



Ana Alexandra Figueiredo Gonçalves Saramago

Licenciada em Biologia

**Estudo de mutações germinais em genes de
susceptibilidade para o Cancro do Colón e Reto
Familiar do tipo X**

Dissertação para obtenção do Grau de Mestre em
Genética Molecular e Biomedicina

Orientadora : Doutora Maria Cristina Mantas Albuquerque Valeroso,
Investigadora, Unidade de Investigação em Patobiologia
Molecular (UIPM), Instituto Português de Oncologia de
Lisboa Francisco Gentil, EPE (IPOLFG)

Júri:

Presidente: Doutora Paula Maria Theriaga Mendes Bernardes Gonçalves

Arguente: Doutora Maria João Aleixo da Silva

Vogal: Doutora Maria Cristina Mantas Albuquerque Valeroso



FACULDADE DE
CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

Setembro, 2014



Ana Alexandra Figueiredo Gonçalves Saramago

Licenciada em Biologia

**Estudo de mutações germinais em genes de
susceptibilidade para o Cancro do Colón e Reto
Familiar do tipo X**

Dissertação para obtenção do Grau de Mestre em
Genética Molecular e Biomedicina

Orientadora : Doutora Maria Cristina Mantas Albuquerque Valeroso,
Investigadora, Unidade de Investigação em Patobiologia
Molecular (UIPM), Instituto Português de Oncologia de
Lisboa Francisco Gentil, EPE (IPOLFG)



FACULDADE DE
CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

Setembro, 2014

Estudo de mutações germinais em genes de suscetibilidade para o Cancro do Colón e Reto Familiar do tipo X

Copyright © Ana Alexandra Figueiredo Gonçalves Saramago, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa

A Faculdade de Ciências e Tecnologia e a Universidade Nova de Lisboa têm o direito, perpétuo e sem limites geográficos, de arquivar e publicar esta dissertação através de exemplares impressos reproduzidos em papel ou de forma digital, ou por qualquer outro meio conhecido ou que venha a ser inventado, e de a divulgar através de repositórios científicos e de admitir a sua cópia e distribuição com objectivos educacionais ou de investigação, não comerciais, desde que seja dado crédito ao autor e editor.

Aos meus pais

Agradecimentos

Queria deixar aqui o meu agradecimento a todos aqueles que me apoiaram na realização desta dissertação, sem os quais esta tarefa teria sido bastante mais difícil.

Agradeço à Dra. Cristina Albuquerque, primeiro por ter aceite a minha participação neste projeto e por toda a motivação, orientação, conhecimentos e esclarecimentos transmitidos ao longo deste ano.

À Mestre Inês Francisco, Dr. Bruno Filipe e Dra. Patrícia Silva um grande agradecimento, por toda a sua ajuda no laboratório, pela disponibilidade em esclarecer qualquer dúvida e pela simpatia com que sempre me receberam.

À Dra. Sofia Fragoso e Dra. Patrícia Machado por toda a boa disposição no laboratório, amizade e simpatia que demonstraram desde o primeiro dia.

À Dra. Branca Cavaco, Coordenadora da Unidade de Investigação em Patobiologia Molecular, do Instituto Português de Oncologia de Lisboa, Francisco Gentil, E.P.E., por ter autorizado a realização deste projeto.

Ao Serviço de Gastrenterologia, Clínica de risco familiar e ao Serviço de Anatomia Patológica do Instituto Português de Oncologia de Lisboa, Francisco Gentil, E.P.E., pela disponibilidade do material biológico e pela informação clínica relativa aos doentes.

Aos meus colegas Gonçalo Pereira, Filipa Fernandez e Marisa Paulino pela companhia, amizade e apoio que me deram. Quero agradecer especialmente à Marisa, sem ela, este projeto não teria sido o mesmo.

Aos restantes membros da Unidade de Investigação em Patobiologia Molecular, pela simpatia com que me receberam.

Um agradecimento especial ao Carlos, por todo o incentivo e paciência, por me ouvir quando mais precisei e pela companhia durante várias horas de escrita desta dissertação.

Por fim, deixo aqui o meu sincero agradecimento aos meus pais e irmão. Por todo o apoio, compreensão e carinho que me dão e por me incentivarem sempre a alcançar os meus objetivos.

Resumo

O cancro do cólon e reto familiar do tipo X (FCCTX) é um síndrome que define as famílias que preenchem os critérios de Amesterdão, mas cujos tumores não apresentam instabilidade de microssatélites e também nas quais não é identificada mutação germinal nos genes de reparação de erros de DNA do tipo *mismatch* (MMR). A sua causa molecular permanece desconhecida. O presente trabalho teve como objetivo avaliar o envolvimento de genes localizados numa região de suscetibilidade previamente identificada para o FCCTX (13q32-33), assim como de mutações identificadas previamente numa família FCCTX, através da análise do exoma por sequenciação de nova geração (NGS). Pretendeu-se ainda melhorar a caracterização molecular de tumores FCCTX.

Foi efetuada análise de mutações germinais nos genes *KDEL1* e *ERCC5* em 15 indivíduos índice de famílias FCCTX e 2 familiares de uma dessas famílias. No caso do gene *TPP2*, foi avaliado o envolvimento de um transcrito expresso alternativamente, previamente identificado, através de análise mutacional e da quantificação da expressão diferencial dos transcritos por *real-time* PCR. Foi ainda efetuada a análise de segregação com a doença na família, de 5 mutações em genes distintos, selecionadas a partir dos resultados da análise do exoma. Foi efetuada a análise de alterações de *copy-number* e de metilação nos genes MMR, *MGMT* e *APC* em 22 tumores FCCTX por MS-MLPA.

Não foram identificadas mutações potencialmente patogénicas nos genes *KDEL1* e *ERCC5*. No entanto, foram identificadas 2 mutações, uma no *ERCC5* (c.2636 A>G) e outra no *KDEL1* (c.455A>T) em relação às quais não se pode excluir a sua patogenicidade. Não foi detetada qualquer mutação no *TPP2* associada à expressão diferencial dos transcritos, no entanto verificou-se que a expressão difere entre tecidos (sangue e cólon). A análise de segregação das mutações selecionadas a partir da análise do exoma, revelou que apenas para um dos genes a alteração poderá ser patogénica. Foram identificados ganhos frequentes, assim como metilação, nos genes MMR e *MGMT*, nos tumores FCCTX, sendo significativamente mais frequentes num subgrupo destes tumores que apresenta perdas em genes supressores de tumor (TSG+), em relação ao grupo que não apresenta estas

alterações.

A metilação no *APC* também apresentou padrões distintos entre os dois subgrupos de tumores FCCTX. Em conclusão, as variantes observadas nos genes *KDELC1*, *ERCC5* e *TPP2*, assim como a alteração identificada no âmbito da análise do exoma não devem ser excluídas, podendo ser possível a sua contribuição para a suscetibilidade para o FCCTX. O padrão de alterações de *copy-number* e de metilação nos tumores FCCTX reforça a existência de pelo menos duas entidades moleculares distintas no FCCTX e sugere mecanismos de tumorigênese específicos para a iniciação tumoral neste síndrome.

Palavras-chave: FCCTX, mutações germinais, genes de suscetibilidade, metilação, *copy-number*

Abstract

Familial colorectal cancer type X (FCCTX) defines families that fulfill the Amsterdam criteria, but in whom no germline mutation is identified in the DNA mismatch repair (MMR) genes and whose tumors do not present microsatellite instability. Its molecular basis is not known yet. This study aimed to evaluate the involvement of genes located in a previously identified susceptibility region (13q32-33) as well as the involvement of mutations previously identified in a FCCTX family by exome sequencing analysis using next generation sequencing (NGS). It was also intended to further improve the molecular characterization of FCCTX tumors.

Analysis of germline mutations in *KDELC1* and *ERCC5* genes was performed in 15 index individuals from FCCTX families and in two family members of one of these families. The involvement of an alternatively expressed transcript of the *TPP2* gene, previously reported, was also evaluated by mutational analysis and by quantification of differential expression by real-time PCR. Five mutations were selected from the results of exome sequencing analysis and a segregation analysis with the disease in the family was performed. The analysis of copy-number alterations and methylation in MMR, *MGMT* and *APC* genes was performed in 22 FCCTX tumors by MS-MLPA.

There were any potentially pathogenic mutations identified in *KDELC1* or *ERCC5* genes. However, for two mutations, one in *ERCC5* (c.2636A>G) and other in *KDELC1* (c.455A>T), cannot exclude a possible pathogenicity. Two mutations were found in the *TPP2* gene, although, not associated with the differential expression of the transcripts. However it was found that the expression of the transcripts differs among tissues (blood and colon). Segregation analysis of selected mutations, obtained from exome analysis, revealed that only one of the alterations may be pathogenic, so far. Frequent gains, as well as methylation, in MMR and *MGMT* genes were identified in FCCTX tumors, being significantly more frequent in a subset of these tumors that show loss of heterozygosity of tumor suppressor genes (TSG+), compared to the group without these changes.

Methylation in *APC* also showed different patterns between these two subgroups

of FCCTX tumors. In conclusion, the variants observed in *KDELC1*, *ERCC5* and *TPP2* genes, as well as the molecular alteration identified through exome sequencing analysis, should not be excluded as possibly contributing to FCCTX susceptibility. The pattern of copy-number changes and tumor methylation in FCCTX reinforces the existence of at least two different molecular entities in FCCTX and suggests specific mechanisms for tumor initiation in this syndrome.

Keywords: FCCTX, germline mutations, susceptibility genes, methylation, copy-number

Índice

1	Introdução	1
1.1	Cancro do cólon e reto	1
1.2	Sequência adenoma-carcinoma	2
1.3	Vias de instabilidade genómica envolvidas no CCR	2
1.4	Síndromes hereditários de CCR	3
1.4.1	Síndromes hereditários de cancro do cólon e reto (CCR) associado a polipose	4
1.4.2	Cancro do cólon e reto hereditário não associado a polipose	6
1.5	Alterações moleculares somáticas nos tumores FCCTX	8
1.5.1	Via de sinalização Wnt	9
1.5.2	Gene <i>APC</i>	10
1.6	Identificação de genes candidatos para suscetibilidade para o FCCTX	12
1.6.1	Genes candidatos de suscetibilidade para o FCCTX na região 13q32-33	13
1.6.2	Genes candidatos de suscetibilidade para o FCCTX na região 21q11	16
1.7	<i>Whole-exome sequencing</i>	17
1.8	Objetivos	18
2	Materiais e métodos	19
2.1	Material biológico	19
2.2	Métodos	20
2.2.1	Extração de ácidos nucleicos	20
2.2.2	Quantificação espectralométrica de ácidos nucleicos	21
2.2.3	<i>Polymerase chain reaction</i> (PCR)	22
2.2.4	Análise de mutações nos genes <i>KDELC1</i> , <i>ERCC5</i> , <i>TPP2</i> e nos genes candidatos obtidos pela análise de WES	25
2.2.5	Sequenciação automática	26
2.2.6	<i>Methylation-Specific Multiplex ligation-dependent probe amplification</i>	29
2.2.7	<i>Real-time PCR</i>	31

3	Resultados	37
3.1	Caracterização molecular dos tumores de indivíduos com FCCTX	37
3.1.1	Genes MMR e <i>MGMT</i>	37
3.1.2	Gene <i>APC</i>	44
3.2	Análise de mutações nos genes <i>KDELC1</i> e <i>ERCC5</i>	48
3.2.1	<i>KDELC1</i>	48
3.2.2	<i>ERCC5</i>	49
3.3	Análise do envolvimento do <i>splicing</i> no potencial exão 13A do gene <i>TPP2</i> na suscetibilidade para o FCCTX	51
3.4	Análise de WES	55
4	Discussão	61
4.1	Caracterização molecular dos tumores de indivíduos com FCCTX e impacto na suscetibilidade para o FCCTX	61
4.1.1	Alterações de <i>copy-number</i> e de metilação nos genes MMR e <i>MGMT</i>	61
4.1.2	Alterações de <i>copy-number</i> e de metilação no gene <i>APC</i>	63
4.2	Análise mutacional dos genes <i>KDELC1</i> e <i>ERCC5</i>	65
4.2.1	<i>KDELC1</i>	66
4.2.2	<i>ERCC5</i>	66
4.3	Análise do envolvimento do <i>splicing</i> no potencial exão 13A do gene <i>TPP2</i> na suscetibilidade para o FCCTX	67
4.3.1	Quantificação da expressão do gene <i>TPP2</i>	68
4.4	Análise de WES	69
A	Anexo A	79
A.1	Extração de RNA a partir de biópsias de tecido congelado - <i>kit RNeasy Mini Kit (Qiagen)</i>	79
B	Anexo B	81
B.1	Análise de mutações dos genes <i>KDELC1</i> , <i>ERCC5</i> , <i>TPP2</i> e nos genes candidatos obtidos pela análise de WES	81
C	Anexo C	87
C.1	Protocolos para preparação dos reagentes utilizados no controlo da eficiência dos produtos de PCR	87
D	Anexo D	89
D.1	Extração e purificação de DNA a partir de banda em gel de agarose - <i>kit QIAquick Gel Extraction (Qiagen)</i>	89

<i>ÍNDICE</i>	xv
E Anexo E	91
E.1 Protocolo de precipitação e purificação do DNA – <i>kit ABI PRISM BigDye[®] Terminator v1.1 Cycle Sequencing (Applied Biosystems)</i>	91
F Anexo F	93
F.1 Análise de MS-MLPA	93
G Anexo G	95
G.1 <i>Real-time</i> PCR	95

Índice de Figuras

1.1	Sequência adenoma-carcinoma	2
1.2	Síndromes hereditários de CCR	4
1.3	Via WNT	10
1.4	Proteína APC e respetivos domínios	11
1.5	Diagrama ilustrativo da interdependência entre os dois eventos genéticos necessários à perda de função do gene <i>APC</i> , de acordo com o número de domínios de regulação da β -catenina	12
1.6	Localização no cromossoma 13 dos genes candidatos para suscetibilidade para o FCCTX	14
1.7	Características estruturais da proteína codificada pelo gene <i>ERCC5</i>	15
2.1	Esquema da reação de PCR	22
2.2	Esquema da reação de MS-MLPA	29
2.3	Esquema da localização das sondas MS-MLPA no gene <i>APC</i>	30
2.4	Representação gráfica das fases do <i>real-time</i> PCR	32
3.1	Resultado da análise de expressão relativa do gene <i>APC</i> no sangue periférico de doentes FCCTX	47
3.2	Eletroforetograma parcial obtido no sequenciador <i>ABI PrismTM 3130 Genetic Analyzer (Applied Biosystems)</i> , representando a mutação c.2636A>G (sequência <i>forward</i> , amostra L449).	51
3.3	Eletroforetograma parcial obtido no sequenciador <i>ABI PrismTM 3130 Genetic Analyzer (Applied Biosystems)</i> , representando a mutação c.2636A>G (sequência <i>reverse</i> , amostra L449).	51
3.4	Eletroforese em gel de agarose 2% (p/v) dos produtos de PCR do fragmento que incorpora os exões 13 ao 17 do gene <i>TPP2</i>	54
3.5	Análise de expressão relativa a dois transcritos do gene <i>TPP2</i> (13/13A e 13/14)	54

3.6	Eletroforese em gel de agarose 2% (p/v) dos produtos amplificados por PCR para o transcrito alternativo 13A	55
3.7	Árvore genealógica da família selecionada para a análise de WES (L56) . .	58
4.1	Mecanismo de reparação de erros no DNA originados por <i>stress</i> oxidativo/agentes alquilantes	63

Índice de Tabelas

1.1	Critérios de Amesterdão I e II	7
1.2	Critérios de Bethesda	7
2.1	Programa utilizado na reação de sequenciação dos fragmentos de cada gene estudado	27
2.2	Programa utilizado na síntese de cDNA	33
3.1	Análise de <i>copy-number</i> aos genes MMR e MGMT nas amostras de tumores e mucosas normais	40
3.2	Frequências obtidas da análise de <i>copy-number</i> aos genes MMR e MGMT nas amostras de tumores e mucosas normais	40
3.3	Análise de metilação aos genes MMR e MGMT nas amostras de tumores e mucosas normais	41
3.4	Frequências obtidas da análise de metilação aos genes MMR e MGMT nas amostras de tumores e mucosas normais	41
3.5	Análise de <i>copy-number</i> aos genes MMR e MGMT nas amostras de sangue periférico	42
3.6	Análise de metilação aos genes MMR e MGMT nas amostras de sangue periférico	43
3.7	Análise de <i>copy-number</i> e de metilação no gene APC em amostras de tumores e mucosa normal de indivíduos FCCTX	45
3.8	Frequências obtidas da análise MS-MLPA ao gene APC nas amostras de tumores e mucosas normais	45
3.9	Análise de <i>copy-number</i> e de metilação no gene APC em amostras de sangue de indivíduos FCCTX	46
3.10	Frequências obtidas da análise MS-MLPA ao gene APC nas amostras de sangue	46
3.11	Identificação das mutações germinais detetadas no gene <i>KDELCL1</i> e respectivas frequências.	48

3.12	Resultados da análise <i>in silico</i> da consequência prevista das mutações germinais encontradas no gene <i>KDELC1</i>	49
3.13	Identificação das mutações germinais detetadas no gene <i>ERCC5</i> e respectivas frequências.	50
3.14	Resultados da análise <i>in silico</i> da consequência prevista das mutações germinais encontradas no gene <i>ERCC5</i>	52
3.15	Identificação das mutações germinais detetadas no gene <i>TPP2</i> e respectivas frequências.	52
3.16	Resultados da análise <i>in silico</i> da consequência prevista das mutações germinais encontradas no intrão 13 do gene <i>TPP2</i>	53
3.17	Análise de segregação das mutações identificadas com a expressão das duas isoformas	53
3.18	Resultado da análise WES nos genes <i>RECQL5</i> , <i>RWDD4</i> , <i>CADM1</i> , <i>PLXNB1</i> e <i>DHCR24</i>	56
3.19	Resultados da análise <i>in silico</i> às variantes germinais encontradas no gene <i>RECQL5</i> com recurso ao <i>software Mutation taster</i>	58
B.1	Programa utilizado na amplificação por PCR utilizando o kit <i>Biotaq (Bioline)</i>	81
B.2	Programa utilizado na amplificação por PCR utilizando o kit <i>Amplitaq Gold (Applied Biosystems)</i>	81
B.3	Sequências dos <i>primers</i> e tamanho do fragmento amplificado para o gene <i>KDELC1</i>	82
B.4	Condições utilizadas na amplificação por PCR dos fragmentos do gene <i>KDELC1</i>	82
B.5	Resultados obtidos através da análise mutacional do gene <i>KDELC1</i>	82
B.6	Sequências dos <i>primers</i> e tamanho do fragmento amplificado para o gene <i>ERCC5</i>	83
B.7	Condições utilizadas na amplificação por PCR dos fragmentos do gene <i>ERCC5</i>	83
B.8	Resultados obtidos através da análise mutacional do gene <i>ERCC5</i>	83
B.9	Sequências dos <i>primers</i> e tamanho do fragmento amplificado para o gene <i>TPP2</i>	83
B.10	Condições utilizadas na amplificação por PCR dos fragmentos do gene <i>TPP2</i>	84
B.11	Resultados obtidos através da análise mutacional do gene <i>TPP2</i>	84
B.12	Sequências dos <i>primers</i> desenhados e tamanho do fragmento amplificado para o gene <i>TPP2</i>	84
B.13	Sequências dos <i>primers</i> desenhados e tamanho do fragmento amplificado para variantes identificadas em genes candidatos obtidos pela análise de WES	84
B.14	Condições utilizadas na amplificação por PCR dos fragmentos dos genes candidatos obtidos pela análise de WES	85

B.15	Sequências dos <i>primers</i> desenhados e tamanho do fragmento amplificado para as variantes identificadas em genes candidatos obtidos pela análise de WES	85
F.1	Reagentes e respectivos volumes para a preparação da <i>mix Ligase A</i> , <i>mix Ligase-65</i> e <i>mix Ligase-Digestion</i>	94
F.2	Reagentes e respectivos volumes para a preparação da <i>mix SALSA</i> e <i>Polimerase mix</i>	94
F.3	Reagentes e respectivos volumes para a preparação da <i>Polimerase mix</i> . . .	94
G.1	Programa utilizado nas reações de <i>real-time</i> PCR no aparelho <i>LightCycler® Plus 480 Multiwell Plates (Roche)</i>	95
G.2	Sequências dos <i>primers</i> utilizados na quantificação do gene <i>TPP2</i>	95
G.3	Programa utilizado nas reações de <i>real-time</i> PCR no aparelho <i>7900HT Sequence Detection System (Applied Biosystems)</i>	95

Acrónimos

A	DEPC dietilpirocarbonato.
APAF polipose adenomatosa familiar atenuada.	DHCR24 <i>24-dehydrocholesterol reductase</i> .
APC <i>adenomatous polyposis coli</i> .	DMSO dimetilsulfóxido.
B	DNA ácido desoxirribonucleico.
BER reparação por excisão de bases.	DNTPs desoxirribonucleótidos fosfatados.
C	DOCK9 <i>dedicator of cytokinesis 9</i> .
CA critérios de Amesterdão.	DTT ditioneitol.
CADM1 <i>Cell adhesion molecule 1</i> .	DVL <i>Dishevelled</i> .
CCR cancro do cólon e reto.	E
CDNA DNA complementar.	ERCC5 <i>excision repair cross complementing group 5</i> .
CID domínio inibidor da β -catenina.	EXO I <i>Exonuclease I</i> .
CIMP fenótipo metilador das ilhas CpG.	F
CIN instabilidade cromossómica.	FAD <i>flavin adenine dinucleotide</i> .
CK1 <i>casein kinase 1</i> .	FCCTX cancro do cólon e reto familiar do tipo X.
CT <i>threshold cycle</i> .	FEN1 <i>flap structure-specific endonuclease 1</i> .
D	FZ <i>Frizzled</i> .
DDNTPs 2',3'-didesoxinucleótidos trifosfatados.	G
	GEF fator de troca de trifosfato de guanina.

GSK3 *glycogen synthase kinase 3.*

H

HNPCC cancro do cólon e reto hereditário não associado a polipose.

HSP70 *heat shock protein 70.*

HSPA13 *heat shock protein 70kDa family, member 13.*

J

JPS síndrome de polipose juvenil.

K

KDEL1 *KDEL (Lys-Asp-Glu-Leu) containing 1.*

L

LOH perda de heterozigotia.

M

MAF *minor allele frequency.*

METTL21C *methyltransferase like 21C.*

MMR reparação de erros do DNA do tipo *mismatch.*

MS-MLPA *Methylation-Specific Multiplex ligation-dependent probe amplification.*

MSI instabilidade de microssatélites.

MSI-H MSI de alto grau.

MSI-L MSI de baixo grau.

MSS microssatélites estável.

N

NER reparação por excisão de nucleótidos.

NGS *next-generation sequencing.*

NLS sinais de localização nuclear.

NRIP1 *Nuclear Receptor Interacting Protein 1.*

P

PAF polipose adenomatosa familiar.

PAM polipose adenomatosa associada ao gene *MUTYH.*

PCNA *proliferating cell nuclear antigen.*

PJS síndrome de Peutz-Jeghers.

PLXNB1 *plexin B1.*

PP2A fosfatase 2A.

PS polipose serrada.

R

RECQL5 *RecQ protein-like 5.*

RWDD4 *RWD domain containing 4.*

S

SAM domínio *Sterile α.*

SAMSN1 *SAM domain, SH3 domain and nuclear localization signals 1.*

SH3 domínio *SRC Homology 3.*

SIFT *Sorting Intolerant From Tolerant.*

SLC10A2 *solute carrier family 10 - sodium/bile acid cotransporter, member 2.*

SNPs *single-nucleotide polymorphisms.*

SSB *single strand DNA binding protein.*

STK24 *serine/threonine protein kinase 24.*

T

TCR reparação acoplada à transcrição.

TEX30 *testis expressed 30.*

TPP2 *Tripeptidyl Peptidase II.*

TSG genes supressores de tumor.

U

UBM motivo de ligação à ubiquitina.

UIPM Unidade de investigação de patobiologia molecular.

V

VEP *Variant Effect Predictor.*

W

WES *whole-exome sequencing.*

WGS *whole-genome sequencing.*

WT *wild-type.*

X

XPG *xeroderma pigmentosum group G.*



Introdução

1.1 Cancro do cólon e reto

O cancro do cólon e reto (CCR) tem uma incidência a nível mundial bastante elevada, sendo o terceiro cancro mais diagnosticado em homens e o segundo em mulheres. As taxas de incidência mais elevadas são encontradas na Austrália e Nova Zelândia, Europa e América do Norte [Siegel et al., 2011].

Em 2014, espera-se que cerca de 71830 homens e 65000 mulheres sejam diagnosticadas com CCR e 26270 homens e 24040 mulheres morram da doença. A nível mundial tem havido um declínio nas taxas de mortalidade desde 1980 nos homens e desde 1947 nas mulheres, sendo estes declínios atribuídos a melhorias no tratamento, mudança de padrões nos fatores de risco para o CCR e aumento do rastreio. Apesar deste progresso dramático na redução da incidência e das taxas de mortalidade do CCR, as disparidades raciais e socioeconómicas permanecem [Siegel et al., 2014].

Os casos de CCR podem dever-se a fatores ambientais presentes na vida quotidiana bem como a fatores genéticos [Zambirinis et al., 2009]. Os fatores de risco ambientais, especialmente nos países desenvolvidos, incluem o tabagismo, sedentarismo, excesso de peso, o consumo excessivo de álcool e de carne vermelha processada [Siegel et al., 2011]. No caso de fatores genéticos, a herança genética do CCR é muito complexa.

Embora a maioria dos CCR surjam esporadicamente, cerca de 25-35% são classificados como CCR familiar, uma vez que existe a contribuição de fatores genéticos herdados para o seu desenvolvimento. Enquanto os perfis genómicos dos CCR esporádicos têm sido estudados extensivamente, poucos estudos analisaram os perfis tumorais dos CCR familiares, fazendo com que apenas 5% dos síndromes hereditários sejam conhecidos [Zambirinis et al., 2009] [van Wezel et al., 2012].

1.2 Sequência adenoma-carcinoma

Em 1990, Fearon e Vogelstein propuseram um modelo genético para a tumorigênese do cólon e reto, denominado sequência adenoma-carcinoma (Figura 1.1). Este modelo associa alterações clínico-patológicas com anomalias genéticas na progressão do CCR [Walther et al., 2009].

Segundo este modelo, o passo inicial na tumorigênese é o da formação do adenoma a partir da mucosa normal do cólon, associado à inativação (por mutação ou perda de heterozigotia (LOH)) do gene *adenomatous polyposis coli* (*APC*) (localizado em 5q). Quando há uma inativação do gene *APC*, não ocorre fosforilação eficaz da β -catenina, resultando na sua acumulação no citoplasma. Posteriormente a β -catenina é translocada para o núcleo e estimula os alvos do *TCF7L2*, como o aumento da proliferação, diferenciação, migração e adesão das células do cólon [Colussi et al., 2013] [Walther et al., 2009].

Mutações no proto-oncogene *KRAS* (localizado em 12p) levam à sua ativação constitutiva, o que permite que a célula escape à apoptose, e LOH em genes localizados na região 18q (*DCC/SMAD4*) resultam na diferenciação do adenoma e progressão. Por fim, a perda bialélica ou inativação do *TP53* (localizado em 17q) medeia a transição adenoma-carcinoma sendo um dos passos fundamentais na carcinogênese do cólon e reto, porque estimula a atividade proliferativa devido à perda de controle do ciclo celular e apoptose [Boland and Goel, 2010] [Colussi et al., 2013].

Estudos posteriores refinaram esta sequência e gerou-se o modelo da via tradicional ou via da instabilidade cromossômica (CIN) [Zambirinis et al., 2009].

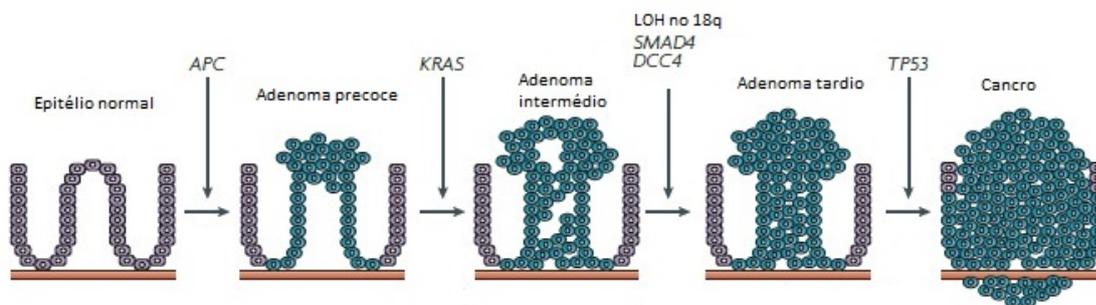


Figura 1.1: Sequência adenoma-carcinoma
(adaptado de Walther et al., 2009)

1.3 Vias de instabilidade genômica envolvidas no CCR

Estudos realizados nos últimos 30 anos têm aumentado a compreensão dos mecanismos envolvidos na iniciação e desenvolvimento do CCR. Os resultados demonstraram a existência de três vias de instabilidade genômica: a via CIN, a da instabilidade de microssatélites (MSI) e a via do fenótipo metilador das ilhas CpG (CIMP).

A via CIN é a mais comum, estando presente na maioria dos CCR (65-70%), sendo caracterizada pela presença de LOH e por uma generalidade de alterações numéricas e estruturais dos cromossomas nas células tumorais. Apesar da base molecular para a via CIN ainda ser desconhecida, muitos genes já lhe foram associados, como referido no ponto 1.2 [Colussi et al., 2013].

A via MSI é responsável por aproximadamente 15% dos CCR esporádicos e pela maioria dos casos de cancro do cólon e reto hereditário não associado a polipose (HNPCC). Esta via é caracterizada pela expansão ou contração de sequências repetitivas devido à inativação do sistema de reparação de erros do DNA do tipo *mismatch* (MMR). Este processo pode dever-se a mutações germinais nos genes MMR (*MSH2*, *MSH6*, *MLH1* e *PMS2*) ou a silenciamento epigenético por metilação dos seus promotores (especialmente do gene *MLH1*). Esta inativação do sistema MMR provoca um aumento de 100 vezes na taxa de mutação ao nível das células da mucosa cólica [Zambirinis et al., 2009].

A via CIMP, que se encontra descrita em cerca de 20-30% dos CCR, tem uma considerável sobreposição com a via MSI. A via CIMP envolve a hipermetilação aberrante de sequências dinucleótidas CpG localizadas nas regiões promotoras de genes envolvidos na regulação do ciclo celular, apoptose, angiogénese, reparação do DNA, invasão e adesão, levando à sua perda de expressão. É de notar que na via CIMP a lesão precursora é o pólipo serreado ao contrário do que acontece nas duas vias anteriores, em que a lesão precursora é o pólipo adenomatoso [Colussi et al., 2013] [Zambirinis et al., 2009].

1.4 Síndromes hereditários de CCR

Ao longo dos últimos 10 a 15 anos, a genética molecular tem tido um impacto significativo na identificação de mutações somáticas e germinais associadas com o desenvolvimento de CCR. Mutações em genes de reparação do DNA (*MLH1*, *MSH2*, *MSH6*, *PMS2* e *MUTYH*), bem como em genes envolvidos na transdução de sinal (*APC* e *SMAD4*) encontram-se associadas a verdadeiros síndromes hereditários. No entanto, a maioria dos casos familiares de CCR são provavelmente poligénicos, e muitas das alterações genéticas envolvidas estão ainda por identificar [Power et al., 2010].

A suspeita clínica da existência de um síndrome hereditário de CCR num indivíduo ou família pode surgir devido a uma idade precoce no diagnóstico, uma forte história familiar, a descoberta de múltiplos pólipos durante uma colonoscopia ou o reconhecimento de outras manifestações clínicas associadas aos síndromes de hereditários conhecidos [Gallagher et al., 2010]. Os síndromes hereditários de CCR estão divididos em duas categorias: o cancro do cólon e reto hereditário associado a polipose e o HNPCC (Figura 1.2).

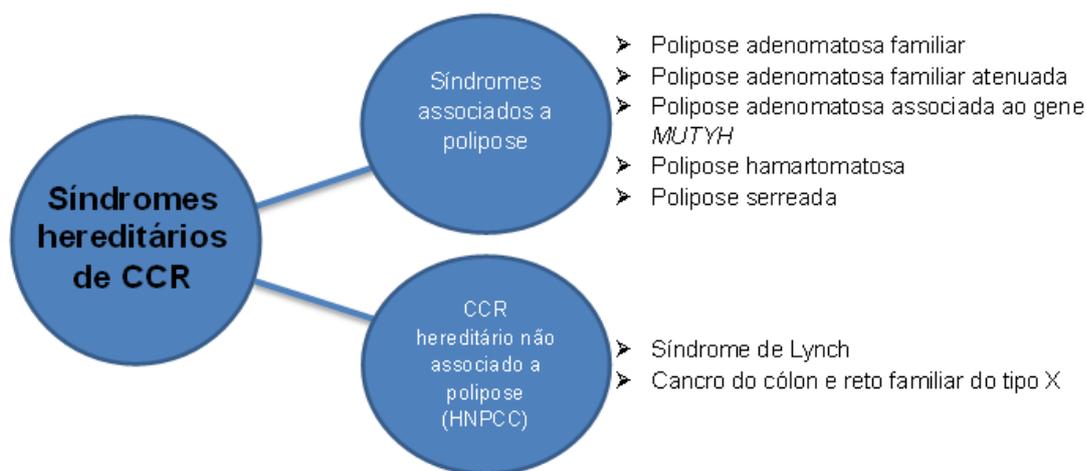


Figura 1.2: Síndromes hereditárias de CCR
(Adaptado de [Lindor et al., 2005])

1.4.1 Síndromes hereditárias de CCR associado a polipose

1.4.1.1 Polipose adenomatosa familiar

A polipose adenomatosa familiar (PAF) é caracterizada por centenas a milhares de pólipos adenomatosos no cólon e reto. Se estes pólipos não forem removidos, o desenvolvimento de cancro é inevitável, ou seja, a penetrância da doença é de 100%. A PAF é herdada de forma autossômica dominante e tem duas variantes estritamente relacionadas, o síndrome de Turcot e o síndrome de Gardner. Surge numa idade precoce, geralmente na segunda ou terceira década de vida e progride para CCR entre os 35 a 40 anos de idade, em praticamente todos os casos. No entanto é um síndrome relativamente raro, representando menos de 1% de todos os casos de CCR [Gallagher et al., 2010] [Power et al., 2010].

A PAF está associada a mutações germinais no gene supressor de tumor *APC*, um gene com 15 exões, localizado na região cromossômica 5q21-q22. Foram identificadas mais de 1000 mutações no gene *APC*, que geralmente originam uma proteína truncada, seja por causa de mutações *frameshift* ou *nonsense* [Patel and Ahnen, 2012].

1.4.1.2 Polipose adenomatosa familiar atenuada

A polipose adenomatosa familiar atenuada (APAF) é uma forma menos grave da PAF, caracterizada por uma média de 30 pólipos adenomatosos cólicos (numa gama de 10-100). A idade de diagnóstico nos indivíduos com APAF é mais tardia e têm tendência a desenvolver neoplasias no cólon proximal numa idade mais avançada (aproximadamente 55 anos) [Jasperson et al., 2010].

A APAF é causada por mutações no gene *APC*, mais especificamente nas extremidades 5' e 3', em oposição à região central do gene, onde as mutações estão associadas a um fenótipo mais agressivo, com um maior número de pólipos. Uma mutação truncante na extremidade 5' do gene anula a síntese da proteína, no entanto, esta pode ser reiniciada

a jusante, a partir de um local de iniciação de tradução alternativo. Assim, as mutações responsáveis pela APAF podem ser vistas como hipomórficas, uma vez que codificam proteínas quase completas, fornecendo uma explicação intuitiva para o seu fenótipo atenuado [Gallagher et al., 2010] [de la Chapelle, 2004]. Este fenótipo pode ser explicado com base no modelo '*just-right*', ou seja, mutações no gene *APC* que levam à inativação completa da regulação da β -catenina são menos vantajosas para a formação do tumor do que a manutenção de um determinado nível de atividade residual [Albuquerque et al., 2002].

1.4.1.3 Polipose adenomatosa associada ao gene *MUTYH*

A polipose adenomatosa associada ao gene *MUTYH* (PAM) caracteriza-se pela presença de mutações germinais bialélicas neste gene (localizado na região cromossômica 1p) e é uma doença com transmissão autossômica recessiva. O gene *MUTYH* é um constituinte da via de reparação por excisão de bases (BER) que está envolvida na reparação de danos oxidativos no DNA [Zambirinis et al., 2009] [Patel and Ahnen, 2012]. Funcionalmente, este gene evita as transversões G:C para T:A causadas pelo *stress* oxidativo. Em todas as populações estudadas, foram detetadas duas mutações *missense* que ocorrem recorrentemente (Y165C e G382D) e que constituem mais de 50% de todas as mutações identificadas no gene *MUTYH* em indivíduos com PAM [de la Chapelle, 2004] [Jasperson et al., 2010].

Em comparação com a PAF, o diagnóstico de PAM é feito numa idade mais avançada, entre os 45 a 56 anos [Lindor, 2009b]. Os doentes com PAM desenvolvem um número variável de pólipos e possuem um risco significativo de progressão para CCR, principalmente no cólon proximal. Apesar de os pólipos adenomatosos serem os precursores mais predominantes, os pólipos serrados também são comuns [Patel and Ahnen, 2012].

Atualmente é evidente que o espectro clínico das mutações germinais no gene *MUTYH* é mais amplo e pode incluir CCR sem polipose [Patel and Ahnen, 2012].

1.4.1.4 Polipose hamartomatosa

Na polipose hamartomatosa estão englobados o síndrome de Peutz-Jeghers (PJS), o síndrome de polipose juvenil (JPS), o síndrome de Cowden, o síndrome de Cronkhite-Canada, o síndrome de Gorlin e o síndrome de Bannayan-Riley-Ruvalcaba [Gallagher et al., 2010]. Destas condições, apenas o PJS e JPS estão associados a um risco aumentado para o CCR e para outras neoplasias [Jasperson et al., 2010].

O PJS é um síndrome autossômico dominante associado a mutações germinais no gene *STK11/LKB1*. Este gene localiza-se na região cromossômica 19p13.3 e a proteína para a qual codifica tem múltiplas funções, incluindo a regulação do ciclo celular, a mediação da apoptose e a polaridade celular [Patel and Ahnen, 2012] [Gallagher et al., 2010]. Este síndrome tem sido associado ao aumento do risco de CCR mas também de cancro de mama, pâncreas, intestino, esófago e ovário [Jasperson et al., 2010].

O JPS é uma doença autossômica dominante rara, caracterizada pelo desenvolvimento de múltiplos pólipos juvenis (10 ou mais) no trato gastrointestinal, que conduz a um risco

aumentado de CCR. As manifestações fenotípicas são detetadas na transição da infância para a adolescência e o termo juvenil refere-se ao tipo histológico do pólip observado. [Power et al., 2010]. O JPS é causado por uma mutação germinal num dos seguintes genes: *SMAD4*, *BMPR1A* e *ENG*, sendo que todos estão relacionados com o fator de crescimento TGF- β [Patel and Ahnen, 2012].

1.4.1.5 Polipose serreada

A polipose serreada (PS) é caracterizada pela presença de múltiplos pólipos serreados (lesão precursora) no cólon e é provavelmente hereditária. A sua base genética ainda não foi estabelecida, mas já foram propostos quer padrões de herança autossómicos dominantes como recessivos [Patel and Ahnen, 2012].

A PS apresenta um risco aumentado para o desenvolvimento de CCR, o qual é diagnosticado em média entre os 50 a 60 anos, podendo ocorrer com frequência tumores síncronos e metacrónicos e as lesões são tendencialmente observadas no cólon proximal. Atualmente é amplamente aceite que existe uma via de carcinogénese serreada para além da sequência adenoma-carcinoma tradicional. A acumulação de alterações somáticas nos pólipos serreados sésses, como mutações ativadoras do *BRAF* e metilação generalizada das ilhas CpG, com ou sem MSI, são eventos importantes nesta via para o desenvolvimento de carcinoma [Jasperson et al., 2010].

1.4.2 Cancro do cólon e reto hereditário não associado a polipose

O HNPCC é a forma familiar de CCR mais comum e divide-se em dois síndromes distintos, o síndrome de Lynch e o cancro do cólon e reto familiar do tipo X (FCCTX). O termo HNPCC foi inicialmente aplicado a grupos de famílias heterogéneas que preenchiam os critérios de Amesterdão I (Tabela 1.1). Posteriormente reconheceu-se a importância dos tumores extra-cólicos associados ao HNPCC, e os critérios anteriores foram modificados nesse sentido, passando a denominarem-se critérios de Amesterdão II (Tabela 1.1) [Vasen et al., 1991] [Vasen et al., 1999].

Para além dos critérios de Amesterdão (CA), foram também criados os critérios de Bethesda (Tabela 1.2), tendo porém um objetivo diferente, o de identificar possíveis casos de síndrome de Lynch para as famílias que não cumprem os CA. Assim, os indivíduos/famílias que não cumprem os CA são avaliados quanto aos critérios de Bethesda, e caso cumpram um deles, têm indicação para pesquisa de MSI nos tumores. Esta pesquisa é efetuada com recurso a um painel de cinco marcadores de microssatélites, denominados marcadores de Bethesda: dois marcadores de mononucleótidos, BAT25 e BAT26, e três marcadores de dinucleótidos, D5S346, D2S123 e D17S250. O número de marcadores de microssatélites que apresentam instabilidade permite classificar o grau de MSI dos tumores em três categorias distintas: quando apenas um marcador apresenta MSI, o tumor é considerado como tendo MSI de baixo grau (MSI-L), se dois ou mais marcadores são alterados, o tumor é considerado como tendo MSI de alto grau (MSI-H) e se o tumor não

apresentar instabilidade em nenhum dos marcadores então é considerado microssatélites estável (MSS) [Umar et al., 2004]. Após a pesquisa de MSI, os indivíduos cujos tumores apresentarem MSI-H devem ser submetidos a testes genéticos para os genes MMR.

Tabela 1.1: Critérios de Amesterdão I e II

Critérios de Amesterdão I
1. Três ou mais parentes com CCR, em que um dos indivíduos afetados é parente em primeiro grau dos outros dois
2. Pelo menos duas gerações afetadas
3. Pelo menos um dos casos de CCR deve ser diagnosticado antes da idade de 50 anos
4. PAF deve ser excluída
Critérios de Amesterdão II
1. Três ou mais parentes com cancro associado a HNPCC (CCR, endométrio, intestino delgado, uréter, pélvis renal, ovário, estômago ou cérebro)

(Adaptado de Umar et al., 2004)

Tabela 1.2: Critérios de Bethesda

Critérios de Bethesda
1. CCR diagnosticado num doente com idade inferior a 50 anos
2. Presença de CCR síncrono, metacrónico ou outros tumores do espectro do HNPCC, independentemente da idade
3. Presença de CCR com histologia de MSI-H (presença de infiltrado linfocitário no tumor, reação Crohn-like, presença de muco e células em forma de sinete), diagnosticado num doente com idade inferior a 60 anos
4. Presença de CCR num indivíduo com um ou mais familiares em 1º grau com um tumor do espectro do HNPCC sendo que um dos tumores foi diagnosticado numa idade inferior a 50 anos
5. Presença de CCR num indivíduo com dois ou mais familiares em 1º ou 2º grau com um tumor do espectro do HNPCC, independentemente da idade

(Adaptado de Umar et al., 2004)

1.4.2.1 Síndrome de Lynch

O síndrome de Lynch é a causa hereditária mais comum de desenvolvimento de CCR, sendo responsável por 2%-4% de todos os casos de CCR. Tem transmissão autossômico dominante e é definido em base genética pela presença de uma mutação germinal num dos genes MMR (*MLH1*, *MSH2*, *MSH6* e *PMS2*) [Boland, 2005], sendo que mais de 90% das mutações germinais se encontram descritas nos dois primeiros genes (*MLH1* e *MSH2*). Os CCRs destes indivíduos são caracterizados pela presença de MSI-H, sendo esta uma característica dos tumores que surgem no contexto da deficiência do sistema MMR [Jasperson et al., 2010].

Recentemente foram descritas deleções germinais no gene *EpCAM* (ou *TACSTD1*), o qual se localiza a montante do *MSH2*, num subconjunto de famílias com síndrome de

Lynch, as quais abrangem normalmente a região promotora e/ou alguns exões deste último, levando à sua inativação somática por metilação do promotor, apresentando assim um mecanismo diferente de inativação de um gene MMR [Ligtenberg et al., 2008].

A nível fenotípico os indivíduos com síndrome de Lynch têm seis vezes mais probabilidade de desenvolver CCR, mas também estão predispostos a vários tipos de tumores extra-cólicos, nomeadamente no endométrio, estômago, rim, ovário, intestino delgado e uréter. O diagnóstico é efetuado numa idade jovem (~ 48,7 anos) e os tumores desenvolvem-se predominantemente no cólon proximal, ao contrário do que acontece com o CCR esporádico. Histologicamente, estes tumores são frequentemente pouco diferenciados, mucinosos e apresentam infiltrado linfocitário [Jasperson et al., 2010] [Ku et al., 2012].

1.4.2.2 Cancro do cólon e reto familiar do tipo X

À medida que mais famílias iam sendo analisadas, verificou-se que existia um grupo de famílias com CCR, que preenchiam os CA mas para os quais não eram identificadas mutações germinais nos genes MMR e cujos tumores não apresentavam MSI [Lynch and de la Chapelle, 2003]. Com base nestas observações e outras do foro clínico, foi sugerido por Lindor et al. em 2005, que estes casos deveriam ser encarados como uma doença hereditária distinta, tendo-se introduzido o termo FCCTX, e posteriormente, reclassificado o termo síndrome de Lynch (indivíduos com mutações germinais nos genes MMR) [Boland, 2005]. As famílias FCCTX têm um padrão de transmissão autossómico dominante, mas a sua etiologia genética ainda se mantém desconhecida [Patel and Ahnen, 2012].

Em comparação com o síndrome de Lynch, o CCR nas famílias FCCTX desenvolve-se mais tarde (~ 60,7 anos), ocorre predominantemente no cólon distal e apresenta poucas características morfológicas distintas, sendo mais semelhante aos CCR esporádicos (média a alta diferenciação, padrões de crescimento infiltrativo e glandular). Estas famílias não apresentam um risco aumentado para o desenvolvimento de tumores extra-cólicos, a taxa de progressão adenoma-carcinoma é mais lenta e apenas apresentam um risco duas vezes superior para o desenvolvimento de CCR [Lindor, 2009a] [Ku et al., 2012].

Apesar de o FCCTX ser uma das principais causas de CCR hereditário, continua a ser pouco investigado. A sua melhor compreensão é necessária, de modo a fornecer pistas importantes sobre a predisposição da doença, contribuir para o diagnóstico molecular e para a identificação de estratégias terapêuticas direcionadas [Dominguez-Valentin et al., 2014].

1.5 Alterações moleculares somáticas nos tumores FCCTX

No ano de 2011, o grupo de gastroenterologia da Unidade de investigação de patologia molecular (UIPM), pertencente ao IPOLFG, E.P.E iniciou um trabalho onde se pretendia avaliar o envolvimento da via CIN num grupo de tumores de famílias diagnosticadas com FCCTX. Desta forma, o grupo identificou 2 identidades moleculares distintas

entre as famílias FCCTX. Uma mais prevalente (72%), em que os tumores apresentam frequente LOH em diversos genes supressores de tumor (TSG) (*APC*, *TP53*, *SMAD4* e *DCC*), e outra entidade (28%) em que os tumores não apresentam evidência de LOH nos TSG. A LOH dos TSG está frequentemente envolvida na tumorigênese do cólon e reto associada à via CIN, o que revela a importância desta via para o FCCTX no grupo mais prevalente. Por outro lado, a diferença de LOH nos dois grupos sugere que estes tumores seguem vias de tumorigênese do cólon e reto diferentes [Francisco et al., 2011].

Ainda se verificou que no primeiro grupo (TGS+) os tumores apresentavam frequentemente mutações somáticas no *KRAS* e *APC* e metilação dos genes *MMR* e *MGMT*, enquanto no segundo grupo (TGS-) estas alterações eram raramente observadas [Francisco et al., 2011]. Dada a informação obtida neste estudo, é importante continuar a caracterização dos tumores e perceber os mecanismos moleculares inerentes ao FCCTX, através do estudo do estado de metilação e das alterações moleculares nos genes *APC*, *MMR* e *MGMT*. Para o gene *APC* é essencial realizar um estudo ao nível do estado de metilação e *copy-number* e para os genes *MMR* e *MGMT* ainda é necessário efetuar a análise ao nível de *copy-number*. Desta forma, no presente trabalho um dos objetivos será continuar a avaliação molecular dos tumores FCCTX.

1.5.1 Via de sinalização Wnt

A maioria dos tumores do cólon e reto adquire mutações que resultam na ativação da via de sinalização Wnt/ β -catenina. A sinalização intracelular desta via constitui um dos principais mecanismos de controlo de processos no desenvolvimento embrionário e homeostase no organismo adulto, através da regulação do equilíbrio da proliferação celular, da diferenciação e da apoptose, do controlo da polaridade celular e da especificação do destino celular [Albuquerque et al., 2011].

Tradicionalmente, a via de sinalização Wnt é classificada em dois tipos: canónica (dependente de β -catenina) e não canónica (independente de β -catenina). A via canónica é a mais estudada e também a mais fortemente implicada no desenvolvimento de CCR [Najdi et al., 2011].

Na via canónica, na ausência de sinalização de um ligando Wnt extracelular, a β -catenina citoplasmática é capturada por um complexo multiproteico intracelular, ou complexo de destruição, que é composto pelas proteínas: APC, proteína Axin, cinases *glycogen synthase kinase 3* (GSK3), *casein kinase 1* (CK1) e fosfatase 2A (PP2A). A β -catenina capturada por este complexo é fosforilada nos seus resíduos N-terminais específicos, o qual provoca a sua ubiquitinação e degradação. Desta forma não ocorre a sua translocação para o núcleo e não há interação com os membros da família dos fatores de transcrição TCF/LEF (Figura 1.3a).

Quando os ligandos Wnt estão presentes, estes ligam-se a um complexo de recetores constituídos pela proteína transmembranar *Frizzled* (FZ) e pelos recetores LRP5/6 (membros da família de recetores de lipoproteínas de baixa densidade). A ligação destes

recetores recruta a proteína *Dishevelled* (DVL), que leva à fosforilação dos recetores LRP5/6 e ao afastamento da proteína Axin e da cinase GSK3 do complexo de destruição. A β -catenina fica livre e ocorre a sua acumulação no meio intracelular, com posterior translocação para o núcleo, onde se associa com os membros da família TCF/LEF. Assim, são ativados os genes alvo da via Wnt que estão envolvidos em vários processos celulares como a apoptose, proliferação e diferenciação (Figura 1.3b) [Najdi et al., 2011] [Albuquerque et al., 2011].

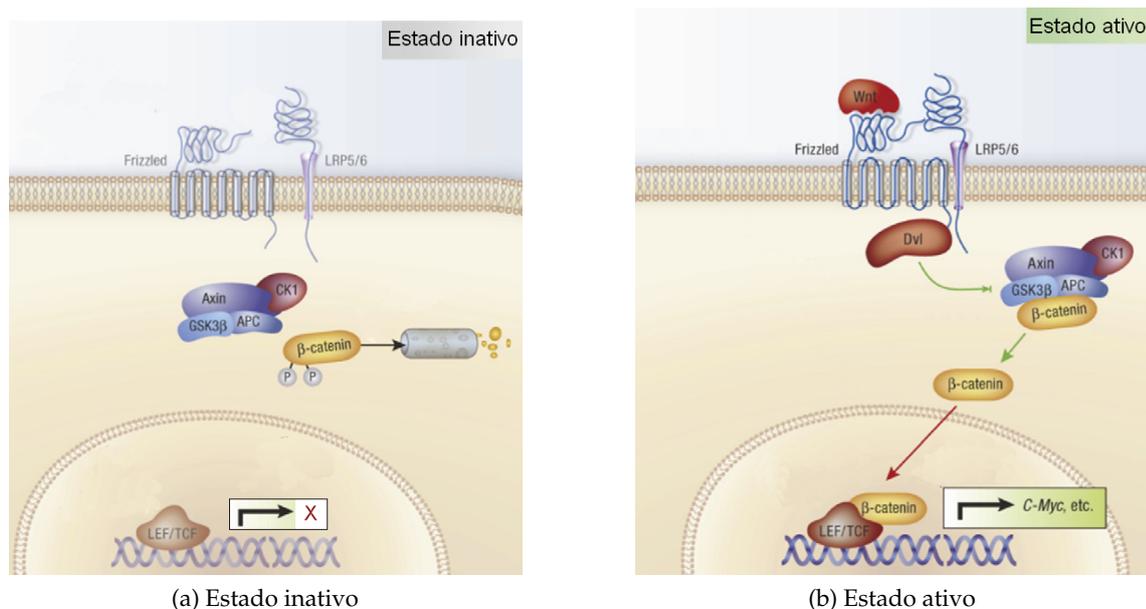


Figura 1.3: Via WNT
(adaptado de Albuquerque et al., 2011)

1.5.2 Gene APC

O gene *APC* codifica para uma proteína homodimérica de 312 kDa que é essencial para a integridade do citoesqueleto, regulação do ciclo celular e apoptose, adesão intercelular e na regulação da via de sinalização Wnt [Coppedè et al., 2014].

Vários motivos no domínio central do gene *APC* são responsáveis pela regulação dos níveis de β -catenina intracelulares. Quatro repetições de 15 aminoácidos (aa) ligam-se à β -catenina, enquanto sete motivos de 20 aa estão envolvidos na ligação e regulação negativa desta. Intercaladas entre essas 20 repetições de aa estão três locais de ligação para as proteínas AXIN1/AXIN2, necessárias para um recrutamento da proteína APC para o complexo de destruição [Albuquerque et al., 2011]. Após a repetição de 20 aa, está o domínio inibidor da β -catenina (CID), que regula a sua sinalização (Figura 1.4). Desta forma, o gene *APC* atua como regulador negativo da sinalização da β -catenina na transformação de células epiteliais do cólon. Quando ocorrem mutações neste gene que provocam a deleção de domínios de ligação e regulação da β -catenina, a capacidade de regulação negativa é perdida e a via de sinalização Wnt é ativada, existindo uma transcrição desregulada dos genes localizados a jusante, os quais promovem a tumorigênese

[Jaiswal and Narayan, 2008]. No entanto, já foi demonstrado que proteínas truncadas que apresentem um, dois ou três domínios de regulação da β -catenina mantêm ainda alguma capacidade de regulação negativa.

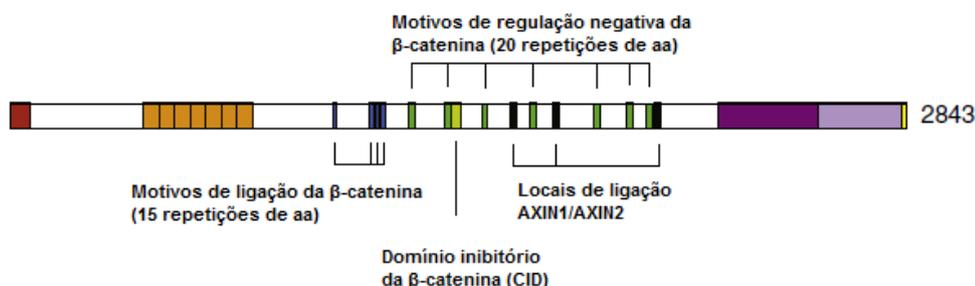


Figura 1.4: Proteína APC e respectivos domínios
(Adaptado de Albuquerque et al., 2011)

Mutações no gene *APC* são responsáveis pela PAF e pela maioria dos CCRs esporádicos. De acordo com a interpretação clássica da hipótese *two-hits* de Knudson para a tumorigênese, os *two-hits* representam dois eventos mutacionais independentes, sendo o resultado final uma perda de função de genes supressores de tumor [Smits et al., 2000]. Com base nesta hipótese, a inativação de ambos os alelos do *APC* ocorre de forma independente, como resultado de dois eventos genéticos, os quais podem ser detetados na maior parte dos tumores intestinais, mesmo nos estádios iniciais de desenvolvimento tumoral [Smits et al., 2000]. No entanto, foi demonstrado posteriormente que o *APC* não segue este modelo, pois os dois eventos mutacionais estão relacionados, ou seja, existe uma interdependência entre a localização da mutação germinal/somática (primeiro evento) e a localização e tipo do segundo evento (somático), o que conduz à formação do tumor. De acordo com o número de domínios de regulação da β -catenina resultantes da mutação inicial podem prever-se três cenários diferentes (Figura 1.5) [Albuquerque et al., 2002]:

1. Se o primeiro evento genético dá origem a uma proteína truncada sem domínios de regulação, o segundo evento origina uma proteína truncada com um ou dois domínios de regulação (menos frequentemente);
2. Se a mutação germinal der origem a uma proteína truncada com apenas um domínio de regulação da β -catenina, o segundo evento observado será a perda alélica, ou então, menos frequentemente, uma proteína APC sem domínios de regulação;
3. Se a mutação germinal resultar numa proteína com dois domínios de regulação da β -catenina, o segundo evento tende a originar uma proteína APC sem domínios de regulação ou então, menos frequentemente, a perda alélica.

Estes resultados indicam que há uma seleção para os genótipos do gene *APC* que são capazes de manter alguma atividade na regulação negativa da sinalização β -catenina, ou seja, que permitem obter um nível de sinalização ótimo, ideal para a formação de tumores [Albuquerque et al., 2002].

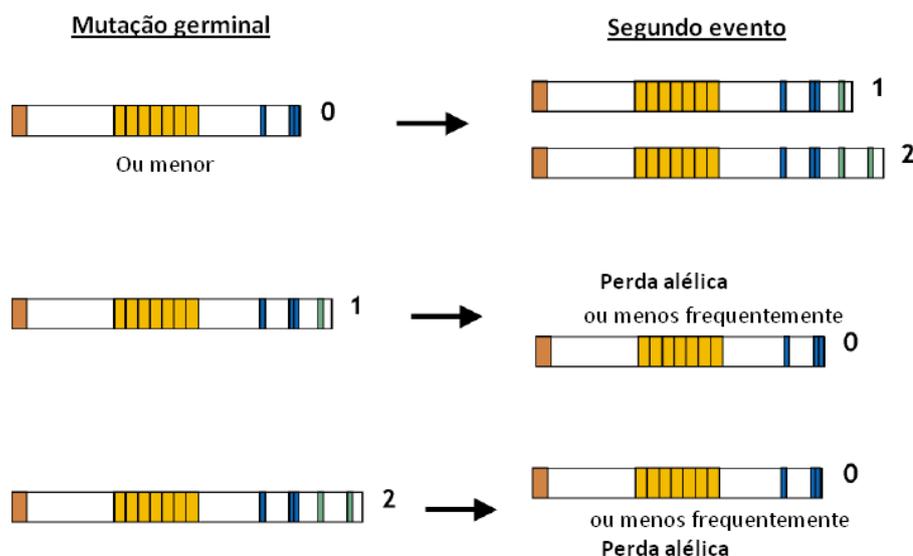


Figura 1.5: Diagrama ilustrativo da interdependência entre os dois eventos genéticos necessários à perda de função do gene *APC*, de acordo com o número de domínios de regulação da β -catenina (adaptado de Albuquerque et al., 2002)

1.6 Identificação de genes candidatos para suscetibilidade para o FCCTX

O mesmo grupo de gastroenterologia, pertencente ao IPOLFG, E.P.E encontra-se a desenvolver um projeto que visa o mapeamento de novos *loci* de suscetibilidade genética para o FCCTX, no seguimento do trabalho descrito no ponto 1.5. Assim, foi efetuado um estudo de *linkage* de todo o genoma utilizando *arrays* de 50000 *single-nucleotide polymorphisms* (SNPs), em 2 famílias FCCTX mais informativas, uma com tumores TGS+ e outra com tumores TGS-, quer em indivíduos afetados como saudáveis. Assim, na família com tumores com TGS+ foram identificadas 3 regiões cromossômicas sugestivas (13q, 16q e 21q) como potenciais regiões de suscetibilidade, enquanto para a outra família não se obtiveram resultados com significado estatístico.

Com o objetivo de delimitar estas regiões, recorreu-se a uma análise de LOH utilizando marcadores de microssatélites, para as regiões acima mencionadas. No entanto, utilizando os mesmos marcadores de microssatélites para efetuar uma análise de *linkage* com a doença, apenas foram confirmadas as regiões 13q e 21q [Belo, H., 2010]. Utilizando mais marcadores de microssatélites para estas regiões, assim como amostras adicionais de tumores FCCTX, foi possível delimitar regiões mínimas de LOH para estas 2 regiões (13q32-33 e 21q11), as quais coincidem com as regiões que apresentavam LOD score mais elevado, previamente identificadas pela tecnologia de DNA *microarrays*. Assim, no cromossoma 13 delimitou-se uma região mínima de LOH de aproximadamente 0,87Mb e no cromossoma 21 de 1,3Mb. Esta delimitação foi importante, porque permitiu que fosse efetuada uma análise mutacional num número reduzido de genes candidatos de suscetibilidade para o FCCTX.

Por análise da histologia dos tumores FCCTX, verificou-se ainda que a LOH no cromossoma 13 era mais frequente em adenomas, enquanto no cromossoma 21 a LOH ocorria frequentemente em carcinomas. Estes dados sugerem que os eventos em 13q ocorrem ao nível da iniciação tumoral, sendo que em 21q os eventos surgem ao nível da progressão tumoral [Belo, H., 2010] [Pereira, C., 2013].

Até ao momento, nenhuma das duas regiões identificadas foi proposta como região de suscetibilidade para o desenvolvimento de CCR. Apenas foi verificada na região 13q a existência de ganhos frequentes em CCR em carcinomas num estadio inicial e em CCR metastático, que englobam frequentemente todo o braço longo do cromossoma [Sayagués et al., 2010] [Middeldorp et al., 2012].

1.6.1 Genes candidatos de suscetibilidade para o FCCTX na região 13q32-33

Foram selecionados 16 genes, dos quais 4 são pseudogenes (*RPL39P29*, *RNY5P8*, *LOC121952* e *RPL7P45*), localizados na região cromossômica 13q32-33 (Figura 1.6) para análise de mutações germinais, tendo em conta a expressão da proteína no cólon, se os seus aspetos funcionais eram sugestivos de poderem estar associados com a tumorigénese (particularmente à do cólon e reto) e se não se encontravam descritas como causa de outras patologias. Até à data já foi efetuada análise mutacional para estes genes, sendo em seguida descrita alguma informação sucinta sobre os mesmos no contexto da suscetibilidade para o FCCTX.

O gene *serine/threonine protein kinase 24* (*STK24*) codifica uma proteína serina/treonina cinase, responsável pela fosforilação de proteínas intra-celulares. Este gene também promove a apoptose em resposta ao *stress* celular por ativação da via das caspases. A sua função e o envolvimento deste tipo de proteínas na predisposição para o síndrome de Peutz-Jeghers, fizeram deste gene um candidato para o desenvolvimento de FCCTX [Hennig et al., 2012].

O gene *dedicator of cytokinesis 9* (*DOCK9*) codifica para um fator de troca de trifosfato de guanina (GEF) que ativa a proteína Cdc42, através da troca de GDP por GTP livre. A Cdc42 medeia vários processos que são essenciais para a homeostase celular, estando a sua desregulação associada a várias patologias, incluindo o cancro [Sinha and Yang, 2008]. Este gene e o *STK24* foram estudados antes de se obter a região mínima de LOH.

O gene *testis expressed 30* (*TEX30*) codifica uma proteína que tem uma classificação hipotética de hidrolase, ou seja, codifica para uma enzima que catalisa a hidrólise de uma ligação química. Este gene foi selecionado para análise de mutações germinais porque atua em vários processos celulares que estão associados à progressão tumoral, tais como a regulação do crescimento e proliferação celular [Liu et al., 2012].

O gene *methyltransferase like 21C* (*METTL21C*) codifica uma metiltransferase, responsável pela transferência de um grupo metil de um dador para um recetor. A proteína codificada pelo gene *METTL21C* tem uma função semelhante à codificada pelo gene *MGMT*, tendo esta última a função de reparação de erros induzidos por *stress* oxidativo

no cólon. Havendo uma relação entre os dois genes, uma alteração na função da proteína METTL21C pode modificar a expressão de vários genes envolvidos na tumorigênese, visto saber-se que o promotor do gene *MGMT* se encontra frequentemente metilado em carcinomas esporádicos do cólon [Shen et al., 2005] [Nagasaka et al., 2008].

O gene *solute carrier family 10 - sodium/bile acid cotransporter, member 2 (SLC10A2)* codifica um co-transportador de ácidos biliares dependente de sódio. Este co-transportador é o mecanismo primário para a absorção de ácidos biliares intestinais por células apicais do íleo distal. Portanto, mutações neste gene podem conduzir à má absorção dos ácidos biliares, sendo que a presença no cólon de ácidos biliares secundários aumenta o risco de cancro [Grünhage et al., 2008].

Alguns destes genes acima descritos já foram excluídos pelo grupo de gastroenterologia, particularmente os genes *STK24* [Pires, S., 2011], *DOCK9* [Zhao, A., 2012], *METTL21C* [Póvoa, V., 2011] e *TEX30* [Pereira, G., 2014] pois não foram detetadas mutações germinais patogénicas. Dos restantes, encontra-se o gene *SLC10A2*, cujo envolvimento no FCCTX ainda não foi excluído e o gene *TPP2*, cujo estudo vai ter seguimento no presente trabalho, juntamente com os genes mais recentemente selecionados, *KDELC1* e *ERCC5* (encontram-se a negrito na Figura 1.6).

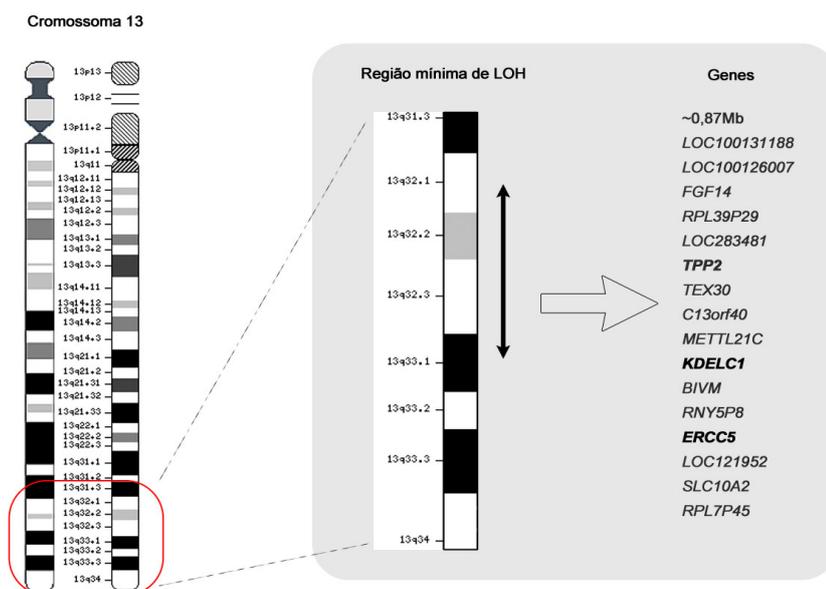


Figura 1.6: Localização no cromossoma 13 dos genes candidatos para suscetibilidade para o FCCTX

1.6.1.1 *KDELC1*

O gene *KDEL (Lys-Asp-Glu-Leu) containing 1 (KDELC1)* codifica uma proteína de 502 aa e é composto por 10 exões. Como membro da família de proteínas do retículo endoplasmático, esta proteína contém um motivo Lisina-Asparagina-Glutamina-Leucina (Lys-Asp-Glu-Leu ou KDEL) localizado na região C-terminal, evitando que as proteínas residentes do retículo endoplasmático sejam secretadas. Assim, as proteínas que

transportam este motivo estão ligadas por um recetor no aparelho de Golgi, de modo a que o complexo ligando-recetor retorne ao retículo endoplasmático [NCBI, 2014c]. Não são conhecidas mais funções para este gene, sendo que existem muito poucos estudos efetuados que o envolvam.

Este gene foi escolhido para análise mutacional devido essencialmente à sua função e porque numa análise prévia de expressão de *microarrays* efetuada em tumores FCCTX, uma das funções alteradas foi observada ao nível do retículo endoplasmático. É previsível que alterações germinais no gene *KDELC1* podem levar à alteração da expressão proteica de proteínas envolvidas no processo de tumorigénese.

1.6.1.2 ERCC5

O gene *excision repair cross complementing group 5 (ERCC5)*, também conhecido como *xeroderma pigmentosum group G (XPG)*, é um dos principais membros da via reparação por excisão de nucleótidos (NER), desempenha um papel fundamental na reparação de danos no ácido desoxirribonucleico (DNA) e na manutenção da integridade do genoma. Este gene é composto por 15 exões.

O gene *ERCC5* codifica uma proteína denominada *DNA repair protein complementing XP-G cells*, que contém 1186 aa e é um membro da família da *flap structure-specific endonuclease 1 (FEN1)*. A estrutura primária da proteína contém os domínios N- e I-, que são altamente conservados e formam o núcleo da nucleasa (Figura 1.7). Também contém um *spacer*, que é muito ácido, para interações proteína-proteína (TFIIH e RPA), um motivo de ligação à ubiquitina (UBM) e um domínio PIP, que medeia as interações com a *proliferating cell nuclear antigen (PCNA)*. A interação entre esta proteína e a PCNA pode estar envolvida no desencadeamento da incisão na via NER. A região C-terminal contém ainda dois sinais de localização nuclear (NLS) putativos nos resíduos 1051-1084 e 1169-1186 [Schärer, 2008].

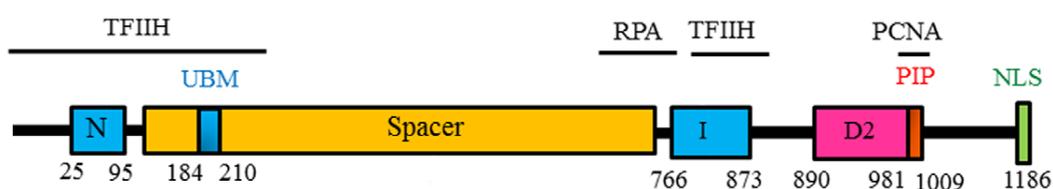


Figura 1.7: Características estruturais da proteína codificada pelo gene *ERCC5*
(Adaptado de Zhu et al., 2012)

A proteína ERCC5 é necessária em duas sub-vias em NER. Uma delas é a reparação acoplada à transcrição (TCR), que preferencialmente remove lesões de bloqueio de alongamento na cadeia de DNA transcrita de genes ativos; a outra é a reparação genómica global (GGR), que remove as lesões em todos os locais do genoma, em qualquer momento na célula [Emmert et al., 2001].

Mutações pontuais claramente patogénicas no gene *ERCC5* causam a desordem *Xeroderma pigmentosum*, no entanto a presença de polimorfismos (Asp1104His ou His46His), que provocam alterações na expressão da XPG, foram associados a um risco aumentado

para o desenvolvimento de diferentes tipos de cancro [Zhu et al., 2012]. Por outro lado, a expressão elevada da proteína foi associada à resistência a agentes quimioterapêuticos, como a oxaliplatina, a fluoropirimidina e a trabectedina [Aracil et al., 2013].

Neste trabalho foi selecionada para estudo apenas a região terminal do gene, mais precisamente a que codifica para o domínio PIP e a região com locais NLS, uma vez que, são regiões de interação com a proteína PCNA. Como a interação entre a proteína ERCC5 e PCNA pode estar envolvida no desencadeamento da incisão 3' em NER, uma mutação nesta região poderá alterar a capacidade de reparação do DNA e consequentemente aumentar o risco de cancro.

1.6.1.3 TPP2

O gene *Tripeptidyl Peptidase II (TPP2)* é composto por 30 exões e codifica uma complexa tripeptidil peptidase II, que é uma aminopeptidase que remove os tripéptidos da região N-terminal livre dos péptidos mais longos, sendo considerada uma enzima *housekeeping*. Este complexo é formado por subunidades repetidas de 138 kDa, reunidas em dois filamentos torcidos que formam um complexo de alto peso molecular (~ 6 MDa) [Preta et al., 2010].

A proteína TPP2 participa no *turnover* das proteínas, na via da ubiquitina-proteossoma, onde opera a jusante do proteossoma 26S na proteólise citosólica. Também tem um papel importante na produção e destruição de antígenos MHC de classe I e na degradação de hormonas peptídicas, como por exemplo, neuropéptidos [Rockel et al., 2012].

Dados recentes indicam que a TPP2 é translocada para o núcleo em resposta a várias formas de *stress* celular, em especial para controlar as respostas aos danos no DNA em células malignas. Desta forma, está implicada na regulação da proliferação e sobrevivência destas células. Esta ligação entre a TPP2 e as respostas ao *stress*, pode ser de grande interesse na biologia do cancro [Preta et al., 2010].

A análise deste gene já esteve inserida em estudos anteriores do projeto de identificação de genes candidatos para suscetibilidade para o FCCTX. Sendo um gene de grandes dimensões, foi dividido em 9 fragmentos com extremidades sobreponíveis para se efetuar a sua análise mutacional no mRNA. Desta forma, observou-se que havia uma inserção de uma isoforma com 72 pb entre o exão 13 e 14 que não se encontra descrita.

Mais recentemente ainda se avaliou a expressão desta isoforma através de *real-time* PCR, onde se observou que a sua expressão é variável, no entanto parece ser transmitida dentro das famílias a estudo. É essencial continuar o estudo deste gene, em específico da isoforma encontrada, pois aparenta estar envolvida na tumorigénese do cólon e reto.

1.6.2 Genes candidatos de suscetibilidade para o FCCTX na região 21q11

Até à data foram analisados 3 genes na região cromossómica 21q11, nomeadamente os genes *Nuclear Receptor Interacting Protein 1 (NRIP1)*, *heat shock protein 70kDa family, member 13 (HSPA13)* e *SAM domain, SH3 domain and nuclear localization signals 1 (SAMS1)*. A seleção destes genes teve por base os mesmos critérios descritos no ponto 1.6.1, à exceção

do *NRIP1*.

O gene *NRIP1* codifica a proteína RIP140, que está envolvida na via do recetor dos estrogénios. Neste caso, o gene foi selecionado com base na presença de cancro de mama em algumas mulheres das famílias FCCTX em estudo, para além do CCR, o que sugere a hipótese do possível envolvimento deste tipo de proteína na tumorigénese de ambos os cancros. A análise mutacional germinal deste gene em indivíduos FCCTX já foi concluída, não tendo sido identificada a presença de qualquer mutação potencialmente patogénica [Pereira, C., 2013].

O gene *HSPA13* codifica uma proteína de choque térmico que pertence à família das *heat shock protein 70* (HSP70). Esta proteína está envolvida em processos celulares de *folding* de proteínas recém-sintetizadas ou desnaturadas, processando proteínas citosólicas e secretoras. Ainda atua na remoção de proteínas desnaturadas ou com estrutura incorreta [Mayer and Bukau, 2005].

O gene *SAMSN1* codifica um membro da família de proteínas de *scaffold* e adaptadores putativos. Esta proteína contém dois domínios de ligação, um domínio *SRC Homology 3* (SH3) e um domínio *Sterile α* (SAM) [Wang et al., 2010]. Mutações neste gene estão associadas a patologias hematopoiéticas [Claudio et al., 2001].

Os genes *HSPA13* e *SAMSN1* foram excluídos como possíveis genes candidatos para o FCCTX [Pereira, G., 2014]. Apesar de não ter sido identificada nenhuma mutação potencialmente patogénica no gene *NRIP1* [Pereira, C., 2013], este ainda não foi excluído em relação ao seu envolvimento no FCCTX.

1.7 *Whole-exome sequencing*

Estudos anteriores de sequenciação do genoma foram constrangidos por limitações tecnológicas, onde foram utilizados métodos tradicionais de sequenciação de baixo rendimento. Atualmente o surgimento das tecnologias de *next-generation sequencing* (NGS) revolucionou a sequenciação do genoma, pelo que estas tecnologias têm sido utilizadas em estudos de *whole-genome sequencing* (WGS) e de *whole-exome sequencing* (WES). Este campo tem avançado muito, sendo acompanhado por outros avanços técnicos, tais como o desenvolvimento de abordagens computacionais e estatísticas.

A WES tem sido amplamente aplicada na identificação de novas mutações germinais subjacentes ao cancro hereditário, podendo ser especialmente importante nos cancros cujas causas genéticas ainda não foram totalmente caracterizadas, como é o exemplo do FCCTX. A sequenciação de todo o exoma também revolucionou a capacidade de identificar novas mutações e genes que predis põem para doenças genéticas anteriormente não identificadas.

Esta análise requer várias etapas de enriquecimento da sequência alvo, seguido de sequenciação paralela em massa. Durante os passos de enriquecimento, as regiões genómicas de interesse (todos os exões) são capturadas, enquanto que as regiões não codificantes são removidos antes da sequenciação, o que leva a uma redução significativa na proporção

do genoma que deve ser sequenciado. O desenvolvimento de uma variedade de métodos para capturar seletivamente todos os exões no genoma humano tornou a WES uma técnica viável, permitindo que o exoma seja sequenciado muito rapidamente e em grande detalhe.

Em comparação com a WGS, a WES é mais rentável, menos desafiadora e mais viável em termos técnicos. Uma vez que apenas 1% a 2% de todo o genoma é sequenciado, há uma maior cobertura de leitura das regiões codificantes, o que tornou a WES uma abordagem mais popular que a WGS [Ku et al., 2013] [Ku et al., 2012].

No seguimento do projeto que visa o mapeamento de novos *loci* de suscetibilidade genética para o FCCTX que está a ser desenvolvido pelo grupo de gastroenterologia, enviaram-se amostras de 4 indivíduos da família com tumores TGS- para uma análise de WES, visto esta não ter obtido resultados significativos na análise dos *arrays* de 50000 SNPs. Com esta análise pretende-se identificar variantes que sejam partilhadas pelos quatro indivíduos, que serão selecionadas para estudos subsequentes de caracterização das mutações, de modo a avaliar a sua patogenicidade e associação com a doença.

1.8 Objetivos

O presente trabalho teve como objetivos o esclarecimento dos mecanismos moleculares envolvidos nas perdas de heterozigotia detetadas frequentemente na região 13q em tumores FCCTX, descritas num estudo prévio, através da análise de *copy-number*, estado de metilação e ao nível da expressão do RNA.

Pretendeu-se efetuar a avaliação do envolvimento de genes na suscetibilidade para o FCCTX, nomeadamente dos genes *ERCC5* e *KDEL1*, localizados na região cromossómica 13q32-33, previamente identificada como uma possível região de suscetibilidade para o FCCTX. Desta forma tentou-se contribuir para o esclarecimento da base molecular deste síndrome, através da análise mutacional dos genes acima mencionados.

Objetivou-se continuar a avaliação do envolvimento na suscetibilidade para o FCCTX do gene *TPP2*, localizado também na região em 13q32-33, através de análise mutacional e ao nível da expressão do RNA.

Ainda se pretendeu efetuar a caracterização das variantes identificadas através da WES, de modo a avaliar o seu carácter patogénico e averiguar o seu envolvimento na suscetibilidade para o FCCTX.



Materiais e métodos

2.1 Material biológico

Para o estudo dos genes *KDELC1* e *ERCC5* foram selecionados 15 indivíduos índice de 15 famílias que cumprem os critérios para FCCTX, seguidas na Consulta de Risco Familiar de Cancro do Cólon e Reto da Clínica de Risco Familiar do IPOLFG, E.P.E. de Lisboa. Foram também incluídos dois familiares de uma das famílias (L55), de modo a efetuar estudos de segregação de algumas variantes genéticas com a doença. Para o estudo complementar do exão 12 do gene *ERCC5* foram ainda analisados 47 indivíduos saudáveis. A análise mutacional foi realizada a partir de DNA genómico extraído a partir de sangue periférico de cada indivíduo.

Na análise mutacional do exão 13/13A do gene *TPP2* foram incluídas 11 amostras, onde 7 são provenientes da família FCCTX L55, 3 são amostras de um grupo de indivíduos saudáveis e uma é proveniente de um indivíduo com síndrome de Lynch. Para a análise do exão 13/14 foram analisadas as mesmas amostras, com adição de 11 amostras de indivíduos com síndrome de Lynch e de 13 amostras do mesmo grupo de indivíduos saudáveis. Ainda para esta última região, foram analisadas 44 amostras de um outro grupo de indivíduos saudáveis como controlos normais. Já na análise mutacional dos exões 13 ao 17 ao nível do DNA complementar (cDNA), utilizaram-se 7 amostras de 2 famílias FCCTX (1 indivíduo da L7 e 6 indivíduos da L55) e 3 amostras de de um grupo de indivíduos saudáveis. O cDNA foi sintetizado a partir de RNA extraído de sangue periférico.

No estudo das mutações detetadas através da análise de WES nos genes *CADM1*, *DHCR24*, *PLXNB1*, *RWDD4* e *RECQL5*, foram analisados 12 indivíduos (afetados e em risco) pertencentes à família L56, a qual cumpre os critérios para FCCTX.

Na análise de MS-MLPA foram utilizadas 24 amostras de DNA genómico extraído a partir de sangue periférico, 26 amostras de DNA extraído a partir de tecido incluído em parafina e 5 amostras de DNA extraído a partir de tecido congelado, todos provenientes das 15 famílias já mencionadas anteriormente. Para esta metodologia é essencial a utilização de amostras normais como referência, como tal, para cada protocolo utilizado na reação de MS-MLPA e por cada tipo de extração de DNA foram utilizadas 3 amostras de referência, o que equivale a 17 amostras adicionais, visto uma mesma amostra ser utilizada em dois métodos de extração diferentes.

Na análise de *real-time* PCR para a quantificação da expressão do gene *APC* foram utilizadas 7 amostras, entre as quais 4 são amostras provenientes de 2 famílias FCCTX (1 indivíduo da L7 e 3 indivíduos da L55 e 3 são amostras utilizadas como referência. Na análise de *real-time* PCR para a quantificação da expressão de dois fragmentos do gene *TPP2* foram analisadas 4 amostras de sangue periférico de indivíduos saudáveis, 2 amostras de carcinomas e 2 amostras de adenomas.

2.2 Métodos

2.2.1 Extração de ácidos nucleicos

A extração de DNA e RNA, correspondentes a cada amostra utilizada neste estudo, já tinha sido efetuada anteriormente, à exceção das amostras de RNA extraídas a partir de biópsias de tecido congelado.

2.2.1.1 Extração de DNA genómico a partir de sangue periférico

A extração de DNA genómico a partir de sangue periférico foi efetuada utilizando o *kit* comercial *CITOGENE*[®] - *Genomic DNA Purification kit* (Citomed), como descrito pelo fabricante. O DNA extraído foi quantificado como referido no ponto 2.2.2 e posteriormente fez-se uma diluição das amostras em água bidestilada, para uma concentração final de 80 ng/ μ l, armazenando-se tanto a alíquota como o DNA *stock* a -20°C.

2.2.1.2 Extração de DNA genómico a partir de tecido incluído em parafina

A extração de DNA genómico incluído em tecido parafinado foi efetuada utilizando o método fenol-clorofórmio. O fenol e o clorofórmio desnaturam as proteínas, ficando estas solubilizadas na fase fenólica que se separa com maior eficácia da fase aquosa, onde se encontra o DNA.

O DNA foi quantificado por eletroforese em gel de agarose 0,8% (p/v) porque alguns dos reagentes utilizados na extração absorvem ao mesmo comprimento de onda que o DNA (260 nm) e para se observar a integridade do mesmo. A quantificação foi efetuada por comparação com padrões de DNA de concentração conhecida e as amostras foram diluídas para uma concentração de trabalho de 80 ng/ μ l. Armazenaram-se tanto a alíquota como o DNA *stock* a -20°C.

2.2.1.3 Extração de RNA a partir de sangue periférico

A extração de RNA foi realizada segundo o protocolo descrito para o reagente *TRI Reagent*[®] *Solution* (Ambion, Life Technologies). A quantificação do RNA foi efetuada como descrito no ponto 2.2.2. Posteriormente, as amostras de RNA foram armazenadas a -80°C.

2.2.1.4 Extração de RNA a partir de biópsias de tecido congelado

A extração de RNA a partir de biópsias de tecido congelado em azoto líquido foi efetuada com recurso ao *kit* comercial *RNeasy Mini Kit* (Qiagen) (Anexo A). Este *kit* proporciona um método rápido e simples para a purificação de RNA a partir de pequenas quantidades de material. O seu procedimento combina as propriedades de ligação seletiva de uma membrana de sílica-gel, com um sistema de tampões de alto teor salino (permitem que o RNA se ligue à membrana de sílica-gel) e ainda com a velocidade da tecnologia de *microspin*.

Inicialmente os tecidos congelados são submetidos a lise e homogeneização na presença de um tampão desnaturante. Este tampão contém guanidina-tiocianato e tem como função inativar as RNases. De seguida é adicionado etanol, que proporciona as condições adequadas de ligação do RNA à membrana, sendo os contaminantes eliminados. O último passo consiste na eluição do RNA, em água tratada com dietilpirocarbonato (DEPC) (Merck).

A quantificação do RNA e controlo da extração foram efetuados como descrito no ponto 2.2.2, sendo no fim as amostras de RNA armazenadas a -80°C.

2.2.2 Quantificação espectrofotométrica de ácidos nucleicos

O DNA extraído foi quantificado num espectrofotómetro *NanoDrop 2000* (Thermo Scientific) pela leitura de absorvância a 260 nm e 280 nm. O quociente entre estes valores indica o grau de pureza dos ácidos nucleicos, sendo que, as amostras são consideradas puras se apresentarem valores entre 1,65 e 2, enquanto valores inferiores indicam contaminação significativa com proteínas e valores superiores indicam contaminação com RNA.

O RNA extraído é quantificado da mesma forma que o DNA, sendo que as amostras são consideradas puras se a razão 260/280 nm estiver entre 1,8 a 2. Um valor inferior indica que há contaminação com proteínas. O RNA extraído também foi sujeito a uma eletroforese em gel de agarose 0,8% (p/v) para se observar a sua integridade e para descartar uma possível contaminação por DNA. No gel de agarose devem de se visualizar duas bandas de RNA ribossomal (23S e 16S). Caso se verifique um arrastamento das bandas, indica que há degradação de RNA e se for observada uma banda acima das duas bandas de RNA ribossomal, indica que há contaminação por DNA genómico.

2.2.3 *Polymerase chain reaction (PCR)*

A amplificação de DNA foi realizada recorrendo à técnica *polymerase chain reaction* (PCR). Esta técnica baseia-se na amplificação enzimática *in vitro* de um fragmento de DNA de interesse, utilizando um par de *primers* que hibridam com as extremidades 3'-OH das cadeias de DNA molde. Duas novas cadeias são sintetizadas a partir da cadeia de DNA molde em cada ciclo completo de PCR, ou seja, ocorre um aumento exponencial, havendo ao fim de n ciclos 2^n vezes mais cópias do que no início. Desta forma, a partir de uma quantidade reduzida de DNA obtêm-se milhares de cópias da cadeia original.

Um ciclo de PCR é constituído por três fases (Figura 2.1) : desnaturação do DNA a temperaturas elevadas, onde ocorre a separação da cadeia dupla do DNA molde através da quebra das pontes de hidrogénio; hibridação ou *annealing* dos *primers* à sua sequência complementar da cadeia simples do DNA, a uma temperatura variável; e extensão, onde a enzima *Taq DNA polymerase* replica a cadeia de DNA, ou seja, a enzima adiciona nucleótidos à cadeia molde, criando uma cadeia dupla a partir de cada uma das cadeias simples. A extensão inicia-se sempre no extremo 3' do *primer* e a cadeia é sintetizada exclusivamente na direção 5' para 3'.

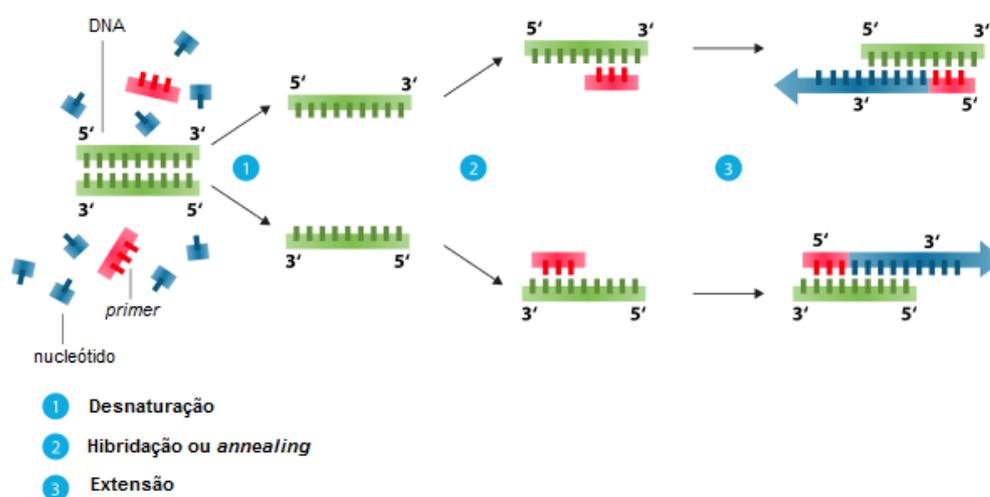


Figura 2.1: Esquema da reação de PCR

(adaptado de http://en.wikipedia.org/wiki/Polymerase_chain_reaction)

Para ocorrer a reação de PCR são necessários vários componentes: DNA genómico, enzima *Taq DNA polymerase* (termoestável e termorresistente, sendo extraída da bactéria termofílica *Thermus aquaticus*), *primers* (curtas sequências sintéticas de nucleótidos, entre 20 a 25 pb, que marcam as extremidades da sequência alvo), uma solução tamponada (mantém um pH ideal para a atividade enzimática), desoxirribonucleótidos fosfatados (dNTPs) (utilizados na síntese das novas cadeias) e íões magnésio (co-fator enzimático).

As reações de PCR foram preparadas para um volume final de 12,5 μ l, apresentando a seguinte composição: 1 μ l de DNA genómico (80 ng); quantidades variáveis de $MgCl_2$, dependendo das condições ótimas de cada reação (50 mM, *Bioline*TM); 1 μ l de dNTPs

(200 μM , *Illustra*TM, *GE Healthcare*); quantidade variável de *primer forward* e *reverse* (10 pmol/ μl); 1,25 μl de solução tamponada de PCR NH_4 10x (*Bioline*TM); 0,08 μl de *Taq DNA polymerase* (5 U/ μl , *Bioline*TM) e água bidestilada para perfazer o volume final.

As reações de PCR foram executadas em termocicladores *Veriti* (*Applied Biosystems*) e *Trio* (*Biometra*), encontrando-se os programas utilizados descritos no Anexo B.

2.2.3.1 Desenho de *primers* utilizados para a amplificação dos fragmentos de DNA de interesse

Os *primers* específicos para a amplificação dos genes *KDELC1* e *ERCC5* já tinham sido desenhados, no entanto, no decurso deste estudo foi necessário desenhar novos *primers* para amplificar os exões dos genes selecionados pela análise WES, e ainda para a região que compreende os exões 12 a 14 e o exão 24 alternativo do gene *TPP2*.

Esta tarefa tem que ser feita com muito cuidado, pois exige que se cumpram alguns critérios de modo a que a reação de PCR seja eficiente, sendo que para tal se recorreu à ferramenta *NetPrimer* [Biosoft, 2013]. A partir da sequência de referência do gene em estudo, escolheram-se pequenas sequências entre os 18 e 25 pb, que se introduziram na ferramenta *NetPrimer* [Biosoft, 2013]. Ao introduzir a sequência temos um retorno de várias informações: temperatura média de desnaturação (T_m), percentagem de guaninas e citosinas (%GC) e se há ou não formação de dímeros e *hairpins*. Desde modo, o conteúdo em GC's não pode ser muito elevado (30 a 50%) e deve ser semelhante entre os dois *primers* (no máximo, de preferência, 5% de diferença). A temperatura média de desnaturação dos dois *primers* deve ser semelhante (até 5°C de diferença preferencialmente) e é de evitar pares de *primers* com complementaridade de bases (dímeros) e também com complementaridade de bases no mesmo *primer* (*hairpins*).

Depois de se obter um *primer forward* e *reverse* que obedeciam a estas regras foi efetuada a confirmação da especificidade dos *primers*, ou seja, que estes apenas hibridam com a sequência de interesse, recorrendo-se à ferramenta *BLAST: Basic Local Alignment Search Tool* [NCBI, 2013].

2.2.3.2 Otimização das reações de PCR

As condições das reações de PCR foram otimizadas para cada fragmento, com o objetivo de se obter apenas amplificação da banda de DNA de interesse, sem o aparecimento de bandas inespecíficas. Na otimização ajustaram-se diferentes condições, como a temperatura de hibridação, a concentração de MgCl_2 e a concentração dos *primers*.

A temperatura de hibridação de um par de *primers* foi calculada com base na fórmula 2.1, em que A indica o número de adeninas na sequência do *primer*, C o número de citosinas, G o número de guaninas e T o número de timinas.

$$T_m = (A + T) \times 2 + (C + G) \times 4 \quad (2.1)$$

Este cálculo é efetuado para o *primer forward* e *reverse*, sendo no fim calculada a média

destes. Desta forma, o primeiro passo na otimização de uma reação de PCR é a alteração da temperatura de hibridação, sendo portanto o parâmetro que mais condiciona a reação. Se a temperatura de hibridação for muito elevada, a hibridação dos *primers* com o fragmento desejado é fraca e o rendimento do produto amplificado é muito baixo. Se por outro lado, a temperatura de hibridação for demasiado baixa, pode ocorrer hibridação inespecífica dos *primers*, resultando assim na amplificação de fragmentos de DNA adicionais. Posto isto, foi necessário aumentar a temperatura de hibridação, caso se observassem produtos inespecíficos, ou diminuir se se observasse pouca ou nenhuma amplificação.

A concentração de $MgCl_2$ também é extremamente importante para a reação de PCR, pois os iões Mg^{2+} funcionam como co-fator da enzima *Taq DNA polimerase*. Assim, uma baixa concentração de Mg^{2+} pode conduzir a baixos rendimentos da reação ou até mesmo à ausência de amplificação. Deste modo, foi sempre efetuada uma titulação de iões Mg^{2+} em solução, para determinar qual a concentração ótima para a amplificação de cada fragmento. De um modo geral, nos casos em que já se tinha aumentado a temperatura de hibridação mas que ainda existiam fragmentos inespecíficos, acabou por se optar pelo aumento da concentração de Mg^{2+} . No entanto, o aumento excessivo de $MgCl_2$ pode, por vezes, provocar o aparecimento de bandas adicionais, pois permite que a *Taq DNA polimerase* funcione de forma inespecífica.

A concentração de *primers* foi alterada quando se observou, em gel de agarose, que se formavam muitos dímeros de *primers*, ou seja, estes ligavam-se entre si, o que tornava a reação de PCR pouco eficiente. Nestes casos, optou-se pela diminuição da concentração de *primers*.

2.2.3.3 Controlo da eficiência da reação de PCR por eletroforese em gel de agarose

A eletroforese em gel de agarose permite a separação dos fragmentos de DNA consoante o seu peso molecular, através da aplicação de um campo elétrico. Deste modo, avaliou-se a eficiência das reações PCR realizando esta metodologia, observando assim se houve amplificação do fragmento de interesse de forma específica.

Para tal, a eletroforese foi efetuada utilizando um gel de agarose 2% (p/v) ou 1,2% (p/v), consoante o peso molecular do fragmento amplificado, em tampão TBE 1x (*National Diagnostics*). Durante a preparação dos géis, foi adicionado brometo de etídio (10 mg/ml, *Biotechnology*), de modo a permitir a visualização das bandas através da incidência de radiação ultravioleta, uma vez que este se intercala com o DNA.

No controlo da eficiência das otimizações das reações PCR, foi utilizado todo o volume da reação de PCR (12,5 μ l), ao qual se adicionou 3 μ l de *Orange G* 5x, sendo o volume total aplicado nos poços do gel de agarose. O *Orange G* é um tampão corado que confere densidade ao produto da reação PCR, tornando assim mais fácil a sua deposição no poço e a visualização da migração das amostras no gel.

Para o controlo da eficiência das reações de PCR das amostras, foi utilizado apenas 3 μ l do produto amplificado, o qual se adicionou a 7 μ l de *Orange G* 1x, tendo-se aplicado

o volume total nos poços do gel de agarose. Em ambas as situações, também se aplicou um marcador de peso molecular num dos poços do gel, o qual é essencial para estimar o tamanho dos fragmentos de DNA, permitindo assim confirmar que o DNA que foi amplificado é o pretendido. Assim, a eletroforese foi efetuada numa tina horizontal, sob voltagem constante de 140V durante pelo menos 30 minutos (este tempo varia com o tamanho dos fragmentos), sendo posteriormente o gel exposto a radiação ultravioleta num transiluminador (*BioDocAnalyze, Biometra*) com captação de imagem.

A preparação dos géis e de todas as soluções utilizadas nesta técnica encontra-se descrita no Anexo C.

2.2.4 Análise de mutações nos genes *KDELC1*, *ERCC5*, *TPP2* e nos genes candidatos obtidos pela análise de WES

O gene *KDELC1* foi analisado em 8 fragmentos (10 exões), enquanto para o gene *ERCC5* foi apenas analisada a região terminal (do exão 11 ao 15), em 5 fragmentos, tendo sido analisadas as sequências codificantes e regiões intrónicas adjacentes.

A análise mutacional do gene *TPP2* compreendeu apenas os exões 13 e 14 e todo o intrão 13. Ao nível do cDNA, foi efetuada ainda a análise da sequência entre os exões 13 e 17.

Em relação aos genes candidatos obtidos pela análise de WES, apenas foram analisadas as regiões que continham as mutações identificadas nos genes *CADM1*, *DHCR24*, *PLXNB1*, *RWDD4* e *RECQL5*.

2.2.4.1 Otimização da reação de PCR e amplificação dos fragmentos dos genes *KDELC1*, *ERCC5*, *TPP2* e nos genes candidatos obtidos pela análise de WES

As otimizações das reações de PCR foram efetuadas como descrito no ponto 2.2.3.2. No entanto, para o exão 12 do gene *ERCC5* ainda se verificou inespecificidade após se realizarem as alterações descritas anteriormente e utilizou-se dimetilsulfóxido (DMSO), um estabilizador de DNA que melhora a desnaturação deste. No entanto, este não contribuiu na eliminação da inespecificidade, tendo sido necessário recorrer-se ao *kit Amplitaq Gold (Applied Biosystems)*.

Para a otimização dos exões 1 e 2 do gene *KDELC1* foi necessário a utilização de uma solução de resolução GC comercial, pertencente ao *kit GC-RICH PCR System (Roche)*, uma vez que as suas sequências possuem um elevado conteúdo em GC, o que dificulta a amplificação específica.

Todas as condições utilizadas nas reações de PCR para amplificação dos vários fragmentos dos genes em estudo, assim como as sequências dos *primers* utilizados encontram-se descritas no Anexo B. Os *primers* que foram desenhados para análise mutacional da região do intrão 12, exões 13, 14 e 24 do gene *TPP2*, bem como os restantes *primers* desenhados para os genes obtidos pela análise WES não chegaram a ser utilizados, ficando para um estudo posterior.

2.2.4.2 Controlo da eficiência do produto de PCR por eletroforese em gel de agarose

O controlo da eficiência da amplificação dos fragmentos dos genes em estudo foi efetuado como referido no ponto 2.2.3.3.

Para o controlo dos fragmentos de maior peso molecular (exão 15 do gene *ERCC5*, fragmento 3/4 e 5/6 do fragmento *KDELC1* e exão 13/14 do gene *TPP2*) foi utilizado um gel de agarose 1,2% (p/v) e o marcador *1 kb DNA Ladder (Promega)* que permite identificar fragmentos de DNA entre os 250 pb e 10000 pb. Para o controlo dos restantes fragmentos, utilizou-se um gel de agarose 2% (p/v) e o marcador *GeneRuler 50 pb DNA Ladder (Thermo Scientific)*, que identifica fragmentos de DNA entre os 50 pb e 1000 pb.

A eletroforese em gel de agarose foi efetuada durante 30 minutos, com exceção dos fragmentos de maior peso molecular, em que esta foi estendida para 40 minutos.

2.2.5 Sequenciação automática

A sequenciação automática é um procedimento que se baseia na sequenciação de Sanger e que permite efetuar a caracterização integral de um fragmento de DNA, pela indicação exata da sequência dos seus nucleótidos constituintes. Neste método, para além dos dNTPs, utilizam-se também 2',3'-didesoxinucleótidos trifosfatos (ddNTPs) na síntese das cadeias de DNA. Estes últimos não apresentam o grupo hidroxilo no carbono 3', formando assim cadeias truncadas, uma vez que não conseguem efetuar a ligação fosfodiéster com outro nucleótido trifosfatado. Deste modo formam-se cadeias que diferem entre si de um nucleótido, permitindo a sua separação em eletroforese capilar, devido ao peso molecular diferente.

Para além disso, os diferentes ddNTPs encontram-se marcados com fluorocromos diferentes, permitindo a distinção das quatro bases azotadas (adenina, guanina, citosina e timina). Assim, durante a eletroforese capilar, o sinal fluorescente de cada ddNTP é detetado, permitindo identificar a sequência da cadeia de DNA.

2.2.5.1 Purificação dos produtos amplificados por PCR

Para a sequenciação automática é necessário efetuar uma purificação dos produtos amplificados por PCR de modo a eliminar resíduos que possam interferir com esta metodologia, nomeadamente, os nucleótidos que não foram incorporados e os *primers* que não hibridaram.

Para a purificação dos produtos de PCR foram utilizados dois métodos distintos: um enzimático, em que se utilizou diretamente o produto amplificado em solução e outro em que se efetuou a excisão da banda de DNA de interesse a partir do gel de agarose. O primeiro método foi utilizado quando foi observado um produto de PCR específico, tendo-se adicionado 0,5 µl da enzima *Exonuclease I (Exo I)* (20 U/µl *Thermo Scientific*) e 1 µl da enzima *Thermosensitive Alkaline Phosphatase (FastAP)* (1 U/µl *Thermo Scientific*) ao produto de PCR. Depois incubaram-se as amostras num termociclador durante 15 minutos a 37°C, seguido de 15 minutos a 85°C para inativação das enzimas.

Nos casos em que não se conseguiu obter apenas uma banda específica, apesar de terem sido testadas várias condições de modo a otimizar a reação de PCR, efetuou-se a excisão da banda de interesse a partir do gel de agarose, seguido de purificação. Também se utilizou este procedimento quanto se efetuou a análise ao nível do cDNA do gene *TPP2*, visto serem observadas duas isoformas e o objetivo ser a sequenciação de ambas em separado. Deste modo, após o controlo da eficiência da reação de PCR, excisou-se a banda de interesse com recurso a um bisturi e procedeu-se à purificação do produto amplificado, utilizando o *kit* de purificação de DNA *QIAquick Gel Extraction (Qiagen)*. Este *kit* baseia-se num método de ligação/lavagem/eluição, em que o DNA adsorve especificamente a uma membrana de sílica contida numa coluna, os contaminantes da reação de PCR e a agarose são removidos e por fim o DNA é eluído numa solução tampão (Anexo D). Após a eluição, o rendimento da purificação foi avaliado através de uma eletroforese em gel de agarose 2% (p/v), como descrito no ponto 2.2.3.3, sendo o restante produto purificado a -20°C.

2.2.5.2 Reação de sequenciação

Seguiu-se o protocolo *ABI PRISM BigDye[®] Terminator v1.1 Cycle Sequencing Ready Reaction Kit (Applied Biosystems)* para a realização das reações de sequenciação. Para cada amostra fez-se uma mistura reacional com os seguintes elementos: 2 µl de *primer forward* ou *reverse* (1,6 pmol/µl); 2 µl de solução tamponada *Buffer Sequencing 5x, BigDye[®] Terminator v1.1 (Applied Biosystems)*; quantidade variável de produto de PCR purificado, a qual depende do rendimento da reação de PCR observado em gel de agarose (como descrito em 2.2.3.3) e do tamanho do fragmento a analisar; mistura reacional de sequenciação *BigDye[®] Terminator v1.1 (Applied Biosystems)* num volume que depende do tamanho do fragmento a sequenciar, e por fim água bidestilada para perfazer o volume final de 20 µl.

De seguida, colocaram-se os tubos contendo as misturas reacionais num termociclador *Veriti (Applied Biosystems)* ou *Trio (Biometra)* (dependendo dos fragmentos para os quais se está a efetuar a reação de sequenciação, ver Anexo B), utilizando o programa referido na Tabela 2.1. A temperatura de hibridação é igual à utilizada na reação de PCR, exceto quando esta é superior a 60°C, sendo nesse caso utilizada a temperatura de 59,5°C, uma vez que neste programa a extensão ocorre a 60°C.

Tabela 2.1: Programa utilizado na reação de sequenciação dos fragmentos de cada gene estudado

Etapa	Temperatura (°C)	Tempo	Ciclos
Desnaturação inicial	96	5 min	1
Desnaturação	95	10 seg	25
Hibridação	Variável	5 seg	
Extensão	60	4 min	
Pausa	4	∞	-

2.2.5.3 Precipitação e purificação do DNA após reação de sequenciação

A precipitação e purificação do DNA é realizada após a reação de sequenciação, de modo a obtermos o DNA purificado para a eletroforese capilar e evitar contaminantes que possam causar algumas interferências. Assim, efetuou-se uma precipitação utilizando Etanol/EDTA/Acetato de Sódio, recomendado pelo protocolo do *kit* de sequenciação *BigDye® Terminator v1.1 Cycle Sequencing* (Applied Biosystems).

Este protocolo consiste na precipitação e purificação do DNA, utilizando três reagentes: o etanol absoluto e o acetato de sódio, que neutralizam as cargas dos ácidos nucleicos e reduzem a sua solubilidade, promovendo assim a precipitação do DNA; e o EDTA, que é um quelante dos íons magnésio, pelo que inibe a enzima utilizada na reação de sequenciação. Ainda é utilizado etanol 70% (v/v) para lavar o *pellet* de DNA, sendo este posteriormente seco a 37°C. Estes reagentes permitem assim a remoção de contaminantes e a formação de um *pellet* de DNA purificado (Anexo E).

Caso não se prossiga de imediato com a eletroforese capilar, as amostras são armazenadas a 4°C.

2.2.5.4 Preparação dos produtos da reação de sequenciação para eletroforese capilar

Os *pellets* secos de DNA foram ressuspensos em 17 µl de formamida Hi-Di (Applied Biosystems), tendo-se homogeneizado em *vortex* e transferido para tubos ou para uma placa de 96 poços, apropriados para os sequenciadores automáticos *ABI Prism™ 310 Genetic Analyzer* (Applied Biosystems) e *ABI Prism™ 3130 Genetic Analyzer* (Applied Biosystems), respetivamente. Posteriormente efetuou-se uma desnaturação das amostras a 95°C, durante pelo menos 5 minutos num termociclador *Biometra (Alfagene)*, seguido de incubação em gelo, para evitar a renaturação das cadeias de DNA das amostras. Nos casos em que se utilizou uma placa de 96 poços, esta foi de seguida centrifugada a 1200 rpm durante 2 minutos. As amostras foram colocadas nos sequenciadores automáticos, onde ocorreu a eletroforese capilar a 50°C, a uma tensão de 15 KV.

No sequenciador automático *ABI Prism™ 3130 Genetic Analyzer* (Applied Biosystems), para os fragmentos de maior peso molecular foi necessário um tempo de eletroforese capilar mais alargado, de aproximadamente 46 minutos para cada amostra, enquanto para os fragmentos menores apenas são necessários 20 minutos. No sequenciador automático *ABI Prism™ 310 Genetic Analyzer* (Applied Biosystems), apenas foram colocados os fragmentos de menor peso molecular, com tempos de eletroforese capilar variáveis entre 20 a 26 minutos.

2.2.5.5 Análise de resultados

As sequências de DNA das amostras foram adquiridas utilizando o programa *Sequencing Analysis 3.4.1* (Applied Biosystems), obtendo-se eletroforetogramas, que posteriormente foram comparados com as sequências de referência de cada gene em estudo. As sequências de referência foram retiradas da base de dados do NCBI.

2.2.5.6 Análise *in silico* das variantes identificadas por sequenciação automática

Foi efetuada a análise *in silico* das variantes detetadas nos genes em estudo, de forma a tentar avaliar a sua eventual patogenicidade, tendo-se para isso recorrido a ferramentas *online*, que prevêm alterações a vários níveis. Para este estudo foram utilizadas duas ferramentas, *Mutation Taster* [Schwarz et al., 2010] e *Variant Effect Predictor (VEP)* [McLaren et al., 2010].

A ferramenta VEP foi utilizada para rapidamente e com precisão se obter a previsão dos efeitos das variantes introduzidas em genes, transcritos, sequências das proteínas e regiões reguladoras. Desta forma, esta ferramenta fornece-nos várias informações, tais como, previsões do *Sorting Intolerant From Tolerant (SIFT)* [Ng and Henikoff, 2003] (prevê se uma substituição de aminoácido irá afetar a função da proteína, baseado na homologia da sequência e propriedades físicas dos aminoácidos) e *Polyphen* [Adzhubei et al., 2010] (prevê o possível impacto de uma substituição de um aminoácido na estrutura e função da proteína, usando considerações comparativas e físicas), indica variantes co-localizadas, publicações que citam a variante e relata dados de *minor allele frequency (MAF)*.

O *Mutation Taster* avalia o potencial que alterações na sequência de DNA têm de causarem doença, incluindo várias análises (conservação evolutiva, alterações de locais de *splicing* ou a fatores de transcrição, perda de funções da proteína, alterações do local de modificação de histonas e alterações que podem afetar a expressão do mRNA).

2.2.6 Methylation-Specific Multiplex ligation-dependent probe amplification

O *Methylation-Specific Multiplex ligation-dependent probe amplification (MS-MLPA)* [Nygren, 2005] é uma variante da técnica de MLPA que combina a determinação de alterações de *copy-number* com o estado de metilação de genes específicos. Nesta reação não é a amostra de DNA que é amplificada durante o PCR, mas sim as sondas MLPA que hibridam com o DNA da amostra. Como apenas é utilizado um único par de *primers* para amplificar todas as sondas este método torna-se muito robusto e pode ser efetuado em *multiplex*.

As etapas fundamentais da técnica de MS-MLPA são: 1) Desnaturação do DNA; 2) Hibridização das sondas com o DNA; 3) Reação de ligação das sondas/Digestão enzimática pela enzima *HhaI*; 4) Reação de PCR (amplificação das sondas) (Figura 2.2).

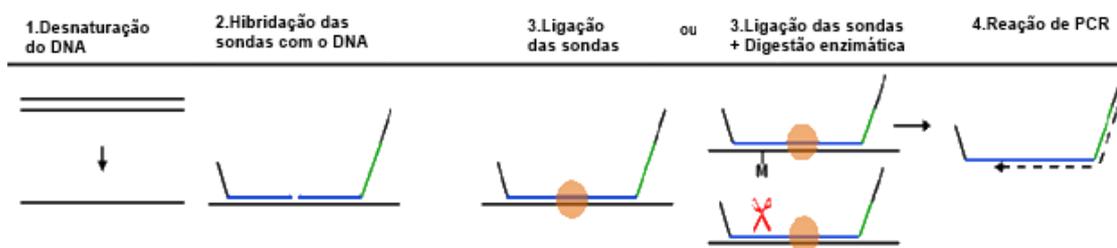


Figura 2.2: Esquema da reação de MS-MLPA (adaptado de *MS-MLPA protocol version MSP-v004*)

Inicialmente é realizada a desnaturação do DNA alvo com posterior hibridização das duas hemi-sondas na sequência de DNA específica. De seguida, a reação é separada em duas partes, de modo a se obter informação sobre alterações de *copy-number* e estado de metilação. Nas amostras utilizadas para a deteção de alterações de *copy-number* é adicionada uma ligase termoestável que liga as hemi-sondas, enquanto nas amostras utilizadas para a deteção do padrão de metilação é adicionada para além desta ligase, a endonuclease *HhaI*. As sondas são desenhadas de forma a que a sequência de DNA com a qual hibridam tenha um local de restrição para esta enzima. Deste modo, a endonuclease apenas cliva as sondas que hibridam com sequências que não se encontram metiladas, permitindo a amplificação por PCR das sondas que hibridam com sequências metiladas. Posteriormente, os produtos da amplificação são separados por eletroforese capilar, uma vez que cada fragmento tem um peso molecular único, e analisados com auxílio de *softwares* computacionais.

A reação de MS-MLPA foi realizada utilizando dois protocolos distintos, um para analisar o gene *APC* (*Two-tube protocol*) e outro para os genes *MMR* e *MGMT* (*One-tube protocol*) (Anexo F). Todos os reagentes utilizados na realização desta metodologia pertencem ao *SALSA MLPA reagent kit (MRC-Holland)*, à exceção da enzima de restrição *HhaI* (10 U/ μ l *Promega*). Para a análise dos genes *MMR* e *MGMT* foi utilizado o *kit* comercial *SALSA MS-MLPA probemix ME011-B2 Mismatch Repair genes (MMR) (MRC-Holland)*, o qual contém 26 sondas distribuídas pelas regiões promotoras dos genes *MSH2*, *MSH6*, *MLH1*, *MSH3*, *PMS2*, *MGMT* e *MLH3* e 13 sondas referência. Para a análise do gene *APC* foi utilizado um *kit custom made* desenhado pelo grupo de gastroenterologia da UIPM (Figura 2.3), o qual contém 19 sondas, das quais 8 se localizam no promotor N, 1A e 1B, nos exões 1, 2, 13 e 14 do gene *APC* e 11 são sondas referência localizadas noutras regiões cromossómicas essenciais para a análise e normalização dos resultados (estas sondas fazem parte do *kit* P200-A1 (*MRC-Holland*)).

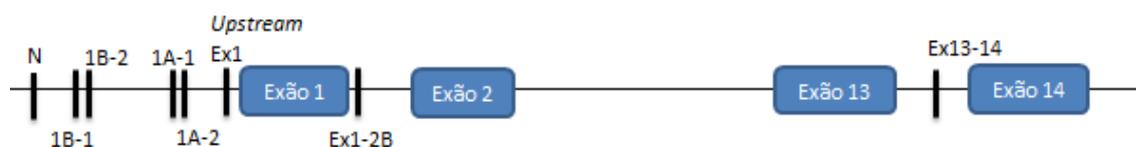


Figura 2.3: Esquema da localização das sondas MS-MLPA no gene *APC*

Foi utilizada uma quantidade de aproximadamente 270 ng das amostras de DNA obtido a partir de sangue periférico e aproximadamente 450 ou 225 ng de DNA das amostras de tecido incluído em parafina, de forma a assegurar a presença de DNA suficiente para a análise. Todos os passos da técnica foram executados com auxílio de um termociclador *Veriti (Applied Biosystems)* com tampa aquecida a 105°C. Após a reação de PCR, o produto amplificado foi controlado num gel de agarose 2% (p/v) como descrito no ponto 2.2.3.3.

2.2.6.1 Preparação das amostras para eletroforese capilar

Para a preparação das amostras amplificadas por PCR para eletroforese capilar, adicionou-se 1 µl de produto amplificado, 0,2 µl de marcador interno de peso molecular *GeneScanTM 500 LIZ[®] Size Standard* (*Applied Biosystems*) e 9 µl de formamida Hi-Di (*Applied Biosystems*). Em alguns casos foi necessário diluir o produto amplificado, uma vez que o sequenciador tem um valor máximo de detecção de fluorescência, o qual se for ultrapassado não permite a obtenção de valores fidedignos para a análise posterior.

A eletroforese capilar foi realizada num sequenciador automático *ABI PrismTM 3130 Genetic Analyzer* (*Applied Biosystems*) e posteriormente analisada com auxílio do *software GeneMapper v4.1* (*Applied Biosystems*), que permite a identificação e visualização dos eletroferogramas. Em cada eletroforese capilar para além das amostras em estudo, foram utilizadas três amostras referência, as quais, apresentam para as regiões de interesse valores de *copy-number* e padrão de metilação normais.

2.2.6.2 Análise de resultados

A análise dos resultados foi realizada recorrendo ao *software Coffalyser.Net* (*MRC-Holland*) [Coffa and van den Berg, 2011], o qual efetua a normalização entre as amostras em análise e as amostras referência, assim como, entre as sondas dos genes de interesse e as sondas referência para cada amostra, de modo a obtermos resultados relativos ao padrão de metilação e de *copy-number* de cada amostra.

2.2.7 Real-time PCR

O *real-time* PCR é uma técnica que possibilita a quantificação em tempo real do DNA amplificado. A sua metodologia baseia-se em técnicas de fluorescência, através da inclusão na reação de uma molécula fluorescente, que produz um sinal de fluorescência diretamente proporcional à quantidade de produto amplificado.

O seu procedimento segue o princípio geral do PCR convencional, apresentando também três fases características (Figura 2.4): a fase de crescimento exponencial, fase de crescimento linear e fase estacionária. Para cada amostra é determinado o seu *threshold cycle* (CT), que corresponde ao ponto a partir do qual a fluorescência detetada ultrapassa o *threshold* (limiar da fase exponencial), sendo este definido pelo *software* do equipamento em função da *baseline* (limiar mínimo de detecção de fluorescência do instrumento).

Os métodos químicos de fluorescência utilizados no *real-time* PCR podem ser divididos em dois grupos: corantes intercalantes e sondas de sequência específica. Os corantes intercalantes (exemplo, *SYBR[®] Green*) são fluorocromos que se intercalam na dupla cadeia dos produtos do PCR permitindo desta forma a sua detecção. Estes corantes não apresentam especificidade para uma sequência de DNA alvo, no entanto o correto desenho dos *primers* confere a este método uma elevada especificidade para a detecção e quantificação das sequências alvo. As sondas de sequência específica (exemplo, *Taqman[®]*),

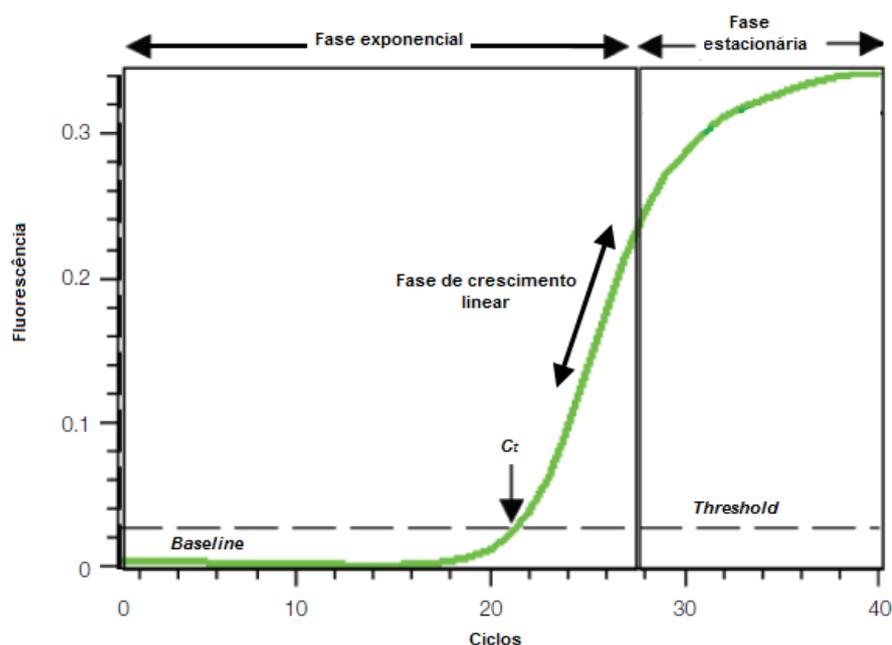


Figura 2.4: Representação gráfica das fases do *real-time* PCR (adaptado de *Real-Time PCR Applications Guide*, Bio-Rad Laboratories)

tal como o nome indica, têm especificidade para uma dada sequência de DNA e detetam-na em todos os produtos de PCR. A quantificação desta sequência é efetuada através de um fluorocromo que está acoplado na extremidade 5' destas sondas.

A técnica de *real-time* PCR permite que possam ser realizados dois tipos de quantificação do DNA: uma quantificação absoluta (método da curva-padrão) ou quantificação relativa. Na quantificação absoluta, é utilizada uma amostra de concentração conhecida que expresse o gene de interesse, para a construção de uma curva padrão (relação linear entre o C_t e o logaritmo da concentração do DNA alvo). Desta forma determina-se a concentração inicial de uma dada amostra de concentração desconhecida a partir desta curva padrão, ou seja, o valor de C_t da amostra de concentração desconhecida é projetado no gráfico que contém a curva padrão, a partir da qual se extrapola a concentração de DNA da amostra em questão.

Na quantificação relativa não se recorre a curvas padrão, sendo a expressão do gene alvo, em todas as amostras em estudo, expressa em relação a uma amostra de referência. Assim, este método consiste na comparação dos valores C_t das amostras do gene alvo, com os valores da amostra de referência. É necessário amplificar em paralelo, para cada amostra, um controlo endógeno que seja expresso de forma ubíqua em todas as células (exemplo, gene *housekeeping*). Desta forma, os valores de C_t de ambos (amostra e referência) são normalizados em relação a este controlo endógeno. Neste trabalho foi apenas efetuada a quantificação relativa.

2.2.7.1 Síntese de cDNA por reação de transcrição reversa do RNA

A reação de transcrição reversa consiste na conversão do RNA em cDNA, tendo sido utilizados *primers random* que se ligam às várias cadeias de mRNAs e uma enzima, a transcriptase reversa, que converte o RNA em cDNA. Posteriormente, a cadeia dupla do cDNA é amplificada numa reação de PCR.

Em cada reação foi utilizado um volume variável de RNA de cada amostra (correspondente a de 1 µg), 0,5 µl de *primers random* (3 µg/µl, Roche) e água tratada com DEPC (Merck) até perfazer o volume de 7,75 µl. De seguida, a mistura reacional foi incubada num termociclador Veriti (Applied Biosystems) a 70°C durante 10 minutos, para os *primers random* hibridarem, e colocada posteriormente em gelo.

Em seguida, foi adicionada uma mistura reacional com um volume final de 12,25 µl composta pelos seguintes reagentes: 4 µl de tampão *5x First-Strand Buffer* (Invitrogen); 4 µl de dNTPs (25 mM, Illustra™, GE Healthcare); 0,75 µl de enzima *RNaseOUT Recombinant Ribonuclease Inhibitor* (40 U/µl, Invitrogen), que inibe a ação das ribonucleases; 2 µl de ditiotreitol (DTT) (0,1 M, Invitrogen), que é um estabilizador enzimático e proteico; 1 µl de enzima transcriptase reversa *Superscript II RT* (200 U/µl, Invitrogen); 0,2 µl de *single strand DNA binding protein* (SSB) (4,47 µg/ml, Promega) que tem a função de se ligar e estabilizar o DNA de cadeia simples, evitando a sua hibridação prematura; e por fim água tratada com DEPC (Merck) até perfazer o volume final de 20 µl.

Os tubos foram de novo colocados no termociclador Veriti (Applied Biosystems), tendo-se continuado o programa descrito na Tabela 2.2. Os cDNAs obtidos foram armazenados a -20°C.

Tabela 2.2: Programa utilizado na síntese de cDNA

Etapa	Temperatura (°C)	Tempo
Desnaturação inicial	70	10 min
Pausa	15	∞
Reação de RT-PCR	42	60 min
	70	15 min
Pausa	15	∞

Posteriormente realizou-se a amplificação do cDNA por PCR dos exões 13 ao 17 do gene *TPP2*, utilizando *primers* exónicos, de forma a confirmar que a síntese foi eficiente. O controlo da eficiência do produto amplificado foi efetuado em gel de agarose 2% (p/v), como descrito em 2.2.3.3 e após confirmação da amplificação específica, o cDNA foi utilizado nas reações de *real-time* PCR.

2.2.7.2 SYBR® Green

Utilizou-se o método do SYBR® Green para quantificar os níveis de expressão de dois transcritos do gene *TPP2*, um que inclui os exões 13 e 13A e outro que inclui os exões 13 e 14. Para cada transcrito foi desenhado um par de *primers* exónicos, cuja sequência se

encontra descrita no Anexo G. O gene *SDHA* foi utilizado como gene *housekeeping*.

Numa placa de 96 poços (*LightCycler® Plus 480 Multiwell Plates (Roche)*) prepararam-se triplicados de cada amostra com a seguinte reação: 7,5 µl de *LightCycler® 480 SYBR Green I Master (2x, Roche)*, 0,75 µl de *primer forward* e *reverse* (5 pmol/µl), 2 µl de cDNA diluído em água bidestilada (1:21 ~ 0,048 µg/µl) e 4 µl de água. O *real-time* PCR foi realizado no *LightCycler® 480 Instrument (Roche)* e para a análise dos resultados utilizou-se o *software LightCycler® 480 SW 1.5 (Roche)*. O programa utilizado encontra-se descrito no Anexo G.

Como a eficiência das reações não era igual entre si, foi necessário calculá-la com base na equação 2.2. O declive foi obtido através da elaboração de uma reta de calibração, ou seja, uma relação linear entre o C_t de uma amostra e o logaritmo da concentração de cDNA dessa mesma amostra (a concentração é desconhecida, por isso utilizaram-se valores fictícios). Esta curva foi construída a partir de uma diluição seriada da amostra, que nos atribuiu diferentes valores de C_t , correspondentes a cada diluição. Posteriormente, com a eficiência de cada par de *primers* foi possível quantificar as amostras a estudo utilizando a fórmula 2.3, onde $E_{\text{gene alvo}}$ é a eficiência do gene alvo, neste caso, o transcrito que se pretende estudar; $E_{\text{gene referência}}$ corresponde à eficiência do gene *SDHA* e ΔC_t equivale à subtração da média dos C_t da amostra controlo com a média dos C_t da amostra a estudo, para os transcritos/genes correspondentes.

$$E = 10^{(-1/\text{declive})} \quad (2.2)$$

$$\text{ratio} = \frac{(E_{\text{gene alvo}})^{\Delta C_t(\text{média controlo} - \text{média amostra})}}{(E_{\text{gene referência}})^{\Delta C_t(\text{média controlo} - \text{média amostra})}} \quad (2.3)$$

2.2.7.3 *Taqman*®

Para a quantificação dos níveis de expressão do gene *APC* utilizou-se o método do *Taqman*®, sendo o gene β -actina utilizado como gene *housekeeping*. As sondas para o gene *APC* contêm na sua extremidade 5' o fluoróforo FAM e as sondas do gene β -actina o fluoróforo VIC.

As reações de *real-time* PCR foram realizadas em triplicados numa placa de 96 poços (*ABI PRISM™ Optical 96-Well Reaction Plate*) e preparadas para um volume final de 20 µl, apresentando a seguinte composição: 10 µl de *Taqman® Universal PCR Master Mix (2x, Applied Biosystems)*, 10 µl de *Taqman® Gene Expression Assay (20x, Applied Biosystems)*, 1,5 µl de cDNA diluído em água bidestilada (1:5 ~ 0,2 µg/µl) e 7,5 µl de água. O *real-time* PCR foi realizado no *7900HT Sequence Detection System (Applied Biosystems)* e para a análise dos resultados utilizou-se o *software SDS 2.4 (Applied Biosystems)*.

A eficiência das reações não foi calculada, porque como apenas foram utilizados reagentes comerciais, a eficiência é elevada e uniforme entre o gene em estudo e o gene *housekeeping*. Desta forma, utilizou-se o método $2^{-\Delta\Delta C_t}$ para a quantificação das amostras. Este método consiste na comparação dos valores C_t das amostras com a média dos C_t

de duas amostras de referência. Os valores de C_t de ambos (amostra e referência) foram normalizados ao gene *housekeeping* β -actina. Na equação 2.4 utilizada, o ΔC_t da amostra é o valor de C_t para qualquer amostra normalizada (ao gene *housekeeping*) e o ΔC_t referência é o valor de C_t para as amostras de referência normalizadas (ao gene *housekeeping*).

$$\begin{aligned} ratio &= 2^{-\Delta\Delta C_t} \\ &= 2^{-(\Delta C_{t\text{amostra}} - \Delta C_{t\text{referência}})} \end{aligned} \quad (2.4)$$



Resultados

3.1 Caracterização molecular dos tumores de indivíduos com FCCTX

Em estudos anteriores o grupo de Gastrenterologia da UIPM, IPOLFG, identificou 2 entidades moleculares distintas entre as famílias FCCTX (ver ponto 1.5). Uma mais prevalente (TSG+), em que os tumores apresentam frequente LOH em diversos TSG, mutações somáticas no *KRAS* e *APC* e metilação dos genes *MMR* e *MGMT*. Na outra entidade (TSG-), menos prevalente, os tumores não apresentam evidência de LOH nos TSG e apresentam menos frequentemente as mutações somáticas e metilação acima referidas. É importante continuar a caracterização destes tumores, sendo este um dos objetivos deste presente trabalho. Desta forma, efetuou-se a análise de metilação num número mais alargado de amostras de tumor, incluiu-se também a mucosa normal respetiva sempre que possível, e avaliaram-se as alterações de *copy-number*.

3.1.1 Genes *MMR* e *MGMT*

Nas amostras de tumor e mucosa normal das famílias TSG+ verificaram-se ganhos frequentes para qualquer gene, sendo mais frequentes para as sondas *upstream* dos genes em estudo. Os ganhos foram detetados mais frequentemente no grupo de tumores TSG+ do que no grupo TSG- (14/16 (88%) vs 1/6 (17%), $P < 0,01$).

Os ganhos nas famílias TSG- foram de facto detetados com menor frequência, observando-se que no caso dos genes *MMR*, apenas a amostra 1541 apresentava ganhos, à exceção de um ganho no gene *MLH1* na amostra 996. No caso do gene *MGMT* detetaram-se ganhos em 4/6 (67%) e 5/6 (83%) das amostras de tumor (nos intrões 1 e 2, respetivamente). É de

salientar a detecção de ganhos frequentes em 6/8 (75%) amostras de mucosa normal nas famílias TSG+, não se observando estes ganhos nas famílias TSG-.

As perdas foram detetadas com maior frequência no grupo de famílias TSG- em comparação com o grupo de famílias TSG+ (3/6 (50%) vs 3/16 (19%) $P=0,28$). Também foram notórias as perdas detetadas na região *upstream* e exão 1 do gene *MSH6*, essencialmente nas famílias TSG- em 4/6 (67%) amostras de tumor para as duas sondas, e em 5/16 (31%) e 3/16 (19%), respetivamente, nas amostras de tumor das famílias TSG+.

Em algumas amostras foram detetadas alterações de *copy-number* em várias sondas de diferentes genes. Foram verificadas alterações em pelo menos 3 sondas dos genes *MMR* ou *MGMT*, onde se obteve um aumento significativo de ganhos em 12/16 (75%) amostras de tumor de famílias TSG+ (maioritariamente nas famílias 24 e 55) em comparação com o grupo de famílias TSG- (1/6 (17%), $P=0,02$). Em relação às perdas verifica-se o contrário, nas famílias TSG- detetaram-se mais perdas que no grupo de famílias TSG+ (3/6 (50%) vs 2/16 (13%).

Em relação ao estado de metilação nos genes *MMR* e *MGMT*, apenas se observa hipermetilação ou um aumento significativo da metilação em relação ao apresentado pelas amostras de referência, nas amostras de tumor ou mucosa normal de indivíduos FCCTX pertencentes ao grupo TSG+ (Tabela 3.3). Não foram detetadas alterações de metilação apenas para três famílias do grupo TSG+ (L71, L84 e L236). Assim, foi detetado um aumento significativo da metilação em 10/16 (63%) amostras de tumor do grupo TSG+, não tendo sido detetado em nenhum caso do grupo TSG- (0/6) ($P<0,01$). Quando se observa apenas as alterações de metilação nos genes *MMR*, estas foram detetadas em 7/16 (44%) amostras de tumor de famílias TSG+, apenas nas famílias L7, L55, L63 e L173.

Em algumas amostras foram detetadas alterações de metilação em várias sondas para diferentes genes. Assim, verificou-se hipermetilação em pelo menos 3 sondas dos genes *MMR* e *MGMT* em 5/16 (31%) amostras de tumor do grupo TSG+ (famílias L55, L63 e L173). No caso das famílias TSG- não se observou metilação em pelo menos 3 sondas em nenhum dos tumores (0/6).

Tendo em conta a elevada frequência de ganhos e perdas nos genes *MMR* e *MGMT* (principalmente de ganhos), não só nas amostras de tumor mas também nas amostras de mucosa normal, colocou-se a hipótese de haver alguma alteração ao nível germinal, nos indivíduos pertencentes a estas famílias FCCTX, que pudesse contribuir para a suscetibilidade para este síndrome. De facto, apesar de ter sido efetuada análise de mutações germinais nos genes *MMR* para estes indivíduos e de não se ter detetado nenhuma alteração, esta análise baseia-se na pesquisa de mutações na região codificante e na análise de grandes deleções, principalmente ao nível da sequência codificante. Adicionalmente, o gene *MGMT* também não foi analisado. O facto dos tumores FCCTX serem MSS, não invalida que uma pequena alteração nestes genes, nomeadamente na região *upstream*, não possa contribuir para o fenótipo sem causar um fenótipo MSI-H. Deste modo, foi efetuada a análise de *copy-number* e de metilação tanto para os genes *MMR* como para o *MGMT* na linha germinal destes indivíduos com FCCTX.

Nesta análise apenas se observaram duas deleções, uma *upstream* ao gene *MGMT*, na amostra L108, e outra no exão 1 do gene *MLH3* na amostra L926. Será necessário confirmar estas deleções por outra técnica, nomeadamente por análise de *copy-number* utilizando *primers* específicos, visto haver a possibilidade de serem devidas a algum polimorfismo em heterozigotia que se localize no local de ligação da sonda. A alteração no gene *MLH3* pode ser uma deleção, mas é menos provável, visto terem sido analisados vários elementos da mesma família (L55) e só a amostra L926 apresentar perda. Existem contudo alguns ganhos, apesar de não significativos, mas na sonda *upstream* do gene *MSH6* estes valores são de averiguar e não se podem excluir por serem elevados. De qualquer forma, como se observam estes valores em 4 indivíduos, esta alteração em princípio deve ter relação com a técnica e não corresponder de facto a um ganho.

Verificou-se ainda a existência de duas amostras com hipermetilação no gene *MGMT*, uma na sonda *upstream* ao gene (amostra L48 – família L24) e outra na sonda que se liga ao intrão 1 (amostra L108 – família L37), o que a confirmar-se deverá merecer investigação adicional.

Tabela 3.1: Análise de *copy-number* aos genes MMR e MGMT nas amostras de tumores e mucosas normais

Familia	DNA	Histologia	Copy-number												MLH1?																																																																																																																																																																																																																
			MSH2			MSH6			MLH1			MSH3				PMS2			MGMT																																																																																																																																																																																																												
			Up	Ex1	Ex7	Up	Ex1	Ex7	Up	Ex1	Ex7	Up	Ex1	Ex7	Up	Ex1	Ex7	Up	Ex1	Ex7	Up	Ex1	Ex7	Int1	Int2	Int3	Int4	Int5	Int6	Int7	Int8	Int9	Int10	Int11	Int12	Int13	Int14	Int15	Int16	Int17	Int18	Int19	Int20	Int21	Int22	Int23	Int24	Int25	Int26	Int27	Int28	Int29	Int30	Int31	Int32	Int33	Int34	Int35	Int36	Int37	Int38	Int39	Int40	Int41	Int42	Int43	Int44	Int45	Int46	Int47	Int48	Int49	Int50	Int51	Int52	Int53	Int54	Int55	Int56	Int57	Int58	Int59	Int60	Int61	Int62	Int63	Int64	Int65	Int66	Int67	Int68	Int69	Int70	Int71	Int72	Int73	Int74	Int75	Int76	Int77	Int78	Int79	Int80	Int81	Int82	Int83	Int84	Int85	Int86	Int87	Int88	Int89	Int90	Int91	Int92	Int93	Int94	Int95	Int96	Int97	Int98	Int99	Int100	Int101	Int102	Int103	Int104	Int105	Int106	Int107	Int108	Int109	Int110	Int111	Int112	Int113	Int114	Int115	Int116	Int117	Int118	Int119	Int120	Int121	Int122	Int123	Int124	Int125	Int126	Int127	Int128	Int129	Int130	Int131	Int132	Int133	Int134	Int135	Int136	Int137	Int138	Int139	Int140	Int141	Int142	Int143	Int144	Int145	Int146	Int147	Int148	Int149	Int150	Int151	Int152	Int153	Int154	Int155	Int156	Int157	Int158	Int159	Int160	Int161	Int162	Int163	Int164	Int165	Int166	Int167	Int168	Int169	Int170	Int171	Int172	Int173	Int174	Int175	Int176	Int177	Int178	Int179	Int180	Int181	Int182	Int183	Int184	Int185	Int186	Int187	Int188	Int189	Int190	Int191	Int192	Int193	Int194	Int195	Int196	Int197	Int198	Int199	Int200
7	437	Ca	1.8	2.82	2	1.3	4.82	1.63	1.64	0.76	1.12	1.36	1.19	1	2.2	1.26	1.3	1.78	2.4	2.42	3.14	1.81	1.1	1.7	1.7	1.2	2.1																																																																																																																																																																																																				
	438	MC	0.88	1.88	1.2	1.4	0.91	0.91	1.14	0.99	1.05	1.43	1.37	1.1	1.3	0.36	1.1	1.05	1.2	1.6	1.44	1.48	1.4	1.1	1.5	1.4	0.36																																																																																																																																																																																																				
	459	ATDAG	1.24	1.57	1.3	0.8	1.66	1.03	1.04	0.88	0.84	1.16	0.94	0.8	1.4	1.04	1	1.16	1.4	1.6	1.05	1.41	1.38	1.1	1.3	1.5	0.5	1.28																																																																																																																																																																																																			
	460	MC	1.33	1.68	1.5	0.9	1.77	1.1	1.1	1.01	0.99	1.36	1.06	1	1.6	1.13	1.1	1.08	1.2	1.5	1.09	1.49	1.41	1.3	1.6	0.8	1.39																																																																																																																																																																																																				
24	947	ATDBG	1.88	2.22	1.6	1	3.98	1.64	1.38	0.93	1.09	1.52	1.42	0.9	1.3	0.37	1.1	1.38	1.7	1.9	2.05	2.34	1.74	1.1	1.7	1.6	0.9	1.75																																																																																																																																																																																																			
	991	AVDAG	1.36	2.03	2	1.1	0.31	1	1.46	1.78	1.97	1.5	2.79	2.1	1.8	1.53	2.3	0.96	1.6	2.1	1.15	1.35	0.8	0.7	1	0.8	1.8																																																																																																																																																																																																				
	992	MC	3.4	1.74	2	1.4	3.53	1.9	2.06	2.48	3.19	2.76	3.05	2.9	3.3	1.38	1.8	3.16	2.4	2.8	1.59	1.48	1.13	1	1.5	1.6	1	1.67																																																																																																																																																																																																			
	993	ATDBG	0.78	1.8	1.4	1.1	0.31	0.64	1.03	1.19	1.3	1.64	1.74	1.7	1.6	1.18	1.4	1.19	1.2	1.3	0.8	0.7	0.86	1	0.7	1.1	0.9	1.1																																																																																																																																																																																																			
55	1775	ATDBG	2.38	4.95	2.9	1.6	4.11	2.28	1.71	1.7	2.18	2.62	2.8	2.5	4.1	1.76	2.2	2.84	2.8	3.8	1.24	2.29	1.1	0.8	1.1	1.3	2.28																																																																																																																																																																																																				
	1777	ATDBG	0.76	0.99	0.9	1.6	0.46	0.82	0.98	0.95	0.93	1	1.06	1.1	0.7	1.2	1.1	1.04	1.1	0.76	0.79	0.69	1.1	0.8	0.8	1.9	0.35																																																																																																																																																																																																				
	1777	ATDBG	1.19	1.33	1.4	2.1	0.78	0.77	0.67	0.86	1.41	1.84	2.12	1.7	1.4	1.31	0.8	1.68	1.3	2.3	0.87	1.13	0.81	1.3	0.9	1.3	1.3	0.67																																																																																																																																																																																																			
	1871	Ca	0.96	2.55	1.7	1	2.47	1.38	1.13	0.75	0.75	1.4	0.95	0.7	1.3	1.07	1.3	0.98	1.5	2.1	1.44	1.77	2.35	1.1	1.3	2.1	0.5	0.85																																																																																																																																																																																																			
63	569	Ca	1.01	1.41	1.3	0.8	1.69	1.08	0.95	0.82	0.78	1.1	0.88	0.8	1.1	1.04	1.1	0.87	1.1	1.2	1.02	1.4	1.55	1.1	1.2	1.5	0.6	0.59																																																																																																																																																																																																			
	570	MC	2.68	1.39	1.3	0.9	3.57	2.46	1.45	1.17	1.67	1.07	1.93	1.3	1.2	0.87	1.3	1.79	1.9	1.7	1.6	1.99	1.63	1	1	0.8	0.88																																																																																																																																																																																																				
	595	Ca	0.42	2.9	1.1	1.1	0.98	0.7	0.82	0.88	0.6	1.4	0.97	0.5	1.2	0.81	0.9	0.85	1.2	2.5	1.2	1.34	1.7	1.2	0.9	1.5	0.7	0.77																																																																																																																																																																																																			
	627	Ca	1.29	1.51	1.2	1	2.73	1.03	0.72	0.92	1.04	1.24	1.22	1	1.7	0.74	0.9	1.38	1.3	2.1	1.2	1.12	0.88	0.7	0.9	0.9	0.6	0.92																																																																																																																																																																																																			
84	628	MC	2.95	1.29	1.2	1.4	1.82	1.21	0.96	1.06	1.65	1.48	1.84	1.5	1.7	1.03	1	1.48	1.5	1.4	0.98	1.36	0.82	0.8	1.1	1.1	1.15																																																																																																																																																																																																				
	148	Ca	2.77	1.88	1.7	1.2	4.05	2.11	1.47	1.41	1.82	1.96	1.87	1.7	2.8	1.88	2.1	1.85	1.5	2.5	1.73	1.38	1.07	0.7	1	0.8	1.93																																																																																																																																																																																																				
	1235	Ca	1.81	1.7	1.6	1.4	0.82	1.06	1.4	1.33	1.56	1.26	2.59	1.6	1.5	0.81	2.1	1.85	1.6	1.8	0.21	1.64	1.23	0.6	0.8	0.7	1.5																																																																																																																																																																																																				
	173	MC	0.75	0.95	1.1	1.3	0.47	0.6	1.01	0.91	1.11	0.71	0.86	0.9	0.7	0.92	1.3	0.58	0.9	1	0.15	1.05	1.24	0.9	0.6	0.8	0.9	1.15																																																																																																																																																																																																			
236	1613	MC	1.28	1.1	1.2	1	1.19	0.95	0.93	0.93	0.92	0.92	0.88	0.9	1	1.2	1.1	1.43	1.4	1.4	0.87	0.87	0.81	0.8	0.9	0.9	0.8	1.01																																																																																																																																																																																																			
	1614	Ca	0.95	0.95	1.3	0.55	0.73	0.81	0.95	1.23	1.4	1.31	1.4	1.3	1.03	0.9	1.2	1.2	1.2	1.2	0.81	1.04	0.65	0.9	0.8	0.8	1.3	0.95																																																																																																																																																																																																			
	L6741	Ca	0.99	0.79	0.7	0.9	0.16	0.51	0.66	0.74	0.85	0.78	0.9	0.7	0.9	0.7	0.77	0.8	0.7	0.81	0.78	0.94	1.4	0.8	1	1.8	0.54																																																																																																																																																																																																				
	995	Ca	0.93	1.15	1.1	1.2	0.82	0.99	1.05	1.08	1.01	1.15	1.14	1.1	0.9	0.99	1	1	1.1	1.18	0.99	1.27	1.3	1.1	1.1	1.6	0.71																																																																																																																																																																																																				
69	996	MC	0.79	1.1	1.2	1.3	0.8	0.97	1.04	1.1	1.09	1.16	1.33	1.3	1.18	1.1	1.27	1	1.3	1.04	1.05	1.07	1.2	1	0.9	1.3	0.84																																																																																																																																																																																																				
	E436T	Ca	0.43	0.79	0.9	0.9	0.21	0.52	0.93	0.77	0.88	0.94	0.81	0.9	0.7	0.92	0.8	0.78	0.8	0.7	0.74	0.73	0.89	1.3	0.8	0.9	1.6	0.63																																																																																																																																																																																																			
	697P	ATDBG	0.44	0.87	0.8	1	0.21	0.58	0.94	0.8	0.86	0.97	0.81	0.9	0.8	0.96	0.8	0.72	0.7	0.7	0.8	0.78	0.97	1.4	0.8	1	1.6	0.73																																																																																																																																																																																																			
	L701T	Ca	0.46	1.06	0.8	1	0.19	0.56	0.95	0.85	0.84	0.95	0.8	0.9	1.08	0.96	0.8	0.72	0.7	0.7	0.86	0.73	0.99	1.4	0.8	1.1	1.5	0.77																																																																																																																																																																																																			
P2209	1541	Ca	2.84	2.06	1.5	0.9	4	1.3	1.44	1.58	2.3	2.83	2.6	2.3	3.4	1.1	1.3	2.09	2	3.1	1.42	2.6	1.16	1	1.4	1.3	1.1	2.03																																																																																																																																																																																																			

Legenda: Ca - carcinoma, MC - mucosa normal, ATDAG - adenoma tubular com displasia de alto grau, ATDBG - adenoma tubuloviloso com displasia de alto grau, ATVDAG - adenoma viloso com displasia de alto grau, ATVDBG - adenoma tubuloviloso com displasia de baixo grau. A cor azul indica a ocorrência de ganhos e a cor vermelha de perdas.

Tabela 3.2: Frequências obtidas da análise de *copy-number* aos genes MMR e MGMT nas amostras de tumores e mucosas normais

Familias	Frequência	Copy-number												MLH1?																																																																																																																																																																																							
		MSH2			MSH6			MLH1			MSH3				PMS2			MGMT																																																																																																																																																																																			
		Up	Ex1	Ex7	Up	Ex1	Ex7	Up	Ex1	Ex7	Up	Ex1	Ex7	Up	Ex1	Ex7	Up	Ex1	Ex7	Up	Ex1	Ex7	Up	Ex1	Ex7	Int1	Int2	Int3	Int4	Int5	Int6	Int7	Int8	Int9	Int10	Int11	Int12	Int13	Int14	Int15	Int16	Int17	Int18	Int19	Int20	Int21	Int22	Int23	Int24	Int25	Int26	Int27	Int28	Int29	Int30	Int31	Int32	Int33	Int34	Int35	Int36	Int37	Int38	Int39	Int40	Int41	Int42	Int43	Int44	Int45	Int46	Int47	Int48	Int49	Int50	Int51	Int52	Int53	Int54	Int55	Int56	Int57	Int58	Int59	Int60	Int61	Int62	Int63	Int64	Int65	Int66	Int67	Int68	Int69	Int70	Int71	Int72	Int73	Int74	Int75	Int76	Int77	Int78	Int79	Int80	Int81	Int82	Int83	Int84	Int85	Int86	Int87	Int88	Int89	Int90	Int91	Int92	Int93	Int94	Int95	Int96	Int97	Int98	Int99	Int100	Int101	Int102	Int103	Int104	Int105	Int106	Int107	Int108	Int109	Int110	Int111	Int112	Int113	Int114	Int115	Int116	Int117	Int118	Int119	Int120	Int121	Int122	Int123	Int124	Int125	Int126	Int127	Int128	Int129	Int130	Int131	Int132	Int133	Int134	Int135	Int136	Int137	Int138	Int139	Int140	Int141	Int142	Int143	Int144	Int145	Int146	Int147	Int148	Int149	Int150	Int151	Int152	Int153	Int154	Int155	Int156	Int157	Int158	Int159	Int160	Int161	Int162	Int163	Int164	Int165	Int166	Int167	Int168	Int169	Int170	Int171	Int172

Tabela 3.5: Análise de copy-number aos genes MMR e MGMT nas amostras de sangue periférico

Família	DNA	Copy-number																									
		MSH2			MSH6			MLH1			MSH3			PMS2			MGMT			MLH3							
		Lp	Ex1	Ex7	Lp	Ex1	Ex1	Ex1	Ex1																		
7	L5	1,05	0,99	1,1	1	0,96	1,02	0,98	1,04	0,98	1,07	1,04	0,97	1,07	0,99	0,97	1,05	1,16	1,11	1,12	1,06	1,01	1,01	1,01	1,04		
		0,99	1,04	1	1	0,97	0,95	1,05	1	1,05	1,06	1,06	1,02	1,06	1,02	1	0,99	1	1	0,91	0,96	0,95	0,92	0,94	0,9	1,04	
		1,08	1,01	1	1	1,07	1,02	0,97	1,02	1,01	1,01	1,01	1,02	1,04	0,97	1	1,02	1,04	1,01	0,54	1,08	1	1,02	1,02	1	1	
55	L295	1,07	1	1	1	1,05	1,02	1,02	1	1	1	0,99	1	0,99	1,01	0,99	0,98	1	1,02	1,02	1	1,02	1	1,01	1,01		
		1,34	1,17	1	1	2	1,36	0,99	1,07	1,02	1,07	1,14	1,02	1,21	0,98	1,07	1,1	1,09	1,22	1,36	1,25	1,08	0,93	1,12	1,06	0,88	0,79
		1,62	1,23	1	1	2,64	1,4	0,99	1,01	1,1	1,05	1,28	1,01	1,29	0,83	1,03	1,24	1,3	1,21	1,42	1,37	1	0,77	1,14	1	1,01	1,24
63	L451	1,07	1,02	1,1	1	1,13	1,07	1,01	1,05	1	1	1,04	1,01	1,03	1,02	1,01	0,98	1,01	1,06	1,09	1,11	1,04	0,99	1,04	1,08	0,93	0,59
		1,01	1,01	1,1	1,1	1,02	1,04	0,97	1,04	0,93	1,08	1,01	0,97	1,03	1,02	1,04	0,99	0,91	1,02	1	1	0,94	1	1	1	0,94	1,03
		1	0,97	1	1	0,99	0,96	1	0,96	1	0,98	0,98	0,98	0,92	1	0,98	1,02	0,98	0,97	0,94	1	1	0,95	0,99	1	0,97	0,97
71	L467	0,99	0,98	1	1	1	0,96	1,03	0,98	0,99	0,96	1,01	0,99	1,04	0,98	1	1	0,96	1,03	0,98	1,02	1	1,01	1,01	1,02	0,96	
		1,37	1,2	1	1	1,92	1,28	0,99	1,04	1,05	1,08	1,2	1	1,21	0,96	1,03	1,11	1,13	1,19	1,26	1,22	1,03	0,92	1,09	1,09	0,94	1,37
		1,64	1,33	1,1	1,1	2,6	1,4	1,01	1,06	1,07	1,14	1,27	1,05	1,36	0,95	1,14	1,21	1,3	1,32	1,5	1,44	1,05	0,84	1,24	1,11	0,95	1,39
84	L484	1	1	1	1	1	1,03	0,97	1,03	1	1,01	1,01	1	1	1,02	0,98	1	1,01	1,02	1,03	1,02	1	0,99	1	1	1,02	
		1	1	1	1	1	1,06	1,03	0,98	1,03	1,02	0,97	1,03	1,01	1,02	1	1,01	1	1,03	1,04	1,03	1,04	0,95	1,07	1,04	0,92	1
		1,18	1,03	1,1	1	1,16	1,12	0,98	1,03	0,94	1,04	1,07	1	1,08	0,99	1,03	1,03	1	1,08	1,04	1,1	1	1,02	1,01	1,03	0,9	1,12
56	L674	1	1,01	1	1	1	1	1	1	1,01	0,99	0,97	1	1	0,97	0,98	0,95	0,99	1	0,97	1	1,04	1	0,99	1,01	1	
		0,93	0,96	1,1	1,1	0,84	0,94	1	0,93	0,97	0,99	1	0,97	0,91	0,98	0,94	1	1	0,94	0,99	1,02	0,99	1,1	0,98	0,93	1,15	0,9
		0,95	1,05	0,9	1,1	0,83	0,96	0,96	1	0,96	0,97	0,99	1	0,87	0,94	1	0,97	0,97	0,95	0,95	0,98	0,98	1,09	1	0,95	1,15	0,9
69	L801	1,18	1,06	1,1	1	1,02	1,03	0,95	1,03	1	1,06	1,08	1,02	1,07	0,96	1,04	1	1,04	1,03	1,08	1,1	1,07	1,02	1,04	1	1,01	1,06
		0,96	1,1	1,1	1,1	0,97	1	1,05	0,98	0,99	0,97	0,91	0,93	1	0,91	0,95	1,07	1,13	1,14	0,95	0,91	0,95	0,94	0,93	0,92	0,86	0,94
		0,99	0,97	1	1	1	0,96	0,98	1,01	0,98	0,96	1,01	0,98	0,99	0,97	1,03	1,01	1	0,96	0,96	0,96	0,99	0,98	0,96	0,99	0,98	0,96
86	698 (s)	0,95	1	1,1	1	0,93	0,93	0,98	1,02	0,97	1	0,95	0,95	0,95	0,97	1	0,95	0,89	1,01	0,95	0,99	1,01	1	0,99	1	1	1
		1,25	1,12	1,1	1	1,28	1,17	1,01	1,08	1,02	1,03	1,12	1,06	1,14	1,01	1,03	1,06	1,07	1,1	1,07	1,13	1,06	1	1,07	1,09	0,93	1,17
		1,01	1,07	1,1	1	1,01	1	1	1	0,98	1	1	1,03	1	1,03	0,98	1,02	1,03	1,06	1,07	1,06	1,03	1,04	1,06	1,08	0,92	1,04

Tabela 3.6: Análise de metilação aos genes MMR e MGMT nas amostras de sangue periférico

Família	DNA	Metilação																																					
		MSH2						MSH6						MLH1						MSH3						PMS2						MGMT						MLH3	
		Up	Ex1	Up	Ex1	Up	Ex1	Up	Ex1	Up	Ex1	Up	Ex1	Up	Ex1	Up	Ex1	Up	Ex1	Up	Ex1	Up	Ex1	Up	Ex1	Up	Ex1	Up	Ex1	Up	Ex1								
7	L5	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%					
24	L48	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%				
37	L108	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%				
	L295	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%				
	L447	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%			
55	L905	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%			
	L926	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		
	L950	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		
63	L451	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		
	L467	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		
71	CAs	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%		
84	L484	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	
148	L984	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	
173	L1038	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	
236	L1200	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	
	L674	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	
	L676	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	
56	L781	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	
	L801	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
69	E436	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
86	L499	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
	698 (s)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
106	L701	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
P2209	L1121	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%

3.1.2 Gene APC

Tendo em conta a elevada frequência de LOH e mutações somáticas no gene *APC*, é importante avaliar o *copy-number* e a metilação, no sentido de caracterizar as alterações moleculares neste gene tão importante na tumorigénese do cólon e reto.

Nas amostras de tumores (Tabela 3.7) verificaram-se ganhos mais frequentemente nas sondas *upstream* ao gene *APC*. Estas alterações encontram-se em igual frequência nas famílias TSG+ e TSG- (8/13 (62%) e 3/5 (60%), respetivamente). Por outro lado, as perdas foram detetadas nas famílias TSG- em 3/5 (60%) amostras de tumores, enquanto nas famílias TSG+ observaram-se estas perdas numa frequência de 6/13 (46%).

Nas famílias TSG+ as alterações de *copy-number* mais evidentes foram a perda na sonda 50 kb *upstream* do exão 1, os ganhos nas sondas 1B-1 e 1A-2, que se verificaram todos em 6/13 (46%) de amostras de tumores. As mesmas alterações verificaram-se nas amostras das famílias TSG- em 2/5 (40%), 1/5 (20%) e 1/5 (20%), respetivamente. Também é notório no grupo TSG+ um ganho em todas as sondas nas amostras 627 e 628. Nas famílias TSG- verificou-se perda na sonda Ex1-2B (3/5 (60%)) e ganhos generalizados nas sondas da região promotora (Tabela 3.8).

Em algumas amostras foram detetadas alterações mais extensivas, nomeadamente alterações em pelo menos três sondas, ou em duas, casos se tratem de sondas localizadas em regiões promotoras. Estas alterações expressam-se de forma muito semelhante nos dois grupos de famílias, onde os ganhos surgem em 7/13 (54%) e 2/5 (40%), e as perdas em 4/13 (31%) e 2/5 (40%) em amostras de tumor de famílias TSG+ e TSG-, respetivamente. É de salientar alterações de *copy-number* na mucosa normal.

Em relação ao estado de metilação no gene *APC*, as famílias TSG+ e TSG- caracterizam-se pela presença de hipermetilação nas regiões promotoras e por hipometilação nas sondas que se localizam a *downstream* dos promotores. Foi detetado um aumento significativo da metilação em 3/15 (20%) amostras de tumor do grupo TSG+ e em 3/5 (60%) amostras de tumor das famílias TSG-. Por outro lado, verificou-se hipometilação em 8/13 (62%) e 1/5 (20%) amostras de tumor dos grupos de famílias TSG+ e TSG-, respetivamente. Desta forma existe uma inversão das frequências nestes dois grupos, enquanto as famílias TSG+ apresentam mais hipometilação das regiões *downstream* aos promotores, as famílias TSG- caracterizam-se pela hipermetilação dos promotores.

Em algumas amostras foram detetadas alterações de metilação em várias sondas do gene *APC*. Observou-se hipermetilação em pelo menos 3 sonda ou 2, caso se trate de regiões promotoras, em 3/15 (20%) e 1/5 (20%) em amostras de tumor TSG+ e TSG- e hipometilação em 1/15 (7%) em amostras de tumor TSG+, sendo que para as famílias TSG- não se observaram hipometilações mais extensas, ou seja, num maior número de sondas. É de salientar alterações de metilação na mucosa normal.

Da mesma forma que para os gene *MMR* e *MGMT*, também foram avaliadas para o gene *APC* as alterações detetadas ao nível da linha germinal, ou seja, nas amostras de sangue periférico.

Tabela 3.7: Análise de *copy-number* e de metilação no gene *APC* em amostras de tumores e mucosa normal de indivíduos FCCTX

Família	DNA	Histologia	<i>Copy-number</i>								Metilação							
			N	1B-1	1B-2	1A-1	1A-2	50kp up Ex1	Ex1-2B	Ex13-14	N	1B-1	1B-2	1A-1	1A-2	50kp up Ex1	Ex1-2B	Ex13-14
7	437	Ca	24,21	2	6,15	1,3	2,37	7,27	2,51	2,01	61%	0%	52%	0%	0%	51%	43%	40%
	438	MC	0,92	1,44	1,13	1,35	1,43	1,26	1,28	1,32	88%	11%	30%	17%	19%	76%	99%	95%
24	459	ATDAG	1,32	1,79	1,2	1,83	1,62	1,55	1,97	1,7	57%	4%	24%	8%	8%	57%	66%	61%
	460	MC	0,3	4,22	1,76	1,73	4,75	0,19	5,96	1,01	51%	45%	42%	52%	51%	152%	45%	48%
55	947	ATDBG	0,2	1,14	1,18	1,11	0,96	0,53	1,11	0,97	89%	6%	25%	18%	19%	78%	100%	112%
	991	AVDAG	2,83	2,25	2	1,32	3,18	0,41	1,6	0,8	70%	35%	46%	45%	51%	99%	85%	40%
	992	MC	5,37	1	1,62	0,79	1,61	0,34	2,02	0,7	78%	73%	95%	105%	40%	80%	19%	81%
	993	ATDBG	2,14	1,48	1,23	0,36	1,05	0,15	1,16	0,56	69%	49%	0%	108%	70%	338%	69%	52%
	1775	ATDBG	4,37	0,91	1,51	0,83	0,71	2,07	0,85	0,76	89%	13%	47%	14%	11%	61%	109%	122%
	1871	Ca	0,88	0,88	1,62	1,45	1,65	1,21	1,11	1,19	95%	5%	0%	17%	14%	127%	64%	90%
	1872	MC	0,46	1,12	1,06	1,35	1,44	0,85	1,33	1,24	81%	5%	41%	14%	11%	105%	88%	104%
63	569	Ca	1,31	1,69	0,72	1	1,86	0,17	1,46	0,52	107%	9%	39%	25%	15%	119%	62%	112%
	570	MC	0,54	1,43	0,44	0,94	1,38	0,18	0,98	0,71	108%	8%	0%	18%	13%	107%	88%	115%
	595	Ca	0,29	1,53	1,05	1,39	1,68	0,46	1,52	1,22	66%	0%	74%	9%	12%	51%	86%	92%
71	801P	ATDBG	1,08	0,69	0,79	0,7	0,79	0,53	0,61	0,73	72%	0%	0%	8%	0%	59%	87%	76%
84	627	Ca	5,01	1,66	3,67	1,97	1,61	3,69	2,03	1,61	67%	0%	37%	9%	11%	53%	89%	95%
	628	MC	5,51	1,81	3,1	2,51	2,23	3,69	2,61	2,26	73%	0%	35%	8%	5%	54%	89%	78%
173	1235	Ca	1,13	1,78	1,57	1,32	1,09	1,05	1,12	0,74	96%	4%	30%	24%	30%	70%	97%	136%
	1236	MC	3,2	1,09	1,25	1,1	0,97	1,58	1,06	0,65	81%	6%	48%	44%	27%	70%	103%	107%
236	1614	MC	2,25	0,85	1,87	0,89	0,75	1,92	1,18	0,98	96%	7%	60%	15%	18%	85%	85%	104%
	1613	Ca	1,35	0,89	1,03	0,97	0,93	0,99	1,01	0,84	86%	7%	42%	7%	10%	72%	107%	106%
56	995	Ca	0,97	1,33	1,14	1,21	1,33	0,92	1,17	0,91	91%	4%	0%	6%	7%	75%	96%	77%
	996	MC	1,56	1,41	1,63	1,27	1,31	1,43	1,16	1,01	106%	8%	33%	9%	10%	74%	105%	83%
69	E436T	Ca	0,95	0,81	0,69	0,63	0,87	0,59	0,65	0,83	97%	0%	58%	26%	28%	57%	110%	109%
86	697P	ATDBG	2,29	0,81	0,9	0,76	0,84	0,95	0,55	0,87	96%	0%	0%	0%	0%	54%	165%	112%
106	L701T	Ca	0,99	0,84	0,81	0,63	0,79	0,53	0,62	0,85	97%	0%	0%	0%	0%	52%	93%	101%
P2209	1541	Ca	6,37	1,31	2,2	1,35	1,14	3,39	1,18	1,07	76%	28%	55%	18%	0%	74%	131%	65%
Amostras de referência			-								83-114%; 0-6%; 0-56%; 7-16%; 8-12%; 79-113%; 97-113%; 88-117%							

Legenda: Ca - carcinoma, MC - mucosa normal, ATDAG - adenoma tubular com displasia de alto grau, ATDBG - adenoma tubular com displasia de baixo grau, AVDAG - adenoma viloso com displasia de alto grau, ATVDBG - adenoma tubuloviloso com displasia de baixo grau. A cor azul indica a ocorrência de ganhos/hipermetilação e a cor vermelha de perdas/hipometilação.

Tabela 3.8: Frequências obtidas da análise MS-MLPA ao gene *APC* nas amostras de tumores e mucosas normais

Famílias	Frequência	<i>Copy-number</i>								Metilação							
		N	1B-1	1B-2	1A-1	1A-2	50kp upstream Ex1	Ex1-2B	Ex13-14	N	1B-1	1B-2	1A-1	1A-2	50kp upstream Ex1	Ex1-2B	Ex13-14
TSG+	ganhos/hipermetilação	6/21 (29%)	10/21 (48%)	6/21 (29%)	3/21 (14%)	11/21 (52%)	3/21 (14%)	7/21 (33%)	5/21 (24%)	0/21 (0%)	4/21 (19%)	1/21 (5%)	5/21 (24%)	5/21 (24%)	1/21 (5%)	0/21 (0%)	1/21 (5%)
TSG-	perdas/hipometilação	2/6 (33%)	2/6 (33%)	1/6 (17%)	0/6 (0%)	2/6 (33%)	1/6 (17%)	0/6 (0%)	0/6 (0%)	0/6 (0%)	1/6 (17%)	0/6 (0%)	1/6 (17%)	1/6 (17%)	0/6 (0%)	1/6 (17%)	1/6 (17%)
TSG+	ganhos/hipermetilação	4/21 (19%)	1/21 (5%)	1/21 (5%)	1/21 (5%)	0/21 (0%)	9/21 (43%)	1/21 (5%)	3/21 (14%)	0/21 (0%)	0/21 (0%)	0/21 (0%)	0/21 (0%)	0/21 (0%)	4/21 (19%)	7/21 (33%)	5/21 (24%)
TSG-	perdas/hipometilação	0/6 (0%)	0/6 (0%)	1/6 (17%)	2/6 (33%)	0/6 (0%)	2/6 (33%)	3/6 (50%)	0/6 (0%)	0/6 (0%)	0/6 (0%)	0/6 (0%)	0/6 (0%)	0/6 (0%)	0/6 (0%)	0/6 (0%)	1/6 (17%)

Assim, na análise MS-MLPA ao gene *APC* nas amostras de sangue (Tabela 3.9) verificaram-se várias alterações de *copy-number*, principalmente perdas. Estas perdas foram verificadas em 7/15 (47%) nas amostras de famílias TSG+ e em 3/9 (33%) nas amostras de TSG-. Por outro lado, os ganhos foram detetados em 5/15 (33%) e em 3/9 (33%) de amostras das famílias TSG+ e TSG- respetivamente. Portanto, as frequências das alterações de *copy-number* mantêm-se dentro dos mesmos valores nos dois grupos de famílias.

As perdas verificam-se maioritariamente no promotor N e na sonda a 50kb *upstream*

do exão 1 (tal como se tinha observado nas amostras de tumores), enquanto os ganhos neste gene verificam-se maioritariamente em regiões promotoras, ou seja, promotor N, 1B, 1A e a *upstream* do exão 1. Estas alterações de *copy-number* observam-se num número elevado de sondas nas várias amostras a estudo. Desta forma, verificaram-se ganhos em pelo menos três sondas ou duas, caso se tratem de regiões promotoras, numa frequência semelhante nas duas famílias (TSG+, 4/15 (27%) vs TSG-, 2/9 (22%). Em relação às perdas, verificou-se que havia uma maior frequência destas alterações nas famílias TSG+ (6/15 (40%) vs 2/9 (22%).

As alterações ao nível de metilação nestas amostras não são significativas, observando-se apenas alterações em duas amostras (L905 e E436). Na amostra L905 apenas se observa hipometilação, que poderá afetar a *copy-number*, visto as sondas que têm redução de metilação também apresentam ganhos no *copy-number*.

Tabela 3.9: Análise de *copy-number* e de metilação no gene *APC* em amostras de sangue de indivíduos FCCTX

Família	DNA	<i>Copy-number</i>								Metilação							
		N	1B-1	1B-2	1A-1	1A-2	50kp <i>upstream</i> Est1	Est1-2B	Est13-14	N	1B-1	1B-2	1A-1	1A-2	50kp <i>upstream</i> Est1	Est1-2B	Est13-14
7	L5	0,82	1,17	0,87	1,15	1,14	0,74	1,04	1,16	101%	0%	0%	4%	0%	77%	102%	106%
24	L48	2,82	1,22	1,13	1,12	1,1	1,35	0,98	1,16	95%	0%	0%	0%	0%	72%	103%	105%
37	L108	2,19	1,35	1,63	1,41	1,38	1,28	1,47	1,17	85%	2%	31%	3%	4%	71%	82%	90%
55	L295	1,82	1,11	0,96	1,14	1,11	1,03	1,09	1,13	102%	0%	0%	4%	0%	75%	94%	114%
	L447	1,15	1,09	1,09	1,02	0,92	0,91	1,12	0,75	81%	0%	0%	4%	0%	64%	93%	90%
	L905	2,22	1,51	1,88	1,6	1,48	1,95	1,77	1,31	71%	2%	26%	3%	4%	44%	70%	64%
	L926	0,2	0,96	0,8	0,93	0,98	0,31	0,88	0,8	85%	2%	27%	2%	2%	74%	91%	98%
63	L950	0,11	1,06	0,78	0,98	1,03	0,2	0,9	0,81	82%	0%	30%	2%	3%	90%	102%	99%
	L451	1,08	1,1	1,05	1,02	1,07	0,87	1,06	1,11	89%	0%	0%	0%	0%	80%	99%	101%
71	L467	0,99	1,09	1,08	1,04	1,15	0,88	1,08	1,05	93%	0%	0%	0%	0%	81%	99%	104%
	CAs 802s	0,54	1,06	1,01	0,93	1	0,64	0,99	0,86	101%	0%	0%	4%	0%	83%	112%	111%
84	L484	0,64	1,43	1,23	1,44	1,5	0,84	1,43	1,21	93%	2%	33%	3%	4%	83%	101%	98%
148	L984	0,1	1,18	0,89	1,13	1,16	0,21	1,05	0,89	115%	2%	33%	3%	4%	90%	101%	105%
173	L1038	0,16	0,99	0,72	0,98	1,05	0,23	0,92	0,82	107%	0%	0%	3%	4%	97%	97%	99%
236	L1200	0,15	1,01	0,76	0,96	1	0,25	0,98	0,83	99%	0%	0%	3%	3%	88%	88%	100%
56	L674	2,23	1,42	1,67	1,39	1,44	1,32	1,55	1,22	82%	2%	40%	4%	4%	71%	84%	95%
	L676	1,02	1,36	1,27	1,29	1,23	0,84	1,26	0,96	87%	2%	31%	4%	4%	69%	104%	93%
	L781	1,13	1,31	1,43	1,24	1,16	0,94	1,25	0,76	87%	3%	30%	4%	4%	78%	109%	105%
	L801	1,15	1,17	1,02	1,11	1,21	0,86	1,09	1,15	101%	2%	32%	4%	4%	93%	102%	104%
69	E436	0,43	0,9	0,79	0,89	0,93	0,47	0,92	0,9	77%	2%	23%	22%	20%	65%	80%	94%
86	L499	1,39	1,17	1,19	1,16	1,22	1,16	1,14	1,12	80%	0%	32%	4%	4%	72%	93%	94%
	698 (s)	1,23	1,04	1,12	0,96	0,91	1,13	1,07	0,86	91%	0%	0%	4%	0%	54%	102%	113%
106	L701	0,04	1,02	0,72	0,96	0,98	0,11	0,87	0,75	88%	0%	0%	2%	0%	105%	104%	101%
P2209	L1121	0,99	1,01	0,7	0,96	0,94	0,59	0,87	0,88	98%	0%	36%	3%	3%	73%	89%	91%

Legenda: A cor azul indica a ocorrência de ganhos/hipermetilação e a cor vermelha de perdas/hipometilação.

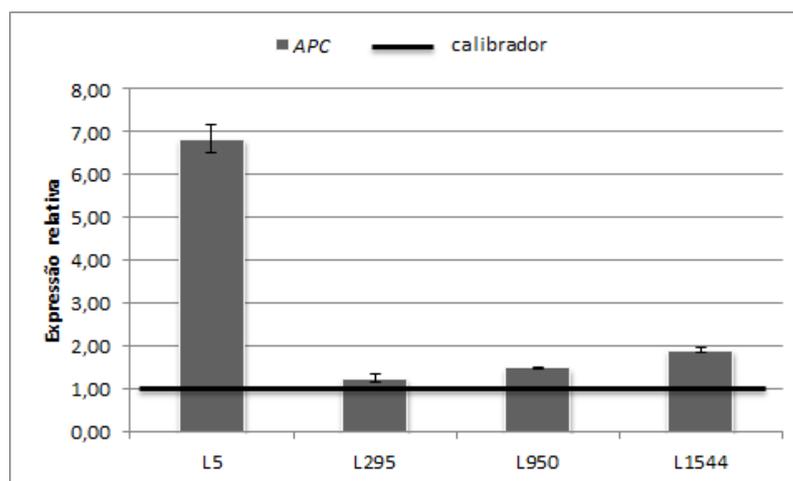
Tabela 3.10: Frequências obtidas da análise MS-MLPA ao gene *APC* nas amostras de sangue

Famílias	Frequência	<i>Copy-number</i>								Metilação							
		N	1B-1	1B-2	1A-1	1A-2	50kp <i>upstream</i> Est1	Est1-2B	Est13-14	N	1B-1	1B-2	1A-1	1A-2	50kp <i>upstream</i> Est1	Est1-2B	Est13-14
TSG+	ganhos/hiper metilação	4/15 (27%)	3/15 (20%)	2/15 (13%)	3/15 (20%)	3/15 (20%)	2/15 (13%)	3/15 (20%)	1/15 (7%)	0/15 (0%)	0/15 (0%)	0/15 (0%)	0/15 (0%)	0/15 (0%)	0/15 (0%)	0/15 (0%)	0/15 (0%)
TSG-	perdas/hipo metilação	1/9 (11%)	3/9 (33%)	2/9 (22%)	1/9 (11%)	1/9 (11%)	1/9 (11%)	0/9 (0%)	0/9 (0%)	0/9 (0%)	0/9 (0%)	0/9 (0%)	1/9 (11%)	1/9 (11%)	0/9 (0%)	0/9 (0%)	0/9 (0%)
TSG+	ganhos/hiper metilação	7/15 (47%)	0/15 (0%)	0/15 (0%)	0/15 (0%)	0/15 (0%)	6/15 (40%)	0/15 (0%)	0/15 (0%)	0/15 (0%)	0/15 (0%)	0/15 (0%)	0/15 (0%)	0/15 (0%)	1/15 (7%)	1/15 (7%)	1/15 (7%)
TSG-	perdas/hipo metilação	2/9 (22%)	0/9 (0%)	0/9 (0%)	0/9 (0%)	0/9 (0%)	3/9 (33%)	0/9 (0%)	0/9 (0%)	0/9 (0%)	0/9 (0%)	0/9 (0%)	0/9 (0%)	0/9 (0%)	0/9 (0%)	1/9 (11%)	0/9 (0%)

Em seguimento da análise de *copy-number* e estado de metilação, foi efetuada a quantificação da expressão do gene *APC* nas amostras FCCTX para as quais existia RNA. Desta forma, pretendeu-se verificar se as alterações detetadas na análise de MS-MLPA tinham alguma implicação ao nível da expressão deste gene. De acordo com o gráfico (Figura 3.1) obtido através desta análise de expressão, observa-se sobreexpressão do gene para a amostra L5 e um ligeiro aumento de expressão para as restantes amostras.

Na análise de *copy-number* verificou-se que a amostra L295 tem um ganho no promotor N e que a amostra L950 tem duas perdas, uma no promotor N e outra *upstream* do exão 1. O ganho da amostra L295 pode existir mas não se traduz na sobreexpressão do gene, visto a sua expressão estar muito próxima do valor do calibrador. Assim, para confirmar este ganho, seria necessário analisar mais amostras com ganhos ao nível da expressão. As perdas observadas na amostra L950 podem ser causadas por uma deleção ou polimorfismo em homozigotia, devido aos reduzidos valores obtidos na análise de *copy-number* (0,11 e 0,2). Se realmente se tratar de uma deleção, é indicativo de que uma deleção nestas regiões não se reflete ao nível da expressão do gene *APC*. A amostra L1544 apresenta a expressão do gene *APC* um pouco aumentada, mas não se pode relacionar, visto não se ter efetuado análise de MS-MLPA para esta amostra. De qualquer forma, a análise de expressão ao gene *APC* nesta amostra deve ser repetida para se confirmar este aumento de expressão.

Em relação à amostra L5, há uma sobreexpressão do gene *APC*, mas não se observa ao nível das sondas estudadas pelo MS-MLPA. De qualquer forma, esta sobreexpressão pode ser uma das causas para o FCCTX nesta família, logo deverá ser efetuada análise de segregação da expressão deste gene noutros familiares.



Calibrador – Média da expressão relativa do gene *housekeeping* utilizado (*β-actina*)

Figura 3.1: Resultado da análise de expressão relativa do gene *APC* no sangue periférico de doentes FCCTX

3.2 Análise de mutações nos genes *KDELC1* e *ERCC5*

3.2.1 *KDELC1*

Os resultados da análise mutacional dos 8 fragmentos e respetivas regiões flanqueadoras do gene *KDELC1*, por indivíduo e por exão, encontram-se no Anexo B, Tabela B.5. Na análise mutacional foram identificadas quatro variantes germinais (Tabela 3.11). Todas as alterações correspondem a substituições de nucleótidos, onde duas são variantes intrónicas na região 5'UTR (c.1-108T>G e c.1-30G>T), uma é do tipo *missense* (c.455A>T) e outra é uma variante silenciosa (c.48A>G), ou seja, trata-se de uma substituição sinónima, onde não há alteração do aminoácido produzido. Todas as variantes foram detetadas com uma frequência alélica de 6%, por isso detetadas em apenas um indivíduo. As variantes intrónicas e a variante silenciosa foram detetadas todas no mesmo indivíduo em estudo (L108).

As quatro alterações estão descritas na base de dados NCBI dbSNP [NCBI, 2014b] como SNVs (Tabela 3.11 - rs). As duas variantes identificadas na região 5'UTR e a variante do exão 1 apresentam uma frequência elevada na população europeia (10%, 10% e 8% respetivamente). Por outro lado, a variante identificada no exão 3 é reportada como existente, mas não apresenta dados sobre a sua frequência nas várias populações. Isto pode dever-se ao facto de esta região não ter sido estudada ou porque realmente esta variante é muito pouco frequente.

Tabela 3.11: Identificação das mutações germinais detetadas no gene *KDELC1* e respetivas frequências.

Fragmento	Variante Germinal	rs	Homozigotia/ Heterozigotia	Frequência em indivíduos a estudo	Frequência descrita*
5'UTR	c.1-108T>G #	rs7993381	Heterozigotia	1/17 (6%)	10%
	c.1-30G>T #	rs7993350	Heterozigotia	1/17 (6%)	10%
Exão 1	c.48A>G #	rs7994595	Heterozigotia	1/17 (6%)	8%
Exão 3	c.455A>T p.N152I	rs200215559	Heterozigotia	1/17 (6%)	-

*Informação retirada da base de dados *1000 Genomes* - Frequência do alelo variante na população europeia

Alteração surge no indivíduo L108

De forma a avaliar a eventual patogenicidade das variantes encontradas recorreu-se a uma análise *in silico*. Utilizando o *software Mutation Taster* verificou-se que as variantes c.1-108T>G, c.1-30G>T e c.48A>G são classificadas como polimorfismos. No entanto, a variante c.1-30G>T poderá conduzir eventualmente a um ganho de um sítio aceitador de *splicing* (0,85) e ao aumento da probabilidade de *splicing* em dois sítios putativos de *splicing* (wt:0,73, mu:0,95; wt:0,25, mu:0,50) enquanto a variante c.48A>G poderá implicar um ganho de um sítio dador de *splicing* (0,30) e aumento da probabilidade de *splicing* também em dois sítios putativos (wt:0,22, mu:0,28; wt:0,82, mu:0,92) (Tabela 3.12). Estas alterações, têm contudo uma baixa probabilidade, uma vez que, principalmente no caso da variante

c.1-30G>T, se localiza na região 5'UTR. As restantes alterações de *splicing* previstas não são relevantes.

Com recurso ao mesmo *software* bioinformático, no caso da variante c.455A>T a previsão é de ser patogénica, principalmente porque se localiza numa região potencialmente reguladora de cromatina aberta. Esta variante poderá alterar a sequência da proteína codificada, onde uma asparagina é substituída por uma serina. Não se verificam outras alterações significativas, como por exemplo, alteração dos locais de *splicing* e o nível da alteração do aminoácido, em termos de estrutura da proteína.

Utilizando o *software* VEP, que inclui a previsão do SIFT e *Polyphen*, não se obteve qualquer previsão ou *score* para as duas variantes intrónicas, porque não há alteração ao nível da proteína. No caso da substituição sinónima, estes *softwares* não fornecem informação sobre a sua previsão, enquanto a variante c.455A>T tem um *score* de 1 no SIFT (tolerada) e de 0,004 no *Polyphen* (benigna) (Tabela 3.12).

Tabela 3.12: Resultados da análise *in silico* da consequência prevista das mutações germinais encontradas no gene *KDELC1*

Variante	VEP		<i>Mutation taster</i>	<i>Mutation Taster</i>		
	SIFT	<i>Polyphen</i>	Previsão	Alterações de <i>splicing</i>		
				Efeito	Posição em relação à variante	Score
c.1-108T>G	-	-	Polimorfismo	sítio dador de <i>splicing</i> ligeiramente aumentado	8 pb <i>upstream</i>	wt: 0.9724 mu: 0.9857
c.1-30G>T	-	-	Polimorfismo	sítio aceitador de <i>splicing</i> ligeiramente aumentado	4 pb <i>upstream</i>	wt: 0.9844 mu: 0.9848
					1 pb <i>downstream</i>	wt: 0.4807 mu: 0.5116
					3 pb <i>downstream</i>	wt: 0.9188 mu: 0.9352
					9 pb <i>upstream</i>	wt: 0.6349 mu: 0.6449
					na mesma posição da variante	wt: 0.9935 mu: 0.9951
				sítio aceitador de <i>splicing</i> aumentado	4 pb <i>downstream</i>	wt: 0.73 mu: 0.95
					2 pb <i>upstream</i>	wt: 0.25 mu: 0.50
ganho de sítio aceitador de <i>splicing</i>	2 pb <i>downstream</i>	0.85				
c.48A>G	-	-	Polimorfismo	sítio aceitador de <i>splicing</i> ligeiramente aumentado	6 pb <i>upstream</i>	wt: 0.3289 mu: 0.3505
				sítio aceitador de <i>splicing</i> aumentado	10 pb <i>downstream</i>	wt: 0.22 mu: 0.28
					8 pb <i>upstream</i>	wt: 0.82 mu: 0.92
ganho de sítio dador de <i>splicing</i>	1 pb <i>upstream</i>	0.30				
c.455A>T	Tolerada	Benigna	Patogénica	-	-	-

Legenda: wt - *wild-type wild-type* (WT) ; mu - mutado

3.2.2 *ERCC5*

Os resultados da análise mutacional dos exões 11 ao 15 e respetivas regiões flanqueadoras do gene *ERCC5*, por indivíduo e por exão, encontram-se no Anexo B, Tabela B.8. Nesta análise foram detetadas três variantes germinais (Tabela 3.13). As alterações correspondem

a substituições de nucleótidos, duas do tipo *missense* (c.2636A>G e c.3310G>C) e uma variante intrónica (c.2878+14C>T). As variantes c.2636A>G e c.3310G>C foram detetadas com uma frequência alélica de 12% na população em estudo e a variante c.2878+14C>T foi detetada tanto em homozigotia como em heterozigotia com uma frequência de 41%.

As três alterações estão descritas na base de dados NCBI dbSNP [NCBI, 2014b] como SNVs (Tabela 3.13 - rs). As duas variantes identificadas nos exões 13 e 15 apresentam ainda uma frequência elevada na população europeia. De mencionar que a alteração c.3310G>C no exão 15 é uma mutação conhecida (HGMD ID CM063964), que já foi analisada num estudo prévio [Zhu et al., 2012]. Esta variante está ainda descrita no ClinVar [NCBI, 2014a], em termos de significado clínico, como alelo provavelmente benigno, não sendo especificado em que contexto. A variante detetada no exão 12 apresenta uma frequência mais elevada nos indivíduos (12%) em estudo comparando com a descrita na base de dados *1000 Genomes* (3,5%). Desta forma, estimou-se a frequência desta alteração na população portuguesa através da análise num conjunto de indivíduos saudáveis e observou-se uma frequência de 2%, o que continua a ser uma frequência inferior à detetada na série de FCCTX.

Tabela 3.13: Identificação das mutações germinais detetadas no gene *ERCC5* e respetivas frequências.

Fragmento	Variante Germinal	rs	Homozigotia/ Heterozigotia	Frequência em indivíduos a estudo	Frequência descrita*	Frequência em indivíduos saudáveis
Exão 12	c.2636A>G p.N879S	rs4150342	Heterozigotia	2/17 (12%)	3,5%	1/47 (2%)
Exão 13	c.2878+14C>T	rs4150360	Homozigotia	7/17 (41%)	26%	-
			Heterozigotia	7/17 (41%)	51%	-
Exão 15	c.3310G>C p.D1104H	rs17655	Heterozigotia	2/17 (12%)	30%	-

*Informação retirada da base de dados *1000 Genomes* - Frequência do alelo variante na população europeia

Através do *software Mutation Taster* foi possível verificar que no caso das variantes c.2878+14C>T e c.3310G>C (substituição de um ácido aspártico por uma histidina) a previsão é de corresponderem a polimorfismos. A variante c.2878+14C>T poderá levar contudo a alterações ao nível do *splicing*, entre elas, um aumento de 5 sítios aceitadores de *splicing* e a variante c.3310G>C poderá conduzir a um aumento de um sítio dador de *splicing* (wt:0,21, mu:0,29) (Tabela 3.14). As restantes alterações de *splicing* detetadas não são relevantes.

Utilizando o mesmo *software*, verificou-se que a alteração c.2636A>G no exão 12 (Figuras 3.2 e 3.3), é classificada como patogénica. Esta variante provoca a substituição de uma asparagina por uma serina e poderá tornar um sítio dador de *splicing* mais forte, em que a sequência wt tem um *score* de 0,82 e a sequência com a alteração terá um *score* de 1,00 (Tabela 3.14).

Os resultados obtidos na utilização do *software* VEP foram ambíguos, pois verificou-se que a variante c.2636A>G é classificada como tolerada e tem um *score* de 0,09 no SIFT,

3. RESULTADOS 3.3. Análise do envolvimento do *splicing* no potencial exão 13A do gene *TPP2* na suscetibilidade para o FCCTX

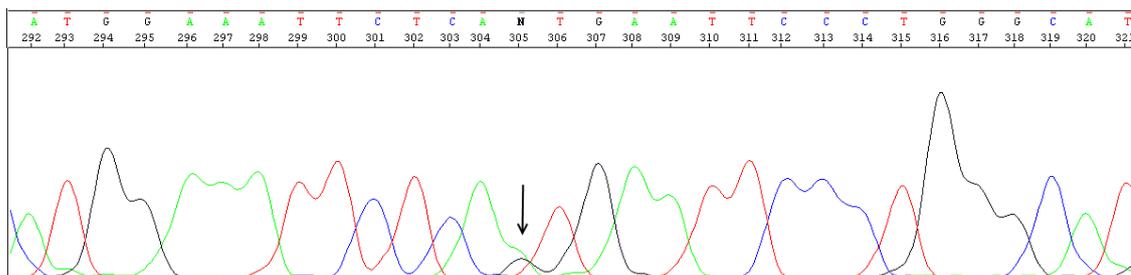


Figura 3.2: Eletroforetograma parcial obtido no sequenciador *ABI Prism™ 3130 Genetic Analyzer* (*Applied Biosystems*), representando a mutação c.2636A>G (sequência *forward*, amostra L449).

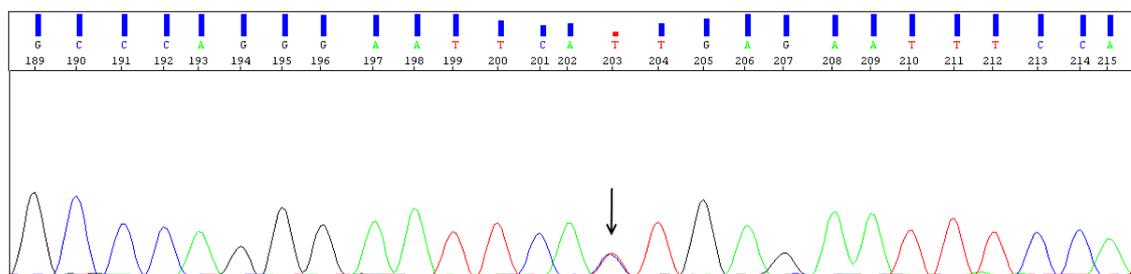


Figura 3.3: Eletroforetograma parcial obtido no sequenciador *ABI Prism™ 3130 Genetic Analyzer* (*Applied Biosystems*), representando a mutação c.2636A>G (sequência *reverse*, amostra L449).

enquanto no *Polyphen* é classificada como provavelmente patogénica, com um *score* de 0,998. Já a variante c.3310G>C é classificada como patogénica e com um *score* de 0,01 no SIFT, sendo que no *Polyphen* é classificada como possivelmente patogénica, com um *score* de 0,864. Estas diferenças prendem-se com a finalidade do *software*, que no caso do SIFT avalia a conservação do aminoácido a nível evolutivo e o *Polyphen*, avalia a conservação do aminoácido, mas a nível de consequência funcional para a proteína. Por fim, a variante c.2878+14C>T, sendo uma alteração intrónica, não tem qualquer informação no SIFT nem no *Polyphen* (Tabela 3.14).

3.3 Análise do envolvimento do *splicing* no potencial exão 13A do gene *TPP2* na suscetibilidade para o FCCTX

No sentido de esclarecer as alterações de expressão nos transcritos wt e 13A do gene *TPP2* identificadas num estudo anterior, envolvendo ou não o *splicing* de um potencial exão 13A, foi efetuada a análise mutacional entre os exões 13 e 14 do gene *TPP2*. Desta forma pretendia-se identificar mutações que pudessem estar associadas a esta alteração de *splicing*. A partir desta análise foram detetadas duas variantes germinais no intrão 13 (Tabela 3.15), neste caso duas variantes intrónicas (c.1679-23A>G e c.1679-94C>T). Na base de dados NCBI dbSNP [NCBI, 2014b] estas duas alterações estão descritas como SNVs (Tabela 3.15 - rs). Como a frequência da variante c.1679-23A>G em heterozigotia encontrada nos indivíduos em estudo é mais elevada que a descrita na base de dados *1000 Genomes* e a variante c.1679-94C>T em homozigotia é mais baixa, estudaram-se estas

3. RESULTADOS 3.3. Análise do envolvimento do *splicing* no potencial exão 13A do gene *TPP2* na suscetibilidade para o FCCTX

Tabela 3.14: Resultados da análise *in silico* da consequência prevista das mutações germinais encontradas no gene *ERCC5*

Variante	VEP		Mutation taster	Mutation Taster		
	SIFT	Polyphen	Previsão	Alterações de <i>splicing</i>		
				Efeito	Posição em relação à variante	Score
c.2636A>G	Tolerada	Provavelmente patogénica	Patogénica	sítio dador de <i>splicing</i> aumentado	1 pb <i>upstream</i>	wt: 0.82 mu: 1.00
c.2878+14C>T	-	-	Polimorfismo	sítio aceitador de <i>splicing</i> ligeiramente aumentado	3 pb <i>downstream</i>	wt: 0.8741 mu: 0.9035
				sítio aceitador de <i>splicing</i> aumentado	5 pb <i>downstream</i>	wt: 0.38 mu: 0.66
					2 pb <i>downstream</i>	wt: 0.53 mu: 0.63
					6 pb <i>downstream</i>	wt: 0.78 mu: 0.89
					4 pb <i>downstream</i>	wt: 0.83 mu: 0.93
					1 pb <i>upstream</i>	wt: 0.29 mu: 0.39
sítio dador de <i>splicing</i> ligeiramente aumentado	3 pb <i>downstream</i>	wt: 0.6186 mu: 0.6225				
c.3310G>C	Patogénica	Possivelmente patogénica	Polimorfismo	sítio dador de <i>splicing</i> aumentado	5 pb <i>downstream</i>	wt: 0.21 mu: 0.29

Legenda: wt - *wild-type*; mu - mutado

alterações num conjunto de indivíduos saudáveis.

Com recurso ao *software Mutation Taster*, foi possível verificar que as variantes c.1679-23A>G e c.1679-94C>T têm a previsão de serem polimorfismos. Os indivíduos que são portadores da variante c.1679-23A>G ganham um sítio dador de *splicing*, com um *score* correspondente a 0,42 (Tabela 3.16). Este sítio dador de *splicing* poderá estar envolvido na alteração da expressão do exão alternativo 13A, dada a sua localização na sequência de DNA (entre o potencial exão alternativo 13A e o exão 14). As restantes alterações de *splicing* detetadas não são relevantes. A análise utilizando o SIFT e *Polyphen* não se aplica, porque as duas variantes são intrónicas.

Tabela 3.15: Identificação das mutações germinais detetadas no gene *TPP2* e respetivas frequências.

Fragmento	Variante Germinal	rs	Homozigotia/ Heterozigotia	Frequência em indivíduos FCCTX e SL	Frequência descrita*	Frequência em indivíduos saudáveis
Intrão 13	c.1679-23A>G	rs9518799	Homozigotia	3/35 (9%)	6%	2/42 (5%)
			Heterozigotia	23/35 (66%)	35%	17/42 (40%)
	c.1679-94C>T	rs2274463	Homozigotia	1/35 (3%)	28%	9/42 (21%)
			Heterozigotia	20/35 (57%)	47%	23/42 (55%)

*Informação retirada da base de dados *1000 Genomes* - Frequência do alelo variante na população europeia

Foi efetuada uma análise de segregação das mutações identificadas com a expressão das duas isoformas (redução do transcrito wt e aumento de expressão do transcrito alternativo, Tabela 3.17). No entanto, verificou-se que as variantes encontradas não parecem ser responsáveis pelas alterações de expressão neste transcrito, visto não se associarem significativamente.

De modo a avaliar a presença de mutações nos exões 13 e 14, que pudessem associar-se

Tabela 3.16: Resultados da análise *in silico* da consequência prevista das mutações germinais encontradas no intrão 13 do gene *TPP2*

Variante	VEP		<i>Mutation taster</i>	<i>Mutation Taster</i>		
	SIFT	<i>Polyphen</i>	Previsão	Alterações de <i>splicing</i>		
				Efeito	Posição em relação à variante	Score
c.1679-23A>G	-	-	Polimorfismo	sítio aceitador de <i>splicing</i> ligeiramente aumentado	6 pb downstream	wt: 0.8320 mu: 0.8841
					5 pb downstream	wt: 0.9523 mu: 0.9546
				ganho de sítio dador de <i>splicing</i>	6 pb upstream	0.42
c.1679-94C>T	-	-	Polimorfismo	sítio dador de <i>splicing</i> ligeiramente aumentado	7 pb upstream	wt: 0.9264 mu: 0.9416
					3 pb upstream	wt: 0.9900 mu: 0.9909

Legenda: wt - *wild-type*; mu - mutado

Tabela 3.17: Análise de segregação das mutações identificadas com a expressão das duas isoformas

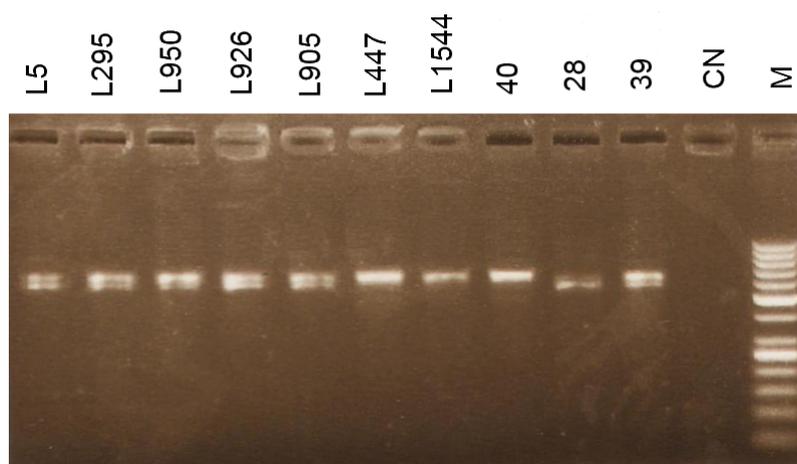
Condições	c.1679-23A>G	<i>One sided p-values</i> $p(O \geq E)$	c.1679-94C>T	<i>One sided p-values</i> $p(O \geq E)$
Indivíduos com redução do transcrito wt	8/11 (73%)	0.56	8/11 (73%)	0.42
Indivíduos sem redução do transcrito wt	18/24 (75%)		15/24 (63%)	
Indivíduos com aumento do transcrito alternativo	15/20 (75%)	0.61	12/20 (60%)	0.63
Indivíduos sem aumento do transcrito alternativo	11/15 (73%)		9/15 (60%)	

à expressão diferencial dos transcritos foi efetuada a amplificação do cDNA compreendendo um fragmento de 577 pb entre os exões 13 e 17. Observou-se em gel de agarose a existência de uma banda correspondente ao transcrito normal e uma outra, de maior peso molecular, correspondente ao transcrito aberrante, que contém o exão hipotético 13A (Figura 3.4). Pretendia-se saber se havia alguma alteração no cDNA, pelo que os dois fragmentos correspondendo aos dois transcritos foram sequenciados separadamente. No entanto, não foram encontradas quaisquer alterações na sequência de cDNA que pudessem ser responsáveis pelo aparecimento do transcrito 13A.

Foi ainda efetuada a análise de expressão por *real-time* PCR dos dois transcritos do gene *TPP2*, wt e 13A (Figura 3.5) em cólon, uma vez que os estudos de expressão tinham sido efetuados em sangue periférico e nada se conhecia sobre a expressão diferencial dos transcritos em cólon. Para isso foram analisadas amostras de sangue de indivíduos saudáveis (A, B, C e D) como referência, 2 carcinomas (C1 e C2) e 2 adenomas (A1 e A2) do cólon. Observou-se uma maior expressão do transcrito alternativo nas amostras de sangue em comparação com as amostras de carcinoma e adenoma. Esta diferença nos valores de expressão parece indicar que o transcrito 13A apresenta níveis de expressão muito reduzidos no cólon ou que não será mesmo expresso. O carcinoma C1 localizava-se no reto e o C2 e adenomas A1 e A2 no cólon proximal. É de notar que este último também apresentava uma expressão do transcrito wt muito baixa em relação aos restantes.

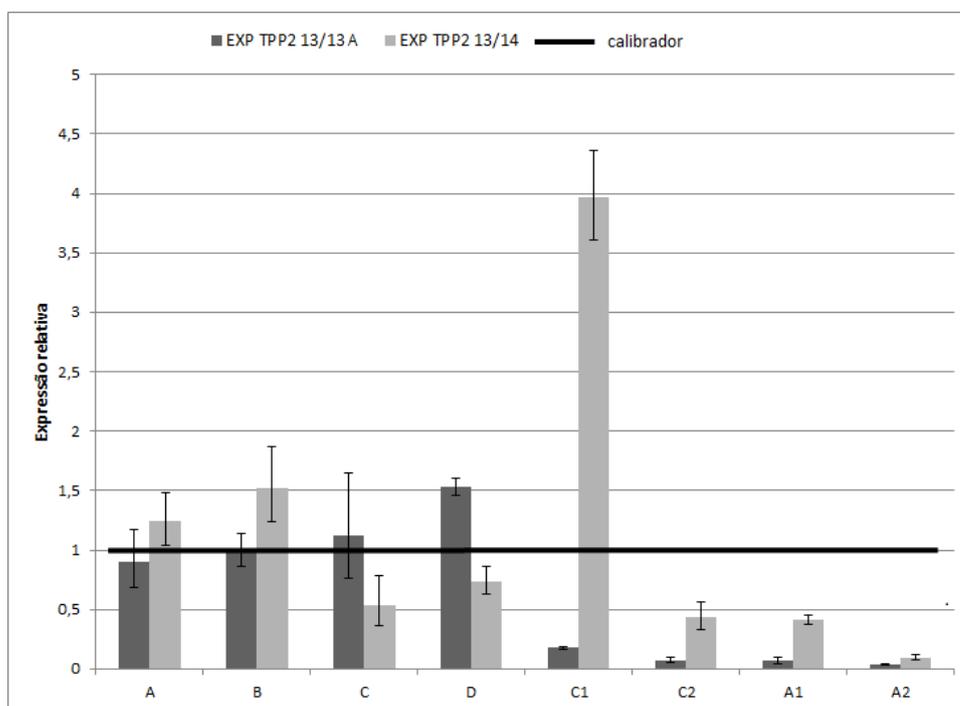
Ainda foi efetuado outro ensaio de análise de expressão destes dois transcritos nas

3. RESULTADOS3.3. Análise do envolvimento do *splicing* no potencial exão 13A do gene *TPP2* na suscetibilidade para o FCCTX



Legenda: CN – controlo negativo; M – marcador *GeneRuler 50 pb DNA Ladder* (*Thermo Scientific*)

Figura 3.4: Eletroforese em gel de agarose 2% (p/v) dos produtos de PCR do fragmento que incorpora os exões 13 ao 17 do gene *TPP2*



Calibrador – Média da expressão relativa do gene *housekeeping* utilizado (*SDHA*); A, B, C e D - amostras de sangue de indivíduos saudáveis; C1 e C2 - carcinomas; A1 e A2 - adenomas

Figura 3.5: Análise de expressão relativa a dois transcritos do gene *TPP2* (13/13A e 13/14)

amostras L447T, L447MV e L447MD (adenoma de risco, mucosa vizinha e mucosa à distância, respetivamente), de um dos indivíduos com FCCTX que apresentava elevada expressão do 13A e quase ausência do wt. A análise por *real-time* PCR terá que ser repetida, uma vez que o ensaio não foi válido. No entanto, a análise em gel de agarose (Figura 3.6b) da amplificação do cDNA destas amostras, que compreende os dois transcritos a

estudo apesar de não ser quantitativa, sugere que as referidas amostras não expressam o transcrito 13A ao nível do cólon e parecem ter uma expressão ao nível do wt muito baixa, como já se tinha observado ao nível germinal.

Esta sugestão é possível, pois uma quarta amostra adicionada (Ad4) como controlo positivo aparece bastante intensa, sendo esta intensidade já observada numa reacção anterior de síntese de cDNA (Figura 3.6a).

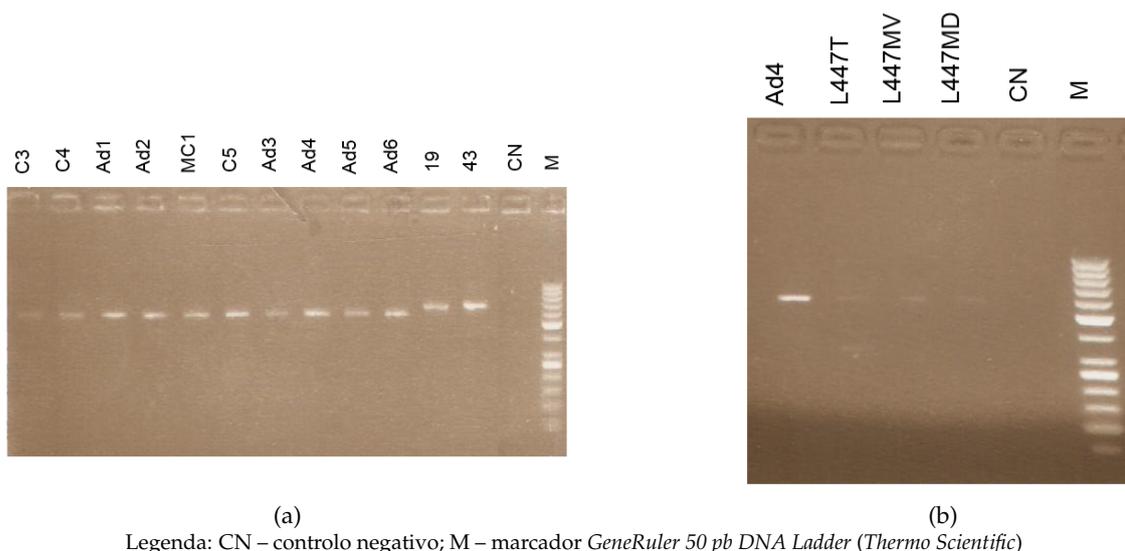


Figura 3.6: Eletroforese em gel de agarose 2% (p/v) dos produtos amplificados por PCR para o transcrito alternativo 13A

3.4 Análise de WES

A partir dos resultados obtidos da análise de WES para 4 indivíduos afetados de uma família FCCTX foram feitas duas análises de filtração de variantes distintas, baseadas apenas na seleção das variantes comuns entre os indivíduos: uma que incluiu 3 indivíduos (L408, L801, L674) e outra que englobou 4 indivíduos (L408, L801, L674 e L1794) (Figura 3.7), isto porque todos os indivíduos desenvolveram carcinoma, excepto o L1794, que só apresentou adenomas com displasia de alto grau aos 39 anos de idade, não se podendo excluir a hipótese, embora pouco provável, de estes serem de origem esporádica. De qualquer forma, este indivíduo encontra-se em risco e o facto de não ter apresentado carcinoma pode dever-se apenas a ser jovem.

Na primeira análise obteve-se um conjunto de 201 variantes: 163 SNVs não-sinónimas, 11 variantes de *splicing*, 5 SNVs com ganho de codão STOP, 4 deleções com alteração da *frameshift*, 3 deleções sem alteração da *frameshift*, 2 substituições com alteração da *frameshift*, 2 substituições sem alteração da *frameshift* e 11 variantes desconhecidas. Numa segunda análise, onde foi adicionado um quarto elemento, este valor foi reduzido para 62 variantes: 58 SNVs não-sinónimas, 2 variantes de *splicing*, 1 deleção com alteração da *frameshift* e 1

deleção sem alteração da *frameshift*.

Após a obtenção de resultados por análise de WES é recomendado efetuar uma validação das variantes mais interessantes através de sequenciação de Sanger. Desta forma, após a avaliação destes resultados, foram selecionados genes candidatos para análise mutacional, com posterior desenho de *primers* (Anexo B, Tabelas B.13 e B.15). O primeiro fator de seleção foi a expressão no cólon, seguido de avaliação do potencial patogénico, com recurso aos *softwares* bioinformáticos (SIFT, *Polyphen* e *Mutation taster*). As alterações mais fidedignas em relação a estes dois fatores foram submetidas a uma posterior seleção, tendo em consideração a função do gene, onde foram incluídos genes envolvidos na carcinogénese, genes relacionados com a apoptose, motilidade celular, proliferação e diferenciação. Por fim, também se teve em consideração o número de leituras (*reads*) e a frequência da mutação nessas leituras, a qual indica se se trata de uma mutação em hétero ou homocigotia ou se devido à baixa frequência pode corresponder a erros durante a sequenciação.

Não foi possível efetuar a análise em todos os genes selecionados para os quais foram desenhados *primers*, pelo que numa primeira fase foi efetuada sequenciação de Sanger nos genes *RecQ protein-like 5 (RECQL5)*, *RWD domain containing 4 (RWDD4)*, *Cell adhesion molecule 1 (CADM1)*, *plexin B1 (PLXNB1)* e *24-dehydrocholesterol reductase (DHCR24)* (Tabela 3.18).

Tabela 3.18: Resultado da análise WES nos genes *RECQL5*, *RWDD4*, *CADM1*, *PLXNB1* e *DHCR24*

Gene	Variante	Tipo de alteração	Número de reads	Frequência	Previsão		
					<i>Mutation taster</i>	SIFT	<i>Polyphen</i>
<i>RECQL5</i>	c.1586-1G>CAG	variante de <i>splicing</i>	72	0,75	Patogénica	-	-
<i>RWDD4</i>	c.364-2T>-	variante de <i>splicing</i>	40	0,53	Patogénica	-	-
<i>CADM1</i>	c.1070T>G p.V357G	variante <i>missense</i>	34	0,47	Patogénica	Patogénica	Patogénica
<i>PLXNB1</i>	c.5197A>C p.T1733P	variante <i>missense</i>	94	0,38	Patogénica	Patogénica	Patogénica
<i>DHCR24</i>	c.983A>C p.Y328S	variante <i>missense</i>	27	0,43	Patogénica	Patogénica	Patogénica

O gene *RECQL5* é um membro da família das helicases *RecQ*, que são enzimas com função de manutenção da estabilidade genómica. Assim, o *RECQL5* está implicado na replicação do DNA, transcrição e processos de reparação e a sua deficiência contribui para a instabilidade genómica. Até ao momento, não há doenças conhecidas ou síndromes associadas com variantes ou mutações no *RECQL5* [Lao et al., 2013]. Estudos anteriores mostraram que modelos de ratos deficientes para *RECQL5* têm um aumento na incidência de tumores formados no cólon. Como o *RECQL5* é altamente conservado entre os ratos e humanos, estes resultados são consistentes para o *RECQL5* com um papel como gene supressor de tumor em CCR humano [Hu et al., 2010].

O gene *RWDD4* interage com o gene *BRCA1*, que codifica uma fosfoproteína nuclear com um papel importante na manutenção da estabilidade genómica e atua como gene

supressor de tumor [NCBI, 2014c]. Esta interação pode levar a que o gene *RWDD4* também esteja envolvido na estabilidade genômica. Também foi observado em estudos anteriores que os domínios RWD são os principais módulos de interação em pelo menos cinco proteínas do cinetócoro diferentes, que sugere que se encontre este elemento estrutural em vários outros componentes do cinetócoro. Alterações no cinetócoro levam à instabilidade cromossômica [Schmitzberger and Harrison, 2012].

O gene *CADM1* codifica uma proteína de membrana de passagem única, que medeia a adesão célula-célula homofílica e adesão célula-célula heterofílica com *CADM3* e *PVRL3*, de uma forma independente de cálcio. Além disso, o *CADM1* afeta a localização de outros recetores de adesão, tais como a E-caderina e a integrina $\alpha6\beta4$ [Moiseeva et al., 2014]. O *CADM1* é essencial para a saúde humana e está implicado em várias doenças, tais como cancro, onde a sua perda da função poderá conduzir à invasão de células cancerosas e/ou metástase. No entanto, o mecanismo subjacente de supressão de tumor pelo gene *CADM1* não está esclarecido [Sakurai-Yageta et al., 2009].

O gene *PLXNB1* faz parte da família de recetores transmembranares para a semaforina Sema4D, que são recetores importantes para a orientação de axónios, angiogénese, mas também em cancro. Estes recetores estão envolvidos em funções celulares que estão frequentemente alterados em células neoplásicas, tais como a adesão, migração e apoptose [Damola et al., 2013]. Em estudos anteriores foram encontradas mutações somáticas *missense* no gene *PLXNB1* em tumores primários e metástases de cancros da mama e da próstata [Tong et al., 2008].

O gene *DHCR24* codifica a redutase 3β -hidroxi-esterol- $\Delta24$, que catalisa a síntese de colesterol a partir de desmosterol e pertence a uma família de óxido-redutases dependente de *flavin adenine dinucleotide* (FAD). Tem sido demonstrado que a *DHCR24* é uma enzima multifuncional residente do retículo, que possui atividades anti-apoptóticas e de síntese de colesterol [Waterham et al., 2001]. A *DHCR24* também interage com o *TP53* e aumenta a sua estabilidade, regulando assim o crescimento celular, senescência e a apoptose. Desta forma, a diminuição da expressão do gene *DHCR24* está associada à apoptose e a sua expressão é necessária para a sobrevivência das células. Por outro lado, a sobreexpressão do *DHCR24* pode proteger a célula da apoptose induzida por *stress* oxidativo [Lu et al., 2008].

A sequenciação de Sanger foi efetuada em 12 membros da família afetados e saudáveis, indicados na árvore genealógica (Figura 3.7), para estudar a segregação das mutações com a doença na família. Alguns indivíduos saudáveis ainda são jovens e podem vir a desenvolver adenomas, mas o indivíduo L1956 já tem uma idade avançada e por isso a probabilidade de não ser portador é elevada, logo poderá servir como "controlo normal".

No gene *RECQL5* a mutação detetada pela análise de WES foi c.1586-1G>CAG. Após a comparação das sequências das amostras em estudo com a sequência de referência e também com recurso à análise *in silico*, verificou-se que esta alteração poderia ser c.1586-2_1586-1insCA, isto porque o efeito seria o mesmo, ocorrer apenas a inserção traduz-se da mesma forma que ocorrer a deleção juntamente com a inserção e também porque se encontra descrita. Na sequenciação de Sanger, para além desta alteração, ainda se

verificaram outras duas, c.1586-6C>A e c.1586-4A>-. A alteração inicial aparece juntamente com a c.1586-6C>A em homozigotia em 8 indivíduos e em heterozigotia em 4 indivíduos. Por outro lado, a alteração c.1586-4A>- surge em todos os indivíduos em heterozigotia (Figura 3.7).

Foi efetuada uma análise de segregação, mas aparentemente as alterações não segregam com a doença. Verifica-se que há indivíduos afetados com carcinoma (L801, L408, L674), com adenoma (L781, L1794), com adenoma e carcinoma noutra órgão (L1104) e um indivíduo aparentemente saudável (L1956) com o mesmo genótipo (Figura 3.7).

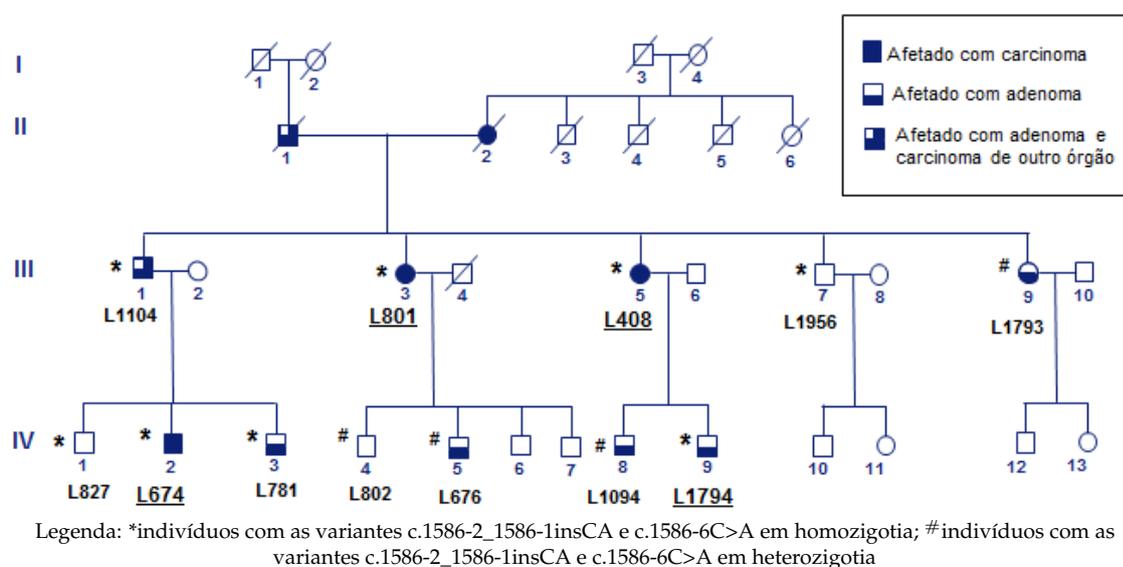


Figura 3.7: Árvore genealógica da família selecionada para a análise de WES (L56)

Tabela 3.19: Resultados da análise *in silico* às variantes germinais encontradas no gene *RECQL5* com recurso ao *software Mutation taster*

Variante	Mutation taster Previsão	Mutation Taster Alterações de <i>splicing</i>		
		Efeito	Posição em relação à variante	Score
		c.1586-2_1586-1insCA	Patogénica	sítio aceitador de <i>splicing</i> perdido
c.1586-6C>A	Polimorfismo	sítio aceitador de <i>splicing</i> aumentado	2 pb upstream	wt: 0.28 mu: 0.91
		sítio aceitador de <i>splicing</i> aumentado	3 pb downstream	wt: 0.28 mu: 0.40
c.1586-4A>-	Patogénica	sítio aceitador de <i>splicing</i> aumentado	1 pb downstream	wt: 0.28 mu: 0.39
		sítio dador de <i>splicing</i> ligeiramente aumentado	2 pb downstream	wt: 0.9975 mu: 0.9976
			5 pb downstream	wt: 0.9470 mu: 0.9570
		ganho de sítio aceitador de <i>splicing</i>	7 pb downstream	0.62
2 pb downstream	0.45			

Legenda: wt - *wild-type*; mu - mutado

Após a sequenciação do gene *RWDD4*, verificou-se que a mutação surge em todos

os elementos da família em estudo. Em relação à análise nos genes *CADM1*, *PLXNB1* e *DHCR24*, as alterações indicadas pela análise de WES não foram encontradas em nenhum dos indivíduos da família, incluindo as amostras que foram enviadas para o estudo do exoma.

4

Discussão

As famílias FCCTX são caracterizadas pela ausência de mutações germinais nos genes MMR, pelo preenchimento dos CA e também por apresentarem tumores sem MSI. No entanto, a causa molecular deste síndrome ainda é desconhecida, sendo deste modo necessário o estudo de genes de suscetibilidade para o FCCTX.

Em estudos prévios, o grupo de Gastreterologia da Unidade de Investigação de Patobiologia Molecular do IPOLFG, E.P.E. de Lisboa, demonstrou a presença de *linkage* com a doença numa família FCCTX, nas regiões 13q e 21q. Posteriormente, estas regiões foram delimitadas a uma região mínima de LOH de 0,87Mb em 13q33 e 1,3Mb em 21q11, com recurso a análise de perdas de heterozigotia utilizando marcadores de microssatélites. Verificou-se que a LOH em 13q ocorre maioritariamente em adenomas e em 21q em carcinomas. Desta forma sugere-se que a iniciação tumoral está associada a alterações na região 13q, enquanto a progressão tumoral se correlaciona com alterações na região 21q.

4.1 Caracterização molecular dos tumores de indivíduos com FCCTX e impacto na suscetibilidade para o FCCTX

4.1.1 Alterações de *copy-number* e de metilação nos genes MMR e *MGMT*

Na análise de *copy-number* efetuada nos tumores e mucosas normais, as famílias TSG+ apresentam um generalizado ganho nos genes MMR e *MGMT*, enquanto o grupo de famílias TSG- quase não apresenta ganhos, apenas se verificam alguns no gene *MGMT*. Também nas amostras de mucosa normal das famílias TSG+ já se observam estes ganhos, indicando que estas serão alterações que surgem precocemente.

As famílias TSG- são caracterizadas predominantemente por perdas nos genes MMR,

à exceção da amostra 1541 da família P2209, que apenas apresenta ganhos. De facto, esta amostra apresenta alterações que se aproximam mais do grupo de famílias TSG+. Também se verificaram perdas na região *upstream* e exão 1 do gene *MSH6*. Apesar de existir a possibilidade de um polimorfismo no local de ligação da sonda poder ser responsável por esta não se ligar e resultar numa falsa deleção, as alterações observadas no gene *MSH6* deverão corresponder a perda, porque as respetivas amostras de sangue periférico não demonstram estas alterações.

Assim observam-se resultados opostos em termos de frequências de perdas e ganhos, nas amostras de tumores das famílias TSG+ e TSG-. Estes dados apoiam o facto de realmente existirem duas entidades moleculares distintas entre as famílias FCCTX.

Observaram-se ganhos consistentes nos genes MMR e *MGMT* (ganhos em pelo menos 3 sondas) nas famílias L7, L24, L55, L63, L84, L148 e L173. Também se efetuou o mesmo estudo, mas neste caso para o estado de metilação, onde se observou metilação em pelo menos 3 sondas nas famílias L55, L63 e L173. Desta forma, dentro das famílias TSG+, parece existirem 2 grupos distintos: famílias que apresentam ganhos e metilação (L55, L63 e L173) e famílias que apresentam apenas ganhos (L7, L24, L84 e L148). É de salientar que a família L63 é incluída no grupo de ganhos frequentes em pelo menos 3 sondas devido às alterações observadas na mucosa normal.

É de notar que estas famílias que apresentam ganhos e hipermetilação nos genes MMR e *MGMT*, muito precocemente, com alterações já na mucosa normal, são as que, desta série de famílias FCCTX, apresentam mais frequentemente adenomas tubulovilosos e adenomas serrados tradicionais (L55, L63 e L173). Num estudo anterior efetuado pelo grupo de Gastrenterologia da UIPM, este tipo de lesões, as quais também apresentavam alterações moleculares semelhantes, foi associado a um mecanismo de tumorigénese relacionado com a resposta a *stress* oxidativo/agentes alquilantes. Este facto, leva a que se coloque uma hipótese para o mecanismo de iniciação tumoral nestas famílias.

Uma mutação causada por um *stress* oxidativo/agentes alquilantes iria recrutar as proteínas codificadas pelo gene *MGMT* para a sua reparação. Estando este metilado, não é capaz de reparar as células mutadas, sendo recrutadas as proteínas codificadas pelos genes MMR (*MSH6/MSH2*). Se estes também se encontram metilados, como se observou em algumas amostras estudadas, ocorrem ciclos fúteis de replicação e reparação, onde as células mutadas proliferam por não haver um sistema de reparação funcional [Liu and Gerson, 2006]. Há desta forma quebras no DNA em dupla cadeia e, posteriormente, alterações pontuais de *copy-number* em vários genes (Figura 4.1). Assim os tumores formados vão-se caracterizar essencialmente por alterações cromossómicas, nomeadamente ganhos. Desta forma, este mecanismo poderá estar na base da iniciação tumoral nestas famílias TSG+ e eventualmente uma deficiência num dos genes envolvidos na resposta a este tipo de erros no DNA poderá contribuir para a suscetibilidade para o FCCTX. O facto destas famílias TSG+ não apresentarem também uma maior frequência de ganhos no gene *APC* em relação às TSG- suporta ainda mais este mecanismo, uma vez que estas alterações não parecem ser generalizadas a qualquer gene.

Assim, entre as famílias TSG+ há várias diferenças e parece haver alguns casos em que esta assinatura molecular é menos evidente do que outros. Sendo este grupo muito heterogêneo, é muito provável a existência de mais que uma entidade molecular nestas famílias. No que diz respeito a alterações de *copy-number* e metilação dos genes MMR e *MGMT* nas amostras de sangue, não se verificaram muitas alterações. As alterações detetadas podem ter sido devido a polimorfismos no local de ligação das sondas, visto não serem muito comuns alterações nestes genes ao nível do sangue. De qualquer forma, estes valores não se podem descartar e deve ser averiguada a sua veracidade através de outra metodologia, nomeadamente em relação às detetadas no gene *MGMT*.

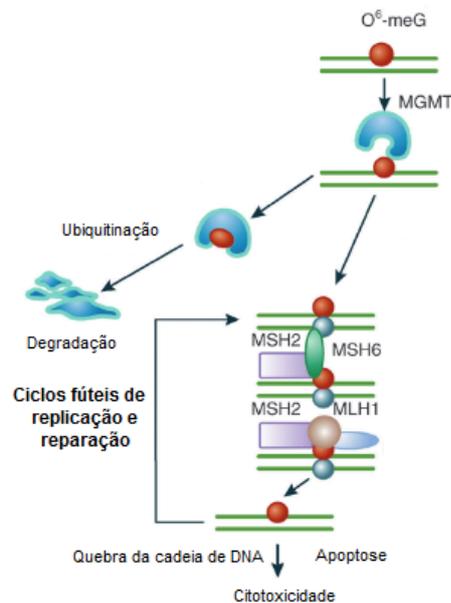


Figura 4.1: Mecanismo de reparação de erros no DNA originados por *stress* oxidativo/agentes alquilantes

(Adaptado de Liu and Gerson, 2006)

4.1.2 Alterações de *copy-number* e de metilação no gene *APC*

Na análise de *copy-number* ao gene *APC* efetuada nos tumores e mucosas normais, não se observou uma grande distinção entre as famílias TSG+ e TSG-, visto as suas frequências de ganhos e perdas rondarem os mesmo valores. Os ganhos neste gene existem predominantemente em regiões promotoras, que podem ter como consequência a sobreexpressão do gene *APC*. Ainda se verifica que estes ganhos ocorrem em lesões precoces, pois já se verificam nas mucosas normais nos pares 437/438, 459/460, 569/570 e 627/628 de amostras de tecido.

A hipermetilação das ilhas CpG na porção a montante do gene *APC* (principalmente promotor 1A e 1B), como potencial mecanismo alternativo de inativação do gene supressor de tumor, está descrita em CCR [Tsuchiya et al., 2000]. As alterações de metilação nas amostras de sangue não são significativas, observando-se maiores alterações nas amostras

de adenomas, carcinomas e mucosas normais. Observa-se principalmente hipermetilação no exão 1A, sendo este o promotor que é o mais vulgarmente ativo. No entanto, estas amostras foram incluídas em parafina e podem conter alguns contaminantes que possam interferir com a reação de MS-MLPA. As amostras 437, 460, 991, 992 e 993 apresentam uma deficiente qualidade do DNA, havendo assim um aumento da percentagem residual de *primers*. Para estas amostras obteve-se uma hipermetilação dos promotores 1A e 1B (exceto o 437), podendo assim haver influência da qualidade do DNA no valor da metilação. Por outro lado, também o facto de o *kit* MS-MLPA utilizado para analisar o gene *APC* conter poucas sondas, vai fazer com que haja um aumento de *primers* residuais, não se podendo excluir desta forma estas metilações, pois tanto podem ser consequência da técnica e do tipo de amostras utilizadas, como podem ser reais e contribuir para a carcinogénese do cólon e reto. De qualquer forma, o padrão de metilação observado nas amostras de tumores de famílias TSG+ é diferente daquele observado nas famílias TSG-. Enquanto os primeiros apresentam predominantemente hipometilação das regiões *downstream* ao promotor, no segundo grupo verificam-se mais hipermetilações nos promotores. Estes resultados parecem sugerir também papéis distintos do gene *APC* nos dois grupos de famílias.

No caso da análise de *copy-number* e metilação ao nível germinal, deve-se valorizar as alterações de *copy-number* (perdas e ganhos) nos promotores N e na sonda a 50kb do exão 1, tanto nas famílias com TSG+ como nas TSG-. Verifica-se uma maior frequência de perdas em valores de homozigotia, que parecem ser devido a polimorfismos que estão descritos nestas regiões. Caso isso se verifique, as sondas não se ligam e o *software* de análise das amostras interpreta como perda alélica. No entanto, para se comprovar essa hipótese, seria necessário observar também essas perdas nos tumores ou mucosas normais correspondentes a cada amostra de sangue. Verificam-se perdas em ambas as amostras do par CAs802s/801P e na sonda *upstream* ao exão 1, nas amostras E436/E436T e L701/L701T, pelo que se pode considerar tratarem-se de polimorfismos.

Esta região pode ser mais importante ao nível somático, porque há tumores que têm deleção e a amostra de sangue correspondente não tem (famílias 24, 55 e 63). De qualquer forma, tendo em conta as frequências de perdas e ganhos nas amostras de sangue, verifica-se que as proporções se mantêm em todas as sondas nos dois grupos de famílias (TSG+ e TSG-).

A análise de expressão do gene *APC* a nível germinal nos indivíduos da família L55 não pareceu confirmar as alterações de *copy-number* nesta família. No entanto, verificou-se uma sobreexpressão na amostra L5, que a confirmar-se, por exemplo, por análise de segregação na família, se pode colocar como hipótese conduzir à inibição da reparação de erros por *stress* oxidativo pela via BER, e, conseqüentemente à suscetibilidade para o desenvolvimento de adenomas/carcinoma.

Encontra-se descrito na literatura que a proteína *APC* interatua com a *DNA polymerase-β* e com a *FEN1*, duas enzimas chave da via BER. Em vários estudos também se observou que pode ocorrer a sobreexpressão do gene *APC* em resposta ao *stress* induzido

por agentes alquilantes do DNA [Jaiswal and Narayan, 2008]. Esta sobreexpressão pode fazer com que o gene *APC* perda a sua função de gene supressor de tumor e como este interatua com duas enzimas da via BER, há um bloqueio da via BER e o processo carcinogénico é estimulado [Jaiswal et al., 2013]. Por outro lado, a inativação do sistema BER está associada ao desenvolvimento de CCR, apoiando esta hipótese como suporte para a sobreexpressão do gene *APC* num indivíduo com FCCTX.

A deficiência do sistema BER é perfeitamente compatível com a assinatura molecular dos tumores correspondentes às amostras em estudo. Esta assinatura molecular passa pela metilação dos genes *MMR* e *MGMT* e alterações cromossómicas frequentes no gene *KRAS*. Pequenas alterações no gene *APC* podem contribuir para o FCCTX noutras amostras e como esta assinatura se estende a mais tumores de outras famílias, deveria ser efetuada análise de expressão para essas restantes famílias.

Outro facto a favor de um papel importante do gene *APC* na tumorigénese associada ao FCCTX é que os carcinomas se localizam preferencialmente no reto e os adenomas no cólon proximal. Os tumores FCCTX apresentam elevada LOH, que é mais vantajosa para a tumorigénese no reto, ou seja, para a formação de carcinomas.

A sinalização de β -catenina necessária para atingir um limiar hipotético para a iniciação do tumor é específica para cada tecido. Anteriormente pensava-se no cólon como um todo, mas verificou-se que diferentes tipos de CCR ocorriam predominantemente no cólon distal/reto ou proximal. Desta forma, essa predominância pode em parte ser explicada por mutações em genes responsáveis pela maior ou menor sinalização da β -catenina, como o *APC*, pois identificaram-se diferenças regionais na sua sinalização ao longo do trato gastrointestinal. De acordo com o modelo de sinalização '*just-right*', apenas as mutações que fornecem vantagem para o crescimento de um determinado tipo de célula, ao mesmo tempo que evitam a indução de apoptose, irão induzir com sucesso à formação de tumores. Assim, seria necessário um determinado grau de deficiência do *APC*, para permitir a acumulação suficiente de β -catenina e ativação dos genes-alvo a jusante para a formação do tumor [Albuquerque et al., 2011].

O reto apresenta menos sinalização basal da via Wnt. Foi observado que tumores do cólon distal/reto preferencialmente adquirem mutações com retenção de uma repetição de 20 aminoácidos no gene *APC* e que apresentam níveis de sinalização maiores da β -catenina, logo conferem uma maior ativação da via Wnt, sendo mais vantajosas no reto.

4.2 Análise mutacional dos genes *KDELC1* e *ERCC5*

A análise mutacional dos genes *KDELC1* e *ERCC5* permitiu a identificação de um total de 7 alterações: 3 variantes *missense*, 1 variante silenciosa e 3 variantes intrónicas. O facto destas variantes polimórficas não terem sido associadas até ao momento ao desenvolvimento de CCR e de algumas possuírem uma frequência bastante elevada na população ibérica e nas amostras analisadas, sugere que não estejam envolvidas na suscetibilidade para o FCCTX. No entanto, não se pode descartar a hipótese de que a

combinação de algumas destas variantes possa conferir risco para o CCR.

4.2.1 *KDELC1*

A análise de mutações germinais no gene *KDELC1* identificou 2 variantes intrónicas na região 5'UTR (1-108T>G e c.1-30G>T), uma do tipo *missense* (c.455A>T) e outra variante do tipo silenciosa (c.48A>G). As variantes intrónicas em princípio não irão alterar o *splicing*, pois localizam-se na região 5'UTR. No entanto podem alterar a expressão do gene *KDELC1*, visto estarem próximas do promotor deste gene.

A variante c.455A>T está descrita como SNV (rs200215559) mas não apresenta dados de frequência associada nas várias populações. Foi identificada em apenas um indivíduo incluído no grupo de 17 indivíduos a estudo e está classificada como patogénica segundo o *software Mutation taster*. Apesar de se obter esta classificação, não há alteração dos locais de *splicing* e as características da proteína codificada mantêm-se, mesmo havendo alteração de um aminoácido. Para além disso, com base nos *softwares* SIFT e *Polyphen*, esta variante é considerada tolerada e benigna, respetivamente. Esta alteração está inserida no domínio *Lipopolysaccharide-modifying protein*. Juntando todos estes fatores, é bastante sugestivo que esta variante não esteja relacionada com a tumorigénese do cólon e reto.

Dado que a frequência das outras alterações identificadas é elevada, nas amostras em estudo como na população descrita, sugere-se que estas não serão patogénicas. Não se pode excluir o gene *KDELC1* em relação à suscetibilidade para o FCCTX. As duas alterações localizadas na região 5'UTR e a localizada no exão 1, estão muito próximas umas das outras e surgem todas no mesmo indivíduo.

4.2.2 *ERCC5*

A análise mutacional do gene de suscetibilidade para o FCCTX *ERCC5* permitiu a identificação de 2 variantes do tipo *missense* (c.3310G>C e c.2636A>G) e 1 variante intrónica (c.2878+14C>T). As variantes c.3310G>C e c.2878+14C>T encontram-se com uma elevada frequência na população europeia e estão descritas como SNVs (rs4150360 e rs17655).

A variante c.3310G>C obteve resultados ambíguos na análise *in silico*. Enquanto os *softwares* SIFT e *Polyphen* classificaram a variante como deletéria e provavelmente prejudicial, o *software Mutation Taster* classificou-a como polimorfismo. Em 2002 foi efetuada uma meta-análise por Zhu et al., com base em 49 estudos de caso-controlo, com um total de 23.490 casos de cancro e 27.168 controlo, onde avaliaram de forma abrangente a associação entre a variante c.3310G>C e o risco de diferentes tipos de cancro, incluindo o CCR. Esta meta-análise incluiu vários tipos de cancros (mama, pele, pulmões, bexiga, cabeça e pescoço, CCR e linfoma não-Hodgkin), diferentes etnias (caucasiana, afro-americana e asiática) e utilizou diferentes fontes de amostras controlo (hospital e população). No entanto este estudo mostrou que o polimorfismo c.3310G>C parece não conferir suscetibilidade aos diversos tipos de cancro em estudo. Desta forma também será de excluir este polimorfismo em relação à suscetibilidade para o FCCTX nas nossas

amostras.

Quando se efetuou a análise *in silico* à variante c.2636A>G verificou-se que esta é classificada como patogénica. Estão previstas alterações ao nível do *splicing* a montante do local da variante e as características da proteína codificada também podem ficar alteradas. Como se verificou que a frequência nos indivíduos em estudo era mais elevada que a descrita na base de dados *1000 Genomes*, analisou-se o exão que contém esta alteração num conjunto de indivíduos saudáveis, onde se obteve uma frequência de 2%, muito semelhante à descrita. Esta alteração inclui-se no domínio *XPG/Rad2 endonuclease*.

Para descartar esta variante em relação ao seu envolvimento na tumorigénese do cólon e reto, será necessário amplificar o cDNA das amostras que contém a alteração. A amplificação deve ser efetuada entre os exões 11, 12 e 13, visto a alteração se encontrar no exão 13, e desta forma irá averiguar-se a existência de alterações ao nível do *splicing*. Também se deve efetuar uma análise de segregação, visto a série de amostras utilizada ser pequena e se ter detetado esta alteração em duas famílias.

Em relação às restantes alterações identificadas neste gene, a sua elevada frequência tanto na população em estudo como na descrita levam a sugerir que não serão patogénicas.

4.3 Análise do envolvimento do *splicing* no potencial exão 13A do gene *TPP2* na suscetibilidade para o FCCTX

Em estudos prévios foi identificado um transcrito alternativo no gene *TPP2* com uma inserção de 72pb, que corresponde à inserção de um possível exão entre os exões 13 e 14 (exão 13A).

A análise mutacional dos exões 13, 14 e hipotético exão 13A do gene *TPP2* efetuada no presente estudo, permitiu a identificação de 2 variantes intrónicas (c.1679-23A>G e c.1679-94C>T). Na comparação das frequências destas variantes nos indivíduos a estudo com as descritas na base de dados *1000 Genomes* para a população europeia, verificou-se que a primeira variante em heterozigotia era mais frequente nos indivíduos a estudo e que a segunda variante em homozigotia era menos frequente. Desta forma, estudaram-se estas alterações num conjunto de indivíduos saudáveis, onde se obtiverem frequências semelhantes às descritas na base de dados *1000 Genomes*.

Provavelmente cada uma destas alterações sozinha pode não conferir vantagem para o *splicing* preferencial na região que conduz à expressão do exão hipotético 13A, mas as duas em conjunto podem adquirir alguma vantagem. Desta forma, com recurso a análises de expressão efetuadas anteriormente nos mesmos indivíduos, fez-se uma análise de segregação da expressão das duas isoformas com as variantes encontradas. No entanto, não se conseguiu encontrar nenhuma relação, haverá outro fator desconhecido que influencia esta expressão diferencial entre os dois transcritos. De qualquer forma, também pode ter ocorrido um enviesamento das frequências das variantes encontradas, pois das 24 amostras de indivíduos com síndrome de Lynch utilizadas, estas provêm

apenas de 7 famílias. Assim, serão de excluir estas variantes em relação à sua influência no *splicing* alternativo desta região.

Após estes resultados, foram desenhados *primers* para análise mutacional da região do intrão 12, exões 13 e 14 deste gene, com o objetivo de encontrar uma mutação responsável por este *splicing* alternativo. A amplificação destas sequências não foi efetuada em tempo útil, sendo necessário a continuação do seu estudo.

Deste modo, a análise mutacional do gene *TPP2* não se encontra completa, não sendo possível excluir a presença de outras mutações potencialmente patogênicas para o desenvolvimento de FCCTX, ou que possam ser responsáveis pelo *splicing* alternativo. Mesmo que se verifique que a expressão do exão hipotético 13A não está relacionada com o desenvolvimento de FCCTX, é importante desvendar os mecanismos inerentes a este *splicing* alternativo, pois trata-se de uma isoforma que ainda não se encontra descrita.

4.3.1 Quantificação da expressão do gene *TPP2*

Com base em estudos anteriores, já se tinha conhecimento que os transcritos wt e alternativo tinham diferentes expressões nas amostras de sangue de famílias FCCTX e síndrome de Lynch. Normalmente quando havia redução do transcrito wt, o transcrito alternativo estava sobreexpresso. No entanto não se tinha conhecimento de como seria a expressão destes transcritos no cólon, ou mesmo se eram expressos. Assim, na análise de quantificação da expressão do gene *TPP2* utilizaram-se amostras de sangue provenientes de indivíduos saudáveis (A, B, C e D), amostras de carcinomas (C1 e C2), adenomas (A1 e A2) e mucosas normais.

Neste estudo observa-se uma maior expressão dos transcritos wt e alternativo nas amostras de sangue comparando com as amostras de tecidos. Nestas amostras o transcrito alternativo tem valores de expressão muito reduzidos, sendo a expressão nesta ordem de valores considerada nula. Desta forma verifica-se que realmente este transcrito alternativo não é expresso no cólon.

A amostra C2 corresponde a um carcinoma de um indivíduo de síndrome de Lynch, enquanto o indivíduo C1 provém de um indivíduo que não tem mutação identificada. Comparando estas duas amostras, observa-se que o C2 tem uma expressão muito reduzida do transcrito wt. Neste caso pode ocorrer uma compensação, ou seja, como este indivíduo tem uma sobreexpressão do transcrito alternativo no sangue, no carcinoma vai ter redução de expressão do transcrito wt.

Desta forma, é necessário repetir a análise de expressão para estas amostras, visto a redução de expressão do transcrito wt poder contribuir para uma suscetibilidade para o desenvolvimento de adenomas.

4.4 Análise de WES

A análise de WES detetou a alteração c.1586-1G>CAG no gene *RECQL5*. No entanto, esta variante parece corresponder apenas a uma inserção de um CA, visto o efeito das duas alterações (G>CAG e insCA) ser o mesmo. Também quando se efetua a análise *in silico* com recurso no *software* bioinformático *Mutation taster* o *output* que este nos fornece é que a alteração introduzida corresponde a c.1586-2_1586-1insCA. Esta alteração está descrita na base de dados dbSNP como DIV (*deletion/insertion variation* - rs142406301).

Nesta região estão descritas várias alterações, entre elas c.1586-3_1586-2dupAC e c.1586-4_1586-3insCA. Assim, a alteração na sequência de DNA podia ser qualquer uma destas 3 variantes, visto o efeito na sequência ser o mesmo. A variante c.1586-3_1586-2dupAC tem uma frequência de 49% em homozigotia e 33% em heterozigotia e as restantes não têm frequência associada.

Para além da alteração inicial descrita pela análise de WES, ao analisar-se as sequências obtidas através da sequenciação de Sanger ainda se verificaram outras duas, c.1586-6C>A e c.1586-4A>-. A primeira foi detetada em heterozigotia e homozigotia, enquanto a segunda apenas surge em heterozigotia.

Na análise da alteração c.1586-2_1586-1insCA (ou c.1586-3_1586-2dupAC) com recurso ao *Mutation Taster*, a previsão é de ser patogénica, onde é perdido um sítio aceitador de *splicing*. Esta alteração não tem frequência associada, mas a variante c.1586-3_1586-2dupAC (que origina o mesmo efeito) está descrita com uma frequência elevada, logo por si só não será patogénica. No entanto, há a possibilidade de as três variantes em conjunto terem um efeito patogénico. Para resolver estes resultados, será necessário analisar esta região ao nível do cDNA, compreendendo os exões 12, 13 e 14, visto as alterações se encontrarem no exão 13, e observar o que acontece ao nível do *splicing* nos vários elementos da família. De qualquer forma, foi constatado que o indivíduo que não expressa a doença (L1956) tem o mesmo fenótipo que os indivíduos afetados, logo à partida supõe-se que alterações não serão patogénicas. No entanto, em alguns casos está descrito haver uma penetrância incompleta, ou seja, pessoas que apresentam uma alteração patogénica podem ser saudáveis, por isso não se pode excluir o envolvimento deste gene na suscetibilidade para o desenvolvimento de FCCTX.

No gene *RWDD4* a alteração indicada pela análise de WES foi confirmada nos restantes elementos da família. Sendo uma alteração ao nível do *splicing* não se obtém previsão por parte do SIFT e *Polyphen*. Pelo facto de o indivíduo L1956 ser portador desta alteração, é uma indicação de que esta variante pode não estar relacionada com o desenvolvimento de FCCTX. Mesmo assim, sendo uma variante que altera o *splicing*, é necessário continuar o seu estudo e analisá-la nas amostras disponíveis da família ao nível do cDNA, sendo que até já foram desenhados *primers* para tal.

Nos genes *CADM1*, *PLXNB1* e *DHCR24*, não foi possível avaliar as alterações detetadas pela análise de WES, porque estas não se encontravam nos indivíduos da família a estudo. Estes falsos positivos podem ocorrer na análise de WES, pelo que nos genes

CADM1 e *DHCR24* pode ter sido devido ao reduzido número de *reads*, que foram 34 e 27 respectivamente.

O estudo dos genes candidatos à suscetibilidade para o FCCTX obtidos pela análise de WES não se encontra completo. Dos genes que foram selecionados e desenhados *primers* para análise das variantes identificadas, apenas foi efetuada análise mutacional em 5 genes. Assim, em estudos posteriores será necessário continuar o estudo das duas variantes de *splicing* já analisadas e das restantes que não foram analisadas em tempo útil.

Bibliografia

Adzhubei, I. A., Schmidt, S., Peshkin, L., Ramensky, V. E., Gerasimova, A., Bork, P., Kondrashov, A. S., and Sunyaev, S. R. (2010). A method and server for predicting damaging missense mutations. *Nat Methods*, 7(4):248–249.

Albuquerque, C., Bakker, E. R., van Veelen, W., and Smits, R. (2011). Colorectal cancers choosing sides. *Biochim Biophys Acta*, (2):219–231.

Albuquerque, C., Breukel, C., van der Luijt, R., Fidalgo, P., Lage, P., Slors, F. J., Leitão, N. C., Fodde, R., and Smits, R. (2002). The 'just-right' signaling model: APC somatic mutations are selected based on a specific level of activation of the beta-catenin signaling cascade. *Hum Mol Genet*, 11(13):1549–1560.

Aracil, M., Dauffenbach, L. M., Diez, M. M., Richeh, R., Moneo, V., Leal, J. F. M., Fernández, L. F. G., Kerfoot, C. A., and Galmarini, C. M. (2013). Expression of XPG protein in human normal and tumor tissues. *Int J Clin Exp Pathol*, 6(2):199–211.

Belo, H. (2010). Papel Do Gene APC e Identificação de Regiões Cromossômicas Envolvidas No Síndrome Familiar de Câncer Do Cólon e Recto Do Tipo X. Tese de Mestrado, Escola Superior de Saúde Egas Moniz.

Biosoft, P. (2013). NetPrimer. www.premierbiosoft.com/netprimer/.

Boland, C. R. (2005). Evolution of the Nomenclature for the Hereditary Colorectal Cancer Syndromes. *Familial Cancer*, 4(3):211–218.

Boland, C. R. and Goel, A. (2010). Microsatellite instability in colorectal cancer. *Gastroenterology*, 138(6):2073–2087.

Claudio, J. O., Zhu, Y. X., Benn, S. J., Shukla, A. H., McGlade, C. J., Falcioni, N., and Stewart, A. K. (2001). HACS1 encodes a novel SH3-SAM adaptor protein differentially expressed in normal and malignant hematopoietic cells. *Oncogene*, 20(38):5373–5377.

- Coffa, J. and van den Berg, J. (2011). Analysis of MLPA data using novel software Coffalyser .NET by MRC-Holland. *Modern approaches to quality control, Dr. Ahmed Badr Eldin (Ed.), InTech*, pages 125–50.
- Colussi, D., Brandi, G., Bazzoli, F., and Ricciardiello, L. (2013). Molecular pathways involved in colorectal cancer: Implications for disease behavior and prevention. *Int J Mol Sci*, 14(8):16365–16385.
- Coppedè, F., Lopomo, A., Spisni, R., and Migliore, L. (2014). Genetic and epigenetic biomarkers for diagnosis, prognosis and treatment of colorectal cancer. *World J Gastroenterol*, 20(4):943–956.
- Damola, A., Legendre, A., Ball, S., Masters, J. R., and Williamson, M. (2013). Function of mutant and wild-type *Plexin-B1* in prostate cancer cells. *The Prostate*, 73(12):1326–1335.
- de la Chapelle, A. (2004). Genetic predisposition to colorectal cancer. *Nat Rev Cancer*, 4(10):769–780.
- Dominguez-Valentin, M., Therkildsen, C., Da Silva, S., and Nilbert, M. (2014). Familial colorectal cancer type X: genetic profiles and phenotypic features. *Mod Pathol*, pages 924–930.
- Emmert, S., Schneider, T. D., Khan, S. G., and Kraemer, K. H. (2001). The human XPG gene: gene architecture, alternative splicing and single nucleotide polymorphisms. *Nucleic Acids Res*, 29(7):1443–1452.
- Francisco, I., Albuquerque, C., Lage, P., Belo, H., Vitoriano, I., Filipe, B., Claro, I., Ferreira, S., Rodrigues, P., Chaves, P., Leitão, C. N., and Pereira, A. D. (2011). Familial colorectal cancer type X syndrome: two distinct molecular entities? *Familial Cancer*, 10(4):623–631.
- Gallagher, D. J., Smith, J. D., Offit, K., and Stadler, Z. K. (2010). Diagnosing hereditary colorectal cancer. *Clin Colorectal Cancer*, 9(4):205–211.
- Grünhage, F., Jungck, M., Lamberti, C., Keppeler, H., Becker, U., Schulte-Witte, H., Plassmann, D., Friedrichs, N., Buettner, R., Aretz, S., et al. (2008). Effects of common haplotypes of the ileal sodium dependent bile acid transporter gene on the development of sporadic and familial colorectal cancer: a case control study. *BMC Med Genet*, 9(1):70.
- Hennig, E. E., Mikula, M., Rubel, T., Dadlez, M., and Ostrowski, J. (2012). Comparative kinome analysis to identify putative colon tumor biomarkers. *J Mol Med*, 90(4):447–456.
- Hu, Y., Lu, X., and Luo, G. (2010). Effect of *Recql5* deficiency on the intestinal tumor susceptibility of *Apcmin* mice. *World J Gastroenterol*, 16(12):1482–1486.
- Jaiswal, A. S. and Narayan, S. (2008). A novel function of adenomatous polyposis coli (*APC*) in regulating DNA repair. *Cancer Lett*, 271(2):272–280.

- Jaiswal, A. S., Panda, H., Pampo, C. A., Siemann, D. W., Gairola, C. G., Hromas, R., and Narayan, S. (2013). Adenomatous polyposis coli-mediated accumulation of abasic dna lesions lead to cigarette smoke condensate-induced neoplastic transformation of normal breast epithelial cells. *Neoplasia*, 15(4):454 – 460.
- Jasperson, K. W., Tuohy, T. M., Neklason, D. W., and Burt, R. W. (2010). Hereditary and familial colon cancer. *Gastroenterology*, 138(6):2044–2058.
- Ku, C. S., Cooper, D. N., and Roukos, D. H. (2013). Clinical relevance of cancer genome sequencing. *World J Gastroenterol*, 19(13):2011–2018.
- Ku, C.-S., Cooper, D. N., Wu, M., Roukos, D. H., Pawitan, Y., Soong, R., and Iacopetta, B. (2012). Gene discovery in familial cancer syndromes by exome sequencing: prospects for the elucidation of familial colorectal cancer type X. *Mod Pathol*, 25(8):1055–1068.
- Lao, V. V., Welch, P., Luo, Y., Carter, K. T., Dzieciatkowski, S., Dintzis, S., Meza, J., Sarvetnick, N. E., Monnat, R. J., Loeb, L. A., and Grady, W. M. (2013). Altered RECQ Helicase Expression in Sporadic Primary Colorectal Cancers. *Trans Oncol*, 6:458–469.
- Ligtenberg, M. J. L., Kuiper, R. P., Chan, T. L., Goossens, M., Hebeda, K. M., Voorendt, M., Lee, T. Y. H., Bodmer, D., Hoenselaar, E., Hendriks-Cornelissen, S. J. B., Tsui, W. Y., Kong, C. K., Brunner, H. G., van Kessel, A. G., Yuen, S. T., van Krieken, Leung, S. Y., and Hoogerbrugge, N. (2008). Heritable somatic methylation and inactivation of *MSH2* in families with Lynch syndrome due to deletion of the 3' exons of *TACSTD1*. *Nat Genet*, 41(1):112–117.
- Lindor, N. M. (2009a). Familial Colorectal Cancer Type X: The Other Half of Hereditary Nonpolyposis Colon Cancer Syndrome. *Surg Oncol Clin N Am*, 18(4):637–645.
- Lindor, N. M. (2009b). Hereditary colorectal cancer: *MYH*-associated polyposis and other newly identified disorders. *Best Pract Res Clin Gastroenterol*, 23(1):75–87.
- Lindor, N. M., Rabe, K., Petersen, G. M., Haile, R., Casey, G., Baron, J., Gallinger, S., Bapat, B., Aronson, M., Hopper, J., et al. (2005). Lower cancer incidence in Amsterdam-I criteria families without mismatch repair deficiency: familial colorectal cancer type X. *Jama*, 293(16):1979–1985.
- Liu, F., Yuan, D., Wei, Y., Wang, W., Yan, L., Wen, T., Xu, M., Yang, J., and Li, B. (2012). Systematic review and meta-analysis of the relationship between *EPHX1* polymorphisms and colorectal cancer risk. *PloS one*, 7(8):e43821.
- Liu, L. and Gerson, S. L. (2006). Targeted modulation of *MGMT*: clinical implications. *Clin Cancer Res*, 12(2):328–331.
- Lu, X., Kambe, F., Cao, X., Kozaki, Y., Kaji, T., Ishii, T., and Seo, H. (2008). 3β -Hydroxysteroid- δ 24 reductase is a hydrogen peroxide scavenger, protecting cells from oxidative stress-induced apoptosis. *Endocrinology*, 149(7):3267–3273.

Lynch, H. T. and de la Chapelle, A. (2003). Hereditary colorectal cancer. *N Engl J Med*, 348(10):919–932.

Mayer, M. P. and Bukau, B. (2005). Hsp70 chaperones: cellular functions and molecular mechanism. *Cell Mol Life Sci*, 62(6):670–684.

McLaren, W., Pritchard, B., Rios, D., Chen, Y., Flicek, P., and Cunningham, F. (2010). Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor. *Bioinformatics*, 26(16):2069–2070.

Middeldorp, A., Van Eijk, R., Oosting, J., Forte, G., van Puijenbroek, M., van Nieuwenhuizen, M., Corver, W., Ruano, D., Caldes, T., Wijnen, J., et al. (2012). Increased frequency of 20q gain and copy-neutral loss of heterozygosity in mismatch repair proficient familial colorectal carcinomas. *Int J Cancer*, 130(4):837–846.

Moiseeva, E. P., Straatman, K. R., Leyland, M. L., and Bradding, P. (2014). *CADM1* controls actin cytoskeleton assembly and regulates extracellular matrix adhesion in human mast cells. *PLoS one*, 9(1):e85980.

Nagasaka, T., Goel, A., Notohara, K., Takahata, T., Sasamoto, H., Uchida, T., Nishida, N., Tanaka, N., Boland, C. R., and Matsubara, N. (2008). Methylation pattern of the O6-methylguanine-DNA methyltransferase gene in colon during progressive colorectal tumorigenesis. *Int J Cancer*, 122(11):2429–2436.

Najdi, R., Holcombe, R. F., Waterman, M. L., et al. (2011). Wnt signaling and colon carcinogenesis: beyond *APC*. *J Carcinog*, 10(1):5.

NCBI (2013). BLAST. www.ncbi.nlm.nih.gov/BLAST/.

NCBI (2014a). ClinVar. <http://www.ncbi.nlm.nih.gov/clinvar/variation/129011/>.

NCBI (2014b). dbSNP. <http://www.ncbi.nlm.nih.gov/projects/SNP/>.

NCBI (2014c). Gene. <http://www.ncbi.nlm.nih.gov/gene/79070>.

Ng, P. C. and Henikoff, S. (2003). SIFT: predicting amino acid changes that affect protein function. *Nucleic Acids Res*, 31(13):3812–3814.

Nygren, A. O. H. (2005). Methylation-Specific MLPA (MS-MLPA): simultaneous detection of CpG methylation and copy number changes of up to 40 sequences. *Nucleic Acids Res*, 33(14):e128.

Patel, S. G. and Ahnen, D. J. (2012). Familial Colon Cancer Syndromes: an Update of a Rapidly Evolving Field. *Curr Gastroenterol Rep*, 14(5):428–438.

- Pereira, C. (2013). Estudo do gene *NR1P1* e de novos *loci* de susceptibilidade para o cancro do cólon e recto familiar do tipo X. Tese de Mestrado, Escola Superior de Saúde Egas Moniz.
- Pereira, G. (2014). Estudo de regiões de susceptibilidade para o cancro do cólon e recto familiar do tipo X: análise de genes candidatos e de ganhos/deleções em tumores. Tese de Mestrado, Escola Superior de Saúde Egas Moniz.
- Pires, S. (2011). Identificação de regiões mínimas de susceptibilidade e envolvimento do gene *STK24* no Cancro do Cólon e Recto do tipo X. Tese de Mestrado, Escola Superior de Saúde Egas Moniz.
- Power, D. G. P., Glogowski, E., and Lipkin, S. M. (2010). Clinical genetics of hereditary colorectal cancer. *Hematology Oncology Clinics of North America*, 24(5):837–859.
- Preta, G., de Klark, R., Gavioli, R., and Glas, R. (2010). The enigma of tripeptidyl-peptidase II: dual roles in housekeeping and stress. *J Oncol*, 2010.
- Póvoa, V. (2011). Estudo de Genes de Susceptibilidade Para o Cancro Do Cólon e Recto Familiar Do Tipo X. Tese de Mestrado, Escola Superior de Saúde Egas Moniz.
- Rockel, B., Kopec, K. O., Lupas, A. N., and Baumeister, W. (2012). Structure and function of tripeptidyl peptidase II, a giant cytosolic protease. *Biochim Biophys Acta*, 1824(1):237—245.
- Sakurai-Yageta, M., Masuda, M., Tsuboi, Y., Ito, A., and Murakami, Y. (2009). Tumor suppressor *CADM1* is involved in epithelial cell structure. *Biochem Biophys Res Commun*, 390(3):977–982.
- Sayagués, J. M., Fontanillo, C., del Mar Abad, M., González-González, M., Sarasquete, M. E., del Carmen Chillón, M., Garcia, E., Bengoechea, O., Fonseca, E., Gonzalez-Diaz, M., et al. (2010). Mapping of genetic abnormalities of primary tumours from metastatic CRC by high-resolution SNP arrays. *PloS one*, 5(10):e13752.
- Schärer, O. D. (2008). *XPG*: its products and biological roles. *Adv Exp Med Biol*, 637:83–92.
- Schmitzberger, F. and Harrison, S. C. (2012). RWD domain: a recurring module in kinetochore architecture shown by a Ctf19-Mcm21 complex structure. *EMBO Rep*, 13(3):216–22.
- Schwarz, J. M., Rodelsperger, C., Schuelke, M., and Seelow, D. (2010). MutationTaster evaluates disease-causing potential of sequence alterations. *Nat Meth*, 7(8):575–576.
- Shen, L., Kondo, Y., Rosner, G. L., Xiao, L., Hernandez, N. S., Vilaythong, J., Houlihan, P. S., Krouse, R. S., Prasad, A. R., Einspahr, J. G., et al. (2005). *MGMT* promoter methylation and field defect in sporadic colorectal cancer. *J Natl Cancer Inst*, 97(18):1330–1338.
- Siegel, R., DeSantis, C., and Jemal, A. (2011). Global Cancer Statistics. *CA Cancer J Clin*, 61(2):69–90.

Siegel, R., DeSantis, C., and Jemal, A. (2014). Colorectal cancer statistics, 2014. *CA Cancer J Clin*, 64(2):104–117.

Sinha, S. and Yang, W. (2008). Cellular signaling for activation of Rho GTPase Cdc42. *Cell Signal*, 20(11):1927–1934.

Smits, R., Hofland, N., Edelmann, W., Geugien, M., Jagmohan-Changur, S., Albuquerque, C., Breukel, C., Kucherlapati, R., Kielman, M. F., and Fodde, R. (2000). Somatic APC mutations are selected upon their capacity to inactivate the β -catenin downregulating activity. *Genes Chromosomes Cancer*, 29(3):229–239.

Tong, Y., Hota, P. K., Hamaneh, M. B., and Buck, M. (2008). Insights into Oncogenic Mutations of *Plexin-B1* Based on the Solution Structure of the Rho GTPase Binding Domain. *Structure*, 16(2):246–258.

Tsuchiya, T., Tamura, G., Sato, K., Endoh, Y., Sakata, K., Jin, Z., Motoyama, T., Usuba, O., Kimura, W., and Nishizuka, S. (2000). Distinct methylation patterns of two APC gene promoters in normal and cancerous gastric epithelia. *Oncogene*, 19(32):3642–3646.

Umar, A., Risinger, J. I., Hawk, E. T., and Barrett, J. C. (2004). Testing guidelines for hereditary non-polyposis colorectal cancer. *Nat Rev Cancer*, 4(2):153–158.

van Wezel, T., Middeldorp, A., Wijnen, J. T., and Morreau, H. (2012). A review of the genetic background and tumour profiling in familial colorectal cancer. *Mutagenesis*, 27(2):239–245.

Vasen, H. F., Mecklin, J. P., Khan, P. M., and Lynch, H. T. (1991). The International Collaborative Group on Hereditary Non-Polyposis Colorectal Cancer (ICG-HNPCC). *Dis Colon Rectum*, 34(5):424–5.

Vasen, H. F., Watson, P., Mecklin, J. P., and Lynch, H. T. (1999). New clinical criteria for hereditary nonpolyposis colorectal cancer (HNPCC, Lynch syndrome) proposed by the International Collaborative group on HNPCC. *Gastroenterology*, 116(6):1453–6.

Walther, A., Johnstone, E., Swanton, C., Midgley, R., Tomlinson, I., and Kerr, D. (2009). Genetic prognostic and predictive markers in colorectal cancer. *Nat Rev Cancer*, 9(7):489–499.

Wang, D., Stewart, A. K., Zhuang, L., Zhu, Y., Wang, Y., Shi, C., Keating, A., Slutsky, A., Zhang, H., and Wen, X.-Y. (2010). Enhanced adaptive immunity in mice lacking the immunoinhibitory adaptor Hacs1. *FASEB J.*, 24(3):947–956.

Waterham, H. R., Koster, J., Romeijn, G. J., Hennekam, R. C., Vreken, P., Andersson, H. C., FitzPatrick, D. R., Kelley, R. I., and Wanders, R. J. (2001). Mutations in the 3β -Hydroxysterol Δ 24-Reductase Gene Cause Desmosterolosis, an Autosomal Recessive Disorder of Cholesterol Biosynthesis. *Am J Hum Genet*, 69(4):685 – 694.

Zambirinis, C. P., Theodoropoulos, G., and Gazouli, M. (2009). Undefined familial colorectal cancer. *World J Gastrointest Oncol*, 1(1):12–20.

Zhao, A. (2012). Estudo Do Gene *DOCK9* e de Novos Loci de Susceptibilidade No Cancro Do Cólon e Recto Familiar Do Tipo X (FCCTX). Tese de Mestrado, Escola Superior de Saúde Egas Moniz.

Zhu, M. L., Wang, M., Cao, Z. G., He, J., Shi, T. Y., Xia, K. Q., Qiu, L. X., and Wei, Q. Y. (2012). Association between the *ERCC5* Asp1104His polymorphism and cancer risk: a meta-analysis. *PLoS One*, 7(7):e36293.



Anexo A

A.1 Extração de RNA a partir de biópsias de tecido congelado - *kit RNeasy Mini Kit (Qiagen)*

1. Colocar a biópsia de tecido no interior de um almofariz arrefecido com azoto líquido. Triturar com o auxílio de um pilão até estar reduzida a pó;
2. Adicionar 600 μl de *Buffer RLT* e decantar a mistura para um tubo *eppendorf* de 1,5 ml arrefecido e devidamente identificado;
3. Centrifugar a 20800g durante 5 minutos e transferir o sobrenadante obtido para um novo tubo *eppendorf*;
4. Adicionar 600 μl de etanol 70% (v/v) ao lisado limpo e pipetar a mistura ($\sim 700 \mu\text{l}$ de cada vez) para uma coluna *RNeasy* inserida num tubo coletor de 2 ml;
5. Incubar durante 2 minutos à temperatura ambiente e centrifugar a 8000g durante 30 segundos. Desprezar o filtrado;
6. Adicionar 700 μl de *Buffer RW1* à coluna *RNeasy* e centrifugar a 8600g durante 30 segundos. Desprezar o filtrado e o tubo coletor;
7. Colocar a coluna num novo tubo coletor e adicionar 500 μl de *Buffer RPE*;
8. Centrifugar a 8600g durante 30 segundos e desprezar o filtrado;
9. Adicionar 500 μl de *Buffer RPE* à coluna, centrifugar a 8000g durante 2 minutos;
10. Colocar a coluna num novo tubo tipo *eppendorf* e centrifugar a 20800g durante 1 minuto. Desprezar o tubo coletor e o filtrado;
11. Colocar a coluna *RNeasy* dentro de um novo tubo coletor e adicionar 30 μl de água com DEPC;
12. Incubar durante 3 minutos e centrifugar a 8000g durante 1 minuto.



Anexo B

B.1 Análise de mutações dos genes *KDELC1*, *ERCC5*, *TPP2* e nos genes candidatos obtidos pela análise de WES

Tabela B.1: Programa utilizado na amplificação por PCR utilizando o kit *Biotaq* (*Bioline*)

Etapa	Temperatura (°C)	Tempo	Ciclos
Desnaturação inicial	95	5 min	1
Desnaturação	94	50 seg	35 e 38*
Hibridação	Variável	30 seg	
Extensão	72	50 seg	
Extensão final	72	7 min	1
Pausa	15	∞	-

*Número de ciclos utilizados na amplificação por PCR dos exões 13 a 17 do gene *TPP2*, utilizando amostras de cDNA

Tabela B.2: Programa utilizado na amplificação por PCR utilizando o kit *Amplitaq Gold* (*Applied Biosystems*)

Etapa	Temperatura (°C)	Tempo	Ciclos
Desnaturação inicial	95	7 min	1
Desnaturação	95	25 seg	40
Hibridação	Variável	25 seg	
Extensão	70	1 min	
Extensão final	70	10 min	1
Pausa	15	∞	-

Tabela B.3: Sequências dos *primers* e tamanho do fragmento amplificado para o gene *KDELC1*

<i>Primer</i>	Sequência (5'→3')	Tamanho do fragmento (pb)
Exão 1-F	GATTCGCTCCCAAATGATGC	600
Exão 1-R	AAACTGCTGTGGAGAAACGAG	
Exão 2-F	CAAGTTTCCAGACTGAAGAGC	560
Exão 2-R	TGCTAATGTATGTCTTCGCG	
Exão 3/4-F	ACCGACATGAGCCAACCTTG	920
Exão 3/4-R	CATGGCACCTAFCTGACACTTTG	
Exão 5/6-F	CTTAGAGGCTGTTTCGAGTATTAG	891
Exão 5/6-R	GCTAACTTCTCCACCTCTTGAAA	
Exão 7-F	CACTTACTAAACGGCAAAGGCA	473
Exão 7-R	ACAGGCCAAAGTCTATTACGTG	
Exão 8-F	CCATGAGACTACCCCCAAAGAT	501
Exão 8-R	CTGTAACCTTTGTAGGATCAGAGGC	
Exão 9-F	TCTAGCCTCTCTCATTGGGTG	312
Exão 9-R	CAGACTGAATTGCTTGGAGACCTC	
Exão 10-F	CAAGCCCTAAAAGCCTCTAGCA	201
Exão 10-R	GTAAAGAAGAGTCTTCAGAGCACC	

Tabela B.4: Condições utilizadas na amplificação por PCR dos fragmentos do gene *KDELC1*

Exão	Temperatura (°C)	MgCl ₂ (μl)	Primers (μl)	Solução GC* (μl)	Ciclos	Termociclador	Kit
1	63	1,25	0,3	2	35	Veriti	Biotaq
2	66	2	0,3	2	35	Veriti	Biotaq
3 e 4	71	1,25	0,15	-	35	Veriti	Biotaq
5 e 6	71	1,75	0,3	-	35	Veriti	Biotaq
7	70,5	1,75	0,15	-	35	Veriti	Biotaq
8	68,5	2	0,3	-	35	Trio	Biotaq
9	69	1,75	0,3	-	35	Biometra	Biotaq
10	69	1,75	0,3	-	35	Biometra	Biotaq

*Esta solução faz parte do *kit GC-RICH PCR System (Roche)*

Tabela B.5: Resultados obtidos através da análise mutacional do gene *KDELC1*

DNA	Família	Mutações germinais - <i>KDELC1</i>							
		Exão 1	Exão 2	Exão 3 e 4	Exão 5 e 6	Exão 7	Exão 8	Exão 9	Exão 10
L5	7	N	N	N	N	N	N	N	N
L48	24	N	N	N	N	N	N	N	N
L108	37	Het. 1/ Het.2/ Het. 3	N	N	N	N	N	N	N
L295	55	N	N	N	N	N	N	N	N
L801	56	N	N	N	N	N	N	N	N
L451	63	N	N	N	N	N	N	N	N
E436	69	N	N	N	N	N	N	N	N
L467	71	N	N	N	N	N	N	N	N
L484	84	N	N	N	N	N	N	N	N
L499	86	N	N	N	N	N	N	N	N
L701	106	N	N	N	N	N	N	N	N
L984	148	N	N	N	N	N	N	N	N
L1038	173	N	N	N	N	N	N	N	N
L1200	236	N	N	N	N	N	N	N	N
L1121	P2209	N	N	Het. 1	N	N	N	N	N
L926	55	N	N	N	N	N	N	N	N
L950	55	N	N	N	N	N	N	N	N

N - normal (sem mutação); Het - heterozigótico; Exão 1: c.1-108T>G (1), c.1-30G>T (2), c.48A>G (3); Exão 3 e 4: c.455A>T (1)

Tabela B.6: Sequências dos *primers* e tamanho do fragmento amplificado para o gene *ERCC5*

<i>Primer</i>	Sequência (5'→ 3')	Tamanho do fragmento (pb)
Exão 11-F	CATTACATGAAGTGGTAGGCAC	577
Exão 11-R	CACCTTTTGTTCACCTGTGGAG	
Exão 12-F	CTTTGGGGGTCAGCGTATAG	563
Exão 12-R	GCACCACTAAGAACTGACTCTG	
Exão 13-F	GGCTTGTGTGATGATTGGGC	612
Exão 13-R	CAACAACCTGAGGAGAGGGAAG	
Exão 14-F	GAGATACAGGGAATGGAATCAAG	295
Exão 14-R	TCAAAATGTCTGCTCTATGCCC	
Exão 15-F	AGGTTGAGCTTGTTGATTGG	773
Exão 15-R	CAAAGACCGTGCCACCAGTTA	

Tabela B.7: Condições utilizadas na amplificação por PCR dos fragmentos do gene *ERCC5*

Exão	Temperatura (°C)	MgCl ₂ (μl)	<i>Primers</i> (μl)	Ciclos	Termociclador	Kit
11	68,5	1,25	0,2	35	<i>Trio</i>	<i>Biotaq</i>
12	66,5	1,75	0,3	40	<i>Veriti</i>	<i>Amplitaq Gold</i>
13	68,5	1,25	0,2	35	<i>Trio</i>	<i>Biotaq</i>
14	68,5	1,75	0,3	35	<i>Trio</i>	<i>Biotaq</i>
15	64	1,5	0,3	35	<i>Veriti</i>	<i>Biotaq</i>

Tabela B.8: Resultados obtidos através da análise mutacional do gene *ERCC5*

DNA	Família	Mutações germinais - <i>ERCC5</i>				
		Exão 11	Exão 12	Exão 13	Exão 14	Exão 15
L5	7	N	N	N	N	N
L48	24	N	N	N	N	N
L108	37	N	N	N	N	N
L295	55	N	N	Het. 1	N	N
L801	56	N	N	Het. 1	N	N
L451	63	N	N	Hom. 1	N	N
E436	69	N	N	Het. 1	N	N
L467	71	N	Het. 1	Het. 1	N	N
L484	84	N	N	Het. 1	N	N
L499	86	N	Het. 1	Hom. 1	N	N
L701	106	N	N	Het. 1	N	Het. 1
L984	148	N	N	Hom. 1	N	N
L1038	173	N	N	Hom. 1	N	N
L1200	236	N	N	Hom. 1	N	N
L1121	P2209	N	N	N	N	Het. 1
L926	55	N	N	Het. 1	N	N
L950	55	N	N	Hom. 1	N	N

N - normal (sem mutação); Het - heterozigótico; Hom - homozigótico;
 Exão 12: c.2636A>G (1); Exão 13: c.2878+14C>T (1); Exão 15: c.3310G>C (1)

Tabela B.9: Sequências dos *primers* e tamanho do fragmento amplificado para o gene *TPP2*

<i>Primer</i>	Sequência (5'→ 3')	Tamanho do fragmento (pb)
qTPP2_ex13-F	GATCATGGCGTTGGCATTGAACC	414
qTPP2_ex?-R	CTAATCGAGCTGCAGTTCTCC	
qTPP2_ex13-F	GATCATGGCGTTGGCATTGAACC	812
qTPP2-R_ex14-R	GCACCACTAAGAACTGACTCTG	
TPP2-5_cDNA-F	AATAACCGTGGCATCTACCT	577
TPP2-5_cDNA-R	CTTCAGTCAGTGTTCCTTTCTC	

Tabela B.10: Condições utilizadas na amplificação por PCR dos fragmentos do gene *TPP2*

Fragmentos	Temperatura (°C)	MgCl ₂ (μl)	Primers (μl)	Ciclos	Termociclador	Kit
13-F/?-R	69	0,75	0,3	35	<i>Trio</i>	<i>Biotaq</i>
13-F/14-R	72	0,85	0,3	35	<i>Trio</i>	<i>Biotaq</i>
5-F/5-R	67,5	2	0,9	35	<i>Veriti</i>	<i>Biotaq</i>

Tabela B.11: Resultados obtidos através da análise mutacional do gene *TPP2*

Mutações germinais - <i>TPP2</i>									
Famílias	DNA	Frag. 13/?	Frag. 13/14	Frag. 5	Famílias	DNA	Frag. 13/?	Frag. 13/14	Frag. 5
FCCTX	L5	N	Hom. 2	N	Saudáveis	20	-	Het. 2	-
	L295	N	Het. 1/ Het. 2	N		21	-	Het. 1	-
	L447	N	Het. 1/ Het. 2	N		22	-	Het. 2	-
	L905	N	Het. 1/ Het. 2	N		23	-	Het. 1	-
	L926	N	Het. 1/ Het. 2	N		24	-	Het. 1/ Het. 2	-
	L950	N	Het. 1/ Het. 2	N		26	-	Het. 1	-
	L1544	N	Het. 1/ Het. 2	N		27	-	Het. 1	-
SL	3	-	Het. 1	-		28	N	Het. 1	N
	4	N	Het. 1	-		29	-	Het. 1	-
	6	-	Het. 1/ Het. 2	-		32	-	Het. 2	-
	9	-	Het. 1/ Het. 2	-		35	-	Het. 2	-
	12	-	Het. 1	-		37	-	Het. 1/ Het. 2	-
	14	-	Het. 1/ Het. 2	-		38	-	Het. 2	-
	18	-	Hom. 1	-		39	N	Het. 1/ Het. 2	N
	19	-	N	-		40	N	Het. 1/ Het. 2	-
	42	-	N	-	41	-	Het. 1/ Het. 2	N	
	43	-	Het. 2	-					
	44	-	Hom. 1	-					
	45	-	Hom. 1	-					

N - normal (sem mutação); Het - heterozigótico; Hom - homozigótico; Frag. 13/14: c.1679-23A>G (1), c.1679-94C>T (2)

Tabela B.12: Sequências dos *primers* desenhados e tamanho do fragmento amplificado para o gene *TPP2*

Primer	Sequência (5'→ 3')	Tamanho do fragmento (pb)
TPP2-int12/ex13-F	GCCAATCATAACGGCATAGGG	524
TPP2-int12/ex13-R	ACACTGGCCCAAACAGCAAT	
TPP2-ex14-F	GTCATACCAGTGCTTGGTGAAG	499
TPP2-ex14-R	CCTGGCAGATGGATAGGTCTA	
TPP2-ex24-F	GCAATGGGGTTGTTGTGGTA	844
TPP2-ex24-R	GAGCAGGAATAAAGTTCGC	

Tabela B.13: Sequências dos *primers* desenhados e tamanho do fragmento amplificado para variantes identificadas em genes candidatos obtidos pela análise de WES

Primer	Sequência (5'→ 3')	Tamanho do fragmento (pb)
RECQL5-c.1586-1G>CAG-F	GAAATGTTCTGGAGGCTTG	262
RECQL5-c.1586-1G>CAG-R	ACAGTGAGACTGGTCGCATG	
RWDD4-c.364-2T>-F	CAACCCACTTCAGGATATAGCTG	184
RWDD4-c.364-2T>-R	TTACGCTTCTGGGCTTTTGA	
CADM1-p.V357G-F	TCCCAGCAGGTGAAGAA	333
CADM1-p.V357G-R	CTCCCAGAGTCCTAATCAGC	
DHCR24-p.Y328S-F	TGAAGACAAACCGAGAGGGC	156
DHCR24-p.Y328S-R	AGTGTTAACTGTCCGGGCTC	
PLXNB1-p.T1733P-F	CCTGTCTCAGTCCCCTTAA	418
PLXNB1-p.T1733P-R	CTAGTCTGATGGGCTCTCTCC	

Tabela B.14: Condições utilizadas na amplificação por PCR dos fragmentos dos genes candidatos obtidos pela análise de WES

Gene	Temperatura (°C)	MgCl ₂ (μl)	Primers (μl)	Ciclos	Termociclador	Kit
RECQL5	64	1	0,3	35	Veriti	Biotaq
RWDD4	65	1,25	0,3	35	Veriti	Biotaq
CADM1	65	0,5	0,3	35	Veriti	Biotaq
PLXNB1	67,5	0,5	0,3	35	Veriti	Biotaq
DHCR24	64	0,5	0,3	35	Veriti	Biotaq

Tabela B.15: Sequências dos primers desenhados e tamanho do fragmento amplificado para as variantes identificadas em genes candidatos obtidos pela análise de WES

Primer	Sequência (5'→ 3')	Tamanho do fragmento (pb)
MUC2-p.N1040T-F	CCAACAACGACTTCACCACG	177
MUC2-p.N1040T-R	CGCTGCTTTTGAGGATGCTG	
RECQL5-c.1586-1G>CAG_cDNA-F	CAAGGGCTACGGGGACTTCA	523
RECQL5-c.1586-1G>CAG_cDNA-R	CGAGTACACATGGGAGGCTG	
RWDD4-c.364-2T>-cDNA-F	CGGGAATTAAGTCCAGTTTC	482
RWDD4-c.364-2T>-cDNA-R	ATCCTTAGAGCCAGTTTTC	
ERN2- p.T229P-F	CTCAGGAAACCCCACTC	274
ERN2- p.T229P-R	GGTCATGGATGCTGTTCAG	
PTCHD2-p.A851V-F	TAGATTTCCCAGGCACCGT	242
PTCHD2-p.A851V-R	GCTTCTCCCTCCCTATTGG	
RNF207-p.R539C-F	GTGGTGCCTGGCTTGAGTAA	302
RNF207-p.R539C-R	CTCAGCCTCCCGAAGTGTTA	
SRCAP-p.T918P-F	AGCAGATGCCCAAAAAGTA	454
SRCAP-p.T918P-R	GTCAGACCCAGTCTAACGA	
UNC5B-p.V385/6/7G-F	CCTTATGTGGTCCCTCTGTCC	244
UNC5B-p.V385/6/7G-R	CCAGACATCCTCGGGTCTTA	
ZBTB48-p.S675A-F	CTCCGCAACCTGATCATC	328
ZBTB48-p.S675A-R	GGACACGGGAACAGACTG	
TGIF1-p.S18L-F	CTTGGGAGGACTGACAGGTCT	229
TGIF1-p.S18L-R	GCCAGAAAAGCCGTGGAA	
PCDH1-p.A712T-F	ATCCAGCCTGAGCTTTGATC	287
PCDH1-p.A712T-R	GTCCATAAGGGTGTGCCACC	
ZFP62-p.N135S-F	GCATCTGAGCCACAGCATA	250
ZFP62-p.N135S-R	AGTCCCTCCACAGTCATCACA	
AGAP3-p.L30R-F	CGGCTCCGAAGCCATGAA	540
AGAP3-p.L30R-R	CAAAGTGTCCCGAACGAGA	
ALPK2-p.1389del-F	CAATGTGAGTCAAGACCAGGA	143
ALPK2-p.1389del-R	CAGCCCTGGTGTACTCTATC	
SYNE2-p.L5144R-F	CTAAGCACCTGTGATGTAGAAAGC	196
SYNE2-p.L5144R-R	CTTGTGCTTCCAGCCAGTTGT	
MAP4-p.S996F-F	CTACAACCCGAAAGCCTGAATCTA	181
MAP4-p.S996F-R	TGAACATTACCACCTCCAGGGA	
CUL7-p.L1672I-F	CTCATCGTCCGAATCTCAAG	383
CUL7-p.L1672I-R	TTACCAAGGCTGACTGACCAAAC	
PCSK5-p.N497T-F	CACAGACCGACAAATCAAGACA	199
PCSK5-p.N497T-R	CAAAAGCTGAGACCTAGTCCAGA	



Anexo C

C.1 Protocolos para preparação dos reagentes utilizados no controle da eficiência dos produtos de PCR

- **Tampão de eletroforese TBE 1x:** diluído a partir da solução comercial TBE 10x (0,89 M *Tris Borato* pH 8.3 + 20 mM Na₂ EDTA; *National Diagnostics*) em água bidestilada para um volume de 2000 ml. Armazenar ao abrigo da luz.
- **Orange G 5x:** preparado a partir de 12,5 ml de *Ficoll (Sigma)*, 125 g de *Orange G (Sigma)* e água bidestilada até perfazer o volume de 50 ml. Armazenar em alíquotas de 2 ml a -20°C.
- **Orange G 1x:** diluído a partir da solução *Orange G 5x* em 1 ml de água bidestilada.
- **Marcador *GeneRuler 50 pb DNA Ladder (Thermo Scientific)*:** Adicionar 50 µl de marcador *GeneRuler 50 pb DNA Ladder (Thermo Scientific)* a 250 µl de *Orange G 5x* e perfazer o volume para 1 ml com água bidestilada.
- **Marcador *1 kb DNA Ladder (Promega)*:** Adicionar 100 µl de marcador *1 kb DNA Ladder (Promega)* a 100 µl de *Orange G 5x* e perfazer o volume para 400 µl com água bidestilada.

Gel de agarose 2% (p/v)

1. Pesar 5 g de agarose (Seakem[®] LE Agarose - Lonza), num *erlenmeyer* de 500 ml;
2. Adicionar 250 ml de tampão TBE 1x;
3. Dissolver esta mistura no microondas e deixar arrefecer um pouco;

4. Adicionar 12,5 µl de brometo de etídeo (10 mg/ml, *MP biomedical*s) e agitar para homogenizar;
5. Colocar a solução num molde com 4 pentes de poços e deixar arrefecer até ocorrer solidificação do gel.

Gel de agarose 1,2% (p/v)

1. Pesar 1,8 g de agarose (*Seakem*[®] *LE Agarose - Lonza*), num *erlenmeyer* de 500 ml;
2. Adicionar 150 ml de tampão TBE 1x;
3. Dissolver esta mistura no microondas e deixar arrefecer um pouco;
4. Adicionar 7,5 µl de brometo de etídeo (10 mg/ml, *MP biomedical*s) e agitar para homogenizar;
5. Colocar a solução num molde com 2 pentes de poços e deixar arrefecer até ocorrer solidificação do gel.



Anexo D

D.1 Extração e purificação de DNA a partir de banda em gel de agarose - *kit QIAquick Gel Extraction (Qiagen)*

1. Pesar um tubo tipo *eppendorf* de 1,5 ml vazio;
2. Excisar a banda de interesse do gel de agarose, com uma lâmina de bisturi limpa e colocá-la dentro do tubo;
3. Pesar novamente o tubo de 1,5 ml contendo a agarose;
4. Adicionar 3 volumes de tampão GC por cada volume de gel (100 mg~100 µl);
5. Incubar a 50°C durante 10 minutos, misturando (agitar em *vortex*) a cada 2-3 minutos, para ajudar a dissolver a agarose;
6. Verificar se a mistura apresenta uma cor amarela. Caso não se verifique, se estiver laranja ou violeta, adicionar 10 µl de acetato de sódio 3 M a pH 5.0;
7. Adicionar 1 volume de isopropanol 100% e agitar em *vortex*;
8. Colocar uma coluna num tubo coletor de 2 ml;
9. Adicionar todo o volume contido no tubo de 1,5 ml à coluna e centrifugar a 13000 rpm, durante 1 minuto;
10. Descartar o volume contido no tubo coletor e recolocar a coluna no mesmo;
11. Adicionar 500 µl de tampão QG à coluna e centrifugar a 13000 rpm durante 1 minuto, para remover qualquer resto de agarose que possa estar agarrado à coluna;
12. Descartar o volume contido no tubo coletor e recolocar a coluna no mesmo;
13. Adicionar 750 µl de tampão PE à coluna, incubar 2 a 5 minutos à temperatura ambiente e centrifugar a 13000 rpm durante 1 minuto. Este tampão contém etanol, o que permite lavar o DNA que está adsorvido à membrana;

14. Descartar o volume do tubo coletor e centrifugar novamente durante 1 minuto para eliminar o resto de tampão PE e secar a membrana;
15. Colocar a coluna num novo tubo tipo *eppendorf* de 1,5 ml;
16. Adicionar um volume variável (20-50 μ l) de tampão de eluição EB ao centro da membrana da coluna, de acordo com a intensidade da banda visualizada no gel de agarose, e incubar durante 1 minuto à temperatura ambiente;
17. Centrifugar durante 1 minuto a 13000 rpm para eluir o produto purificado.



Anexo E

E.1 Protocolo de precipitação e purificação do DNA – *kit ABI PRISM BigDye[®] Terminator v1.1 Cycle Sequencing (Applied Biosystems)*

1. Num tubo de 2 ml efetuar uma mistura reacional contendo, por amostra, 2 µl de EDTA (125 mM), 2 µl de acetato de sódio (3M, pH 4,6) e 50 µl de etanol absoluto. Agitar em *vortex* e centrifugar brevemente;
2. Distribuir 54 µl do sobrenadante da mistura por tubos de 1,5 ml devidamente identificados, adicionar o produto da reação de sequenciação e agitar em *vortex*;
3. Incubar durante 15 minutos à temperatura ambiente;
4. Centrifugar a 14000 rpm durante 30-40 minutos a 4°C;
5. Retirar completamente o sobrenadante com uma micropipeta;
6. Adicionar 100 µl de etanol 70%(v/v) e agitar ligeiramente em *vortex*;
7. Centrifugar a 14000 rpm durante 15 minutos a 4°C;
8. Retirar completamente o sobrenadante com uma micropipeta;
9. Secar o *pellet* a 37°C, num banho seco, durante 5-10 minutos.



Anexo F

F.1 Análise de MS-MLPA

Two-tube protocol

1. Identificar dois conjuntos de tubos 0,2 ml;
2. Diluir as amostras em TE para um volume final de 6 μ l, de acordo com as quantidades de DNA estabelecidas;
3. Colocar os tubos num termociclador e desnaturar as amostras a 98°C durante 10 minutos;
4. Preparar uma mistura reacional (*mix* de hibridação) contendo por amostra: 1,5 μ l de *MLPA Buffer*, 1 μ l de *Probemix P200* e 0,5 μ l de *Synthetic probemix*. Colocar 3 μ l da *mix* de hibridação nos novos tubos e adicionar 5 μ l do DNA desnaturado;
5. Colocar os tubos no termociclador a 95°C durante 1 minuto, seguido de 16 horas (no mínimo, a incubação pode estender-se até um máximo de 20 horas) a 60°C;
6. Marcar um novo conjunto de tubos 0,2 ml (com *);
7. Preparar a *mix Ligase A*, *mix Ligase-65* e *mix Ligase-Digestion* de acordo com a tabela F.1;
8. Retirar os tubos do termociclador, adicionar 13 μ l de *mix Ligase A* a cada tubo e transferir 10 μ l da mistura para os novos tubos marcados com *;
9. Incubar os dois conjuntos de tubos a 49°C durante 1 minuto, num termociclador;
10. Enquanto a 49°C, adicionar 10 μ l de *mix-Ligase 65* aos tubos iniciais e 10 μ l de *mix Ligase-Digestion* aos tubos marcados com *;
11. Incubar durante 30 minutos a 49°C, seguido de 5 minutos a 98°C;
12. Identificar dois conjuntos de tubos 0,2 ml (UC - correspondentes aos tubos iniciais e

- C - correspondentes ao tubos marcados com *);
13. Transferir 5 µl das reações anteriores para os tubos novos correspondentes (apenas ligação - UC/ ligação + digestão - C);
 14. Pré-aquecer o termociclador a 72°C;
 15. Preparar a *mix SALS*A e *Polimerase mix* de acordo com a tabela F.2;
 16. Adicionar 15 µl de *mix SALS*A e 5 µl de *Polimerase mix* a cada tubo. Colocar os tubos no termociclador e começar de imediato a reação de PCR de acordo como o programa descrito no protocolo.

Tabela F.1: Reagentes e respetivos volumes para a preparação da *mix Ligase A*, *mix Ligase-65* e *mix Ligase-Digestion*

<i>Mix Ligase A</i>	<i>Mix Ligase-65</i>	<i>Mix Ligase-Digestion</i>
Água - 10 µl	Água - 8,25 µl	Água - 7,75 µl
<i>Ligase Buffer A</i> - 3 µl	<i>Ligase-65 Buffer B</i> - 1,5 µl	<i>Ligase-65 Buffer B</i> - 1,5 µl
	<i>Ligase-65 Enzime</i> - 0,25 µl	<i>Ligase-65 Enzime</i> - 0,25 µl
		<i>HhaI Enzime</i> - 0,5 µl

Tabela F.2: Reagentes e respetivos volumes para a preparação da *mix SALS*A e *Polimerase mix*

<i>Mix SALS</i> A	<i>Polimerase mix</i>
Água - 13 µl	Água - 2,75 µl
<i>SALS</i> A PCR Buffer - 2 µl	<i>SALS</i> A PCR Primers - 1 µl
	<i>SALS</i> A Enzyme Dilution Buffer - 1 µl
	<i>SALS</i> A Polimerase - 0,25 µl

One-tube protocol

Este protocolo é idêntico ao descrito anteriormente, com as seguintes exceções:

- no ponto 4 a *mix* de hibridação contém por amostra 1,5 µl de *MLPA Buffer*, 1,5 µl de *Probemix ME011* e 0,5 µl de *Synthetic probemix*. A *mix* de hibridação (3,5 µl) e o DNA desnaturado são transferidos para tubos de 0,2 ml já identificados com UC;
- no ponto 6 os tubos de 0,2 ml são marcados com C;
- os pontos 12 a 14 não existem;
- no ponto 15 apenas é preparada a *Polimerase mix* de acordo com a tabela F.3;
- no ponto 16 é adicionado 5 µl de *Polimerase mix* a cada tubo.

Tabela F.3: Reagentes e respetivos volumes para a preparação da *Polimerase mix*

<i>Polimerase mix</i>
Água - 3,75 µl
<i>SALS</i> A PCR Primer Mix - 1 µl
<i>SALS</i> A Polimerase - 0,25 µl



Anexo G

G.1 Real-time PCR

Tabela G.1: Programa utilizado nas reações de *real-time* PCR no aparelho *LightCycler® Plus 480 Multiwell Plates* (Roche)

Etapa da reação	Ciclos	Temperatura (°C)	Duração	Ramp Rate (°C/seg)	Aquisições /seg	Modo de aquisição	Modo de análise
Pré-incubação	1	95	10 min	4,4	-	-	-
		95	10 seg	4,4	-	-	-
Amplificação	45	61	15 seg	2,2	-	-	Quantification
		72	25 seg	4,4	-	Single	
		95	5 seg	4,4	-	-	
Melting curve	1	65	1 min	2,2	-	-	Melting Curves
		97	-	-	10	Continuous	
Arrefecimento	1	40	10 seg	1,5	-	-	-

Tabela G.2: Sequências dos *primers* utilizados na quantificação do gene *TPP2*

Primer	Sequência (5' -> 3')
qTPP2_ex13-F	GATCATGGCGTTGGCATTGAACC
qTPP2_ex?-R	GGAGAACTGCAGCTCGATTAG
qTPP2_ex13/14-F	CCGGAGAACACAGAAAACCTCTG
qTPP2_ex14-R	GTCAGTGTCCCAGCCATTG

Tabela G.3: Programa utilizado nas reações de *real-time* PCR no aparelho *7900HT Sequence Detection System* (Applied Biosystems)

Parâmetro	Incubação	Ativação da polimerase	PCR (40 ciclos)	
			Desnaturação	<i>annealing</i> / extensão
Temperatura	50°C	95°C	95°C	60°C
Tempo	2 mins	10 mins	15 s	1 min