tabela 3.3, o polimorfismo FGG 10034 C>T não apresenta diferenças quer na frequência alélica, quer na frequência genotípica, quando se compara o subgrupo TEV com o grupo controlo ($pvalue=0,552$ e $pvalue=0,309$,

Tabela 3.3: Análise Alélica e Genotípica dos polimorfismos FGG 10034 C>T e FGG 9340 T>C

Polimorfismos	Subgrupo TEV	Grupo Controlo	χ^2	Pvalue ^a	OR ^b
FGG 10034 C>T					
C	0,807	0,782	0,352	0,552	1,164
T	0,193	0,218			0,858
C/C	0,614	0,771			1,020
C/T	0,385	0,493			0,979
T/T	0	0,036			ND
FGG 9340 T>C					
T	0,717	0,731	0,100	0,751	0,929
C	0,283	0,269			1,075
T/T	0,530	0,627			1,215
T/C	0,373	0,650			0,596
C/C	0,096	0,030			5,653

Polimorfismos	Subgrupo TEV -32	Grupo Estudo Saudáveis +42	χ^2	Pvalue ^a	OR ^b
FGG 10034 C>T					
C	0,770	0,804	0,194	0,659	0,815
T	0,230	0,196			1,226
C/C	0,541	0,652			0,627
C/T	0,459	0,304			1,942
T/T	0	0,043			ND
FGG 9340 T>C					
T	0,757	0,826	0,803	0,370	0,655
C	0,243	0,174			1,527
T/T	0,595	0,652			0,782
T/C	0,324	0,348			0,900
C/C	0,081	0			ND

^a Calculado através do teste Qui-Quadrado de Pearson

^b O.R. – Odds Ratios, calculado como explicado nos métodos

ND – Não definido; em negrito o alelo predisponente

respectivamente). Através do cálculo de OR é possível inferir o risco relativo de ocorrência de TV na presença do alelo ou genótipo em estudo. Relativamente ao polimorfismo FGG 10034 C>T, podemos verificar que os valores são muito próximos de 1, o que significa que este polimorfismo não parece influenciar o risco de TV. Comparando o subgrupo TEV com o primeiro acidente trombótico com idades inferiores a 32 anos e o grupo de estudo constituído por indivíduos saudáveis com mais de 42 anos de idade, não foi possível verificar diferenças entre as frequências alélicas nem nas frequências genotípicas (*pvalue*=0,659 e *pvalue*=0,251, respectivamente). Contudo, quando consideramos os OR, verifica-se que o polimorfismo FGG 10034 C>T quando em heterozigotia (C/T) parece aumentar o risco de TVP (OR = 1,942).

Relativamente ao polimorfismo FGG 9340 T>C, parece haver uma diferença nas frequências genotípicas apresentadas pelo grupo TEV e pelo grupo controlo. Estas diferenças podem ser consideradas significativas (*pvalue*=0,025), quando se considera o intervalo de confiança de 95%. Estes resultados parecem demonstrar que existe uma maior frequência de homozigóticos (C/C) mutados na população de estudo. No entanto quando se comparam as frequências alélicas, já não é perceptível nenhuma diferença entre os grupos, sendo as frequências alélicas bastante semelhantes. O cálculo de OR parece indicar uma diminuição do risco para os portadores

deste polimorfismo em heterozigotia (T/C) e um aumento significativo do risco (OR = 5,653) para os portadores da mutação (C/C). Esta conclusão provém do facto deste polimorfismo estar mais representado na população com acidente trombótico do que na população em geral.

Quando se compara o subgrupo TEV com o primeiro acidente trombótico com idades inferiores a 32 anos e o grupo de estudo constituído por indivíduos saudáveis com mais de 42 anos de idade, não se verificam diferenças entre as frequências alélicas ou genóticas. Neste caso o OR parece apontar uma diminuição do risco para os indivíduos que não apresentam a mutação (T/T). Estes resultados advêm do facto do genótipo T/T estar mais representado na população que aos 42 anos de idade ainda se apresenta saudável.

3.5.1 *Efeito dos variantes na idade do primeiro acidente trombótico*

Para inferir a influência que os diferentes genótipos poderiam estar a exercer sobre a idade do primeiro acidente trombótico (possível modulação do acidente trombótico), foram realizadas as médias da idade do primeiro acidente trombótico consoante os genótipos que os indivíduos apresentavam. Com o intuito de saber se estas apresentavam diferenças significativas foi aplicado o teste *t* de Student.

Para o polimorfismo FGG 10034 C>T, é possível verificar um aparente decréscimo da idade do primeiro acidente trombótico na presença do alelo T. Assim, a idade média do primeiro acidente trombótico para indivíduos CC é 38 anos (n=63, *pvalue*=0,064), para os indivíduos CT a idade média apresenta-se aos 33 anos (n=37, *pvalue*=0,134), enquanto que para os indivíduos TT a média é 26 (n=2, *pvalue*=0,206). Sabendo que a idade média de ocorrência do primeiro acidente trombótico é aos 37 anos, estes resultados apontam para a presença de modulação negativa por parte do polimorfismo FGG 10034T. Esta conclusão não tem poder estatístico, limite imposto pela pequena dimensão e grande variância da amostra. Contudo, é uma indicação positiva a ter em conta, um sinal de que resultados semelhantes podem ser obtidos, com um grau de significância maior, quando o número de indivíduos analisados for aumentado.

Para o polimorfismo FGG 9340 T>C, é possível verificar o efeito inverso ao explicado anteriormente. Na presença do alelo C a idade do primeiro acidente trombótico aumenta. Para os indivíduos TT a idade do primeiro acidente trombótico é de 34 anos (n=57, *pvalue*=0,060), para os indivíduos TC a média apresenta-se aos 39 anos (n=36, *pvalue*=0,318) enquanto que os indivíduos com CC apresentam uma idade média de 37 anos (n=9, *pvalue*=0,395). Mais uma vez estes resultados não têm significância estatística devido ao pequeno número de amostras, mas apresen-

tam uma indicação de que este polimorfismo poderá estar a exercer uma modulação do risco trombótico, aumentando a idade em que a patologia começa a ocorrer nos indivíduos.

Apenas o aumento do número de indivíduos analisados permitirá retirar conclusões assertivas. Todavia, através desta análise, é possível inferir que, apesar deste polimorfismo estar mais representado na população com acidente trombótico, a sua função parece ser moduladora, aumentando a idade em que começam a ocorrer os episódios trombóticos.

Quando se incluem os factores secundários na análise da modulação da idade do primeiro acidente trombótico pelos polimorfismos em estudo, os resultados parecem indicar que o alelo predisponente do polimorfismo FGG 10034 C>T parece acentuar o risco trombótico de forma mais rigorosa na ausência de factores secundários. Assim, para a média da idade da primeira ocorrência de patologia trombótica com factores secundários temos: FGG 10034 CC a apresentar uma média de 34 anos (n=29) e o CT uma média de 33 (n=19). Contudo, na ausência de factores secundários as médias são de 37 (n=22) e 26 (n=13) respectivamente. Para esta análise não foram encontrados indivíduos com a presença do alelo T em homozigotia.

Por seu lado o polimorfismo FGG 9340 T>C parece modular o risco trombótico quando se consideram os factores secundários. Este facto é evidenciado pela média da idade do primeiro acidente trombótico, que na presença do FGG 9340 TT é de 28 anos (n=22) e na presença do polimorfismo em heterozigotia sobe para 42 anos (n=21), efeito protector. Contudo, na ausência de factores secundários parece verificar-se o contrário, descendo a idade do primeiro acidente trombótico de 34 anos (n=22) para 24 anos (n=10), na presença de FGG 9340 TT e CT respectivamente. O polimorfismo FGG 9340 CC estava representado em relativamente poucos indivíduos, em 5 indivíduos com factores secundários e em 3 indivíduos sem factores secundários, sendo a média da idade (24 e 46 anos, respectivamente) pouco informativas devido à grande variância apresentada.

3.6 Haplótipos

As frequências dos haplótipos determinados para cada um dos grupos de estudo, estão descritas na tabela 3.4. Quando se compara o subgrupo TEV com o grupo controlo verifica-se que as frequências dos haplótipos são muito semelhantes entre si, sendo o valor do teste qui-quadrado muito baixo ($pvalue=0,886$). Quando se calcula o risco relativo que cada um dos haplótipos representa, verifica-se que os valores são muito próximos de 1, o que significa que nenhum dos haplótipos parece influenciar o risco de TVP.

Contudo, quando se comparam os indivíduos do subgrupo TEV que tiveram o primeiro acidente trombótico com menos de 32 anos (subgrupo TEV - 32) e os indivíduos saudáveis do grupo de estudo que atingiram os 42 anos de idade sem apresentar acidente trombótico (grupo estudo saudáveis +42), é possível verificar que os grupos apresentam diferenças significativas (quando considerado um intervalo de confiança de 95% ($pvalue = 0,02$)). Assim, o haplótipo PTH1 parece estar mais representado neste grupo de indivíduos saudáveis, indicando uma diminuição do risco relativo de acidente trombótico nos indivíduos portadores de FV Leiden ($OR=0,400$). Por outro lado o haplótipo PTH2 parece estar mais representado nos indivíduos do grupo TEV que tiveram acidente trombótico mais precocemente, o que aponta na direcção deste haplótipo aumentar o risco relativo de ocorrência de acidente trombótico, e de este ocorrer em idades mais precoces ($OR = 3,236$). Contudo, estes dois subgrupos são constituídos por poucos indivíduos o que poderá estar a influenciar os testes estatísticos. Estes resultados são apresentados na tabela 3.4

Tabela 3.4: Comparação entre Haplótipos

Haplótipo	Subgrupo TEV	Grupo Controlo	χ^2	Pvalue ^a	OR ^b
PTH1	0,603	0,587	0,241	0,886	0,882
PTH2	0,234	0,257			1,066
PTH3	0,163	0,156			1,0587
PTH4	0	0			ND

Haplótipo	Subgrupo TEV - 32	Grupo estudo saudáveis +42	χ^2	Pvalue ^a	OR ^b
PTH1	0,591	0,688	7,228	0,02	0,400
PTH2	0,212	0,156			3,236
PTH3	0,197	0,156			1,008
PTH4	0	0			ND

^a Calculado através do teste Qui-Quadrado de Pearson

^b O.R. – Odds Ratios, calculado como explicado nos métodos

ND – Não definido

Ao comparar, nos vários haplótipos, a idade do primeiro acidente trombótico é possível verificar que os resultados apontam para uma modulação do risco trombótico, através de uma influência exercida nesta mesma idade. Deste modo o haplótipo protector/neutro (PTH2/PTH1) aumenta a idade do primeiro acidente em cerca de 4 anos, quando comparativamente com o haplótipo neutro (PTH1). Por sua vez o haplótipo potenciador/neutro (PTH3) parece diminuir a idade do primeiro acidente em cerca de 5 anos, quando comparado com o haplótipo neutro. Estes resultados apresentam-se de forma clara no gráfico da figura 3.10.

É interessante notar que, quando se divide o grupo de TEV em indivíduos com e sem factores secundários, os haplótipos também parecem apresentar influência na idade do primeiro acidente trombótico na presença de factores secundários. Nos indivíduos que apresentam o haplótipo neutro a idade do primeiro acidente trombótico

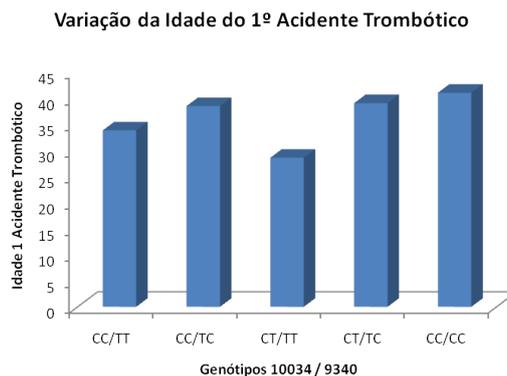


Figura 3.10: Variação da idade do primeiro acidente trombótico consoante a presença dos haplótipos.

não varia muito calculando-se uma média de 35 e 36 anos respectivamente para os indivíduos com e sem factores secundários. Como seria de esperar nos indivíduos que apresentam factores secundários a idade do primeiro acidente trombótico é a mais baixa. Todavia o haplótipo protector/neutro (PTH2/PTH1) parece exercer uma maior influência nos indivíduos que apresentam factores secundários, visto que a idade do primeiro acidente trombótico sobe de 38 para 41 anos. O haplótipo causador/neutro (PTH3/PTH1) também parece influenciar de modo mais severo os indivíduos que apresentam factores secundários, visto que a idade desce de 28 para 22 anos. No gráfico 3.11 são ainda apresentadas as idades médias dos indivíduos que não apresentam factores secundários, com a correspondência aos haplótipos presentes. Contudo, devido ao baixo número de indivíduos presentes, as conclusões não podem ser consideradas significativas.

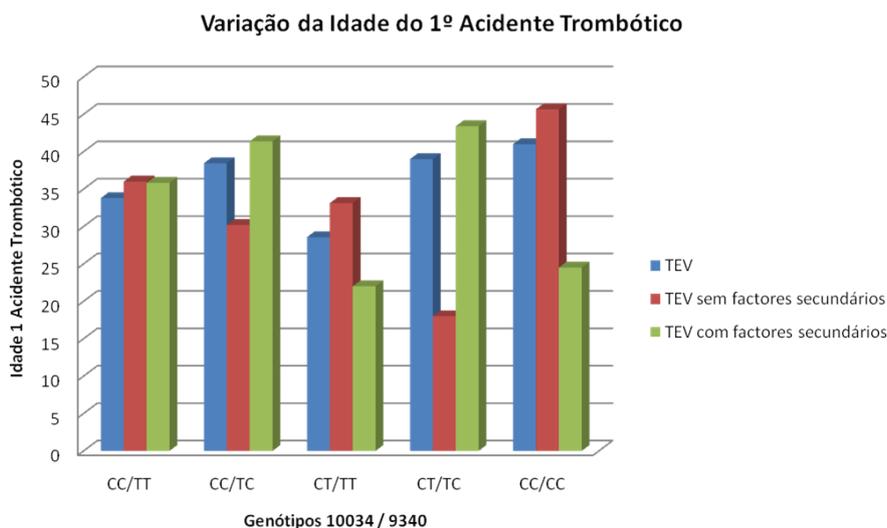


Figura 3.11: Variação da Idade do primeiro acidente trombótico considerando os diferentes haplótipos e a presença ou ausência de factores secundários.

Capítulo 4

Conclusão

A mutação FV Leiden pode considerar-se como um dos factores genéticos mais importantes que conduzem à ocorrência de TV. Contudo, existe uma grande variação em termos de manifestações clínicas e na idade dos primeiros acidentes trombóticos. Este facto sugere a existência de factores moduladores, genéticos e ambientais, que contribuem para a etiologia da TV. Assim, este trabalho teve como objectivo verificar o potencial efeito modulador de dois polimorfismos existentes no Fibrinogénio γ , FGG 10034 C>T e FGG 9340 T>C, descritos na literatura como factores que influenciam o risco de TV em populações caucasianas [67, 48, 68].

As frequências alélicas e genótípicas dos polimorfismos FGG 10034 C>T e FGG 9340 T>C do grupo controlo e do subgrupo de estudo TEV não são significativamente diferentes. Da mesma forma, as frequências alélicas e genótípicas não podem ser consideradas diferentes quando se compara o subgrupo de indivíduos saudáveis portadores de FV Leiden com mais de 42 anos (grupo estudo saudáveis + 42) e os indivíduos com patologia trombótica portadores de FV Leiden com menos de 32 anos (subgrupo TEV - 32). Estes resultados parecem apontar no sentido de que o potencial efeito dos dois polimorfismos não será suficiente para introduzir uma alteração das frequências alélicas, não sendo um factor de risco principal na ocorrência de TV.

Esta conclusão está de acordo com o que seria expectável. Como o FV Leiden é um factor de risco *major*, o efeito de um outro factor de risco *minor* poderá ser mascarado. Põe-se a hipótese de que o Fibrinogénio poderá ser um factor modulador ou, apenas apresentar um efeito mais pronunciado na ausência do FV Leiden. Todavia, seria necessário estudar um grupo de indivíduos com TV e sem FV Leiden para avaliar o potencial do Fibrinogénio na ocorrência de patologia trombótica.

Considerando a amostra da população portuguesa estudada, parece haver uma modulação da idade em que ocorre o primeiro acidente trombótico pelos polimor-

fismos FGG 10034 C>T e FGG 9340 T>C . O polimorfismo 9340TC aponta para um aumento da idade em que ocorre o primeiro acidente trombótico, actuando como um genótipo protector. Pelo contrário, o polimorfismo 10034CT parece actuar como potenciador levando à diminuição da idade do primeiro acidente trombótico. Estes dados não atingem uma significância estatística porque o número de indivíduos que constitui a amostra é pequeno, mas são bons indicadores que de facto poderá ocorrer uma modulação do risco trombótico por estes polimorfismos, através do aumento ou da diminuição da idade do primeiro acidente trombótico.

A influência sobre a idade do primeiro acidente trombótico que começou por ser observada na análise genotípica de ambos os polimorfismos separadamente, é evidenciada quando são considerados os haplótipos. A comparação das frequências haplotípicas parece conduzir a uma conclusão contraditória, visto que o haplótipo neutro está sobre-representado no grupo de indivíduos saudáveis enquanto o haplótipo supostamente protector está sobre-representado no subgrupo de indivíduos com TEV, chegando mesmo a atingir um valor elevado para o risco relativo (OR) de 3,236. Contudo, esta conclusão contraditória pode ser esclarecida ao estabelecer a média do primeiro acidente trombótico. Apesar do haplótipo PTH2 estar sobre-representado nesta população, a sua presença parece aumentar a idade em que ocorre o primeiro acidente trombótico. De facto, quando analisado sobre esta perspectiva o haplótipo protector modela de uma forma positiva o acidente trombótico, pois os indivíduos atingem uma idade mais avançada sem apresentar patologia. Por sua vez, o haplótipo potenciador/neutro (PTH3/PTH1) parece de facto diminuir a idade em que o primeiro acidente ocorre, modelando de forma negativa o acidente trombótico.

Quando se pondera a presença de factores de risco secundários, estes haplótipos parecem exercer a mesma função, visto que o haplótipo protector/neutro (PTH2/PTH1) aumenta ainda mais a idade em que ocorre o primeiro acidente trombótico e o haplótipo potenciador/neutro (PTH3/PTH1) diminui de forma mais acentuada a idade do primeiro acidente trombótico.

Na população caucasiana, os polimorfismos FGG 10034C>T e 9340T>C são os marcadores de dois dos haplótipos mais comuns, H2 e H3 respectivamente [30]. Estes haplótipos têm sido extensivamente estudados e os dados sugerem que são estes os polimorfismos que apresentam o efeito sobre o risco de TV. Neste contexto o haplótipo H2 leva ao aumento do risco de TV, enquanto que o H3 parece estar associado com a sua diminuição [52].

O haplótipo PTH2 está incluído no haplótipo FGG-H3 descrito na literatura como protector relativamente ao risco trombótico. De facto, o haplótipo PTH2 parece diminuir o risco dos seus portadores apresentarem um acidente trombótico

em idade precoce, de acordo com o descrito para o haplótipo FGG-H3 [30].

Apesar do grupo de estudo inicial ser constituído por 280 indivíduos, devido à falta de dados clínicos completos, o número ficou bastante reduzido, dificultando a análise estatística. No entanto, os resultados preliminares obtidos neste estudo, indicam uma associação destes polimorfismos do FGG na modulação do risco trombótico. Deste modo, com base nestes resultados, poderá estender-se o estudo de uma forma mais confiante, aumentando o número de indivíduos e considerando outros grupos (por exemplo, sem a presença da mutação FV Leiden ou com trombose arterial).

A abordagem apresentada neste estudo representa uma mais-valia na informação clínica no que respeita à idade do primeiro acidente trombótico e às estratégias mais apropriadas na prevenção de acidentes trombóticos adicionais nos pacientes portadores de FV Leiden e seus familiares. A continuação e aprofundamento deste estudo, e a posterior tradução dos resultados obtidos para a parte clínica permitirá uma melhoria dos cuidados prestados aos doentes com TV e aos seus familiares.

Bibliografia

- [1] EM. Scott, RA. Ariëns, and PJ. Grant. Genetic and environmental determinants of fibrin structure and function: Relevance to clinical disease. *Arterioscler Thromb Vasc Biol*, 24:1558–1566, 2004.
- [2] H. Stormorken. The discovery of factor V: a tricky clotting factor. *Journal of Thrombosis and Haemostasis*, 1:206–213, 2003.
- [3] DA. Lane and PJ. Grant. Role of hemostatic gene polymorphisms in venous and arterial thrombotic disease. *Blood*, 95:1517–1532, 2000.
- [4] FR. Rosendaal and PH. Reitsma. Genetics of venous thrombosis. *Journal of Thrombosis and Haemostasis*, 7 suppl.1:301–304, 2009.
- [5] ID. Bezemer, LA. Bare, JM. Doggen, and et al. Gene variants associated with deep vein thrombosis. *JANA*, 11:1306–1314, 2008.
- [6] M. Franchini and PM. Mannucc. Multiple gene interaction and modulation of hemostatic balance. *Clin Chem Lab Med*, 47:1455–1460, 2009.
- [7] M. Cushman. Epidemiology and risk factors for venous thrombosis. *Seminars in Hematology*, 44:62–69, 2007.
- [8] F. Anderson and F. Spencer. Risk factors for venous thromboembolism. *Circulation: Journal of the American Heart Association*, 107:9–16, 2003.
- [9] G. Yenicesu, M. Cetin, O. Ozdemir, and et al. A prospective case-control study analyzes 12 thrombophilic gene mutations in turkish couples with recurrent pregnancy loss. *American Journal of Reproductive Immunology*, 63:126–136, 2010.
- [10] E. Varga. Inherited thrombophilia: Key points for genetic counseling. *J Genetic Counseling*, 16:261–277, 2007.
- [11] Ensembl, 10 2010. <http://www.ensembl.org/index.html>.

- [12] EA. Vucic, KL. Thu, AC. Williams, and et al. *Genetic Variation: Methods and Protocols, Methods in Molecular Biology*, volume 428. Humana Press, first edition, 2010. Chapter 6, pages 103–117.
- [13] J. Suntharalingam, K. Goldsmith, V. van Marion, and et al. Fibrinogen A α Thr312Ala polymorphism is associated with chronic thromboembolic pulmonary hypertension. *European Respiratory Journal*, 31:736–741, 2008.
- [14] A. Zivelin, J. Griffin, and X. Xu. A single genetic origin for a common caucasian risk factor for venous thrombosis. *Blood*, 15:392–402, 1997.
- [15] A. Bersano, E. Ballabio, N. Bresolin, and L. Candelise. Genetic polymorphisms for the study of multifactorial stroke. *Human Mutation*, 6:776–795, 2008.
- [16] EF. Tizzano, JS. Soria, I. Coll, and et al. The prothrombin 20210A allele influences clinical manifestations of hemophilia A in patients with intron 22 inversion and without inhibitors. *Haematologica*, 87:279–285, 2002.
- [17] RM. Bertina, RC. Koeleman, T. Koster, and et al. Mutation in blood coagulation factor V associated with resistance to activated protein c. *Nature*, 369:64–67, 1994.
- [18] V. De Stefano and L. Giuseppe. Resistannce to activated protein c due mutated factor v as a novel cause of inherited thrombophilia. *Haematologica*, 80:344–356, 1995.
- [19] FR. Rosendaal, T. Koster, JP. Vandenbroucke, and PH. Reitsma. High risk of thrombosis in patients homozygous for factor V leiden (activated protein C resistance). *Blood*, 85:1504–1508, 1995.
- [20] SR. Poort, FR. Rosendaal, PH. Reitsma, and RM. Bertina. A common genetic variation in the 3'-untranslated region of the prothrombin gene is associated with elevated plasma prothrombin levels and an increase in venous thrombosis. *Blood*, 88:3698–3703, 1996.
- [21] ID. Bezemer, CJM. Doggen, HL. Vos, and FR. Rosendaal. No association between the common MTHFR 677C>T polymorphism and venous thrombosis. *Arch Intern Med*, 167:497–501, 2007.
- [22] S. Khan and JD. Dicherman. Hereditary thrombophilia. *Thrombosis Journal*, 4:1–17, 2006.
- [23] I. Weisberg, P. Tran, B. Christensen, S. Sibani, and R. Rozen. A second genetic polymorphism in methyltetrahydrofolate reductase (MTHFR) associated with decreased enzyme activity. *Mol Genet Metab*, 68:511–512, 1999.

- [24] J. Townend, J. O'Sullivan, and JT. Wilde. Hyperhomocysteinemia and vascular disease. *Blood Reviews*, 12:23–34, 1998.
- [25] V. De Stefano, I. Castorelli, E. Rossi, B. Zappacosta, and G. Leone. Interaction between hyperhomocysteinemia and inherited thrombophilic factors in venous thrombosis. *Semin Thromb Haemost*, 26:305–311, 2000.
- [26] EV. Theodoraki, T. Nikopentis, J. Suhrutsenko, and et al. Fibrinogen beta variants confer protection against coronary artery disease in a greek case-control study. *BMC Medical Genetics*, 11:28–35, 2010.
- [27] K. Jood, J. Danielson, D. Ladvall, and et al. Fibrinogen gene variation and ischemic stroke. *Journal of Thrombosis and Haemostasis*, 6:897–904, 2008.
- [28] JS. Danik, G. Paré, DI. Chasman, and et al. Novel loci, including those related to crohn disease, psoriasis, and inflammation, identified in a genome-wide association study of fibrinogen in 17686 women: The womens genome health study. *Circ Cardiovasc Genet*, 2:134–141, 2009.
- [29] M. van Oijen, EY. Cheung, CE. Geluk, and et al. Haplotypes of the fibrinogen gene and cerebral small vessel disease: the rotterdam scan study. *J Neurol Neurosurg Psychiatry*, 79:799–803, 2008.
- [30] S. Uitte de Willige, MCH. De Visser, J. Houwing-Duistermaat, and et al. Genetic variation in the fibrinogen gamma gene increases the risk for deep venous thrombosis by reducing plasma fibrinogen γ' levels. *Blood*, 106:4176–4183, 2005.
- [31] Fibrinogen/fibrin (sigma), 10 2010. <http://www.sigmaaldrich.com/life-science/metabolomics/enzyme-explorer/learning-center/structural-proteins/fibrinogen-fibrin.html>.
- [32] B. Siegerink, FR. Rosendaal, and A. Algra. Genetic variation in fibrinogen; its relationship to fibrinogen levels and the risk of myocardial infarction and ischemic stroke. *Journal of Thrombosis and Haemostasis*, 7:385–390, 2008.
- [33] R. Asselta, S. Duga, and ML. Tenchini. The molecular basis of quantitative fibrinogen disorders. *Journal of Thrombosis and Haemostasis*, 4:2115–2129, 2006.
- [34] MN. Mannila, RS. Lovely, SC. Kazmierczak, and et al. Elevated plasma fibrinogen γ' concentration is associated with myocardial infarction: effects of variation in fibrinogen genes and environmental factors. *Journal of Thrombosis and Haemostasis*, 5:766–773, 2007.

- [35] D. Vu and M. Neerman-Arbez. Molecular mechanisms accounting for fibrinogen deficiency: from large deletions to intracellular retention of misfolded proteins. *Journal of Thrombosis and Haemostasis*, 5:125–131, 2007.
- [36] AP. Reiner, CL. Carty, CS. Carlson, and et al. Association between patterns of nucleotide variation across the three fibrinogen genes and plasma fibrinogen levels: the coronary artery risk development in young adults (CARDIA) study. *Journal of Thrombosis and Haemostasis*, 4:1279–1287, 2006.
- [37] SN. Roy, G. Mukhopadhyay, and CM. Redman. Regulation of fibrinogen assembly. *The Journal of Biological Chemistry*, 265:6389–6393, 1990.
- [38] WB. Kannel, PA. Wolf, WP. Castelli, and et al. Fibrinogen and risk of cardiovascular disease. the framingham study. *Journal of American Medical Association*, 258:1183–1186, 1987.
- [39] TW. Meade, S. Mellows, M. Brozovic, and et al. Haemostatic function and ischaemic heart disease: principal results of the northwick park heart study. *Lancet*, 2:533–537, 1986.
- [40] J. Danesh, S. Lewington, SG. Thompson, and et al. Plasma fibrinogen level and the risk of major cardiovascular diseases and nonvascular mortality: an individual participant meta-analysis. *Journal of American Medical Association*, 294:1799–1809, 2005.
- [41] B. Jacquemin, C. Antoniadou, F. Nyberg, and et al. Common genetic polymorphisms and haplotypes of fibrinogen alpha, beta, and gamma chains affect fibrinogen levels and the response to proinflammatory stimulation in myocardial infarction survivors. *Journal of the American College of Cardiology*, 52:941–952, 2008.
- [42] GD. Lowe, M. Woodward, A. Rumley, and et al. Associations of plasma fibrinogen assays, C-reactive protein and interleukin-6 with previous myocardial infarction. *Journal of Thrombosis and Haemostasis*, 1:2312–2316, 2003.
- [43] M. Verhovsek, K. Moffat, and C. Hayward. Laboratory testing for fibrinogen abnormalities. *American Journal of Hematology*, 83:928–931, 2008.
- [44] N. Proudfoot, A. Furger, and M J. Dye. Integrating mRNA processing with transcription. *Cell Press*, 108:501–512, 2006.
- [45] J. Kant, A. Fornace Jr, D. Saxe, and et al. Evolution and organization of the fibrinogen locus on chromosome 4: Gene duplication accompanied by transposition and inversion. *Proc. Natl. Acad. Sci.*, 82:2344–2348, 1985.

- [46] A. Fornace Jr, D. Cummings, C. Comeau, and et al. Structure of the human γ -fibrinogen gene. *The Journal of Biological Chemistry*, 259:12826–12830, 2006.
- [47] RS. Lovely, LA. Falls, HA. Al-Mondhiry, and et al. Association of gammaA/gamma' fibrinogen levels and coronary artery disease. *Thromb. Haemost*, 1:26–31, 2002.
- [48] S. Uitte de Willige, IM. Rietveld, MCH. De Visser, and et al. Polymorphism 10034C>T is located in a region regulating polyadenylation of FGG transcripts and influences the fibrinogen γ' / γ a mRNA ratio. *J Thromb Haemost*, 5:1243–1249, 2007.
- [49] LA. Lange, AP. Reiner, CL. Carty, and et al. Common genetic variants associated with plasma fibrin D-dimer concentration in older european- and african-american adults. *Journal of Thrombosis and Haemostasis*, 6:654–659, 2008.
- [50] K. Beyer, T. Dandekar, and W. Keller. Rna ligands selected by cleavage stimulation factor contain distinct sequence motifs that function as downstream elements in 3'-end processing of pre-mRNA. *The Journal of Biological Chemistry*, 272:26769–26779, 2006.
- [51] C. MacDonald, J. Wilusz, and Thomas Shenk. The 64-kilodalton subunit of the cstF polyadenylation factor binds to pre-mRNAs downstream of the cleavage site and influences cleavage site location. *Molecular and Cellular Biology*, 1994.
- [52] W. Koch, P. Hoppmann, J. Biele, and et al. Fibrinogen genes and myocardial infarction: A haplotype analysis. *Arterioscler Thromb Vasc Biol*, 28:758–763, 2008.
- [53] N. Guzmán, F. Lanás, and LA. Salazar. Influence of amerindian mitochondrial dna haplogroups on thrombosis susceptibility and frequency of four genetic prothrombotic variants in southern chilean subjects. *Clinica Chimica Acta*, 411:444–447, 2010.
- [54] Seattlesnps variation discovery resource, 10 2010. <http://pga.gs.washington.edu>.
- [55] G. Reed, J. Kent, and C. Wittwer. High-resolution DNA melting analysis for simple and efficient molecular diagnostics. *Future Medicine*, 8:597–608, 2006.
- [56] MN. Mannila, P. Eriksson, Leander K., and et al. The association between fibrinogen haplotypes and myocardial infarction in men is partly mediated through pleiotropic effects on the serum IL-6 concentration. *Journal of Internal Medicine*, 261:138–147, 2006.

- [57] S. Uitte de Willige, CJM. Doggen, MCH. De Visser, and et al. Haplotypes of the fibrinogen gamma gene do not affect the risk of myocardial infarction. *Journal of Thrombosis and Haemostasis*, 4:474–476, 2006.
- [58] Qiaamp DNA blood midi/maxi kit, 10 2010. <http://www.qiagen.com/products/genomicdnastabilizationpurification>.
- [59] QIAamp[®] DNA Blood Midi/ Maxi Handbook – QIAGEN[®].
- [60] NanoDrop[®] ND-1000 Spectrophotometer V3.0.1 Users Manual.
- [61] λ DNA/hind iii fragments, 10 2010. <http://products.invitrogen.com/ivgn/product>.
- [62] Hyperladders selection, 10 2010. <http://www.bioline.com>.
- [63] Taqman[®] SNP genotyping assays, 10 2010. <https://products.appliedbiosystems.com>.
- [64] CR. Newton, A. Granham, LE. Heptinstall, and et al. Analysis of any point mutation in DNA. the amplification refractory mutation system (ARMS). *Nucleic Acids Research*, 17:2503–2516, 1989.
- [65] Lightcycler[®] 480 high resolution melting master, 10 2010. <http://www.roche-applied-science.com/proddata/>.
- [66] D. Pestana and S. Velosa. *Introdução à Probabilidade Estatística*, volume 1. Fundação Calouste Gulbenkian, 2002.
- [67] S. Uitte de Willige, M. Pyle, H. Vos, and et al. Fibrinogen gamma gene 3'- end polymorphisms and risk of venous thromboembolism in the african-american and caucasian population. *Thrombosis and Haemostasis*, 101:1078–1084, 2009.
- [68] G. Grunbacher, W. Weger, E. Marx-Neuhold, and et al. The fibrinogen gamma (FGG) 10034C NT polymorphism is associated with venous thrombosis. *Thrombosis Research*, 121:33–36, 2007.

Apêndice A

Anexos

A.1 Anexo A – Procedimento para extracção de DNA utilizando QIAamp Midi

Sempre que necessário e de forma a acertar os volumes das amostras para 2mL adiciona-se PBS 1x. Antes de começar:

- As amostras de sangue e os tampões de extracção devem estar à temperatura ambiente (15-25°C).
- Preparar o banho a 65°C para usar no passo 4 do procedimento.

Procedimento:

- Pipetar 200 μ L de Protease no fundo de um tubo de centrifugação de 15mL.
- Adicionar 2mL de sangue e misturar brevemente.
- Adicionar 2,4mL de Buffer AL e misturar invertendo o tubo 15 vezes, seguido de uma agitação vigorosa durante 1 minuto. Para assegurar uma lise adequada, a amostra deve ser misturada com o Buffer AL até a solução ser homogénea. Utilizar seringas com agulhas de calibre baixo (27G) (ou dobrar ligeiramente a ponta da agulha) e aspirar e largar o sangue 5 vezes. Este passo induz mecanicamente a lise das células, através da sua passagem por um diâmetro apertado.
- Incubar a 65°C durante 10 minutos. O DNA atinge a lise máxima após 10 minutos a 65°C, mas uma incubação mais prolongada não afecta negativamente o rendimento.
- Adicionar 2mL de etanol (96-100%) à amostra, e misturar invertendo o tubo 10 vezes, seguido de uma agitação vigorosa. Para assegurar que a ligação é eficaz,

é essencial que a amostra seja misturada após a adição de etanol, formando uma solução homogénea.

- Transferir metade do volume da solução para o tubo Falcon de 15mL com a coluna. Centrifugar durante 3 minutos a 3000 rpm.
- Retirar a coluna, descartar o filtrado e colocar novamente a coluna no tubo Falcon de 15mL. Carregar com a solução restante e centrifugar novamente durante 3 minutos a 3000 rpm. Se a solução não passou completamente pela membrana, centrifugar novamente a uma rotação superior.
- Retirar a coluna, descartar o filtrado e colocar novamente a coluna no tubo Falcon de 15mL. Se o filtrado não for descartado, o bocal da coluna ficará submerso no filtrado e a eficácia das lavagens será reduzida.
- Cuidadosamente, adicionar 2mL do Buffer AW1 à coluna. Centrifugar durante 1min a 5000rpm. Neste passo não é necessário descartar o filtrado.
- Cuidadosamente adicionar 2mL do Buffer AW2 à coluna. Centrifugar durante 15min a 5000rpm. O aumento do tempo de centrifugação tem como objectivo remover todos os vestígios do Buffer AW2 da coluna antes da eluição do DNA. É necessário que esta remoção seja efectuada pois os vestígios do Buffer AW2 e também de etanol podem causar inibição do PCR, levando a resultados negativos falsos.
- Colocar a coluna num tubo Falcon de 15mL limpo e descartar o tubo que contém o filtrado.
- Pipetar 200 μ L de Buffer AE directamente sobre a membrana da coluna. Incubar durante 5 minutos à temperatura ambiente e centrifugar a 5000rpm durante 2 minutos.
- Transferir o DNA para um tubo eppendorf, etiquetar com o número de amostra, e guardar a 4°C.

A.2 Anexo B – Tabela de Primers

Figura A.1: Tabela de Primers

Polimorfismo	Primer	Sequência Oligonucleotídica (5' → 3')	Comprimento (bases)	Porcentagem Conteúdo GC
FGG 10034	FGG E10-1	ATA TTC ACA GGC TGG AGA CG	20	50 %
FGG 10034	FGG E10-R	GCA TAA AAT GAG GAC TCA ATT AC	23	34,8%
FGG 9340	FGG E9-1	GAG AAG GAC AGC AAC ACC AC	20	55%
FGG 9340	FGG E9-R	TTA ATT TCC ATT GAA GGC TAA A	22	27,3%
FGG 9340	FGG E9-2	CTG CGG AAA CAG AAT ATG ACT	21	42,9%
FGG 9340	FGG_9348C	GAT TTG TAG AAA ATT AAC TGC TAA C	25	28 %
FGG 9340	FGG_9348T	GAT TTG TAG AAA ATT AAC TGC TGA T	25	28%
FGG 9340	FGG_9348Cb	GAT TTG TAG AAA ATT AAC TGC TGA C	25	32 %

A.3 Anexo C – SSCP

Antes da preparação do gel de electroforese

- O vidro mais pequeno deverá ser tratado com Repel Silan a cada 6-12 meses (consoante a sua utilização), de modo a repelir o gel aquando da desmontagem dos vidros, após a electroforese. Coloca-se uma pequena quantidade de Repel Silan sobre o vidro e espalha-se uniformemente com o auxílio de papel de mãos. Deixa-se secar completamente ao ar durante cerca de 1 hora.

Preparação das soluções de deposição

- Solução de deposição desnaturante: Pipetar 800 μ L de Formamida, 100 μ L Azul de Bromofenol (10mg/mL), 100 μ L Xilenocianol (10 mg/mL), 2 μ L NaOH 5M, 4 μ L EDTA 0,25M. Acertar para pH 7,5.
- Solução de deposição não-desnaturante: Pipetar 800 μ L Formamida, 100 μ L Azul de Bromofenol (10mg/mL), 100 μ L Xilenocianol (10 mg/mL).

Preparação das soluções para coloração do gel

- Etanol 10% - Medir 100mL Etanol (EtOH) Absoluto e juntar 900mL Água bidestilada (ddH₂O).
- Ácido Nítrico 1% - Pipetar 15,5mL Ácido Nítrico (HNO₃) a 65% e juntar ddH₂O até perfazer 1L. Esta solução deve ser preparada na hotte.
- Nitrato de Prata 0,012M - Pesar 2,02g de Nitrato de Prata (AgNO₃) em pó e juntar ddH₂O até perfazer 1L. Esta solução por ser sensível à luz, deve ser armazenada protegida da luz (por exemplo num shoot de vidro escuro).
- Carbonado de sódio 0,28M com Formaldeído 35% - Pesar 29,6g de Carbonato de sódio (Na₂CO₃) e juntar ddH₂O até perfazer 1L. Dissolver num agitador. Quando a solução estiver uniforme, juntar 320 μ L de Formaldeído, na hotte. Misturar por inversão e agitação. Esta solução também deve ser armazenada num shoot de vidro escuro.
- Ácido Acético 10% - Medir 100mL de ácido acético (CH₃COOH) glacial e juntar 900mL de ddH₂O.

Preparação do gel de electroforese

- Limpar os vidros, espaçadores e pente de electroforese (já devidamente lavados e secos) com etanol absoluto. Este passo destina-se a retirar quaisquer impurezas que estejam presentes nos vidros, nomeadamente gorduras e que possam interferir com a migração das amostras.
- Montar os vidros em cima de um suporte (tampa de esferovite ou similar). Colocar o vidro maior em cima da superfície de esferovite, colocar os espaçadores de 8mm (já com as respectivas borrachas de vedação na extremidade) de lado, alinhados com o vidro. Colocar o vidro mais pequeno e ajustar

tudo, de modo a ficar alinhado com o vidro maior e com as borrachas dos espaçadores bem justapostas. Prender de lado com molas de metal, de modo a fixar todo o sistema. Colocar o vedante de borracha bem ajustado, retirando as molas de metal. O gel de poliacrilamida é um gel que corre verticalmente, mas todos os passos anteriores à solidificação do gel são feitos horizontalmente, tendo em atenção que o sistema se encontra perfeitamente equilibrado, usando um nível.

- Preparar o gel de electroforese numa proveta graduada:
- Gel 6% TBE 1x:
- Pipetar 10mL de ddH₂O para um gobelet
- Pipetar para a proveta:
 - 20 mL de ddH₂O,
 - 6,5 mL de Acrilamida:Bisacrilamida (37:5:11),
 - 11 mL TBE 10x,
 - Completar com ddH₂O até perfazer 100 mL.
 - Adicionar 400 μ L de APS 10%
 - Adicionar 40 μ L TEMED
- Juntar no gobelé e uniformizar bem a solução por agitação.
- Aplicar entre o sistema dos dois vidros, com a ajuda de uma seringa. Ter cuidado para evitar a formação de bolhas. No final, colocar cuidadosamente o pente e prendê-lo com molas de metal de modo a não ficar gel entre cada vidro e o pente.
- Deixar polimerizar à temperatura ambiente, 1 a 2 horas. Pode verificar-se o estado de polimerização pelas sobras de gel que ficaram na seringa ou no gobelet.
- Quando o gel estiver polimerizado, retirar as molas. Retirar o pente muito cuidadosamente para que os poços não fiquem deformados. Retirar o vedante de borracha e limpar os vidros com papel de mãos.
- Montar o sistema de vidros com o gel na tina de electroforese vertical, com a face do vidro pequeno e espaçadores alinhados. O lado superior é o que inclui os poços. Prender os vidros com as quatro molas da tina.
- Fechar a torneira que se encontra na lateral direita da tina e colocar o tampão TBE 1x nos compartimentos superior e inferior da tina criados após a colocação dos vidros.
- Verificar se não há fugas de tampão.
- Lavar os poços com tampão TBE 1x com ajuda de uma seringa e respectiva agulha.
- Programar a fonte de voltagem com o programa a utilizar durante a corrida de electroforese (40W, durante o tempo apropriado consoante o número de

aplicações).

Preparação das amostras e da electroforese

- Marcar os tubos eppendorf.
- Para cada tubo pipetar $5\mu\text{L}$ de solução de deposição desnaturante e $5\mu\text{L}$ de produto de PCR.
- Pipetar $5\mu\text{L}$ de solução de deposição não-desnaturante e $5\mu\text{L}$ de produto de PCR. As amostras preparadas com a solução de deposição não desnaturante têm como objectivo fazer uma referência do padrão de mobilidade electroforética das cadeias duplas de DNA. Este mesmo procedimento é realizado para o marcador molecular.
- Para a desnaturação das amostras, colocar os tubos a 95°C durante 5 minutos e de imediato em gelo. Os tubos com solução de deposição não desnaturante são directamente colocados no gelo, para não haver desnaturação dos fragmentos amplificados.
- Pipetar as amostras para os respectivos poços no gel.
- Fechar a tampa da tina. Ligar os eléctrodos à tina e à fonte. Ligar a fonte e verificar se passa corrente e se os parâmetros programados estão correctos.

Coloração e secagem do gel de poliacrilamida

- Depois da corrida de electroforese, desligar a fonte e retirar os fios. Abrir a torneira lateral da tina, de modo a escoar o tampão do reservatório superior. Lavar a tina com água destilada sem esfregar os eléctrodos, de modo a que não se danifiquem.
- Retirar os vidros e colocá-los horizontalmente em cima da bancada, com o vidro mais pequeno virado para cima. Retirar os espaçadores e depois o vidro mais pequeno, com ajuda de uma espátula de metal.
- Colocar cerca de 500mL de EtOH 10% na tina de coloração (de vidro ou inox) e colocá-la sobre um agitador orbital.
- Retirar o gel para a tina de coloração contendo EtOH a 10%. Deixar imerso durante 5 minutos, com agitação suave.
- Retirar a solução utilizando o sistema de vácuo ligado à torneira de água, cuidadosamente para que o gel não se parta nem seja aspirado pelo sistema.
- Colocar cerca de 500mL de solução de HNO_3 1%. Deixar imerso durante 3 minutos, com agitação suave. Retirar com o sistema de vácuo.
- Colocar ddH₂O. Deixar imerso cerca de 1 minuto, com agitação suave. Retirar com o sistema de vácuo.
- Colocar cerca de 500 mL de AgNO_3 0,012M. Deixar imerso durante 20 min-

utos, com agitação suave. Retirar a solução para um Kitasato, servindo-se do sistema de vácuo.

– *Nota:* o nitrato de prata é classificado como uma substância corrosiva e perigosa para o ambiente. É necessário extremo cuidado no seu manuseamento e a sua eliminação deve ser feita através de um protocolo de tratamento adequado e próprio, não sendo eliminada para o esgoto.

– Lavar duas vezes com ddH₂O, como descrito anteriormente.

– Colocar cerca de 500mL da solução de Na₂CO₃ com Formaldeído. Deixar imerso com agitação suave até as bandas de DNA serem visíveis. Quanto mais tempo, maior será a intensidade das bandas, mas menor será a definição das bandas (e como tal menor será a informação retirada do padrão de bandas devido a eventuais sobreposições das mesmas). É necessário atingir um equilíbrio. Retirar com o sistema de vácuo.

– Colocar cerca de 500mL de CH₃COOH 10% para parar a reacção de revelação. Deixar imerso durante 5 minutos, com agitação suave. Retirar com o sistema de vácuo.

– Colocar ddH₂O. Deixar imerso cerca de 5 minutos com agitação suave. Retirar com o sistema de vácuo. Repetir este passo, para retirar bem a solução anterior.

– Cortar um pedaço de papel de filtro 3mm e fazê-lo aderir ao gel (já sem nenhuma solução). Levantar cuidadosamente, de modo a que o gel fique direito, aderente ao papel.

– Cobrir o gel com película aderente.

– Secar num secador de géis ligado a um sistema de vácuo, durante 1-2 horas.

– Retirar o gel seco e guardar.

A metodologia de SSCP é validada quando um fragmento apresenta um padrão de migração idêntico, quando analisado sob as mesmas condições e/ou quando fragmentos que se saiba terem sequências diferentes apresentem padrões diferentes de migração electroforética.

A.4 Anexo D – Colunas de purificação

(Microcon, Millipore - Centrifugal Filter Devices)

- Colocar a coluna dentro do tubo eppendorf de 1,5mL (fornecidos pelo kit).
- Pipetar $50\mu\text{L}$ do produto de PCR e adicionar $450\mu\text{L}$ de água destilada autoclavada, sem tocar na membrana com a ponta da pipeta.
- Centrifugar durante 12 minutos a 500xg. A membrana não deve ficar seca. Ter em atenção que a membrana não deve secar. Durante este primeiro passo de centrifugação verificar o volume que fica na coluna.
- Descartar o filtrado e juntar cerca de $500\mu\text{L}$ de água destilada autoclavada.
- Centrifugar durante 12 minutos a 500xg
- Descartar o filtrado e juntar cerca de $500\mu\text{L}$ de água destilada autoclavada.
- Centrifugar durante 12 minutos a 500xg. Neste passo é importante que o volume na coluna seja cerca de $30\mu\text{L}$, para que a concentração de DNA seja a adequada.
- Inverter a coluna num novo tubo eppendorf de 1,5mL.
- Centrifugar durante 3 minutos a 1000xg para transferir o concentrado para o eppendorf.

A.5 Anexo E – Tabela Geral das Famílias dos Indivíduos