



Marisa Sofia Marçalo Paulino

Licenciatura em Biologia

**Identificação de biomarcadores de
prognóstico e marcadores preditivos de
resposta à terapia no Cancro do Cólon e
Recto**

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

Orientador: Doutora Cristina Albuquerque, Unidade de
Investigação em Patobiologia Molecular (UIPM), Instituto
Português de Oncologia de Lisboa Francisco Gentil, EPE
(IPOLFG, EPE)

Júri:

Presidente: Prof. Doutora Paula Maria Theriaga
Mendes Bernardo Gonçalves

Arguente: Prof. Doutora Ana Luísa Ribeiro da Silva



FACULDADE DE
CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

Setembro 2014



Marisa Sofia Marçalo Paulino

Licenciatura em Biologia

**Identificação de biomarcadores de
prognóstico e marcadores preditivos de
resposta à terapia no Cancro do Cólon e
Recto**

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

Orientador: Doutora Cristina Albuquerque, Unidade de
Investigação em Patobiologia Molecular (UIPM), Instituto
Português de Oncologia de Lisboa Francisco Gentil, EPE
(IPOLFG, EPE)

Júri:

Presidente: Prof. Doutora Paula Maria Theriaga
Mendes Bernardo Gonçalves

Arguente: Prof. Doutora Ana Luísa Ribeiro da Silva



Setembro 2014

Identificação de biomarcadores de prognóstico e marcadores preditivos de resposta à terapia no cancro do cólon e recto

Copyright© Marisa Sofia Marçalo Paulino, 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.

Agradecimentos

Gostaria de fazer um agradecimento especial à Doutora Cristina Albuquerque pela oportunidade de realizar investigação no grupo de Gastrenterologia do Instituto Português de Oncologia, Lisboa, Francisco Gentil, E.P.E, pelos seus ensinamentos e esclarecimentos ao longo deste projecto, pela sua disponibilidade e simpatia. Agradeço à Professora Doutora Paula Gonçalves pela sua disponibilidade e dedicação. À Doutora Branca Cavaco, Coordenadora da Unidade de Investigação em Patobiologia Molecular do IPOLFG-E.P.E., por ter concedido a oportunidade de participar neste projecto. Ao Serviço de Gastrenterologia, Clínica de Risco Familiar e ao Serviço de Anatomia Patológica do IPOLFG-E.P.E., pela disponibilização do material biológico e informação clínica dos doentes.

Quero também agradecer à Mestre Inês Francisco por me ter ensinado as técnicas essenciais para a realização deste projecto, por todo o apoio e dedicação e pela ajuda na revisão da dissertação. À Dra. Patrícia Silva pelo apoio prestado na técnica de MS-MLPA e pela disponibilidade em me responder a questões. Ao Dr. Bruno Filipe pelo seu apoio e disponibilidade no laboratório, pela sua simpatia e paciência. Às Dra. Sofia Fragoso, Dra. Patrícia Machado e Dra. Sidónia por todo o apoio, carinho e motivação ao longo deste ano. Aos restantes colaboradores do UIPM pela simpatia e por me terem recebido tão bem.

Um agradecimento especial à minha colega Ana Saramago pela amizade e companhia, por estar sempre disponível para ouvir os meus desabafos e a pela incentivação ao longo deste ano. Aos meus colegas Filipa Fernandez e Gonçalo Pereira pela companhia e simpatia, especialmente à Filipa por ter transformado muitos momentos do laboratório com a sua boa-disposição.

Agradeço, também, aos meus queridos amigos por estarem sempre presentes e ajudarem-me nesta etapa com carinho, paciência, compreensão e amor e pelos momentos de partilha, nos quais me transmitiram confiança e bom ânimo nos momentos menos bons.

Por fim, agradeço às pessoas mais importantes da minha vida e sem as quais tudo isto não seria possível, aos meus pais e irmã. Obrigada por todo o amor e por estarem sempre ao meu lado e incentivarem-me a alcançar os meus sonhos sem nunca desistir. Agradeço a Deus por ter colocado cada uma destas pessoas na minha vida, por ser o meu melhor amigo e pelo seu amor incondicional.

Resumo

Mais de 50% dos novos casos de cancro do cólon e recto (CCR) diagnosticados desenvolvem metástases, as quais apresentam uma elevada resistência às terapias convencionais, tornando o CCR uma das principais causas de morte por cancro. Por estas razões, é necessário identificar biomarcadores moleculares de prognóstico e preditivos da resposta ao tratamento. A maioria dos tumores colorectais apresentam mutações em genes que codificam para componentes da via de sinalização WNT. O presente trabalho teve como objectivo estudar o papel de mutações específicas no gene *TCF7L2*, um importante factor de transcrição desta via, como potenciais factores de prognóstico no cancro do cólon e recto e avaliar a presença de alterações (epi)genéticas e a sua relação com a resposta à quimiorradioterapia pré-operatória no cancro do recto, de modo a identificar marcadores preditivos de resposta ao tratamento.

Foi efectuada a análise de mutações em exões específicos do *TCF7L2* em 68 amostras de CCR, por DGGE, *Genescan* e sequenciação automática. Foi realizada a análise de perda de heterozigotia por *Genescan* e de *copy-number* e metilação por MS-MLPA, para o mesmo gene, em 16 tumores do recto. Foram confirmados alguns destes resultados, ao nível do DNA e RNA, por qPCR.

Foram detectadas mutações no gene *TCF7L2* em 7/68 (10%) CCR, as quais foram menos frequentes em tumores sem instabilidade de microssatélites (MSS) do que em tumores instáveis (5/56, 11% vs 8/12, 67%). Na maioria dos tumores do recto resistentes à terapia foram detectados ganhos da região *upstream* e 5' do gene, os quais foram mais frequentes do que em tumores sensíveis (4/6 vs 1/8, $p=0,063$). Os tumores do recto que responderam melhor à quimiorradioterapia apresentaram mais frequentemente ganhos na região 3' do gene. Em conclusão, as mutações no gene *TCF7L2* são pouco frequentes em tumores estáveis, sendo necessário mais estudos para avaliar relação com o prognóstico. As alterações de *copy number* poderão ser potenciais marcadores de resposta à quimiorradioterapia pré-operatória no recto, no entanto os resultados ainda são preliminares, sendo necessária a análise de um maior número de casos.

Palavras-chave: Cancro colorectal, via de sinalização WNT, gene *TCF7L2*, marcadores preditivos, *copy-number*

Abstract

Over 50% of the colorectal cancer cases (CRC) develop metastases, which have a high resistance to conventional therapies, making death associated to CRC one of the leading causes of cancer death. For these reasons, it is necessary to identify molecular biomarkers of prognosis and predictive of response to treatment. The majority of colorectal tumors have mutations in genes encoding components of the Wnt signaling pathway. The present work aimed to study the role of specific mutations in the *TCF7L2* gene, an important transcription factor of this pathway, as potential prognostic factors in CRC and analyze the role of (epi) genetic changes as predictive biomarkers of response to preoperative chemo-radiotherapy in rectal cancer.

The analysis of mutations in specific exons of the *TCF7L2* gene was performed in 68 CRC samples by DGGE, Genescan and automatic sequencing. Sixteen rectal tumor samples were analyzed for loss of heterozygosity by Genescan and for copy-number variations and methylation by MS-MLPA. Some of these results were confirmed, at the DNA and RNA level, by qPCR. Mutations in the *TCF7L2* gene were detected in 7/68 (10%) CRC and were less common in tumors without microsatellite instability (MSS) than in microsatellite instable tumors (5/56, 11% vs. 8/12, 67%). Most rectal tumors that were resistant to pre-operative chemo-radiotherapy presented gains in the upstream region and in the 5' region of the gene. These gains were more frequent in the latter tumors than in tumors that were sensitive to pre-operative chemotherapy (4/6 vs 1/8, $p=0,063$). Expression analysis in two samples showing these gains revealed increased levels of gene expression. Rectal tumors that responded to pre-operative chemo-radiotherapy presented more frequently gains in the 3' region of the gene. In conclusion, mutations in the *TCF7L2* gene are infrequent in microsatellite stable tumors, although more studies are needed to assess the relationship with prognosis. Copy number variations may be potential predictive biomarkers of response to pre-operative chemo-radiotherapy in the rectum, however the results are still preliminary and the analysis of a larger number of cases is required.

Keywords: Colorectal cancer, Wnt signaling pathway, *TCF7L2* gene, predictive biomarkers, copy number

Índice Geral

1. Introdução	1
1.1. Cancro do cólon e recto	1
1.2. Classificação do estadio tumoral	1
1.3. Tratamento no CCR	3
1.4. Sequência adenoma-carcinoma	4
1.5. Vias de instabilidade genómica associadas à tumorigénese do cólon e recto	5
1.6. Síndromes hereditários do cancro do cólon e recto	6
1.7. Via de Transdução de Sinal WNT/ β -catenina	9
1.8. Gene <i>TCF7L2</i>	10
1.8.1. Papel do gene <i>TCF7L2</i> na tumorigénese colorectal	12
1.9. Gene <i>TCF7L2</i> como marcador preditivo de resposta ao tratamento em CCR metastizado e tumores do recto localmente avançados	14
1.10. Objectivos	14
2. Materiais e Métodos	15
2.1. Amostras Biológicas	15
2.2. Extracção de ácidos nucleicos	15
2.2.1. Extracção de DNA genómico	15
2.2.1.1. Extracção de DNA genómico a partir de tecido incluído em parafina	16
2.2.1.1.1. Quantificação de DNA por electroforese em gel de agarose 0,8% (p/v)	16
2.2.1.2. Extracção de DNA genómico a partir de sangue periférico	17
2.2.1.2.1. Quantificação de DNA por espectrofotometria	17
2.2.1.3. Extracção de DNA genómico a partir de tecido congelado	17
2.2.1.4. Extracção de RNA a partir de sangue periférico	18
2.3. <i>Polymerase Chain Reaction</i> (PCR)	18
2.3.1. Optimização das condições das reacções de PCR	18

2.3.2.	Amplificação de alguns exões do gene <i>TCF7L2</i> por PCR	19
2.3.2.1.	Controlo da eficiência dos produtos amplificados por electroforese em gel de agarose	20
2.4.	Sequenciação Automática	21
2.4.1.	Purificação enzimática dos produtos da reacção de PCR	21
2.4.2.	Reacção de Sequenciação	22
2.4.3.	Precipitação e purificação dos produtos da reacção de sequenciação	22
2.4.4.	Preparação das amostras para a electroforese capilar	23
2.4.5.	Análise de Resultados	23
2.5.	<i>Denaturing Gradient Gel Electrophoresis</i> (DGGE)	23
2.5.1.	Amplificação de alguns exões do gene <i>TCF7L2</i> por reacção de PCR	25
2.5.2.	Preparação dos géis de poliácridamida	25
2.5.3.	Preparação das amostras para aplicar no gel de DGGE	26
2.5.4.	Electroforese dos géis de DGGE	26
2.5.5.	Análise dos géis de DGGE	26
2.6.	Análise da mutação <i>frameshift</i> no exão 17 no gene <i>TCF7L2</i> utilizando o <i>software GeneScan</i>	26
2.6.1.	Amplificação do fragmento 17A por PCR	26
2.6.2.	Preparação das amostras para a electroforese capilar, utilizando o <i>software GeneScan</i>	27
2.6.3.	Análise de resultados	27
2.7.	Análise <i>in silico</i> das mutações somáticas identificadas	27
2.8.	Análise de perda de heterozigotia	28
2.8.1.	Amplificação dos marcadores de microssatélites por PCR para análise de LOH	28
2.8.2.	Preparação das amostras para a electroforese capilar, utilizando o <i>software GeneScan</i>	28
2.8.3.	Análise de resultados	28
2.9.	Análise de metilação e <i>copy number</i> pela técnica de MS-MLPA	29
2.9.1.	Preparação das amostras	30

2.9.2.	Reacção de MS-MLPA	30
2.9.3.	Preparação das amostras para análise de MS-MLPA pelo <i>software GeneMapper</i>	32
2.9.4.	Análise de Resultados	32
2.10.	Análise dos níveis de expressão por PCR quantitativo em tempo real (qPCR) ..	32
2.10.1.	<i>SYBR Green</i> [®]	34
2.10.2.	<i>TaqMan</i> [®]	35
2.10.2.1.	Síntese de cDNA por transcrição reversa seguida de PCR (RT-PCR)	36
2.11.	Análise Estatística	37
3.	Resultados	39
3.1.	Análise de mutações somáticas no gene <i>TCF7L2</i> em CCR	39
3.1.1.	Análise do <i>copy number</i> das amostras de CCR por MS-MLPA	44
3.2.	Análise de perda de heterozigotia em tumores do recto	47
3.3.	Análise do <i>copy-number</i> do gene <i>TCF7L2</i> em tumores do recto por MS-MLPA .	48
3.4.	Análise do <i>copy number</i> do gene <i>TCF7L2</i> por qPCR	53
3.5.	Quantificação do nível de expressão do gene <i>TCF7L2</i> por qPCR.....	55
4.	Discussão	57
4.1.	Caracterização molecular do gene <i>TCF7L2</i> em tumores colorectais	57
4.2.	Caracterização molecular do gene <i>TCF7L2</i> em tumores do recto e relação com a resposta à quimiorradioterapia pré-operatória	58
5.	Conclusão	61
6.	Bibliografia	63
7.	Anexos	69

Índice de Figuras

1.1. Esquema representativo do estadio de CCR com base no sistema TNM.....	2
1.2. Esquema representativo da sequência adenoma-carcinoma.....	4
1.3. Representação esquemática dos tipos de CCR, com descrição dos síndromes hereditários.....	6
1.4. Esquema representativo da via WNT/ β -catenina.....	9
1.5. Esquema representativo do cromossoma 10 com a localização do gene <i>TCF7L2</i>	10
1.6. Esquema representativo da proteína TCF7L2 (em cima) e dos exões sujeitos a <i>splicing</i> alternativo (em baixo).....	11
1.7. Esquema representativo da proteína TCF7L2.....	11
1.8. Esquema representativo das isoformas do gene <i>TCF7L2</i>	12
1.9. Esquema representativo das isoformas do gene <i>TCF7L2</i> com as variantes de <i>splicing</i> C-terminal identificadas.....	12
2.1. Esquema representativo de uma reacção de MS-MLPA.....	29
2.2. Esquema da localização das sondas MS-MLPA no gene <i>TCF7L2</i>	30
2.3. Sistema representativo do método qPCR.....	33
3.1. Electroforetograma parcial obtido no sequenciador <i>ABI Prism™ 3130 Genetic Analyzer</i> , representando as mutações nos exões 13 delT (A), 15 T>C (B) e 17 del A(C) no sequenciador (sequência <i>forward</i>).....	40
3.2. Electroforetograma correspondente a um caso (tumor/normal) com perda de heterozigotia no primeiro alelo para o marcador de microssatélites D10S562.....	47
3.3. Resultado obtido através da amplificação do cDNA entre os exões 2 e 7 do gene <i>TCF7L2</i> , para a amostra L1499.....	53
3.4. Resultado da análise do <i>copy number</i> do exão 1 do gene <i>TCF7L2</i> para as amostras, nas quais foram detectados ganhos do <i>copy number</i> pela técnica MS-MLPA.....	54
3.5. Resultados da análise do <i>copy number</i> do intrão e exão 4 do gene <i>TCF7L2</i> para as amostras, nas quais foram detectadas perdas do <i>copy number</i> pela técnica MS-MLPA.....	54
3.6. Resultados da análise do <i>copy number</i> exão 14 do gene <i>TCF7L2</i> para a amostra 1970, na qual foi detectada perda do <i>copy number</i> pela técnica MS-MLPA.....	55

3.7. Resultados do ensaio de expressão relativa dos exões 3-4 do gene <i>TCF7L2</i> em amostras com alteração do <i>copy number</i> no intrão 4.....	56
3.8. Resultados do ensaio de expressão relativa dos exões 10-11 do gene <i>TCF7L2</i> em amostras com alteração do <i>copy number</i>	56

Índice de Tabelas

1.1. Esquema representativo do estadio de CCR com base no sistema TNM.....	2
1.2. Esquema representativo da sequência adenoma-carcinoma.....	4
1.3. Representação esquemática dos tipos de CCR, com descrição dos síndromes hereditários.....	6
1.4. Esquema representativo da via WNT/ β -catenina.....	9
1.5. Esquema representativo do cromossoma 10 com a localização do gene <i>TCF7L2</i>	10
1.6. Esquema representativo da proteína TCF7L2 (em cima) e dos exões sujeitos a <i>splicing</i> alternativo (em baixo).....	11
1.7. Esquema representativo da proteína TCF7L2.....	11
1.8. Esquema representativo das isoformas do gene <i>TCF7L2</i>	12
2.1. Esquema representativo de uma reacção de MS-MLPA.....	29
2.2. Esquema da localização das sondas MS-MLPA no gene <i>TCF7L2</i>	30
2.3. Sistema representativo do método qPCR.....	33
3.1. Electroforetograma parcial obtido no sequenciador <i>ABI Prism™ 3130 Genetic Analyzer</i> , representando as mutações nos exões 13 delT (A), 15 T>C (B) e 17 del A(C) no sequenciador (sequência <i>forward</i>).....	40
3.2. Electroforetograma correspondente a um caso (tumor/normal) com perda de heterozigotia no primeiro alelo para o marcador de microssatélites D10S562.....	47
3.3. Resultado obtido através da amplificação do cDNA entre os exões 2 e 7 do gene <i>TCF7L2</i> , para a amostra L1499.....	53
3.4. Resultado da análise do copy number do exão 1 do gene <i>TCF7L2</i> para as amostras, nas quais foram detectados ganhos do <i>copy number</i> pela técnica MS-MLPA.....	54
3.5. Resultados da análise do <i>copy number</i> do intrão e exão 4 do gene <i>TCF7L2</i> para as amostras, nas quais foram detectadas perdas do <i>copy number</i> pela técnica MS-MLPA.....	54

3.6. Resultados da análise do <i>copy number</i> exão 14 do gene <i>TCF7L2</i> para a amostra 1970, na qual foi detectada perda do <i>copy number</i> pela técnica MS-MLPA.....	55
3.7. Resultados do ensaio de expressão relativa dos exões 3-4 do gene <i>TCF7L2</i> em amostras com alteração do <i>copy number</i> no intrão 4.....	56
3.8. Resultados do ensaio de expressão relativa dos exões 10-11 do gene <i>TCF7L2</i> em amostras com alteração do <i>copy number</i>	56

1. Introdução

1.1. Cancro do cólon e recto

O cancro do cólon e recto (CCR) é uma das principais causas de morbilidade e mortalidade (Centelles, 2012), sendo o quarto cancro mais frequente e o terceiro mais mortal a nível mundial. Em relação ao género, é o terceiro mais frequente e o quarto cancro mais mortal nos homens e o segundo mais frequente e o terceiro mais mortal nas mulheres (GLOBOCAN, 2012).

Em relação aos dados de 2012, estima-se que tenham sido diagnosticados 136 100 novos casos de CCR e 694 000 mortes provocadas por esta doença (GLOBOCAN, 2012). Este cancro é mais frequente em países como Austrália, Nova Zelândia, Estados Unidos e no continente Europeu comparativamente com os países africanos ou asiáticos, o que permite associar este tipo de cancro aos países desenvolvidos (Jemal *et al.*, 2011; World Health Organization, 2012). Em Portugal, o CCR é o terceiro cancro mais frequente e o segundo mais mortal, sendo que nos homens a taxa de incidência foi de 41,8 casos por 100 000 habitantes, nas mulheres 23,6 casos por 100 000 habitantes e a taxa de mortalidade foi de 19/100 000 nos homens e 9,4/100 000 nas mulheres (World Health Organization, 2012).

O CCR apresenta uma etiologia multifactorial, uma vez que, tanto os factores ambientais e os factores genéticos podem desempenhar um papel importante no desenvolvimento deste cancro (Hagggar e Boushey, 2009). De entre os factores ambientais, a dieta influencia bastante o risco de desenvolver CCR (consumo elevado de carne vermelha vs dieta rica em fibras e vegetais), mas também o tabagismo e a ingestão de álcool estão associados ao aumento do risco (Boyle e Langman, 2000). Cerca de 70% dos CCR ocorre num contexto esporádico (desenvolvimento tardio de CCR, por volta dos 70 anos de idade; sem história familiar e surge pela acumulação de mutações somáticas), sendo que, os restantes 25-30% apresentam história familiar (desenvolvimento de CCR precoce, com idade inferior aos 50 anos de idade), o que pressupõe o envolvimento de factores genéticos para além dos ambientais. (Jasperson *et al.*, 2010; Centelles, 2012). Contudo, apenas 5% das formas familiares estão bem caracterizadas, nas quais estão incluídos os síndromes hereditários que apresentam mutações em genes de elevada penetrância. Os restantes 25% correspondem a mutações em genes ainda desconhecidos, provavelmente genes de baixa penetrância, alguns ainda não caracterizados e que podem interagir com factores ambientais, dando início ao desenvolvimento tumoral (Zambirinis *et al.*, 2009; Jasperson *et al.*, 2010).

1.2. Classificação do estadio tumoral

O CCR é classificado histologicamente de acordo com o estado de invasão do local primário e pela presença de metástases nos nódulos linfáticos e em outros órgãos numa escala entre 0 e IV (de

acordo com o grau de gravidade do tumor), com base no sistema TNM (Figura 1.1. e Tabela 1.1.) (Centelles, 2012).

O sistema TNM foi desenvolvido pelo American Joint Committee on Cancer (AJCC) e é o sistema mais utilizado, pois é considerado o mais preciso e descritivo. O T refere-se ao grau de crescimento do tumor primário no epitélio cólico, o N representa a presença de células tumorais nos nódulos linfáticos e o M descreve a presença de metástases noutros órgãos.

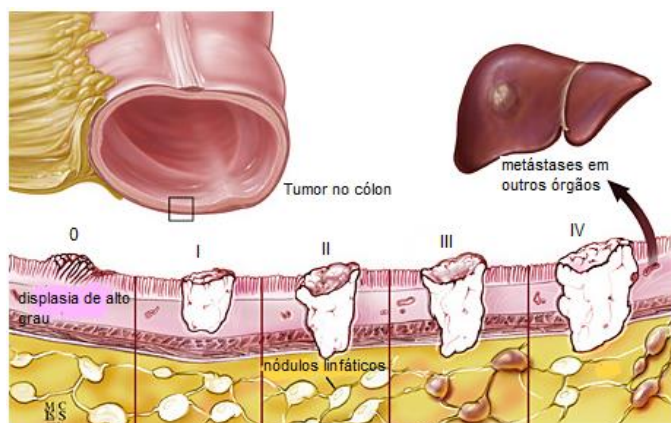


Figura 1.1. Esquema representativo do estadiamento de CCR com base no sistema TNM (adaptado de http://www.hopkinscoloncancercenter.org/CMS/CMS_Page.aspx?CurrentUDV=59&CMS_Page_ID=EAA2CD91-3276-4123-BEEB-BAF1984D20C7).

✓ T (Tumor primário):

- Tx – tumor primário não pode ser avaliado;
- T0 – nenhuma evidência de tumor primário;
- Tis – carcinoma *in situ*: intra-epitelial ou invasão da lâmina própria;
- T1 – tumor invade submucosa;
- T2 – tumor invade o músculo;
- T3 – tumor invade a membrana externa do epitélio cólico ou atinge órgãos ou estruturas próximas do intestino;
- T4 – presença de metástases noutros órgãos.

✓ N (Nódulos linfáticos):

- Nx – os nódulos linfáticos não podem ser avaliados;
- N0 – ausência de células tumorais nos nódulos linfáticos;
- N1 – presença de células tumorais em 1 a 3 nódulos próximos do intestino;
- N2 – presença de células tumorais em mais de quatro nódulos linfáticos.

✓ M (metástases à distância):

- Mx – a metastização à distância não pode ser avaliada;
- M0 – sem metástases à distância;
- M1 – presença de metástases à distância.

Tabela 1.1. Estádios dos tumores do cólon com base no sistema TNM (adaptado de Centelles, 2012).

Estadio	T	N	M
0	Tis	N0	M0
I	T1 ou T2	N0	M0
II	T3 ou T4	N0	M0
III	T1 - T2/ T3 ou T4	N1 ou N2	M0
IV	Qualquer T	Qualquer N	M1

1.3. Tratamento no CCR

O primeiro fármaco desenvolvido para o tratamento de CCR foi o 5-fluorouracilo (5-FU), o qual apresenta uma estrutura semelhante às pirimidinas do DNA e do RNA e actua de duas formas: 1) um dos seus metabolitos forma um complexo estável com o enzima timidilato sintase e inibe a produção de desoxitimidina monofosfato, a qual é essencial para a replicação e reparação do DNA e a sua ausência causa citotoxicidade; 2) a sua incorporação nas macromoléculas (DNA e RNA) causa erros nestas moléculas, tendo consequências ao nível do metabolismo e viabilidade celular. Este fármaco permitiu aumentar o tempo médio de sobrevivência dos doentes de 5-6 meses para 10-12 meses (Longley *et al.*, 2003).

A combinação de 5-FU com novos fármacos, tais como a oxaliplatina (FOLFOX) e o irinotecano (FOLFIRI) permitiu um aumento de sobrevivência dos doentes para os 20-24 meses, sendo esta, actualmente, a primeira linha de terapia. A oxaliplatina forma aductos de platina na cadeia de DNA, provocando a paragem do ciclo celular e, posteriormente, a apoptose. O segundo fármaco é um composto análogo à camptotecina que inibe o enzima topoisomerase I, originando quebras na dupla cadeia de DNA e a consequente morte celular, especificamente na fase S do ciclo celular (Wilke *et al.*, 2003).

A evolução clínica tem sido melhorada através da inclusão de agentes biológicos para o tratamento de CCR (Lenz, 2007). O bevacizumab é um anticorpo monoclonal anti-angiogénico dirigido contra o factor de crescimento endotelial vascular (VEGF – *vascular endothelial growth factor*), o qual juntamente com os fármacos de primeira linha melhorou em alguns casos a resposta à terapia. Este anticorpo inibe a formação de vasos sanguíneos pela ligação ao factor VEGF-A, que é responsável pela angiogénese quer das células normais quer das células tumorais (Hurwitz *et al.*, 2004). O cetuximab foi o primeiro agente desenvolvido contra o receptor do factor de crescimento epidérmico (EGFR – *epidermal growth factor receptor*). Este fármaco impede a ligação do ligando ao receptor, inibindo a activação das cascatas de sinalização, responsáveis pelo aumento da proliferação celular, da angiogénese, da invasão e da metastização, como também na diminuição da apoptose (Graham *et al.*, 2004).

Contudo, a cirurgia é o único tratamento curativo para doentes com cancro colorectal. A maioria dos casos é diagnosticada numa fase avançada em estadio III ou IV (de acordo com o sistema TNM), ou seja, com metástases nos nódulos linfáticos e/ou em outros órgãos, respectivamente. Em doentes com CCR em estadio III, verifica-se que a quimioterapia adjuvante (após a cirurgia) com 5-fluorouracilo (5-FU) aumenta o tempo sobrevivência do doente. Nos tumores do recto, a quimiorradioterapia pré-operatória representa o tratamento padrão para estes doentes, no entanto, a frequência com que estes doentes apresentam ausência de resposta a este tratamento é ainda elevada.

Mais de 50% dos novos casos diagnosticados desenvolvem metástases, as quais apresentam uma elevada resistência às terapias convencionais, tornando o CCR uma das principais causas de morte por cancro (Chu, 2008; Carethers, 2008; Yalçın, 2009). Por estas razões, é necessário identificar biomarcadores moleculares e farmacogénómicos e/ou outros agentes biológicos, para

prever a resposta clínica ao tratamento, de modo a desenvolver quimioterapias dirigidas para os doentes com CCR.

1.4. Sequência adenoma-carcinoma

Em 1990, Fearon e Vogelstein propuseram um modelo genético e histológico capaz de explicar a evolução das células do epitélio cólico normal para carcinoma, conhecido por sequência adenoma-carcinoma (Figura 1.2.) (Fearon e Vogelstein, 1990).

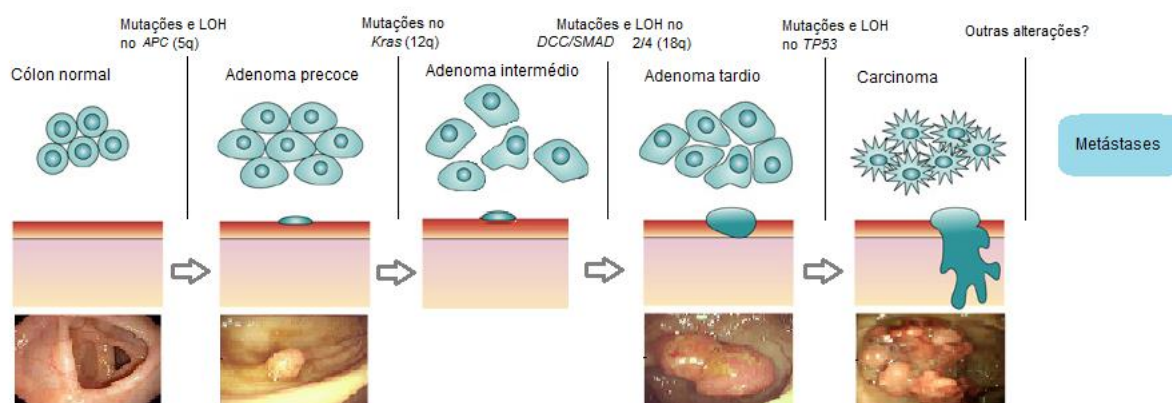


Figura 1.2. Esquema representativo da sequência adenoma-carcinoma (adaptado de Fearon e Vogelstein, 1990; Kerr, 2003 e Pino e Chung, 2010). Legenda LOH – Perda de heterozigotia.

Este modelo é ainda hoje aceite para a maioria dos carcinomas e caracteriza-se pela associação de uma sucessão de alterações histológicas com a presença de alterações genéticas bem caracterizadas, tais como a activação do oncogene (*KRAS*) e a inactivação de genes supressores tumorais, como o *APC* e o *TP53* e a perda de heterozigotia do braço longo cromossoma 18 (LOH 18q) (Fearon e Vogelstein, 1990; Armaghany *et al.*, 2012)

O primeiro passo consiste na inactivação do gene supressor tumoral *APC* - *adenomatous polyposis coli* (localizado em 5q21), o qual leva à formação de um adenoma a partir da mucosa cólica normal, dando início à tumorigénese (Fearon e Vogelstein, 1990). Este gene codifica para uma proteína com o mesmo nome, que regula vários processos celulares: diferenciação, adesão, polaridade, migração, desenvolvimento e apoptose, assim como a segregação dos cromossomas. No entanto, a sua principal função é a regulação negativa dos níveis intracelulares de β -catenina na via de transdução de sinal WNT/ β -catenina (Pino e Chung, 2010). Este gene encontra-se mutado em cerca de 60 a 80% de todos os CCR esporádicos e adenomas (Narayan e Roy, 2003).

Em seguida, ocorrem mutações no oncogene *KRAS* - *kirsten rat sarcoma viral oncogene homolog* (localizado em 12p12.1), com a consequente activação do mesmo, as quais são responsáveis pela progressão dos adenomas (Fearon e Vogelstein, 1990). Este gene codifica para uma proteína com o mesmo nome, que participa na via de sinalização RAS/MAPK e está envolvida na divisão, diferenciação e apoptose celular (Campbell *et al.*, 1998). Cerca de 30 a 50% de todos os CCR apresentam mutações neste gene (Pino e Chung, 2010).

A perda de heterozigotia (LOH) da região 18q21 é frequente em CCR num estadio avançado. Os genes *DCC* (*deleted in colorectal carcinoma*), *SMAD2* (*mothers against decapentaplegic homolog 2*) e *SMAD4* (*mothers against decapentaplegic homolog 4*) localizam-se nessa região e a sua inactivação contribui para a progressão e diferenciação dos adenomas (Fearon e Vogelstein, 1990; Pino e Chung, 2010; Armaghany *et al.*, 2012).

Por último, a perda bi-alélica ou inactivação do gene supressor tumoral *TP53* (localizado em 17q13.1) é responsável pela transição de adenoma tardio para carcinoma, uma vez que provoca uma proliferação descontrolada das células tumorais do epitélio cólico, devido à perda da capacidade de controlar o ciclo celular e indução de apoptose (Pino e Chung, 2010; Armaghany *et al.*, 2012).

Assim, ao longo do desenvolvimento tumoral, são alterados diversos processos celulares, tais como a adesão celular, a sinalização, os factores de crescimento, a apoptose e a angiogénese, que são seleccionadas por se tratar de características vantajosas para a tumorigénese e assim contribuírem para a formação do carcinoma (Kerr, 2003).

1.5. Vias de instabilidade genómica associadas à tumorigénese do cólon e recto

Para a progressão tumoral são necessários múltiplos eventos genéticos e a instabilidade genómica é actualmente reconhecida como uma característica celular essencial que acompanha a aquisição dessas alterações (Pino e Chung, 2010).

No CCR estão descritas três vias distintas de instabilidade genómica: a da instabilidade cromossómica (CIN – *chromosomal instability*); a da instabilidade de microssatélites (MSI – *microsatellite instability*) e a do fenótipo metilador das ilhas de CpG (CIMP – *CpG island methylator phenotype*) (Pino e Chung, 2010).

A via CIN é a mais comum nos tumores sólidos, sendo uma característica de cerca de 85% dos CCR, e descreve um estado dinâmico no qual as células ganham ou perdem, total ou parcialmente, os cromossomas, originando alterações estruturais e numéricas (Centelles, 2012). Como consequência, as células apresentam aneuploidia, amplificação genómica sub-cromossómica e uma elevada frequência de perda de heterozigotia (LOH), assim como mutações em genes supressores de tumor e oncogenes específicos (Pino e Chung, 2010) como referido no ponto 1.4.

A via da instabilidade de microssatélites (MSI) é caracterizada pelo aumento ou diminuição (“instabilidade”) no número de repetições nas sequências microssatélites ao longo do genoma devido à inactivação do sistema de reparação de erros no DNA do tipo *mismatch* (MMR) (Geiersbach e Samowitz, 2011). As sequências de microssatélites (também conhecidas por *short tandem repeats* – STRs) correspondem a pequenas sequências repetitivas, normalmente, de 1 a 6 nucleótidos, as quais estão particularmente sujeitas a erros durante a replicação do DNA (Boland e Goel, 2010; Geiersbach e Samowitz, 2011). A MSI é uma característica dos tumores de síndrome de Lynch, devido a mutações germinais nos genes do sistema MMR (*MLH1*, *MSH2*, *MSH6* e *PMS2*) mas

também ocorre em cerca de 15% dos CCR esporádicos, devido, principalmente, à inactivação somática do gene *MLH1* por hipermetilação do promotor (Armaghany *et al.*, 2012).

Em 1999, Toyota e os seus colaboradores propuseram o fenótipo metilador das ilhas CpG (CIMP), o qual associa a hipermetilação de ilhas CpG localizadas nas regiões promotoras de genes com o desenvolvimento tumoral (Toyota *et al.*, 1999).

Em células normais, as ilhas CpG encontram-se, normalmente, não metiladas, pelo que o gene é expresso. Pelo contrário, quando o promotor está metilado, pela adição de um grupo metil à citosina, pelas enzimas DNA metiltransferases, o gene é silenciado. O silenciamento, por hipermetilação, de genes supressores tumorais ou de genes do sistema MMR contribui para o desenvolvimento tumoral, sendo que 80% dos CCR esporádicos com MSI apresentam hipermetilação das ilhas CpG do promotor do gene *MLH1* (Boland e Goel, 2010; Armaghany *et al.*, 2012).

1.6. Síndromes hereditários do cancro do cólon e recto

Como referido anteriormente, cerca de 25-30% dos CCR tem uma origem familiar mas apenas 5% se encontram bem caracterizados clínica e molecularmente. Os síndromes hereditários de CCR são definidos com base nas suas características clínicas, patológicas e/ou genéticas e encontram-se organizados em dois grupos: os síndromes associados a polipose e o cancro do cólon e recto não associado a polipose (HNPCC) (Jasperson, *et al.*, 2010) (Figura 1.3).

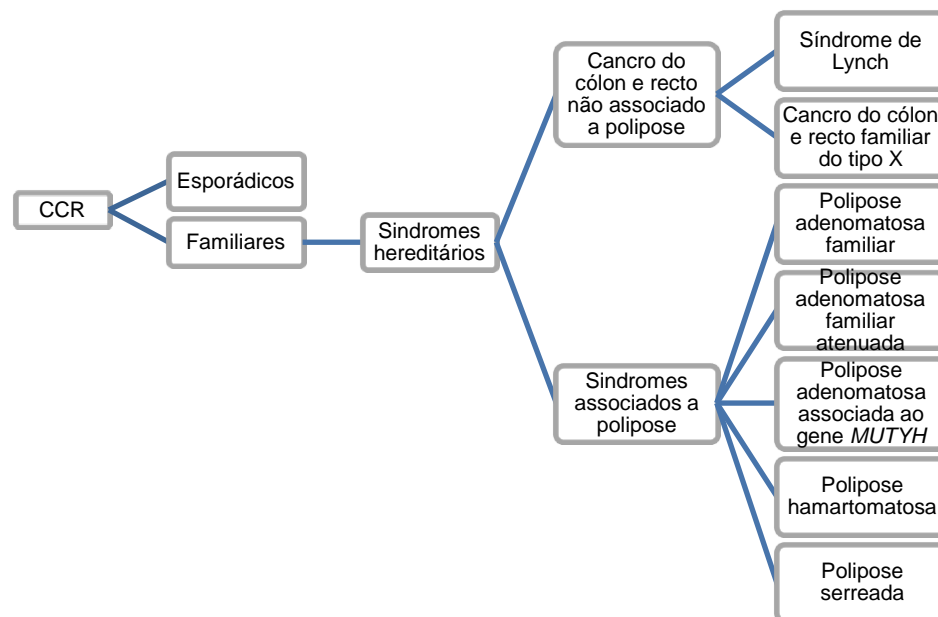


Figura 1.3. Representação esquemática dos tipos de CCR, com descrição dos síndromes hereditários (adaptado de Lindor *et al.*, 2005).

A polipose adenomatosa familiar (PAF) apresenta um padrão de transmissão autossómica dominante, causada por mutações germinais truncantes no gene *APC* (localizado em 5q21) (Centelles, 2012). Este síndrome caracteriza-se clinicamente pela presença de centenas a milhares de pólipos adenomatosos ao longo do cólon, os quais levam ao desenvolvimento inevitável de CCR

por volta dos 40 anos de idade caso o cólon não tenha sido removido (Jasperson *et al.*, 2010). A polipose adenomatosa familiar atenuada (APAF) é uma forma menos severa da PAF, que se caracteriza clinicamente pela presença de 10-99 adenomas no cólon e um desenvolvimento mais tardio de CCR (por volta dos 50 anos de idade) (Jasperson *et al.*, 2010). Na APAF, as mutações germinais tendem a localizar-se principalmente nas regiões 5', 3' e no exão 9 do gene *APC* (Heinen, 2010).

A polipose adenomatosa associada ao gene *MUTYH* (PAM) mimetiza clinicamente a APAF ao número de pólipos mas apresenta um padrão de transmissão autossômico recessivo, devido a mutações germinais bi-alélicas no gene *MUTYH* (localizado em 1p34). Os indivíduos com estas mutações apresentam um risco aumentado de desenvolver CCR por volta dos 40 anos de idade, no entanto, ainda não é claro o risco associado aos indivíduos com mutações mono-alélicas (heterozigóticos) neste gene (Jasperson *et al.*, 2010; Centelles, 2012, Kashfi *et al.*, 2013).

Os principais síndromes associados à polipose hamartomatosa são o síndrome Peutz-Jeghers (PJS) e o síndrome de polipose juvenil (JPS), os quais estão associados a um risco aumentado para o desenvolvimento de CCR e outros tumores. O PJS apresenta um padrão de transmissão autossômico dominante, causado por mutações no gene *STK11* (localizado em 19p13.3) e caracteriza-se clinicamente pela presença de múltiplos pólipos hamartomatosos no tracto gastrointestinal e lesões cutâneas nos lábios e mucosa oral. Por seu lado, o JPS é caracterizado por um padrão de transmissão autossômico dominante raro, associado a mutações nos genes *BMPR1A* ou *SMAD4* (localizados em 10q22.3 e em 18q21.2, respectivamente) e caracteriza-se clinicamente pela presença de pelo menos 3 pólipos juvenis no cólon (Jasperson *et al.*, 2010; Centelles, 2012).

Por fim, o síndrome de polipose serreada (SPS) é uma condição rara e caracteriza-se pela presença de múltiplos pólipos no cólon, distintos dos pólipos adenomatosos, por apresentarem uma histologia serreada. A idade de diagnóstico é por volta dos 50-60 anos e é frequentemente observada a presença de tumores síncronos e metacrónicos (Jasperson *et al.*, 2010). No entanto, a causa molecular que predispõe para esta doença não é ainda conhecida, tendo sido propostos padrões de transmissão tanto autossômica dominante como recessiva (Leggett e Whitehall, 2010; Patel e Ahnen, 2012).

O HNPCC (câncer do cólon e recto hereditário não associado a polipose) é o síndrome hereditário de CCR mais comum e é definido por critérios clínicos denominados de critérios de Amesterdão (CA). Inicialmente estes critérios apenas contemplavam famílias com CCR, o que os tornava bastante restritos, tendo mais tarde sido revistos de modo a incluírem os tumores extra-cólicos frequentemente associados a estas famílias (Umar *et al.*, 2004) (Tabela 1.2). Este síndrome apresenta um padrão de transmissão autossômica dominante e inclui o síndrome de Lynch e o cancro do cólon e recto familiar do tipo X (FCCTX).

Tabela 1. 2 Critérios de Amesterdão (Umar *et al.*, 2004).

Critérios
<ul style="list-style-type: none"> • Três ou mais familiares afectados com CCR ou tumores extra-cólicos do espectro do síndrome de Lynch (endométrio, ovário, estômago, intestino delgado, tracto hepatobiliar, tracto urinário, cérebro e pele), sendo um deles parente em 1º grau dos outros dois. • Pelo menos duas gerações afectadas com CCR. • Pelo menos um indivíduo com CCR, diagnosticado antes dos 50 anos. • Exclusão do síndrome de polipose adenomatosa familiar (PAF).

O síndrome de Lynch (SL) representa 2-4% de todos os CCR e é definido, actualmente, em base genética pela presença de mutações germinais nos genes do sistema de reparação de erros no DNA do tipo *mismatch* (MMR), nomeadamente nos genes *MLH1*, *MSH2*, *MSH6* e *PMS2* (Boland, 2005). Recentemente, foram descritas deleções germinais do gene *EpCAM* como estando associadas ao SL. Este gene localiza-se a montante do gene *MSH2*, e as deleções genómicas acabam por, invariavelmente, abranger o promotor e/ou os primeiros exões deste último, tratando-se assim de um novo mecanismo de inactivação do *MSH2* (Ligtenberg *et al.*, 2008). A idade de diagnóstico no SL é inferior a 50 anos. Os tumores localizam-se preferencialmente no cólon proximal e para além do CCR, estão também associados tumores extra-cólicos (endométrio, ovário, estômago, intestino delgado, tracto hepatobiliar, tracto urinário, cérebro e pele) (Jasperson *et al.*, 2010). Uma característica específica dos CCR dos doentes com SL é a presença de instabilidade de microssatélites (MSI), devido à *slippage* da DNA polimerase durante a replicação, em particular, em sequências repetitivas (sequências microssatélites), as quais, neste caso, não são reparadas uma vez que estas células possuem um sistema MMR deficiente (Heinen, 2010). Em 1997, foram estabelecidos os critérios de Bethesda, que aproveitavam esta característica dos CCR, com o objectivo de identificarem novos casos de SL em casos que não preenchem os CA. Estes critérios baseiam-se em características clínicas e patológicas, tendo sido revistos em 2003 (Tabela 1.3.) (Umar *et al.*, 2004).

Tabela 1.3. Critérios de Bethesda (adaptado de Umar *et al.*, 2004).

Critérios
<ol style="list-style-type: none"> 1. CCR diagnosticado numa idade inferior a 50 anos; 2. Presença de CCR síncrono ou metacrónico ou outros tumores relacionados com HNPCC (estômago, bexiga, uretra, pélvis renal, tracto biliar, cérebro – glioblastoma, adenomas nas glândulas sebáceas, intestino delgado) independentemente da idade; 3. CCR com instabilidade de microssatélites de alto grau diagnosticado antes dos 60 anos de idade; 4. CCR diagnosticado em doentes com um ou mais parentes em primeiro-grau com CCR ou outros tumores relacionados com HNPCC. Um dos tumores tem de ser diagnosticado antes dos 50 anos de idade; 5. CCR diagnosticado em doentes com dois ou mais parentes com CCR ou tumores relacionados com HNPCC, independentemente da idade.

Assim, a pesquisa de instabilidade de microssatélites é recomendada nos tumores dos indivíduos que preenchem pelo menos um destes critérios, a qual é efectuada com recurso a um painel de 5 marcadores de microssatélites (marcadores de Bethesda). Consoante o número de marcadores de microssatélites onde é detectada instabilidade, assim é classificado o tumor: apresentando instabilidade de alto grau (MSI-H) quando 2 ou mais dos marcadores apresentam

instabilidade; de baixo grau (MSI-L), quando apenas um dos marcadores apresenta instabilidade ou microsatélites estável (MSS) quando não é detectada instabilidade no conjunto dos marcadores (Heinen, 2010).

Por outro lado, as famílias com cancro do cólon e recto familiar do tipo X (FCCTX) são caracterizadas por preencherem os critérios de Amesterdão mas não apresentarem mutações germinais nos genes MMR e os tumores serem MSS (Lindor *et al.*, 2005). Para além disso, comparativamente com o SL, a idade de diagnóstico é mais tardia, os tumores localizam-se preferencialmente no cólon distal ou recto, desenvolvem um maior número de adenomas e os tumores extra-cólicos são raros (Lindor, 2009). Os indivíduos destas famílias apresentam um risco aumentado para desenvolver CCR, mas inferior ao de indivíduos com tumores MSI-H e ainda não se encontra estabelecida qual a base genética deste síndrome (Patel e Ahnen, 2012).

1.7. Via de Transdução de Sinal WNT/ β -catenina

A homeostase normal no tracto intestinal é regulada por um fino equilíbrio entre a proliferação, a diferenciação e a apoptose nas células epiteliais. A desregulação da homeostase conduz à formação tumoral. Uma das vias mais importantes para a regulação deste equilíbrio nas células epiteliais é a via de sinalização WNT canónica, também conhecida como via de sinalização APC/ β -catenina/TCF7L2. Esta via tem como principal função a regulação dos níveis intracelulares de β -catenina (Figura 1.4.) (Reya e Clevers, 2005). WNTs são uma família de glicoproteínas com funções na especificação do destino celular, padronização do sistema nervoso central e da divisão celular assimétrica (Centelles, 2012).

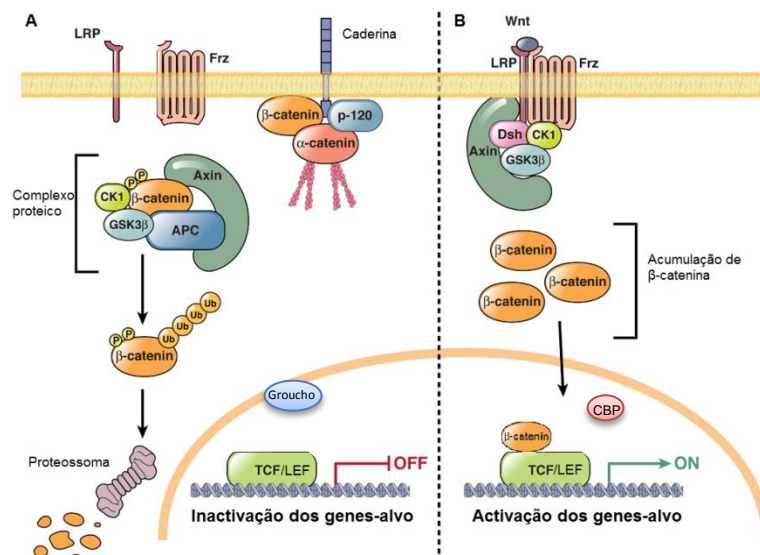


Figura 1.4. Esquema representativo da via WNT/ β -catenina: **A** – via na sua forma inactiva: na ausência de sinalização Wnt, a β -catenina é fosforilada pelo complexo proteico (APC, Axin, GSK3 β e Dsh) e degrada pelo proteossoma. **B** – via na sua forma activa: na presença da ligação das proteínas Wnt ao receptor Frizzled, ocorre a destabilização do complexo proteico e a translocação da β -catenina para o núcleo, onde juntamente com os factores de transcrição TCF/LEF promove a activação de genes alvo da via Wnt (adaptado de Pino e Chung, 2010 e Albuquerque, 2011).

A via WNT/ β -catenina encontra-se na sua forma inactiva na ausência de proteínas WNT (Figura 1.4. – A) e, por isso, a β -catenina é fosforilada por um complexo multiproteico, o qual é formado pela Axin, APC, GSK3- β (*glycogen synthase kinase 3 β*) e Dsh (*Dishevelled*). A β -catenina fosforilada pelo GSK3- β é então reconhecida pela ubiquitina, a qual promove a ubiquitinação e a degradação da β -catenina pelo sistema ubiquitina-proteossoma. Deste modo, os níveis citoplasmáticos de β -catenina mantêm-se baixos, pelo que não há translocação para o núcleo e não há activação dos genes-alvo da via WNT (Pino e Chung, 2010; Centelles, 2012).

A ligação da WNT ao receptor *Frizzled* (Fz) e co-receptor LRP (especificamente LRP-5 e LRP-6) activa a via WNT (Figura 1.4. – B). Uma vez formado o complexo ligando-receptor, o Dsh é activado e recruta a proteína Axin e GSK3- β até à membrana. Neste complexo, o Dsh está activo e o GSK3- β está inactivo, impedindo a fosforilação da β -catenina. Desta forma, a β -catenina acumula-se no citoplasma e é translocada para o núcleo, onde se liga aos factores de transcrição TCF/LEF (*T-cell factor/lymphoid enhancing factor*), regulando a expressão dos genes-alvo da via WNT (Centelles, 2012).

A via WNT/ β -catenina pode ser activada, de uma outra forma, pela presença de mutações em genes desta via, que provocam a sua activação constitutiva. De facto, a maioria dos tumores do cólon e recto apresentam mutações em genes tais como *APC*, *CTNNB1* (codifica β -catenina) ou *AXIN2* (Albuquerque *et al.*, 2010; Albuquerque *et al.*, 2011). Estas mutações têm como consequência a acumulação de β -catenina no citoplasma e a sua translocação para o núcleo, onde a formação do complexo TCF7L2/ β -catenina leva à activação dos genes-alvo da via WNT e, conseqüentemente, à proliferação descontrolada e ao desenvolvimento tumoral (Polakis, 2000; Giles *et al.*, 2003).

1.8. Gene *TCF7L2*

A indução ou repressão dos genes-alvo da via WNT é controlada pelo complexo TCF7L2/ β -catenina, o qual é formado após a translocação da β -catenina para o núcleo. Nesse momento, o co-repressor Groucho liberta-se e os co-activadores são recrutados, activando a expressão dos genes alvo da via WNT. Desta forma, o complexo TCF7L2/ β -catenina controla o equilíbrio entre a proliferação e a diferenciação celular em células saudáveis e tumorais do epitélio intestinal (Polakis, 2000; Giles *et al.*, 2003; Albuquerque *et al.*, 2011).

Os factores de transcrição TCF/LEF constituem uma grande família de reguladores transcripcionais, caracterizada pela presença do domínio HMG (*High Mobility Group*) de ligação ao DNA bastante conservado. Os membros fundadores desta família são Tcf1 e Lef1, os quais foram originalmente identificados em linfócitos e assim, denominados de factor de célula T1 (Tcf1 – *T-cell factor 1*) e factor de ligação potenciador de linfóide (Lef1- *lymphoid enhancer-binding factor 1*) (Travis *et al.*, 1991; van de Wetering *et al.*, 1991).

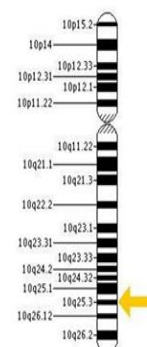


Figura 1.5 Esquema representativo do cromossoma 10 com a localização do gene *TCF7L2* (seta amarela) (adaptado de <http://www.genecards>).

Existem quatro genes que codificam as proteínas de TCF/LEF: *TCF7* (TCF1); *LEF1*, *TCF7L1* (TCF3) e *TCF7L2* (TCF7L2), sendo o último o mais estudado, dos quatro, e o mais importante no cólon e recto como efector da via de sinalização WNT (Duval *et al.*, 2000).

O gene *TCF7L2* (*Transcription Factor 7-like 2*) localiza-se em 10q25.2 (Figura 1.5.) e codifica a proteína TCF7L2, a qual desempenha um papel importante na via de sinalização WNT/ β -catenina.

Este gene é constituído por 17 exões e apresenta diversos locais sujeitos a *splicing* alternativo, principalmente nos exões 4, 13, 14, 15 e 16 (Figura 1.6). Para além destes exões, podem existir variações adicionais nos locais aceitadores e dadores de *splicing* nos exões 7, 8 e 9, respectivamente. A seguir aos exões 8, 15, 16 e 17 podem-se formar codões *stop*, assim, dependendo da sequência de mRNA, formar-se-ão diferentes proteínas (Weise *et al.*, 2010).

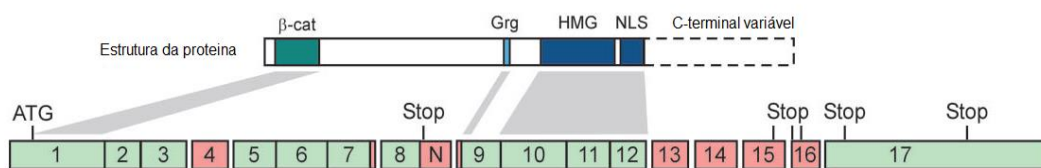


Figura 1.6 Esquema representativo da proteína TCF7L2 (em cima) e dos exões sujeitos a *splicing* alternativo (em baixo). Os exões sujeitos a *splicing* estão representados a vermelho e os exões constitutivos a verde, a posição do codão *start* (ATG) indicada no exão 1 como também os vários codões *stop* (adaptado de Weise *et al.*, 2010).

A proteína TCF4 (Figura 1.7.) é formada por diversos domínios: a região N-terminal é responsável por mediar a interação com a β -catenina; o domínio Grg (*Groucho related gene*), ao qual se liga o co-repressor Groucho; o domínio HMG, o qual é altamente conservado e medeia a ligação específica da sequência de proteínas TCF/LEF ao DNA; como também o local de sinalização nuclear (NLS) adjacente ao HMG; o domínio CRARF (C-cisteína; R-arginina; A-alanina; F-fenilalanina) (também denominado de domínio C-Clamp) que tem como função de auxiliar a ligação ao DNA e os domínios CtBP (*C-terminal-binding protein*), que são um local de ligação adicional para o co-repressor (Duval *et al.*, 2000; Hecht e Stemmler, 2003).

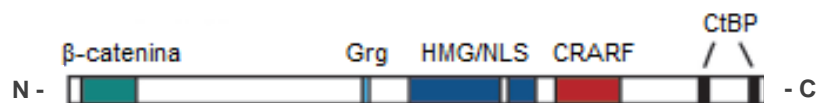


Figura 1.7. Esquema representativo da proteína TCF4. Estão indicados os domínios de ligação à β -catenina (verde), aos co-repressores Grg e CtBP (azul claro e preto, respectivamente), a HMG e o NLS (azul escuro) e o domínio CRARF (vermelho) (adaptado de Hecht e Stemmler, 2003).

Como já foi referido anteriormente, o gene *TCF7L2* apresenta diversas isoformas, as quais sofrem alterações na região C-terminal e, de um modo geral, podem agrupar-se em quatro grupos (Figura 1.8.): a isoforma TCF4E que contém todos os elementos representados na Figura 1.6.; a TCF4N composta apenas pelo local de ligação à β -catenina, não contendo por isso o domínio HMG e todos os elementos localizados a jusante; a TCF4M que não contém o local de ligação à CtBP e pode ter ou não o motivo C-clamp; e por fim, a isoforma TCF4S que é igual à TCF4M com excepção do motivo C-clamp que está incompleto (Duval *et al.*, 2000; Hecht e Stemmler, 2003; Weise *et al.*, 2010).

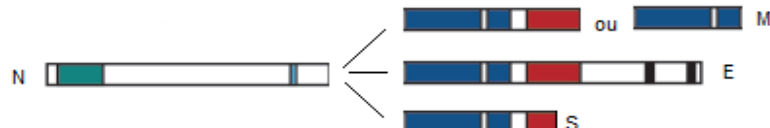


Figura 1.8. Esquema representativo das isoformas do gene *TCF7L2*. TCF4N composta apenas pelo local de ligação à β -catenina; a TCF4M que não contém o local de ligação à CtBP e pode ter ou não o motivo C-clamp ; a TCF4S com motivo C-clamp que está incompleto. Estão indicados os domínios de ligação à β -catenina (verde), aos co-repressores Grg e CtBP (azul claro e preto, respectivamente), a HMG e o NLS (azul escuro) e o domínio CRARF (vermelho) (adaptado de Duval *et al.*, 2000; Hecht e Stemmler, 2003; Weise *et al.*, 2010).

Diferentes combinações dos exões 13, 14, 15 e 16 dão origem a diferentes isoformas (Figura 1.9.), as quais regulam a expressão de diferentes genes da via WNT, pelo que actuam de diferentes modos para o desenvolvimento tumoral. Por norma, as formas curtas da proteína estão associadas a uma maior agressividade do tumor (Weise *et al.*, 2010).

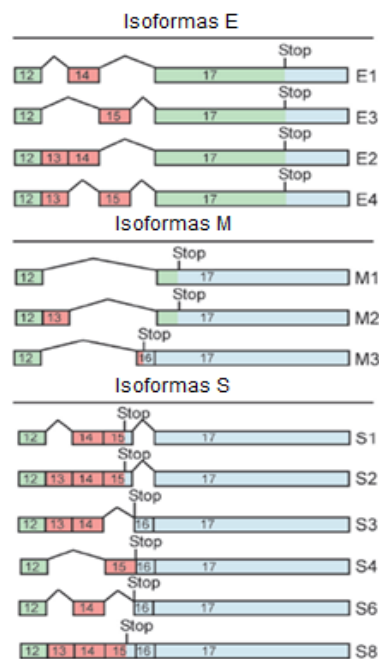


Figura 1.9. Esquema representativo das isoformas do gene *TCF7L2* com as variantes de *splicing* C-terminal identificadas. Os exões sujeitos a *splicing* estão representados a vermelho, os exões constitutivos a verde e a azul são as regiões não traduzidas e os codões *stop* também estão identificados (adaptado de Weise *et al.*, 2010).

1.8.1. Papel do gene *TCF7L2* na tumorigénese colorectal

O papel do gene *TCF7L2* na tumorigénese colorectal é controverso, pois enquanto alguns estudos indicam que este gene funciona como um oncogene outros têm proposto uma função de supressor tumoral.

Em 1998, Korinek e seus colaboradores demonstraram que a deficiência do gene *TCF7L2* resultava na morte pré-natal em ratinhos, devido à ausência do compartimento de proliferação celular no epitélio intestinal (Korinek *et al.*, 1998). Este gene foi descrito como um activador da via de sinalização WNT, apresentando, assim uma função oncogénica (Duval *et al.*, 1999). Mais tarde,

verificou-se que a expressão de uma forma dominante negativa do *TCF7L2* induzia a paragem do ciclo celular em linhas celulares de CCR (van de Wetering *et al.*, 2002). Estas observações estão de acordo com a função oncogénica do *TCF7L2* na cascata de sinalização da via Wnt.

Por outro lado, alguns estudos sugerem que o *TCF7L2* pode actuar como um gene supressor de tumor. Foi demonstrado que a isoforma TCF4E tem propriedades únicas de transactivação dos genes *AXIN2* e *Cdx1* (localizado em 5q32), que contribuem para a supressão tumoral (Hecht e Stemmler, 2003). Assim, parece haver uma relação entre as isoformas longas e a função de supressor de tumor. Também se demonstrou que a haplo-insuficiência do *TCF7L2* juntamente com a hiperactivação da via de sinalização WNT, devido à mutação do *APC*, tinha como resultado o desenvolvimento de tumores no cólon, o que sugere que a proteína TCF7L2 funcione normalmente como supressor tumoral, pela modulação da proliferação (Angus-Hill *et al.*, 2011). Alguns estudos relacionam a inibição do gene *TCF7L2* com o aumento da expressão do HGF (*hepatocyte growth factor* – factor de crescimento de hepatócitos), o qual faz parte da via de sinalização HGF/MET que tem como função a regulação da invasão e metastização tumoral. Desta forma, este resultado apoia a hipótese do *TCF7L2* estar associado à supressão tumoral, através da inibição da metastização e invasão (Tang *et al.*, 2008). Além disso, Varnat e os seus colaboradores, demonstraram que a inibição do *TCF7L2* está associada a um maior desenvolvimento de metástases em ratinhos (Varnat *et al.*, 2010). Uma vez que, estas descobertas sugerem que a diminuição da expressão do *TCF7L2* pode ser responsável pelo desenvolvimento de metástases, pela indução do HGF, as alterações moleculares neste gene poderão ser potenciais biomarcadores do desenvolvimento tumoral e da metastização.

Contudo, estas funções opostas podem reflectir as actividades específicas dos promotores de diferentes variantes de *splicing* do *TCF7L2*, regulando a expressão de grupos distintos de genes-alvo da via WNT (Weise *et al.*, 2010).

Num estudo prévio realizado pelo grupo de Gastrenterologia da UIMP, IPOLFG, EPE., verificou-se que a deleção de uma adenina na região repetitiva de 9 adeninas [(A)9] do exão 17 do gene *TCF7L2* (Duval *et al.*, 1999) (região coincidente com o domínio *C-clamp*) em tumores colorectais com deficiência no sistema MMR, isto é, apresentavam instabilidade de microssatélites de alto grau (MSI-H), aumentava a sua frequência ao longo da sequência adenoma-carcinoma (33% vs 65%, $P=0,047$), sugerindo, assim, um possível papel desta alteração no desenvolvimento tumoral. Nos CCR com deficiência no sistema MMR, estas mutações foram significativamente mais frequentes nos tumores localizados no cólon proximal ($P=0,046$). Portanto, estes resultados sugerem, à semelhança do que tem sido demonstrado para um tipo específico de mutações no gene supressor tumoral *APC* e no gene *CTNNB1* (Albuquerque *et al.*, 2010; Albuquerque *et al.*, 2011), que mutações na região (A)9 no gene *TCF7L2* podem contribuir para a formação tumoral. Deste modo, é importante analisar os restantes exões do gene *TCF7L2* para identificar outras mutações como potenciais biomarcadores moleculares.

Além disso, foi identificada uma alteração específica no gene *TCF7L2* em 22% dos CCR, a qual revelou associação com o estadio IV, ou seja, nos tumores com metástases síncronas

(metástases diagnosticadas até 6 meses após o diagnóstico do CCR). Estes resultados sugerem a associação de alterações no *TCF7L2* com molecular progressão tumoral.

1.9. Gene *TCF7L2* como marcador preditivo de resposta ao tratamento em CCR metastizado e tumores do recto localmente avançados

Ghadimi e seus colaboradores verificaram que a sobre-expressão do gene *TCF7L2* estava associada a uma maior resistência à quimioterapia com 5-FU em tumores do recto, comparando tumores resistentes e sensíveis de doentes que tinham sido tratados com quimioterapia pré-operatória (Ghadimi *et al.*, 2005). Recentemente, a redução da expressão do *TCF7L2* foi associada a uma maior sensibilidade à quimiorradioterapia em linhas celulares de cancro do recto (Kendziorra *et al.*, 2011).

A sobre-expressão do gene *TCF7L2* em tumores do recto localmente avançados resistentes à quimiorradioterapia e o facto do silenciamento deste gene aumentar a sensibilidade do CCR à terapia, sugere que a sua proteína pode actuar em diferentes fases do desenvolvimento tumoral com diferentes impactos que podem ser dependentes do nível de expressão e do *splicing* alternativo, o que terá consequências distintas em termos de níveis de sinalização WNT. Este fenómeno parece assemelhar-se à selecção de mutações com base no nível de sinalização WNT óptimo para a tumorigénese, o qual é dependente do tempo e do tecido à proposta para as mutações nos genes *APC* e *CTNNB1* de acordo com o modelo de sinalização *just-right* (Albuquerque *et al.*, 2002; Albuquerque *et al.*, 2011). Desta forma, a análise de possíveis alterações (epi)genéticas no *TCF7L2*, em tumores do recto localmente avançado, e a sua correlação com a resposta à quimiorradioterapia pré-operatória é extremamente importante de modo a identificar marcadores preditivos de resposta, os quais terão implicações significativas no tratamento destes doentes.

1.10. Objectivos

O presente trabalho teve como principal objectivo estudar alterações moleculares no gene *TCF7L2*, e avaliar o seu eventual papel como factores de prognóstico (em particular biomarcadores de metastização) no cancro do cólon e recto e avaliar alterações (epi)genéticas e a sua relação com a resposta à quimiorradioterapia pré-operatória no cancro do recto.

Com esta finalidade pretendeu-se:

- 1) Realizar a pesquisa de mutações em exões específicos do *TCF7L2* e respectivas regiões intrónicas flanqueadoras em tumores do cólon e recto. Relacionar os resultados com as características clínicas dos doentes e histopatológicas dos tumores indicativas de mau prognóstico.
- 2) Analisar alterações (epi)genéticas no gene em estudo (perda de heterozigotia, variações de *copy number* e metilação) em tumores do recto e correlacionar com a resposta à quimioterapia pré-operatória, de modo a identificar marcadores preditivos de resposta à terapia neste tipo de tumores.

2. Materiais e Métodos

2.1. Amostras Biológicas

Neste projecto, foram seleccionados doentes que são seguidos no serviço de Gastrenterologia e na Consulta de risco familiar de CCR da Clínica de Risco Familiar do Instituto Português de Oncologia de Lisboa Francisco Gentil, E.P.E. (IPOLFG, E.P.E.).

Para a pesquisa de mutações no gene *TCF7L2*, foram incluídas 68 amostras de tumor: 60 de doentes com CCR esporádico, dos quais 4 apresentavam instabilidade de microssatélites e 56 eram microssatélites estáveis, e 8 tumores provenientes de doentes com síndrome de Lynch. Esta análise foi realizada a partir de DNA genómico extraído a partir da mucosa tumoral de cada indivíduo.

Foram ainda incluídas 17 amostras de DNA tumoral e respectivas mucosas normais, de igual número de doentes com tumores do recto, de modo a correlacionar as alterações (epi)genéticas com a resposta à quimiorradioterapia pré-operatória. Destes 17 indivíduos, 6 tiveram resposta completa após quimiorradioterapia pré-operatória (sem células neoplásicas); 5 tiveram resposta moderada (tumor residual viável) e 5 não responderam à terapia. Nesta análise foi utilizado DNA genómico, extraído a partir de tecido (mucosa normal e tumoral) incluído em parafina e tecido congelado. As amostras de tecido tumoral e normal foram provenientes do Serviço de Anatomia Patológica do IPOLFG, E.P.E.. Nos doentes em que não foi possível obter mucosa normal, foi utilizado DNA extraído a partir de sangue periférico, como DNA de referência. Para a análise do número de cópias (*copy number*) foram utilizadas 3 amostras normais adicionais como referência para cada tipo de material utilizado.

Na análise de *real-time* PCR para avaliar o efeito das alterações de *copy-number* ao nível da expressão, foram utilizadas 5 amostras de RNA, obtido a partir de sangue periférico, para sintetizar DNA complementar (cDNA), entre as quais 3 apresentavam alterações de *copy number* e 2 pertenciam a indivíduos saudáveis.

2.2. Extracção de ácidos nucleicos

2.2.1. Extracção de DNA genómico

Neste estudo, o DNA genómico foi extraído a partir de três tipos de amostras biológicas diferentes: sangue periférico, tecido incluído em parafina e tecido congelado. O DNA genómico da maioria das amostras utilizadas neste estudo já tinha sido previamente extraído.

2.2.1.1. Extracção de DNA genómico a partir de tecido incluído em parafina

O DNA genómico extraído a partir de tecido incluído em parafina foi obtido através de uma extracção fenol-clorofórmio, a qual se encontra descrita no anexo 1 (ponto 1.1).

Resumidamente foram utilizados cerca de 10 cortes de 10µm de tecido incluído em parafina por amostra, os quais foram submetidos a uma digestão, *over-night* (o/n) a 56°C, com uma solução de lise celular, juntamente com proteinase K. Após a digestão do tecido, foi efectuada a desnaturação e precipitação das proteínas, por adição de uma solução de fenol/clorofórmio, seguida de centrifugação. À fase aquosa, foi então, adicionada uma solução de clorofórmio/álcool isoamílico de modo a remover completamente o fenol e evitar que a formação de espuma. O DNA foi precipitado com etanol absoluto, o qual também contribui para a remoção de vestígios de sais, de fenol e clorofórmio, e por fim, ressuspendido em TE pH8 (*Citogene*, Portugal).

2.2.1.1.1. Quantificação de DNA por electroforese em gel de agarose 0,8% (p/v)

O DNA genómico obtido a partir de tecido incluído em blocos de parafina não deve ser quantificado espectrofotometricamente, uma vez que alguns dos reagentes utilizados na extracção absorvem ao mesmo comprimento de onda do DNA (260nm). Assim, a quantificação do DNA foi efectuada por electroforese em gel de agarose 0,8% (p/v) numa solução tamponada de TBE1x pH8 (TBE10x, *National Diagnostic*, E.U.A.), utilizando três amostras de DNA com concentração conhecida (400ng/µl; 800ng/µl e 1600 ng/µl) para estimativa da concentração das demais amostras.

Para a quantificação, adicionou-se 1µl da amostra de DNA extraído a 9µl de solução de deposição *Orange G* 1x, o qual confere densidade, permitindo a deposição da amostra no poço do gel, assim como a visualização da mesma no decorrer da electroforese. Para além das amostras, aplicaram-se também dois marcadores de peso molecular: marcador *GeneRuler 50 bp DNA Ladder* (com um padrão de bandas entre 50-1000) (*Thermo Scientific*, E.U.A.) e marcador *Lambda DNA/HindIII* (com um padrão de bandas entre 125-23130pb) (*Fermentas*, E.U.A.).

Na preparação do gel, adicionou-se brometo de etídeo (10mg/ml, *Invitrogen*), o qual se intercala com a molécula de DNA durante a electroforese e quando exposto a radiação ultravioleta (UV), permite a sua visualização. Esta electroforese, realizou-se numa tina horizontal, sob voltagem constante de 140V durante, aproximadamente, 60 minutos.

Após a electroforese, colocou-se o gel num transiluminador, que contém uma câmara fotográfica e um sistema informático acoplado (*BioDocAnalyse*, *Biometra*, Alemanha), de modo a observar os padrões das amostras de DNA quando expostas a radiação UV. Após a quantificação, foram efectuadas alíquotas de trabalho para uma concentração de 80ng/µl, as quais foram armazenadas a -20°C juntamente com as amostras de DNA *stock* extraído.

A preparação de todas as soluções utilizadas nesta técnica encontra-se descrita no anexo 2.

2.2.1.2. Extracção de DNA genómico a partir de sangue periférico

A extracção de DNA genómico foi efectuada a partir de cerca de 3ml de sangue periférico colhidos para um tubo de hemograma contendo EDTA (ácido etilenodiaminotetracético – agente quelante de cálcio que evita a coagulação do sangue) tendo-se utilizado o *kit* comercial *CITOGENE*[®] – *Genomic DNA Purification Kit* (*Citomed*, Portugal), de acordo com o descrito pelo fabricante (Anexo 1, ponto 1.2).

O primeiro passo deste protocolo consistiu na adição de uma solução de lise de eritrócitos ao sangue periférico, a qual foi efectuada por duas vezes de modo a garantir a eliminação total destas células. Em seguida, foi efectuada a lise dos leucócitos, para a obtenção do DNA genómico, incubou-se a solução com uma RNase de modo a eliminar o RNA e precipitaram-se as proteínas. O DNA genómico foi precipitado com isopropanol 100% e o *pellet* de DNA foi lavado com etanol 70% (v/v). As amostras foram, por fim, ressuspensas em TE pH8 e guardadas a 4°C até serem quantificadas.

2.2.1.2.1. Quantificação de DNA por espectrofotometria

De modo, a quantificar o DNA obtido a partir de sangue periférico, realizou-se um doseamento espectrofotométrico. Esta análise baseia-se na lei de Lambert-Beer, a qual descreve que a absorvância de uma determinada amostra, a um comprimento de onda específico e a uma espessura fixa é directamente proporcional à concentração do material absorvente (Fórmula 2.1), ou seja:

$$A = \epsilon bc$$

Fórmula 2.1. Expressão matemática que traduz a lei de Lambert-Beer.

Legenda: A - absorvância, ϵ - coeficiente de extinção molar ($M^{-1}cm^{-1}$), b - distância percorrida pela radiação através da solução (cm) e c - concentração molar da amostra (M).

Uma vez que as purinas e as pirimidinas, que constituem o DNA, absorvem a radiação ultravioleta (UV), as leituras espectrofotométricas realizadas para calcular a concentração de ácidos nucleicos na amostra foram realizadas a 260 nm (Sambrook e Russel, 2001) utilizando um *Nanodrop 2000* (*Thermo Scientific*, EUA), de acordo com as especificações do fabricante.

Após a quantificação, efectuou-se uma alíquota de trabalho, em água bidestilada, para uma concentração final de 80ng/ μ l, sendo as amostras por fim armazenadas a -20°C.

2.2.1.3. Extracção de DNA genómico a partir de tecido congelado

Para a extracção de DNA genómico a partir de tecido congelado, partiu-se de cerca de 150mg de tecido congelado e utilizou-se o *kit* comercial *CITOGENE* – *Genomic DNA Purification Kit* (*Citomed*), tendo sido realizado o protocolo descrito pelo fabricante com algumas alterações (Anexo 1). Este protocolo baseia-se nos mesmos princípios descritos anteriormente (ver ponto 2.1.2.). Após a extracção, o DNA foi quantificado por espectrofotometria (ver ponto 2.1.2.1.) e efectuada uma alíquota de trabalho, em água bidestilada, para uma concentração final de 80ng/ μ l, sendo as amostras armazenadas a -20°C.

2.2.1.4. Extração de RNA a partir de sangue periférico

O RNA foi extraído a partir de linfócitos que tinham sido anteriormente isolados a partir de sangue periférico. O protocolo utilizado foi o descrito para o *TRI Reagent*[®] (Ambion, Life Technologies) (ver anexo 1, ponto 1.4). Após quantificação espectrofotométrica (ver ponto 2.2.1.2.1), os tubos contendo o RNA foram armazenados a -80°C. O RNA extraído também foi sujeito a uma electroforese em gel de agarose 0,8% (p/v) para verificar a integridade e para descartar uma possível contaminação por DNA. No gel de agarose deve-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 com maior peso molecular das duas bandas de RNA ribossomal, indica que há contaminação por DNA genómico.

2.3. Polymerase Chain Reaction (PCR)

A técnica de *Polymerase Chain Reaction* (PCR) ou Reacção em Cadeia da Polimerase consiste na amplificação enzimática *in vitro* de fragmentos específicos de DNA, através da utilização de dois *primers* (curtas sequências sintéticas de nucleótidos, entre 20 a 25pb, que marcam as extremidades da sequência alvo) de sequência conhecida e específica para o fragmento a amplificar, que hibridam pela extremidade 3'-OH com a região complementar DNA e flanqueiam a região de interesse (Elrich, 1989).

Para além da quantidade mínima de DNA genómico e do par de *primers* (*forward* - 5' e *reverse* - 3'), para que ocorra a reacção de PCR é necessário: uma enzima para a síntese de DNA, como a *Taq* polimerase (esta enzima foi isolada a partir da bactéria *Thermus aquaticus*, sendo termoestável e termorresistente, consegue manter a sua actividade durante cada ciclo de PCR); de uma solução tampão (mantém pH óptimo para a actividade da enzima); de dNTP's (desoxirribonucleótidos fosfatados utilizados na síntese das novas cadeias) e de iões magnésio (co-factor da enzima *Taq* polimerase, o qual forma um complexo solúvel com os dNTP's) (Sambrook e Russel, 2001).

Uma reacção de PCR consiste numa série de ciclos, sendo que cada um é composto por três etapas distintas: desnaturação da molécula de DNA; hibridação dos *primers* com as sequências alvo em cadeia simples e extensão dos *primers* à temperatura óptima para a síntese, por parte da DNA polimerase. Desta forma, os produtos sintetizados num ciclo servem de molde ao próximo ciclo, levando a uma amplificação exponencial do DNA (Elrich, 1989).

2.3.1. Optimização das condições das reacções de PCR

Inicialmente optimizaram-se as condições de PCR para cada fragmento a ser amplificado, de modo a obter um produto específico correspondente à região de interesse. Para isso, foram ajustados diversos parâmetros, tais como: a temperatura de hibridação dos *primers*; a concentração de Mg²⁺ e o número de ciclos da reacção.

A temperatura de hibridação dos *primers* tem um papel muito importante na eficiência da reacção de PCR, tendo-se recorrido à regra de Wallace (Fórmula 2.2) para a sua determinação. Esta regra descreve que a temperatura de *melting* (T_m) é a temperatura, à qual 50% dos *primers* se ligam especificamente à cadeia simples de DNA, ou seja:

$$T_m (^{\circ}\text{C}) = 4(\text{G} + \text{C}) + 2(\text{A} + \text{T})$$

Fórmula 1.2 Expressão matemática que traduz a lei de Wallace.

Legenda: T_m - temperatura de *melting*, G - guanina, C - citosina, A - adenina e T - timina

Ao calcular a média da T_m do par de *primers* (*forward* e *reverse*), obteve-se a temperatura de hibridação teórica destes *primers*, tendo-se efectuado em simultâneo, uma reacção de PCR de optimização à temperatura calculada e outra com menos 2°C.

A concentração de ião magnésio também desempenha um papel crítico no PCR, pois influencia directamente o rendimento da reacção, uma vez que a *Taq* polimerase é dependente deste ião, assim como a sua especificidade. Inicialmente, nas reacções de optimização de PCR, foi efectuada uma titulação de ião magnésio com três volumes diferentes (0,5µl; 1µl e 1,5µl).

Outro parâmetro que pode ser alterado é o número de ciclos por reacção, geralmente entre 30 a 40, o que influencia a quantidade de produto específico correspondente à região de interesse.

De um modo geral, para a optimização de uma reacção de PCR na qual não se verificou amplificação do fragmento de interesse, diminuiu-se a temperatura de hibridação dos *primers* e/ou aumentou-se a concentração de magnésio e/ou o número de ciclos. Por outro lado, quando se observou a existência de fragmentos inespecíficos, aumentou-se a temperatura de hibridação dos *primers* e/ou diminuiu-se a concentração de magnésio e/ou o número de ciclos. Por fim, escolheram-se as condições para as quais se verificou uma melhor eficiência do produto de PCR para cada fragmento a analisar.

Para além destes parâmetros, visto que a maioria do DNA genómico utilizado neste estudo foi obtido através de tecido incluído em parafina e, por isso, encontrar-se mais fragmentado, o que dificulta a amplificação de alguns fragmentos, recorreu-se nestes casos à utilização da *AmpliTaq Gold[®] DNA polymerase* (*Applied Biosystems*, E.U.A.), a qual se apresenta na forma inactiva, necessitando de um *hotstart* térmico a 95°C para iniciar a sua actividade. Esta característica torna-a numa enzima mais robusta, o que permite melhorar a eficiência e rendimento do produto de PCR uma vez que impede a síntese de produtos inespecíficos que normalmente se formam a temperaturas sub-optimas de amplificação, e conseqüentemente, disponibiliza os *primers* para se ligarem apenas à região de interesse.

2.3.2. Amplificação de alguns exões do gene *TCF7L2* por PCR

Dos 17 exões do gene *TCF7L2*, pretenderam-se analisar 10 (1, 4, 7, 8, 9, 13, 14, 15, 16 e o 17) e as respectivas regiões intrónicas flanqueadoras, pelo facto de para estes exões estar descrita uma maior frequência de mutações somáticas ou serem sujeitos a *splicing* alternativo (Duval *et al.*, 2000).

Como a maioria das amostras utilizadas neste estudo provieram de tecido incluído em parafina, os fragmentos amplificados não devem exceder os 200pb e, por isso, os *primers* utilizados foram desenhados de modo a obter produtos de PCR inferiores a este valor. Por esta razão, não foi possível amplificar o exão 17 (418pb) num único fragmento, pelo qual este foi dividido em quatro fragmentos sobreponíveis (A-D).

Neste projecto, foram analisados alguns exões por sequenciação automática directa (1, 15, 16, 17B, 17C e 17D), outros exões por DGGE (4, 13 e 14; não foi possível determinar as condições de amplificação dos restantes exões) e o fragmento 17A por análise de fragmentos utilizando o *software Genescan*, uma vez que a região amplificada contém uma sequência repetitiva, a qual varia de tamanho quando mutada sendo facilmente distinguível por este método. Devido às metodologias utilizadas os *primers* tiveram de ser sintetizados com algumas modificações: o *primer forward* utilizado, para a amplificação dos produtos de PCR usados na análise por DGGE, contém uma cauda rica em GC (*GC-clamp*), enquanto o *primer forward* do fragmento 17A utilizado para a análise de fragmentos foi marcado com um fluorocromo (FAM). Para os fragmentos analisados por DGGE foi encomendado um *primer forward* sem cauda *GC-clamp* ou desenhado um novo *primer forward* ou um novo par de *primers* para permitir a sequenciação automática das variantes detectadas nos géis de DGGE. As sequências de todos os *primers* utilizados para a amplificação das regiões de interesse encontram-se descritas no anexo 3.

Cada reacção de PCR foi preparada para um volume final de 12,5µl em tubo tipo *ependorf* de 0,2ml, utilizando 1µl de DNA genómico como *template* (80ng) e 1µl de dNTPs (200 µM, *Illustra*TM, GE *Healthcare*), sendo que as condições utilizadas na amplificação dos diversos fragmentos em estudo encontram-se descritas no anexo 4 (Tabela 1).

Para além das amostras a amplificar, realizou-se simultaneamente um controlo negativo para cada reacção de PCR, de modo a garantir a inexistência de qualquer contaminação. Este controlo foi sujeito às mesmas condições de PCR, continha a mesma mistura reaccional, mas em lugar do DNA genómico foi adicionado 1µl de água bidestilada.

As reacções de PCR foram colocadas num termociclador de acordo com os programas descritos no anexo 4.

2.3.2.1. Controlo da eficiência dos produtos amplificados por electroforese em gel de agarose

De modo a verificar a eficiência da reacção de PCR, efectuou-se o controlo desta através da realização de uma electroforese em gel de agarose 2% (p/v), o qual permite a separação de fragmentos de DNA de acordo com o seu peso molecular por aplicação de um campo eléctrico.

Para o controlo da eficiência da optimização das reacções de PCR, adicionou-se ao volume total da reacção (12,5µl) 3µl de solução de deposição (*Orange G* 5x), enquanto no caso das amostras a analisar utilizou-se 3µl do produto de PCR ao qual se adicionou 7µl de solução de deposição (*Orange G* 1x). As amostras foram aplicadas no gel juntamente com o marcador de peso molecular *GeneRuler 50 bp DNA Ladder* (*Thermo Scientific*, EUA), de modo a se verificar se o tamanho da

banda observada correspondia ao peso molecular do produto de PCR esperado. A electroforese realizou-se a uma voltagem constante de 140V durante, aproximadamente, 30 minutos. Posteriormente, a visualização das bandas dos produtos de PCR foi efectuada num transiluminador com um sistema de imagem acoplado (*BioDocAnalyze*, *Biometra*, Alemanha) sob radiação UV.

A preparação de todas as soluções utilizadas nesta técnica encontra-se descrita no anexo 2.

2.4. Sequenciação Automática

A sequenciação automática é um método que se baseia no método de Sanger, também conhecido por método de terminação de cadeias, o qual permite determinar a sequência exacta do fragmento de DNA (Sanger *et al.*, 1977).

Esta técnica envolve a síntese *de novo* de uma série de cadeias simples de DNA, que possuem nucleótidos marcados numa das extremidades e que diferem entre si apenas por um nucleótido, usando como molde o fragmento de DNA que se quer sequenciar. A síntese das cadeias inicia-se por um *primer* (*forward* ou *reverse*) e termina por incorporação de nucleótidos terminadores. Estas moléculas são derivadas de desoxinucleótidos (dNTPs) mas não possuem o grupo hidroxilo na posição 3' da desoxirribose (2',3'-didesoxinucleótidos trifosfatos - ddNTPs) e, por isso, impedem as ligações fosfodiéster ao DNA. Consequentemente, termina a incorporação de nucleótidos à cadeia de DNA que está a ser sintetizada pela DNA polimerase, o que resulta na formação de uma cadeia truncada.

Durante a electroforese capilar, os fragmentos de DNA sintetizados são separados de acordo com o seu tamanho (em pb) e quando o laser incide sobre estes, é emitido um sinal correspondente ao ddNTP incorporado (cada nucleótido está marcado com um fluorocromo de cor diferente) em cada uma das cadeias de DNA truncadas, o qual ao ser detectado pelo detector de fluorescência, vai permitir a identificar toda a sequência de bases (Lodish *et al.*, 2004).

Esta informação é depois transformada, por um *software*, num electroforetograma, no qual cada pico corresponde a um nucleótido, sendo possível determinar a sequência do DNA de interesse.

2.4.1. Purificação enzimática dos produtos da reacção de PCR

A preparação das amostras para a sequenciação é um passo extremamente importante, para a eficiente leitura das sequências. Por este motivo, antes da realização da reacção de sequenciação submeteu-se cada produto de PCR a uma purificação enzimática, para eliminar a quantidade de *primers* que não hibridaram e outros reagentes da mistura reaccional que possam interferir na reacção de sequenciação.

Para a purificação, preparou-se, em gelo, uma mistura reaccional para cada amostra, com 1µl de *FastAP - Thermosensitive Alkaline Phosphatase* (1U/µl *Thermo Scientific*, E.U.A.) a qual catalisa a remoção dos grupos fosfato do DNA e 0,5µl de *Exo I - Exonuclease I* (20U/µl *Thermo Scientific*, E.U.A.) de modo a degradar o DNA em cadeia simples e agitou-se em *vórtex*. Em seguida, distribuiu-se 1,5µl da mistura reaccional por cada tubo tipo *ependorf* de 0,2ml e colocaram-se as amostras

num termociclador para uma incubação a 37°C durante 15 minutos seguida da inactivação das enzimas a 85°C durante 15 minutos. Após a purificação, as amostras podem ser utilizadas de imediato ou armazenadas a -20°C.

2.4.2. Reacção de Sequenciação

A reacção de sequenciação foi efectuada em duplicado para cada amostra, de modo a que fosse possível determinar a sequência nucleotídica em ambas as cadeias. Assim, prepararam-se, em tubos tipo *ependorf* de 0,2ml, duas misturas reaccionais para cada fragmento a sequenciar (uma com o *primer forward* e outra com o *primer reverse*) para um volume final de 20µl com os seguintes componentes: uma quantidade variável de DNA genómico purificado (dependente da intensidade da banda obtida na electroforese em gel de agarose de 2% (p/v) como descrito no ponto 2.3.3.1.); 2µl de solução tamponada de sequenciação *Buffer Sequencing 5x, BigDye™ Terminator v1.1 (Applied Biosystems)*; 2µl do respectivo *primer* (5' ou 3') (1,6pmol/µl) do fragmento a analisar; quantidade variável da mistura reaccional de sequenciação *BigDye™ Terminator v1.1 (Applied Biosystems)* (dependente do tamanho em pb de cada fragmento e da intensidade da banda obtida na electroforese em gel de agarose de 2% (p/v) e água bidestilada para perfazer o volume total. Por fim, colocaram-se os tubos num termociclador *Veriti (Applied Biosystems)* e iniciou-se a reacção de sequenciação utilizando o programa referido na tabela 2.1..

Após a reacção de sequenciação, armazenaram-se os tubos a 4°C até posterior utilização (período máximo de 24h). A temperatura de hibridação é igual à utilizada na reacção de PCR, excepto quando esta é superior a 60°C, nesses caso realiza-se esta etapa a 59,5°C, uma vez que neste programa a extensão ocorre a 60°C.

Tabela 2.1. Programa utilizado para a reacção de sequenciação no termociclador *Veriti (Applied Biosystems)*.

Etapa	Temperatura	Tempo	Ciclos
Desnaturação inicial	96°C	5 minutos	1
Desnaturação	95°C	10 segundos	
Hibridação	Variável- até 59,5°C	5 segundos	25
Extensão	60°C	4 minutos	
Pausa	4°C	∞	-

2.4.3. Precipitação e purificação dos produtos da reacção de sequenciação

Após a reacção de sequenciação foi necessário purificar os produtos da reacção, de modo a remover ddNTPs não-incorporados assim como vestígios dos restantes reagentes utilizados na reacção e precipitar o DNA. Para tal, seguiu-se o protocolo Etanol/EDTA/Acetato de Sódio

recomendado pelo *kit* de sequenciação *BigDye[®] Terminator v1.1 Cycle Sequencing (Applied Biosystems)* (ver anexo 5).

A cada produto de reacção, adicionou-se uma mistura reaccional contendo: EDTA – quelata os iões de magnésio que estão na reacção; acetato de sódio – neutraliza as cargas dos ácidos nucleicos e etanol absoluto – juntamente com o acetato de sódio promove a precipitação do DNA; e incubou-se à temperatura ambiente. Em seguida centrifugaram-se os tubos, descartou-se o sobrenadante e lavou-se o *pellet* de DNA com etanol 70% (v/v), para remover o excesso de sais. Centrifugou-se novamente, retirou-se completamente o sobrenadante e secou-se o *pellet* de DNA a 37°C. Por fim, armazenaram-se os tubos a 4°C até se proceder à electroforese capilar.

2.4.4. Preparação das amostras para a electroforese capilar

De modo a preparar as amostras para a electroforese capilar, ressuspenderam-se os *pellets* de DNA em 17µl de formamida *Hi-Di (Applied Biosystems)* (agente desnaturante), agitaram-se os tubos em *vórtex* e transferiu-se todo o conteúdo para tubos ou para uma placa de 96 poços (*Platemax, Axygen, EUA*) apropriados para o sequenciador automático *ABI Prism[™] 310 Genetic Analyzer (Applied Biosystems)* ou *ABI Prism[™] 3130 Genetic Analyzer (Applied Biosystems)*, respectivamente. Em seguida, desnaturaram-se as amostras num termociclador a 95°C durante, pelo menos, 5 minutos e colocaram-se em gelo entre 2 a 3 minutos. No caso da placa de 96 poços, foi em seguida, centrifugada a 1200 rpm durante 2 minutos.

Por fim, colocaram-se as amostras no sequenciador automático *ABI Prism[™] 310 Genetic Analyzer (Applied Biosystems)* ou no *ABI Prism[™] 3130 Genetic Analyzer (Applied Biosystems)*. No sequenciador, cada amostra foi submetida a uma electroforese capilar durante, aproximadamente, 20 minutos a 50°C, a uma voltagem constante de 15KV (*ABI Prism[™] 310 Genetic Analyzer*) ou 13,2KV (*ABI Prism[™] 3130 Genetic Analyzer*).

2.4.5. Análise de Resultados

A sequência de DNA de cada uma das amostras foi obtida recorrendo ao *software Sequencing Analysis 3.4.1 (Applied Biosystems)*, que converte os sinais de fluorescência em electroforetogramas. As sequências foram lidas manualmente, por comparação com as sequências de DNA referência (base de dados *Ensembl*), de modo a identificar qualquer possível alteração.

2.5. Denaturing Gradient Gel Electrophoresis (DGGE)

A técnica *Denaturing Gradient Gel Electrophoresis (DGGE)* ou electroforese em gel de poliacrilamida com gradiente desnaturante baseia-se na migração de moléculas de DNA em cadeia dupla ao longo de géis de poliacrilamida, os quais contêm agentes desnaturantes (formamida e ureia)

em concentrações crescentes. Este método permite detectar com grande sensibilidade e eficácia mutações pontuais na sequência de DNA (Fodde e Losekoot, 1994).

Um parâmetro bastante importante nesta metodologia é a temperatura de *melting* (T_m), que se define por ser a temperatura à qual cada base da molécula de DNA está em perfeito equilíbrio entre o estado helicoidal e o estado desnaturado, isto é, quando a cadeia dupla de DNA está meio desnaturada. Assim, num fragmento de DNA podem existir várias regiões com diferentes temperaturas de *melting*, designadas por domínios de desnaturação, os quais podem variar entre 50 a 300pb. Quando a molécula de DNA migra ao longo do gel de poliacrilamida com gradiente desnaturante, os seus domínios de desnaturação vão sendo progressivamente desnaturados em cadeia simples. Este fenómeno depende da estabilidade da dupla hélice, a qual é determinada pela composição da sequência em guaninas e citosinas e pela interacção entre pares de bases adjacentes.

Em indivíduos heterozigotos, durante a amplificação por PCR do fragmento de interesse, a desnaturação e a renaturação contínua das cadeias de DNA resulta na formação de moléculas homoduplex (emparelhamento de duas cadeias simples de DNA normais ou mutadas) e heteroduplex (emparelhamento entre uma das cadeias simples de DNA normal e uma mutada). Assim, quando esse fragmento de DNA atinge uma determinada posição ao longo do gradiente desnaturante, que iguala a T_m do seu domínio de desnaturação mais baixo, ocorre uma desnaturação parcial e, conseqüentemente, uma diminuição na mobilidade do fragmento. Desta forma, os fragmentos de DNA, que diferem apenas num único nucleótido no seu domínio de desnaturação mais baixo, podem ser distinguidos, uma vez que a sua desnaturação e, conseqüente, retardação na mobilidade electroforética ocorrem em posições diferentes no gel. Visto que a migração dos heteroduplex é mais lenta, aparecem duas bandas adicionais no gel de gradiente desnaturante, facilitando a detecção de mutações.

Inicialmente, não era possível identificar os indivíduos heterozigóticos quando a substituição nucleotídica num determinado fragmento de DNA se encontrava no domínio de desnaturação com T_m mais elevado, pois ocorria a dissociação completa das cadeias de DNA e era perdida a capacidade de migração dependente da sequência. De modo a resolver este problema, um dos *primers* contém um domínio termoestável rico em GC, denominado por *GC-clamp*, o qual possui um elevada T_m , evitando a completa desnaturação da molécula e permitindo a detecção de alterações nucleotídicas em todos os domínios de desnaturação.

A capacidade de separação entre as cadeias homoduplex e heteroduplex é dependente do gradiente do gel e do tempo da electroforese. Como primeira tentativa, os géis foram realizados com uma diferença de percentagem de agentes desnaturantes de 30-70% (v/v) desde a base do gel até ao topo e submetidos a uma electroforese de 4 horas, de modo a identificar em que intervalo de percentagem ocorria a separação destas cadeias para cada exão em estudo. A temperatura a que se realizou a electroforese, 60°C, foi determinada empiricamente de modo a exceder a temperatura de desnaturação de um fragmento de DNA rico em adeninas e timinas, na ausência de agentes desnaturantes.

Neste projecto, a técnica de DGGE foi utilizada como *pré-screening* de mutações pontuais para alguns exões do gene *TCF7L2*, pois após optimização é mais fácil e rentável determinar em qual das amostras de DNA se encontra a alteração e, posteriormente, sequenciar apenas essas para confirmar e caracterizar a mutação. A descrição de todas as soluções utilizadas para esta técnica encontra-se no anexo 6.

2.5.1. Amplificação de alguns exões do gene *TCF7L2* por reacção de PCR

A amplificação dos exões 4, 7, 8, 9, 13 e 14 do gene *TCF7L2* para posterior electroforese nos géis de DGGE foi efectuada com o descrito no ponto 2.3.2. O programa utilizado neste PCR encontra-se descrito no anexo 4 (ver tabela 4) e caracteriza-se por ter uma rampa de diminuição gradual da temperatura após a extensão final (de 10 em 10°C), para maximizar a formação dos heteroduplex entre as cadeias.

Após a reacção de PCR, de modo a verificar a sua eficiência, as amostras foram controladas em gel de agarose 2% (p/v), de acordo com o ponto 2.3.2.1..

2.5.2. Preparação dos géis de poliácridamida

Inicialmente limpou-se todo o material utilizado (vidros, espaçadores e pentes) com etanol a 70% (v/v), colocaram-se os espaçadores entre os vidros, ajustaram-se e colocaram-se as molas para segurar os vidros. O molde foi colocado numa base apropriada para esta técnica e colocaram-se os pentes.

Em seguida, aplicou-se no molde, com o auxílio de uma seringa, 2ml de solução de acrilamida a 80% (v/v) desnaturante, com 50µl de APS 10% (p/v) – agente polimerizante e 5µl de TEMED (*Merk*) – catalisador da polimerização (*gel sealing*), com a finalidade de evitar fugas do gel desnaturante que se adiciona posteriormente.

Os gradientes foram formados, utilizando volumes variáveis de duas soluções de poliácridamida: uma solução 0% (v/v) em agentes desnaturantes e outra solução 80% (v/v) em agentes desnaturantes, sendo o volume final de cada gradiente de 10ml.

Após a polimerização do *gel sealing*, adicionou-se 152 µl de APS 10% (p/v) e 8,9 µl de TEMED (*Merk*) a cada uma das soluções com as diferentes concentrações de agente desnaturante, agitou-se rapidamente por inversão e verteu-se cada solução num dos vasos comunicantes do “misturador de gradientes”. Colocou-se, especificamente, a solução com maior concentração de agente desnaturante no vaso por onde é efectuada a saída do gel para o molde, de modo a formar um gradiente de concentração crescente, o qual é obtido por acção da gravidade, sob agitação. Após a deposição do gel, deixou-se polimerizar durante, aproximadamente, 40 minutos.

Após polimerização, retiraram-se os pentes cuidadosamente e lavaram-se os poços com uma solução tamponada TAE1 x, para retirar vestígios de solução de acrilamida que poderiam prejudicar a deposição das amostras nos poços.

2.5.3. Preparação das amostras para aplicar no gel de DGGE

A cada amostra, adicionou-se 3,7µl de solução de deposição, a qual confere densidade às amostras e permite controlar a migração das mesmas ao longo do gel durante a electroforese.

Para além das amostras a analisar, depositou-se simultaneamente um controlo positivo (amostra heterozigótica para uma mutação conhecida) em cada gel, de modo a garantir que o gradiente foi formado correctamente e para servir de comparação com as demais amostras, de modo a identificar possíveis mutações nas amostras analisadas.

Por fim, preencheu-se, cuidadosamente, o restante volume dos poços com solução tamponada TAE 1x, de modo a não haver passagem das amostras para os poços adjacentes.

2.5.4. Electroforese dos géis de DGGE

A electroforese realizou-se numa tina vertical, contendo TAE 1x, sob agitação, com uma voltagem de 180V durante 4 horas a 60°C (à tina está acoplado um sistema de recirculação de água constituído por um banho com temperatura controlada).

2.5.5. Análise dos géis de DGGE

Após a electroforese, retirou-se cuidadosamente as molas e colocou-se o gel (sob um dos vidros) a corar num recipiente com solução de brometo de etídeo (10mg/ml, *Invitrogen*) durante, aproximadamente 10 minutos. Em seguida, o gel foi lavado com água para tirar o excesso de brometo de etídeo e soltar o gel do vidro. Por fim, colocou-se o gel num transiluminador, que contém uma câmara fotográfica e um sistema informático acoplado (*BioDocAnalyse, Biometra*), no qual se observou os padrões das amostras de DNA quando expostas a radiação UV e por comparação com o controlo positivo, determinou-se quais as amostras que poderiam ter possivelmente uma alteração.

As amostras, em que se detectaram alterações nos géis de DGGE, foram analisadas por sequenciação automática, tendo sido novamente amplificadas por PCR, com as condições descritas no ponto 2.3.2.. (ver anexo 3 e anexo 4 – tabela 3), e sequenciadas de acordo com o ponto 2.4., para confirmação do resultado e caracterização da mutação detectada.

2.6. Análise da mutação *frameshift* no exão 17 no gene *TCF7L2* utilizando o software *GeneScan*

2.6.1. Amplificação do fragmento 17A por PCR

A amplificação do fragmento 17A do gene *TCF7L2* foi efectuada de acordo com o descrito no ponto 2.3.2., tendo sido utilizado o programa que se encontra descrito no anexo 4 (ver tabela 2). Após a reacção de PCR, foi efectuada o controlo das amostras em gel de agarose 2% (p/v), de acordo com o ponto 2.3.2.1, de modo a verificar a sua eficiência.

2.6.2. Preparação das amostras para a electroforese capilar, utilizando o software GeneScan

Para tal, preparou-se, por amostra, uma mistura reaccional com 14,5µl de formamida *Hi-Di* (*Applied Biosystems*) (agente desnaturante) e 0,5µl de marcador de peso molecular de *GeneScan™ TAMRA 500™ Size Standard* (*Applied Biosystems*), que permite identificar o peso molecular da amostra, tendo como base pesos moleculares conhecidos. Distribui-se 15µl por cada tubo e adicionou-se 0,5µl de uma diluição do produto de PCR correspondente, a qual foi efectuada como base na intensidade das bandas obtidas na electroforese em gel de agarose 2% (p/v). Para esta análise, foi também incluído um controlo positivo (amostra com a mutação *frameshift* no exão 17) e um controlo normal (amostra sem a mutação *frameshift*) para posterior comparação do padrão com os resultados das amostras em estudo. As amostras foram desnaturadas num termociclador *Biometra* (*Alfagene*), durante cerca de 5 minutos a 95°C, e colocadas de imediato em gelo.

Por último, realizou-se a electroforese capilar no sequenciador automático *ABI Prism™ 310 Genetic Analyzer* (*Applied Biosystems*), a uma temperatura de 60°C com uma voltagem de 15KV e durante 21 minutos para cada amostra.

2.6.3. Análise de resultados

Os resultados foram obtidos utilizando o software *GeneScan*, versão 3.1.2 (*Applied Biosystems*), sob a forma de electroforetograma. Este software fornece, entre outros parâmetros, os valores correspondentes ao tamanho molecular dos picos do fragmento analisado, sendo assim possível inferir o resultado das amostras em estudo por comparação com os controlos (o controlo normal apresenta um pico aos 175pb, enquanto o controlo positivo apresenta um pico com peso molecular inferior).

2.7. Análise *in silico* das mutações somáticas identificadas

De modo a avaliar a eventual patogenicidade das mutações detectadas, realizou-se uma análise *in silico*, recorrendo à utilização de dois softwares: *Variant Effect Predictor* (VEP) (McLaren *et al.*, 2010) e *Mutation Taster* (Schwarz *et al.*, 2010). O software VEP permite obter a previsão dos efeitos das variantes introduzidas em transcritos, sequências de aminoácidos e regiões reguladoras. Este software também fornece a informação do *Sorting Intolerant From Tolerant* (SIFT) (Ng e Henikoff, 2003), o qual prevê se uma substituição de aminoácido irá afectar a função da proteína, baseado na homologia da sequência e propriedades físicas dos aminoácidos; e do *Polyphen* (Adzhubei *et al.*, 2010), o qual prevê o possível impacto de uma substituição de um aminoácido na estrutura e função da proteína, pela análise comparativa e física. O software *Mutation Taster* avalia o potencial da alteração mutacional causar doença, com base na conservação evolutiva, alteração de locais de *splicing* ou a factores de transcrição, perdas de funções da proteína e alterações do local de modificação de histonas.

2.8. Análise de perda de heterozigotia

A análise de perda de heterozigotia (LOH) foi efectuada em 21 amostras de DNA tumoral e respectiva mucosa normal.

Os marcadores de microssatélites utilizados foram seleccionados de acordo com os seguintes critérios: a sua localização na região de interesse, a percentagem de heterozigotia (de preferência superior a 70%) e o tamanho do produto amplificado (não deve ser superior a 200pb, uma vez que a maioria das amostras de DNA foram extraídas a partir tecido incluído em parafina, encontrando-se bastante fragmentado).

Neste estudo, utilizaram-se três marcadores de microssatélites: D10S1269, D10S168 e D10S562, sendo que o marcador D10S168 localiza-se no intrão 4 do gene *TCF7L2* e é flanqueado a montante e a jusante pelos demais marcadores. A caracterização dos marcadores de microssatélites encontra-se descrita no anexo 7.

2.8.1. Amplificação dos marcadores de microssatélites por PCR para análise de LOH

A amplificação dos marcadores de microssatélites foi efectuada pela técnica de PCR para amostras de DNA tumoral (a partir de tecido incluído em parafina ou congelado) e respectivo DNA normal (a partir de tecido incluído em parafina, congelado ou sangue periférico) em simultâneo (ver anexo 7 - Tabela 3 e 4). O controlo da eficiência da reacção de PCR foi efectuada por electroforese em gel de agarose 2% (p/v), tal como descrito no ponto 2.3.2.1.

2.8.2. Preparação das amostras para a electroforese capilar, utilizando o software GeneScan

A análise de LOH efectuou-se por electroforese capilar no sequenciador automático *ABI Prism™ 310 Genetic Analyzer (Applied Biosystems)*, tendo-se preparado as amostras de acordo com o descrito no ponto 2.6.2.

2.8.3. Análise de resultados

Os resultados foram obtidos utilizando o *software GeneScan*, versão 3.1.2 (*Applied Biosystems*), sob a forma de electroforectograma. Este *software* fornece informação relativa ao tamanho molecular (em pb) e às áreas dos picos das amostras em estudo, os quais correspondem aos alelos das amostras de DNA tumoral e normal. Desta forma, o cálculo da perda de heterozigotia foi efectuada com base na fórmula 2.3, utilizando os valores correspondentes às áreas dos picos respeitantes aos alelos da amostra de DNA de mucosa normal e tumoral. Segundo esta fórmula, todos os valores inferiores a 0,67 (indicativo de perda do alelo de menor peso molecular) ou

superiores a 1,5 (indicativo de perda do alelo de maior peso molecular) são indicativos de LOH para um determinado marcador de microssatélite.

$$LOH = \frac{AN2/AN1}{AT2/AT1}$$

Fórmula 2.3 - Expressão matemática que permite calcular a perda de heterozigotia.

Na fórmula apresentada:

AN1 – área do pico respeitante ao alelo de menor peso molecular da amostra de DNA normal;

AN2 – área do pico respeitante ao alelo de maior peso molecular da amostra de DNA normal;

AT1 – área do pico respeitante ao alelo de menor peso molecular da amostra de DNA tumoral;

AT2 – área do pico respeitante ao alelo de maior peso molecular da amostra de DNA tumoral.

Nos casos em que o indivíduo era homozigótico para o marcador de microssatélites ou em que o DNA tumoral apresentava instabilidade de microssatélites não se pôde avaliar a perda de heterozigotia.

2.9. Análise de metilação e *copy number* pela técnica de MS-MLPA

A técnica de MS-MLPA (*Methylation-Specific Multiplex Ligation-dependent Probe Amplification*) (Nygren, 2005) permite analisar as alterações no número de cópias (*copy number*) e o padrão de metilação de amostras de DNA numa única reacção de PCR em *multiplex*, utilizando um único par de *primers*, o que torna o método bastante robusto. Cada reacção de MS-MLPA resulta num conjunto único de fragmentos de PCR entre 65-500 nt, os quais são separados por electroforese capilar.

Esta técnica baseia-se na hibridação de sondas com o DNA alvo, sendo que algumas dessas contêm um local de reconhecimento (CpG) para a enzima de restrição sensível à metilação *HhaI*, o que torna possível a discriminação entre as sequências metiladas e não metiladas. Após a hibridação das sondas (constituídas por dois oligonucleótidos adjacentes) com o DNA, a reacção é dividida em duas partes iguais, sendo que numa parte ocorre apenas a ligação destas (análise de *copy-number*) enquanto na outra é efectuada a ligação das sondas e também a digestão com esta enzima (análise de metilação) (Figura 2.1.).

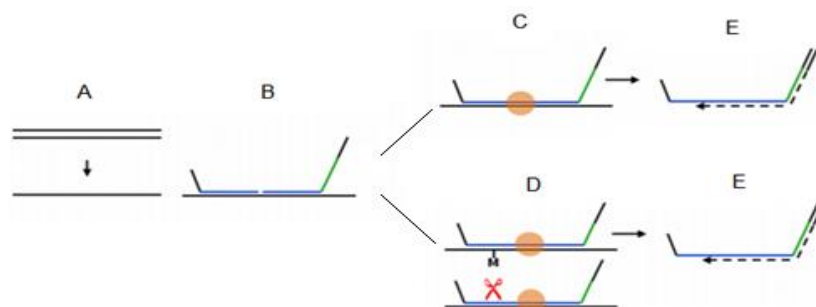


Figura 2.1. Esquema representativo de uma reacção de MS-MLPA (adaptado do protocolo *MS-MLPA® Two-Tube Protocol (MRC-Holland)*). Legenda: A - Desnaturação da cadeia dupla de DNA; B - Hibridação das sondas; C - Ligação; D - ligação e Digestão; E - Amplificação das sondas por PCR.

Quando a sequência está metilada, o grupo metil impede a digestão da sonda pela endonuclease e esta é amplificada por PCR, o mesmo não acontecendo se o DNA não se encontrar metilado, uma vez que a sonda é digerida e não ocorre amplificação deste fragmento. Consequentemente, após a electroforese capilar só os fragmentos amplificados serão visualizados no electroforetograma, sob a forma de um pico.

Enquanto na análise de metilação, a amostra é comparada entre si (parte da reacção sujeita a digestão vs não digerida), na análise de *copy-number* é necessário incluir amostras normais (de cada um dos tipos de tecido: incluído em parafina, congelado ou sangue, consoante o tipo de tecido que se está a analisar) que funcionam como referência para a normalização dos resultados e permitem detectar alterações no número de cópias das amostras em estudo.

Para a análise de alterações de *copy-number* e de metilação no gene *TCF7L2* foi utilizada uma *mix* de sondas sintéticas específicas para a região *upstream* do gene e exão 1 e para algumas regiões com elevado conteúdo em GC potencialmente relevantes (4 localizam-se na potencial região promotora, 2 nos exões – exões 1 e 14 e 3 em intrões (intrões 2, 5 e 16)). Estas sondas tinham sido previamente desenhadas pelo grupo de Gastrenterologia da Unidade de Investigação de Patobiologia Molecular (UIPM). A esta *mix* foi adicionada uma *mix* comercial de sondas referência localizadas noutras regiões cromossómicas essenciais para a análise e normalização dos resultados (P200-*MRC-Holland*).

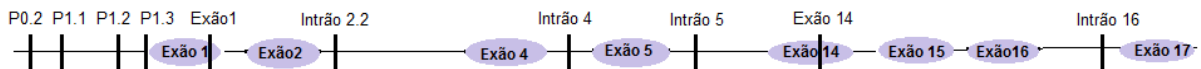


Figura 2.2. Esquema da localização das sondas MS-MLPA no gene *TCF7L2*. As barras pretas representam a localização das sondas.

2.9.1. Preparação das amostras

Para esta análise, foi efectuada nova quantificação dos DNA extraídos a partir de tecido incluído em parafina e dos DNA extraídos a partir de tecido congelado, como descrito nos pontos 2.1.1.1. e 2.1.2.1., respectivamente, para confirmação da sua concentração, uma vez que algumas das amostras são bastante antigas e é necessário uma quantidade exacta de cada uma, a qual depende do tipo de conservação do material biológico.

Em seguida, as amostras de DNA foram diluídas para um volume final de 6µl em TE pH8 (*Citogene*, Portugal), sendo que das amostras extraídas a partir de tecido incluído em parafina foi utilizada uma quantidade de, aproximadamente, 450ng e nas amostras de tecido congelado e sangue utilizaram-se, aproximadamente, 270ng.

2.9.2. Reacção de MS-MLPA

O protocolo de MS-MLPA é constituído por três etapas: desnaturação do DNA e hibridação das sondas; reacção de ligação e digestão e reacção de PCR, tendo sido utilizado neste estudo o MS-

MLPA® Two-Tube Protocol (MRC-Holland), segundo as indicações do fabricante. Todos os reagentes utilizados neste protocolo pertencem ao *SALSA MLPA EK1 reagent kit (MRC-Holland)* à excepção da enzima de restrição *HhaI (Promega)*.

Desnaturação do DNA e hibridação das sondas

As amostras, diluídas anteriormente, foram desnaturadas num termociclador *Veriti (Applied Biosystems)* durante 10 minutos a 98°C. Em seguida, e após a temperatura diminuir para os 25°C, retiram-se os tubos do termociclador e transferiu-se 5µl para novos tubos tipo *ependorf* de 0,2ml. À temperatura ambiente, adicionou-se 3µl de uma mistura reaccional contendo 1,5µl de *MLPA buffer*, 1µl de *Probemix P200 (MRC-Holland)* e 0,5µl de *Synthetic probemix* (contendo as sondas específicas para o gene *TCF7L2*) a cada tubo. Posteriormente, incubaram-se os tubos durante 1 minuto a 95°C seguindo-se de pelo menos 16 horas (máximo 20 horas) a 60°C, para a hibridação das sondas com o DNA-alvo.

Reacção de ligação e digestão

À temperatura ambiente, adicionou-se 10µl de uma mistura reaccional contendo 3µl de *Ligase-buffer A* com 10µl de água bidestilada, em cada tubo e ressuspendeu-se. Em seguida, transferiu-se 10µl para um segundo conjunto de tubos e incubaram-se ambos os conjuntos de tubos a 49°C durante, pelo menos, um minuto no termociclador. Ao fim desse período, à mesma temperatura, adicionou-se no primeiro conjunto de tubos (amostras que serão submetidas apenas à ligação - análise de *copy-number*) 10µl de uma mistura reaccional composta por 8,25µl de água bidestilada, 1,5µl de *Ligase-65 buffer B* e 0,25µl de *Ligase-65 enzyme* e no segundo conjunto de tubos (amostras que serão submetidas a ligação e digestão – análise de metilação) adicionou-se 10µl de uma mistura reaccional contendo 7,75µl de água bidestilada, 1,5µl de *Ligase-65 buffer B*, 0,25µl de *Ligase 65 enzyme* e 0,5µl da enzima de restrição *HhaI*. Posteriormente incubaram-se os tubos durante 30 minutos a 49°C, seguido de 5 minutos a 98°C para inactivação das enzimas.

Reacção de PCR

Retiraram-se os tubos do termociclador e transferiu-se 5µl da reacção para dois novos conjuntos de tubos tipo *ependorf* de 0,2ml. Nos novos tubos, adicionou-se 15µl de uma mistura reaccional composta por 13µl de água bidestilada e 2µl de *SALSA PCR buffer* e colocaram-se os tubos em gelo. Por fim, adicionou-se a cada tubo 5µl de uma mistura reaccional contendo 2,75µl de água bidestilada, 1µl de *SALSA PCR-primers*, 1µl de *SALSA Enzyme dilution buffer* e 0,25µl de *SALSA Polymerase*, colocaram-se os tubos no termociclador pré-aquecido a 72°C e iniciou-se de imediato a reacção de PCR, com o programa descrito na tabela 2.2.

Tabela 2.2 Programa utilizado para a amplificação das sondas de MS-MLPA por reacção de PCR.

Temperatura	Tempo	Ciclos
95°C	30 segundos	35
60°C	30 segundos	
72°C	60 segundos	
72°C	20 minutos	1
15°C	∞	-

2.9.3. Preparação das amostras para análise de MS-MLPA pelo software *GeneMapper*

Para a preparação das amostras para a electroforese capilar, a eficiência da reacção de PCR foi verificada por electroforese em gel de agarose 2% (p/v) (ver ponto 2.3.2.1.) e observou-se a intensidade das bandas obtidas de modo a determinar a diluição do produto de PCR necessária para a análise pelo software *GeneMapper* versão 4.1 (*Applied Biosystems*).

Em seguida, preparou-se uma mistura reaccional com 9µl de formamida *Hi-Di* (*Applied Biosystems*) e 0,2µl de marcador de peso molecular *GeneScan™ LIZ 500™ Size Standard* (*Applied Biosystems*), que permite identificar o peso molecular da amostra, tendo como base pesos moleculares conhecidos. Distribuiu-se 9,2µl por cada poço de uma placa de 96 poços e adicionou-se, de um modo geral, 1µl do produto de PCR correspondente (algumas das amostras tiveram de ser diluídas de modo a não se atingir o valor máximo de detecção de fluorescência estipulado para o sequenciador automático utilizado). Para além das amostras em estudo, prepararam-se, do mesmo modo, 3 amostras normais de referência. Estas são derivadas de indivíduos saudáveis (sangue) ou de mucosas normais de indivíduos afectados com CCR, mas com um *copy number* e perfil de metilação normais para as regiões de interesse. Em seguida, desnaturou-se a placa num termociclador *Biometra* durante 5 minutos a 95°C, colocou-se a placa em gelo durante 2 minutos e centrifugou-se a 1200 rpm durante 2 minutos. Por fim, colocou-se a placa no sequenciador automático *ABI Prism™ 3130 Genetic Analyzer* (*Applied Biosystems*), para realizar a electroforese capilar com uma voltagem de 13KV e durante 24 minutos para cada amostra, para a separação dos fragmentos amplificados por PCR de acordo com o seu peso molecular.

2.9.4. Análise de Resultados

Os resultados foram obtidos sob a forma de electroforetogramas e a análise dos resultados foi realizado recorrendo ao software *Coffalyser.Net* (*MRC-Holland*, Holanda) (Coffa e van den Berg, 2011), o qual efectua a normalização entre as amostras em estudo e as referência mas também, em cada amostra, entre as sondas específicas do gene em estudo e as sondas referência. Os valores obtidos para o *copy number* (incluindo o desvio padrão) foram interpretados como normais se se situavam entre 0,71-1,29, como perda do número de cópias se inferiores a 0,71 e como ganho de cópia do material genético se superiores a 1,29.

2.10. Análise dos níveis de expressão por PCR quantitativo em tempo real (qPCR)

O PCR quantitativo (qPCR) permite quantificar os níveis de expressão de um determinado gene. Esta técnica utiliza moléculas *reporter* fluorescentes para monitorizar a quantidade de produto amplificado em cada ciclo da reacção de PCR (Holland *et. al.*, 1991), podendo tratar-se de sondas

específicas ou corantes intercalantes. Neste projecto foram utilizados os dois métodos de detecção dos produtos de PCR (*TaqMan*[®] e *SYBR Green*[®]).

O método *TaqMan*[®] utiliza uma sonda constituída por uma molécula *reporter* fluorescente na extremidade 5' e uma molécula *quencher* na extremidade 3'. Enquanto a sonda está intacta, a proximidade do *quencher* diminui a fluorescência emitida pelo *reporter*, quando a sequência-alvo está presente na amostra em estudo, a sonda liga-se e é clivada pela polimerase à medida que ocorre a extensão do *primer*. A clivagem da sonda provoca a separação entre o *reporter* e o *quencher*, aumentando o sinal de fluorescência do *reporter* a cada nova cadeia sintetizada (Figura 2.3.-1).

O fluorocromo *SYBR Green*[®], por sua vez, intercala-se na cadeia dupla de DNA e aquando da amplificação da região de interesse, liga-se às novas cópias de DNA em cadeia dupla, resultando num aumento exponencial na intensidade da fluorescência observada (Figura 2.3.-2).

Em ambos os métodos, o aumento da intensidade da fluorescência é directamente proporcional à quantidade de produto de PCR produzido.

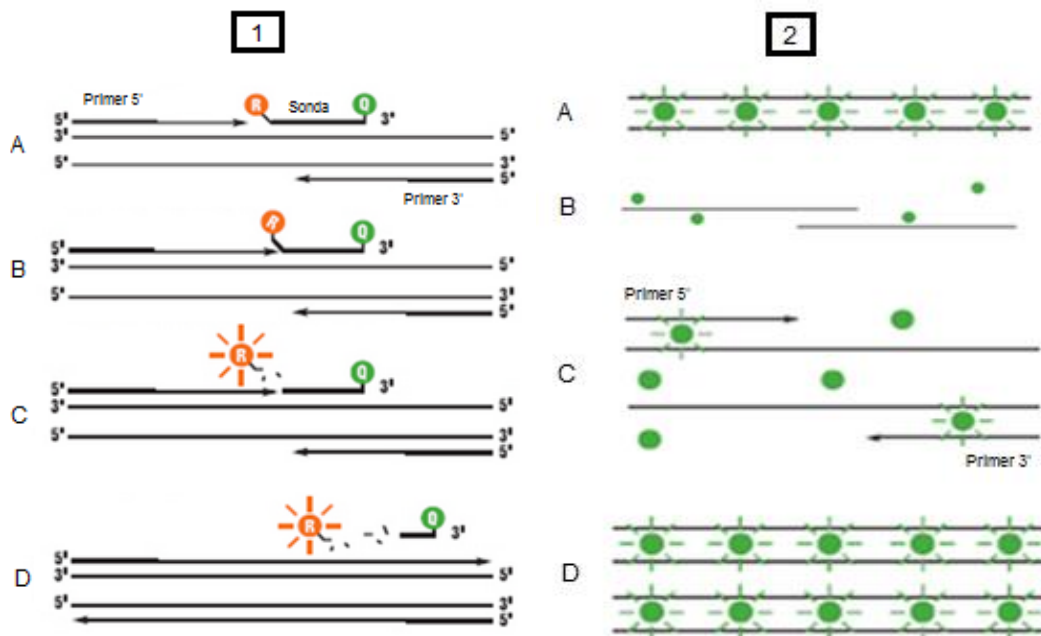


Figura 2.3- Sistema representativo do método qPCR. 1 – Sistema representativo do método *Taqman*: 1A – Sonda *TaqMan* com as moléculas *reporter* e *quencher* nas extremidades 5' e 3', respectivamente; 1B – A sonda está intacta e o *reporter* não emite fluorescência; 1C – Durante cada ciclo, a DNA polimerase cliva o *reporter* da sonda; 1D – Separação do *quencher* e emissão de fluorescência pelo *reporter*. 2 - Sistema representativo do método *SYBR Green*: 2A – *SYBR Green* ligado à dupla cadeia de DNA, emitindo fluorescência; 2B – Quando o DNA é desnaturado, o *SYBR Green* liberta-se da molécula de DNA, diminuindo a fluorescência; 2C – Formação de produtos de PCR durante a extensão; 2D – No final da polimerização, o *SYBR Green* liga-se à nova cadeia dupla de DNA, originando o aumento de fluorescência (adaptado de <http://www.biosyn.com/tew/taqman-vs-sybr-green-chemistries.aspx>).

Durante os primeiros ciclos de PCR, o sinal de fluorescência emitido (*TaqMan*[®] e *SYBR Green*[®]) é normalmente muito fraco para se registar acima da linha de base (normalmente definida entre o 3^o e o 15^o ciclos), sendo que o ciclo em que este sinal transpõe esta linha se designa por ciclo de *threshold* (Ct).

Esta técnica permite efectuar dois tipos de quantificação: uma quantificação absoluta ou uma quantificação relativa.

Na quantificação absoluta (método da curva-padrão), é seleccionada uma amostra que expressa o gene de interesse, para a construção de uma curva-padrão, ou seja, uma relação linear entre o Ct e o logaritmo da concentração da amostra padrão. Esta curva é obtida através de diluições seriadas desta amostra, a qual serve como referência para a extrapolação da concentração da amostra em estudo. Nesta quantificação, também se amplifica em paralelo um controlo endógeno para cada amostra.

Na quantificação relativa, é utilizada uma amostra referência, a qual funciona como calibrador, ou seja, o valor de expressão do gene de interesse nas amostras em estudo é obtido em relação ao calibrador. Nesta quantificação, é necessário também amplificar em paralelo um controlo endógeno para cada amostra, ou seja, um gene *housekeeping* (expresso de uma forma ubíqua em todas as células). Neste trabalho apenas se efectuou quantificação relativa.

2.10.1. SYBR Green®

O método do SYBR Green® foi utilizado para confirmar as alterações de *copy number* detectadas pela técnica de MS-MLPA nos exões 1, 4 e 14 e no intrão 4 no gene *TCF7L2*, tendo-se utilizado o gene *SDHA housekeeping*. Para além das amostras de DNA em estudo, amplificou-se ainda três amostras que não apresentavam as alterações de *copy number* que se pretendiam analisar, de modo a servirem de controlo normal e um controlo negativo (sem DNA), para cada gene em análise. Para assegurar um maior rigor na quantificação realizaram-se triplicados de cada amostra em estudo.

As reacções foram realizadas seguindo o protocolo – *LightCycler® 480 SYBR Green I Master* (Roche, Applied Science) de acordo com o descrito pelo fabricante, mas com algumas alterações. Assim, num placa de 96 poços (*LightCycler® 480 Multiwell Plates 96*, Roche Applied Science) em gelo, distribuiu-se para cada gene, 13µl de uma mistura reaccional composta por 4µl de água bidestilada, 0,75µl de *primer forward* e *reverse* (5pmol/µl) e 7,5µl de *LightCycler® 480 SYBR Green I Master*, 2x concentrada (com fluoróforo SYBR Green I, dNTPs, *FastStart Taq DNA Polimerase*, tampão da reacção e MgCl₂) (Roche, Applied Science) à qual se adicionou 2µl de DNA (10ng). Centrifugou-se a 1200rpm durante 2 minutos e colocou-se a placa no *LightCycler® 480 Instrument II* (Roche Applied Science) iniciando-se o programa descrito no anexo 8 (ver tabela 1) e a análise dos resultados foi realizada utilizando o *software LightCycler® 480 SW 1.5* (Roche Applied Science).

As sequências dos *primers* utilizados encontram-se descritas no anexo 8 (ver tabela 2).

Calculou-se a eficiência das reacções, uma vez que não era igual entre si, com base na fórmula 2.3. O declive foi obtido através da elaboração de uma recta de calibração, ou seja, uma relação linear entre o Ct de uma amostra e o logaritmo da concentração de cDNA dessa mesma amostra. Esta curva foi construída a partir de uma diluição seriada da amostra, a qual atribuiu diferentes valores de Ct, correspondentes a cada diluição. Posteriormente, com a eficiência de cada

par de *primers* foi possível quantificar as amostras a estudo utilizando a fórmula 2.4, na qual $E_{\text{gene alvo}}$ é a eficiência do gene alvo, neste caso, o exão que se pretende estudar; $E_{\text{gene referência}}$ corresponde à eficiência do gene SDHA e ΔC_t equivale à subtração da média dos ΔC_t da amostra controlo com a média dos ΔC_t da amostra em análise, para os exões correspondentes.

$$E = 10^{\left(-\frac{1}{\text{declive}}\right)}$$

Fórmula 2.3. Expressão matemática para calcular a eficiência da reacção.

$$\text{ratio} = \frac{(\text{Egene alvo})^{\Delta C_t(\text{média do controlo}-\text{média da amostra})}}{(\text{Egene referência})^{\Delta C_t(\text{média do controlo}-\text{média da amostra})}}$$

Fórmula 2.4. Expressão matemática para determinar o *copy number*.

2.10.2. TaqMan®

O método do *TaqMan*® foi utilizado para quantificar os níveis de expressão de três amostras de cDNA (síntese descrita no ponto 2.10.2.1.) que apresentaram alterações de *copy number* (ganho do exão 1 e deleção do intrão 4) no gene *TCF7L2* pela técnica de MS-MLPA. Para tal, realizaram-se dois ensaios, sendo que num deles se amplificou a região correspondente ao exão 3 e 4 e no outro se amplificou a região entre os exões 10 e 11, tendo sido utilizado o gene β -actina como *housekeeping*. Este último ensaio, permite realizar uma quantificação total da expressão. Para além das amostras de cDNA em estudo, amplificaram-se ainda duas amostras normais para o gene *TCF7L2*, para servirem de controlo e um controlo negativo (sem cDNA) em cada um dos ensaios, para os quais foram realizados triplicados para cada amostra.

A sonda para o gene *TCF7L2* contém na sua extremidade 5' o fluoróforo FAM e a sonda do gene β -actina contém o fluoróforo VIC.

As reacções foram realizadas seguindo o protocolo – *TaqMan*® *Universal PCR Master Mix* (*Applied Biosystems, Roche*), de acordo com o descrito pelo fabricante, mas com algumas alterações. Assim, numa placa de 96 poços (*ABI PRISM™ Optical 96-Well Reaction Plate*), em gelo, distribuiu-se para cada ensaio 18,5µl de uma mistura reaccional contendo: 7,2µl de água bidestilada, 1µl (*Taqman Gene Expression Assay 20x, Applied Biosystems*) e 10µl de *TaqMan*® *Universal PCR Master Mix, 2x* concentrada (com sonda *TaqMan*® específica para cada um dos ensaios, dNTPs, *AmpliQ Gold DNA Polimerase* e tampão da reacção) (*Applied Biosystems, Roche*) à qual se adicionou 1,5µl de cDNA (≈200ng). Centrifugou-se a 1200rpm durante 2 minutos e colocou-se a placa no *7900HT Sequence Detection System* (*Applied Biosystems*) iniciando-se o programa descrito no anexo 8 (ver tabela 3), sendo a análise dos resultados efectuada pelo *software SDS 2.4* (*Applied Biosystems*).

Ao contrário do ponto anterior, a eficiência das reacções não foi calculada, uma vez que apenas foram utilizados reagentes comerciais, assim, a eficiência é elevada e uniforme entre o gene

em estudo e o gene *housekeeping*. Recorreu-se ao 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 , quer da amostra em análise quer da amostra referência, foram normalizados em relação ao gene β -actina (*housekeeping*). Na fórmula 2.5., 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*).

$$ratio = 2^{-(\Delta C_t amostra - \Delta C_t referência)}$$

Fórmula 2.5. Expressão matemática para a quantificação dos níveis de expressão do gene em estudo.

2.10.2.1. Síntese de cDNA por transcrição reversa seguida de PCR (RT-PCR)

O RNA extraído a partir de linfócitos (ver ponto 2.2.1.4.) foi utilizado para a síntese de DNA complementar de cadeia simples (cDNA), recorrendo à técnica RT-PCR. Esta técnica baseia-se numa reacção de transcrição reversa, na qual se utiliza *primers* aleatórios (*random primers*) que hibridam com os diferentes mRNA e a sua extensão é realizada pela enzima transcriptase reversa para a síntese de cDNA de cadeia simples e, posteriormente converte o híbrido RNA-cDNA em moléculas de cadeia dupla.

Para a síntese de cDNA, primeiramente, adicionou-se um volume variável de RNA de cada amostra (correspondente a uma concentração final de 200ng de RNA), 0,5µl de hexâmeros de *primers* (3µg/µl, Roche) – mistura de hexanucleótidos com terminações aleatórias 5' e 3' com grupo hidroxilo, e água tratada com dietilpirocarbonato (DEPC, Merck), que inactiva as enzimas RNase, para um volume final de 7,75µl. Posteriormente incubou-se a 70°C durante 10 minutos no termociclador *Veriti* (Applied Biosystems) e, depois colocou-se em gelo. De seguida, adicionou-se 12,25µl de uma mistura reaccional com 0,3µl de água tratada com DEPC, 4µl de tampão *5x First-Strand Buffer* (Invitrogen), 4µl de dNTPs (25mM, *Illustra™*, GE Healthcare), 2µl de ditiotretol (DTT) (0,1M, *Invitrogen*) – estabilizador enzimático e proteico; 0,75µl de *RNaseOUT™ Recombinant Ribonuclease Inhibitor* (40U/µl, *Invitrogen*) – inibe as ribonucleases; 1µl de enzima transcriptase reversa *Superscript II RT* (200U/µl, *Invitrogen*) e 0,5µl de enzima *single strand binding protein* (SSB) (4,47µg/ml, *Promega*) - liga-se às cadeias simples de RNA, evitando a sua hibridação prematura. Por fim, realizou-se a reacção de transcrição reversa num termociclador *Veriti* (Applied Biosystems). Os tubos foram novamente colocados no termociclador *Veriti* (Applied Biosystems). Os tubos com cDNA foram armazenados a -20°C. O programa utilizado para a síntese de cDNA encontra-se descrito na Tabela 2.3..

Tabela 2.3. Programa de RT-PCR utilizado na síntese de cDNA num termociclador *Veriti* (*Applied Biosystems*).

Etapa	Temperatura	Tempo
Desnaturação inicial	70°C	10 minutos
	15°C	∞
Reacção de RT-PCR	42°C	60 minutos
	70°C	15 minutos
	15°C	∞

Para além da utilização do cDNA para o ensaio de quantificação de expressão, realizou-se em uma das amostras a amplificação de cDNA por PCR dos exões 2 ao 7 do gene *TCF7L2* para verificar a presença de um transcrito aberrante. O controlo da eficiência do produto amplificado foi efectuado em gel de agarose 2% (p/v), como descrito no ponto 2.3.2.1.

2.11. Análise Estatística

Em relação à análise realizada em amostras de tumores do recto, recorreu-se ao teste de Fisher para correlacionar os resultados obtidos com a resposta à quimiorradioterapia pré-operatória e outras características clínicas dos doentes. O valor de p inferior a 0,05 foi considerado com significado estatístico.

3. Resultados

3.1. Análise de mutações somáticas no gene *TCF7L2* em CCR

Dos 10 exões a analisar foi otimizada a análise para 7 exões e respectivas regiões flanqueadoras do gene *TCF7L2* (1, 4, 13, 14, 15, 16, 17), por indivíduo e por exão, cujos resultados encontram-se no anexo 9. Foram detectadas 7 mutações somáticas no conjunto de amostras analisadas (7/68 tumores; \approx 10%), com uma frequência, cada uma, entre 2% a 17%. (Tabela 3.1.). Destas sete mutações, duas são mutações *frameshift* pela deleção de um nucleótido e as restantes são mutações de substituição. Não foram detectadas mutações nos exões 1, 4 e 14, como também nas respectivas regiões intrónicas flanqueadoras dos mesmos.

As mutações foram caracterizadas com base na sequência genómica e utilizando vários transcritos do gene *TCF7L2*, pois não foi possível seleccionar apenas um transcrito para caracterizar as mutações, uma vez que o gene *TCF7L2* tem diversos locais sujeitos a *splicing*, originando diferentes transcritos que expressam alternadamente os exões. Desta forma, não existe nenhum transcrito completo (com os 17 exões), descrito nas bases de dados (*Ensembl* - <http://www.ensembl.org/index.html>; NCBI - <http://www.ncbi.nlm.nih.gov/>), pelo que foram seleccionados três transcritos que continham a sequência de interesse e, preferencialmente, também os exões 4, 13, 14 e 17, por estarem associados a locais de *splicing* mais frequentes. Por este motivo, recorreu-se à localização das mutações a nível genómico. Para a análise das mutações nos exões 13 e 17 foi utilizado o transcrito ENST00000355995 (HGNC -TCF7L2-203 - *Ensembl*) (expressa todos os exões excepto os exões 15 e o 16), para a análise do exão 15 foi seleccionado o transcrito ENST00000543371 (HGNC-TCF7L2-208- *Ensembl*) (expressa todos os exões excepto os exões 13, 14 e o 16) e para a localização da região intrónica do exão 16 foi seleccionado o transcrito ENST00000352065 (HGNC-TCF7L2-201- *Ensembl*) (expressa todos os exões excepto os exões 4 e 15). Para além disso, não foi possível averiguar se todas as mutações detectadas já tinham sido descritas e as suas frequência na população. Contudo, sabe-se, que duas das mutações detectadas (mutações *frameshift*) estão descritas na literatura (Duval *et al.*, 1999; Duval *et al.*, 2000)

Das duas mutações *frameshift*, uma delas ocorre, em heterozigotia, na região flanqueadora a montante do exão 13 pela deleção de um nucleótido numa região repetitiva de 8 timinas (1/51 tumores; \approx 2%) (Figura 3.1-A) e a outra mutação ocorre no exão 17 pela deleção de um nucleótido numa região repetitiva de 9 adeninas (Figura 3.1-B). Duas amostras apresentavam a mutação em homozigotia (2/36 tumores; \approx 6%) e as restantes em heterozigotia (6/36 tumores; \approx 17%). Para além disso, no exão 15 foram detectadas duas mutações diferentes, uma substituição de uma guanina (G) por uma adenina (A) (1/20 tumores; \approx 5%) e de uma timina (T) por uma citosina (C) (1/20 tumores; \approx 5%) (Figura 3.1-C); no exão 17 foram detectadas, também, duas mutações, uma substituição de uma timina (T) por uma citosina (C) (1/44 tumores; \approx 2%) e de uma citosina (C) por uma timina (T) (2/39

tumores; $\approx 5\%$) e na região intrónica a jusante do exão 16 foi encontrada uma substituição de uma citosina (C) por uma timina (T) (1/55 tumores; $\approx 2\%$).

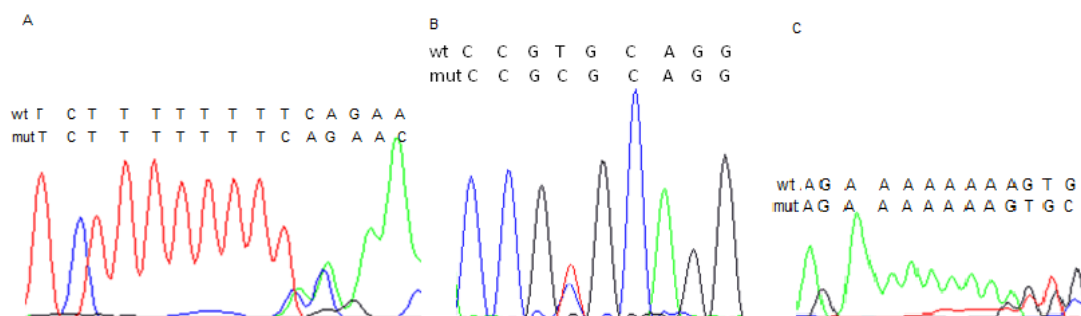


Figura 3.1. Electroforetograma parcial obtido no sequenciador *ABI PrismTM 3130 Genetic Analyzer*, representando as mutações nos exões 13 delT (A), 15 T>C (B) e 17 del A(C) (sequência *forward*).

Tabela 3.1 Identificação das mutações somáticas detectadas no gene *TCF7L2* e respectiva frequência.

Localização Exão/Intrão	Posição genómica	Heterozigotia/Homozigotia	Frequência das mutações nas amostras analisadas
Intrão 13	113,158,663 del T	Heterozigotia	2%
Exão 15	113,160,646 G>A	Heterozigotia	5%
	113,160,688 T>C	Heterozigotia	5%
Intrão 16	113,161,621 C>T	Heterozigotia	2%
	113,165,559 del A	Heterozigotia	6%
Exão 17	113,165,559 del A	Heterozigotia	17%
	113,165,759 T>C	Heterozigotia	2%
	113,165,957 C>T	Heterozigotia	5%

Os tumores analisados foram estratificados em estáveis (MSS) e instáveis (MSI) e observou-se que as mutações *frameshift* apenas ocorreram em tumores instáveis. A mutação *frameshift* do exão 17 foi detectada em 8 dos 12 tumores instáveis analisados ($\approx 67\%$), no entanto a frequência da mutação no exão 13 foi menor, tendo sido apenas detectada em 1 dos 12 tumores analisados ($\approx 8\%$). Nos tumores estáveis, foram apenas detectadas mutações do tipo substituição (5/56 tumores; $\approx 11\%$).

De forma a avaliar a eventual patogenicidade destas mutações, realizou-se uma análise *in silico*, utilizando-se o software *Variant Effect Predictor* (VEP) do *Ensembl* (Tabela 3.2), o qual fornece informação a nível da patogenicidade da mutação de acordo com os softwares *SIFT* e *Polyphen*. Para confirmar a patogenicidade recorreu-se ao software *Mutation Taster* (<http://www.mutationtaster.org/>), neste software foi apenas analisado um transcrito, o mesmo em que se baseou a caracterização da mutação.

O efeito da mutação varia de acordo com o transcrito produzido, sendo apresentadas na tabela 3.2 as várias consequências obtidas com essa alteração.

Pela análise *in silico*, foi possível verificar que a mutação *frameshift* do exão 13 parece afectar o *splicing*, pois a alteração ocorre na região de *splicing* a 3 bases do exão, produzindo um mRNA com um codão *stop* prematuro (proteína truncada) na maioria dos transcritos. A análise pelo *software Mutation Taster* revelou que esta alteração é patogénica por aumentar o local aceitador de *splicing* com um *score* maior do que o *score wild-type*. A substituição do nucleótido guanina pela adenina no exão 15 origina, maioritariamente, um transcrito que não altera a sequência aminoacídica, preservando-se a presença do aminoácido glicina, no entanto, em dois transcritos a mutação origina um aminoácido diferente (troca de glicina por ácido glutâmico). A análise desta mutação através do *SIFT* classificou-a como patogénica e através do *Polyphen* como provavelmente ou possivelmente patogénica dependendo do transcrito. A análise pelo *software Mutation Taster* confirmou a patogenicidade, devido ao aumento ligeiro do local dador de *splicing*, no entanto o *score* não é significativo. Assim, esta mutação deverá ser considerada sem efeito provável. A segunda mutação detectada neste exão altera a sequência de aminoácidos, havendo uma substituição do aminoácido fenilalanina pela serina (variante *missense*). A análise desta mutação através do *SIFT* classificou-a como tolerada, no entanto, o *software Polyphen* classificou-a como benigna ou possivelmente patogénica, dependendo do transcrito. Também se verifica, que esta alteração pode causar uma variante 3' UTR, pelo facto de o transcrito não codificar na totalidade este exão. A análise pelo *software Mutation Taster* considera esta mutação como patogénica, por levar à perda do local dador de *splicing*.

A substituição encontrada na região intrónica flanqueadora a jusante do exão 16 não causa qualquer consequência na proteína, sendo classificada como uma variante intrónica. Por outro lado, o resultado do *software Mutation Taster* revelou que esta alteração poderia ser patogénica, pois aumenta ligeiramente o local dador de *splicing*, contudo a probabilidade deste processo ocorrer é muito reduzida. Desta forma, esta alteração foi considerada sem efeito provável. Este facto é suportado pelo estudo realizado por Duval, o qual refere que esta mutação não foi detectada numa linha celular com níveis extremamente elevados de *splicing* alternativo (Duval *et al.*, 2000).

A deleção do nucleótido adenina na região repetitiva de 9(A) no exão 17, na maioria dos transcritos altera a grelha de leitura, produzindo uma proteína truncada. Em alguns casos, é uma variante 3' UTR, pelo facto de o transcrito não codificar este exão. As substituições detectadas neste exão na maioria dos transcritos não alteram a sequência de aminoácidos, no caso da substituição da timina pela citosina não se verifica alteração do aminoácido (isoleucina) e na substituição da citosina pela timina também não se verifica alteração do aminoácido (treonina). Os resultados também indicam que estas substituições correspondem a uma variante 3' UTR, pelo facto de alguns transcritos não codificarem na totalidade o exão 17. A análise pelo *software Mutation Taster* demonstra apenas que a deleção do nucleótido adenina, aumenta o local dador, aumentando a probabilidade de *splicing* nesse local (aumento do *score* significativo em relação à sequência *wild-type*) enquanto as substituições aumentam ligeiramente o local aceitador do *splicing*. Assim, as mutações de substituição foram consideradas sem efeito provável. Por outro lado, também se verificou que nos três tipos de mutações detectadas neste exão a sequência variante pode estar numa região reguladora de cromatina aberta sobre a qual não existe ainda informação.

Tabela 3.2 Resultados da análise *in silico* das mutações detectadas no gene *TCF7L2* em amostras de CCR.

Exão	Posição genómica	Transcrito <i>Ensembl</i>	HGNC – <i>TCF7L2</i>	Caracterização da mutação	VEP		<i>Mutation Taster</i>
					<i>SIFT</i>	<i>Polyphen</i>	
Exão 13	113,158,663 del T	ENST00000355995	203	c.1319-4delT	Variante de região de <i>splicing</i>		Patogénica (aumenta o local aceitador de <i>splicing</i> -3') wt: 0.50 / mu: 0.66
		ENST00000543371	208	c.1361 G>A (p.G454E)	Patogénica (0,01)	Variante <i>missense</i> Provavelmente patogénica (0,928)	Patogénica (aumenta ligeiramente o local dador de <i>splicing</i> -5') wt: 0.9333 / mu: 0.9642
Exão 15	113,160,646 G>A	ENST00000542695	207	c. 1142 G>A (p.G454E)	Patogénica (0,01)	Variante <i>missense</i> Possivelmente patogénica (0,718)	-
		ENST00000545257	209	c.1416 G>A (p.G472G)	Variante silenciosa		-
		ENST00000543371	208	c.1403 T>C (p.F468S)	Tolerada (0,19)	Variante <i>missense</i> Benigna (0,146)	Patogénica Perda do local dador
		ENST00000542695	207	c.1184 T>C (p.F395S)	Tolerada (0,2)	Variante <i>missense</i> Possivelmente patogénica (0,549)	-
		ENST00000369386	006		Variante 3' UTR		
Exão 16	113,161,621 C>T	ENST00000352065	201	c.1380+36	Variante intrónica		Patogénica (aumenta ligeiramente o local dador de <i>splicing</i> -5') wt: 0.7332 / mu: 0.7533

Tabela 3.2 (Continuação) Resultados da análise *in silico* das mutações detectadas no gene *TCF7L2* em amostras de CCR.

Exão	Posição genómica	Transcrito <i>Ensembl</i>	HGNC – <i>TCF7L2</i>	Caracterização da mutação	VEP		<i>Mutation Taster</i>
					<i>SIFT</i>	<i>Polyphen</i>	
Exão 17	113,165,559 del A	ENST00000355995	203	c.1447delA (p.K483*23)	Variante <i>frameshift</i>		Patogénica (aumenta o local dador de <i>splicing</i> -5') wt: 0.22 / mu: 0.99
		ENST00000470254	013		Variante 3' UTR		-
		ENSR00000362043			Variante de região reguladora		-
	113,165,759 T>C	ENST00000355995	203	c.1647 T>C (p.I537I)	Variante silenciosa		Patogénica (aumenta o local aceitador de <i>splicing</i> -3') wt: 0.31 / mu: 0.42
		ENST00000545257	209		Variante 3' UTR		-
		ENSR00000362044			Variante de região reguladora		-
	113,165,957 C>T	ENST00000355995	203	c.1845 C>T (p.T615T)	Variante silenciosa		Patogénica (aumenta o local aceitador de <i>splicing</i> -3') wt: 0.64 / mu: 0.71
		ENST00000545257	209		Variante 3' UTR		-
		ENSR00000362044			Variante de região reguladora		-

Legenda: wt – score da sequência *wild type*; mu – score da sequência mutada

3.1.1. Análise do *copy number* das amostras de CCR por MS-MLPA

Após a análise mutacional das amostras de CCR, efectuou-se a análise de *copy number* por MS-MLPA nas amostras em que foram detectadas mutações. Apesar das mutações terem sido detectadas em 7 amostras de tumores, só foi possível analisar o *copy number* em 6, pois não havia DNA *stock* suficiente para realizar esta técnica na amostra com a mutação 113,165,957 C>T. Contudo, esta mutação foi detectada em mais uma amostra (595), a qual foi possível analisar. Para esta análise é necessário utilizar DNA da mucosa tumoral e DNA de mucosa normal. Mas em dois casos não existia DNA da mucosa normal disponível, tendo sido utilizado DNA extraído a partir de sangue periférico.

Pela análise dos resultados de MS-MLPA pelo *software Coffalyser.Net* (MRC-Holland, Holanda) (resultados descritos na tabela 3.3) verifica-se que a maioria das amostras apresenta deleção no exão 14, por isso, comparou-se estes resultados com resultados de *copy number* para amostras que não apresentavam mutações (esta análise de *copy number* foi realizada num estudo anterior) (Tabela 3.4.). Verificou-se que os tumores colorectais com mutação parecem apresentar maior frequência da deleção no exão 14 do que os tumores colorectais sem presença de mutações no gene *TCF7L2* [5/6 tumores com mutações ($\approx 83,3\%$) vs 1/6 tumores sem mutações ($\approx 16,6\%$), $p=0,04$].

Tabela 3.3. Resultados da análise de *copy number*, por MS-MLPA, nas amostras de CCR com mutações detectadas e respectivo estadiamento do tumor.

Exão/Intrão	Mutações	Amostra	Copy number										Estadiamento
			P0,2	P1,1	P1,2	P1,3	Exão 1	Intrão 2,2	Intrão 4	Intrão 5	Exão 14	Intrão 16	
Intrão16	113,161,621 C>T	595	0,71	2,28	1,84	1,25	2,47	0,8	1,08	1,2	0,55	1,69	M0
Exão 17	113,165,957 C>T	596	1,05	1,31	1,18	1,11	1,27	0,91	1,17	1,26	1,06	1,19	
Exão 15	113,160,646 G>A	611	0,66	1,66	2,03	1,15	2,27	0,81	1,02	1,53	0,52	1,61	M0
		612	0,77	2,08	2,08	1,32	2,43	0,77	1,17	1,55	0,66	1,92	
Exão17	113,165,759 T>C	1987	0,78	3,32	2,77	1,7	3,02	0,98	1,3	1,66	0,51	0,81	M0
		1988	1,03	1,3	1,12	1,12	1,34	0,96	1,13	1,27	1,05	1,22	
Exão 15	113,160,688 T>C	1977	0,98	1,22	1,3	1,05	1,18	1,46	1,06	1,28	1,03	1,27	T3N0M1
		1978	0,79	0,9	1,03	0,91	1,16	1,16	0,95	1,32	0,89	1,32	
Intrão 13	113,158,663 del T	1911	0,94	1,3	1,04	1,14	1,3	0,71	0,78	1,18	0,67	1,39	T3N0M0;
Exão 17	113,165,559 del A	1912	1,14	0,92	1,1	1,07	1,27	0,86	0,48	1,14	1,19	1,47	
Exão 17	113,165,559 del A	2229	1,22	2,51	2,01	1,73	2,42	1,26	1,25	1,69	0,86	1,09	T3N0M0
		2230	0,84	2,04	1,57	1,49	2,06	0,94	1,15	1,3	0,62	1,17	

Legenda: Cor azul – ganhos em relação as amostras de referência (a intensidade da cor é proporcional ao valor do ganho); Cor vermelha - perda em relação às amostras de referência (a intensidade da cor é proporcional ao valor da perda); P- região promotora.

Tabela 3.4 Resultados da análise de *copy number*, por MS-MLPA, em amostras de CCR sem mutação e respectivo estadiamento do tumor.

Amostra	Copy Number										Estadiamento
	P0,2	P1,1	P1,2	P1,3	Exão 1	Intrão 2,2	Intrão 4	Intrão 5	Exão 14	Intrão 16	
1639	2,36	6,28	1,16	2,21	2,72	1	2,63	3,12	0,7	2,84	T3N2M1
1640	3,06	5,83	1,14	2,14	3,81	1	1,82	2,77	0,55	2,51	
1705	0,89	1,25	1,02	1,11	1,34	1,28	1,1	1,32	1,26	1,07	T4N2M1
1706	0,89	1,18	0,95	1,17	1,44	0,77	1,25	1,25	0,97	1,19	
1707	0,95	1,79	0,54	2,17	2,24	1,93	0,58	1,21	1,46	1,29	T3N2M1
2031	1,16	2,66	2,75	1,75	2,34	2,25	1,17	1,46	0,8	1,21	T3N2M0
2032	1,52	1,6	1,25	1,22	1,37	1,46	1,12	1,3	0,79	1,32	
2115	1,03	1,23	2,3	1,28	1,09	1	1,18	1,3	1,24	1,36	T2N0M0
2116	1,57	1,25	1,71	1,68	1,23	0,91	1,21	1,08	1,25	0,98	
2197	1,79	3,23	1,78	2,22	2,29	2,3	1,08	2,12	1,02	1,77	T3N0M0
2198	1,23	1,33	1,1	1,16	1,2	0,81	1,11	0,97	1,31	1,61	

Legenda: Cor azul – ganhos em relação as amostras de referência (a intensidade da cor é proporcional ao valor do ganho); Cor vermelha - perda em relação as amostras de referência (a intensidade da cor é proporcional ao valor da perda); P- região promotora

3.2. Análise de perda de heterozigotia em tumores do recto

A análise de perda de heterozigotia (LOH) do gene *TCF7L2*, no cromossoma 10, foi realizada em 16 tumores do recto, com diferentes tipos de resposta à quimiorradioterapia pré-operatória. Para esta análise, recorreu-se a três marcadores de microssatélites (D10S1269, D10S168 e D10S562), uma vez que são analisadas sequências flanqueadoras do gene em estudo (marcadores D10S1269 e D10S562) e, também, uma sequência no intrão 4 do gene *TCF7L2* (marcador intragénico D10S168).

Inicialmente foram seleccionadas 21 amostras, mas em 5 amostras não foi possível amplificar especificamente, por PCR, a região de interesse dos marcadores para análise. Das 16 amostras de analisadas, 14 foram consideradas informativas, as quais apresentavam pelo menos um resultado normal ou perda de heterozigotia. Uma amostra foi considerada não informativa (caso 1927) e no caso 1937, a amostra tumoral analisada não era do recto mas sim da sigmóide, pelo que não se teve em conta esta amostra na análise estatística. Assim, verificou-se a presença de LOH (em pelo um dos marcadores analisados) em 10 dos 14 tumores analisados (10/14; $\approx 71\%$) (Figura 3.2). Considerou-se os casos E431T, 2525 e 2965, para as quais não foi possível analisar um dos marcadores, mas que apresentavam um resultado normal ou um padrão de instabilidade, sem perda de LOH, uma vez que na maioria dos casos a perda de heterozigotia verifica-se em pelo menos 2 marcadores de microssatélites. No entanto, será necessário complementar a análise de LOH.

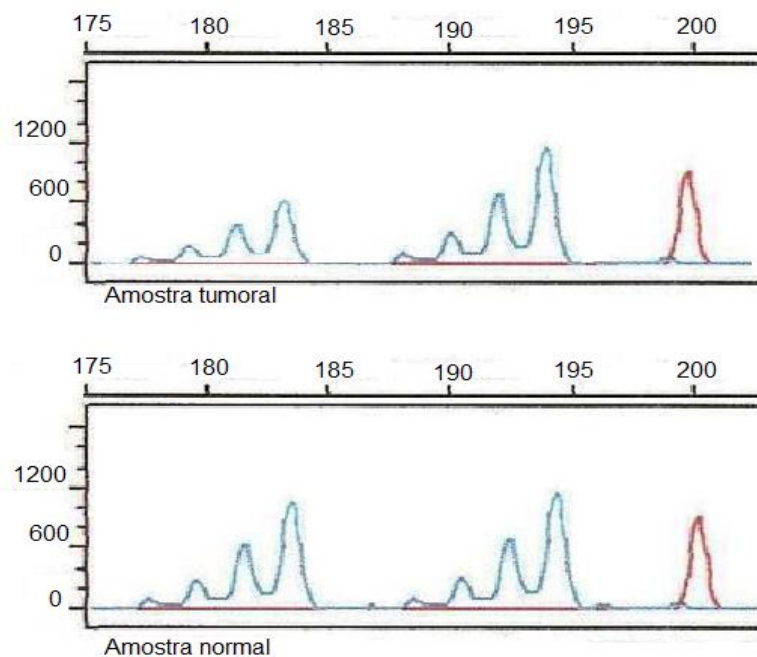


Figura 3.2. Electroforetograma correspondente a um caso (tumor/normal) com perda de heterozigotia (perda do alelo de menor peso molecular), para o marcador de microssatélites D10S562.

Os resultados obtidos encontram-se na tabela 3.5., na qual se descreve o resultado da perda de heterozigotia segundo a fórmula descrita no ponto 2.8.3, para cada um dos marcadores de microssatélites analisados.

Tabela 3.5. Resultados da análise de perda de heterozigotia (LOH) no gene *TCF7L2* em amostras do recto e a resposta à QRT.

Amostra		Marcadores de microssatélites			LOH	Resposta à QRT
Tumor	Normal	D10S1269	D10S168	D10S562		
2241	A1414	NI	MSI	N	MSI	(+)
2435	2436	NI	LOH	LOH	LOH	(+)
2525	2526s	-	MSI	MSI	MSI	(+)
1923	1924s	LOH	-	-	LOH	(+)
2965	2966	N	-	MSI	MSI	(+)
1937	1938	MSI	N	-	MSI	(+)
E431T	E431	N	N	-	N	(+/-)
1969	1970s	LOH	N	LOH	LOH	(+/-)
2837	2838	LOH	LOH	MSI	LOH	(+/-)
2887	2887	LOH	-	N	LOH	(+/-)
815T	816s	N	LOH	LOH	LOH	(-)
1913	1914s	NI	-	LOH	LOH	(-)
1927	1928	NI	NI	NI	NI	(-)
2817	2818s	LOH	LOH	LOH	LOH	(-)
2845	2846	LOH	NI	NI	LOH	(-)
3013	3014	LOH	-	MSI	LOH	(-)

Legenda: LOH – perda de heterozigotia; QRT – quimiorradioterapia; N – normal (sem perda de heterozigotia); MSI – Instabilidade de microssatélites; NI – não informativo (devido à homozigotia); - não amplificou; (+) resposta completa; (+/-) resposta moderada; (-) sem resposta; s – DNA extraído a partir de sangue periférico; E - DNA extraído a partir de tecido congelado.

A correlação entre a presença de perda de heterozigotia e a resposta do tumor do recto à quimiorradioterapia (QRT) pré-operatória revelou que a presença de LOH parece mais frequente em tumores que não responderam à terapia (inclui os tumores que não responderam e responderam moderadamente à terapia) do que em tumores que responderam totalmente à QRT pré-operatória [8/9 tumores (≈88,8%) vs. 2/5 tumores (40%), $p=0,094$].

Para além disso, estes resultados revelaram que a maioria dos tumores do recto que responderam completamente à quimiorradioterapia apresenta, em pelo menos um dos marcadores, instabilidade de microssatélites. No entanto, o seu valor não é estatisticamente significativo [3/5 tumores (60%) vs. 2/9 tumores (≈22,2%), $p=0,202$].

3.3. Análise do *copy-number* do gene *TCF7L2* em tumores do recto por MS-MLPA

Nesta análise, avaliou-se o *copy number* em 15 amostras de tumores do recto que responderam de diferentes formas à quimiorradioterapia pré-operatória. A avaliação do *copy number* permite completar os resultados obtidos pela análise da perda de heterozigotia. Desta forma, mesmo que o resultado indique perda de heterozigotia, pode não significar, exactamente, a perda de um dos alelos, mas sim o ganho de um dos alelos. Por isso, é necessário recorrer à técnica de MS-MLPA para confirmar o *copy number*, ou seja, se a perda de heterozigotia corresponde a um ganho ou a

uma perda. Para além disso, esta análise permite obter o resultado de um dos casos que pela análise de LOH não foi possível determinar a presença de LOH, uma vez que o resultado obtido foi não informativo.

Os electroforetogramas de todos os doentes em estudo foram analisados em relação às amostras de referência (amostras sem alteração de *copy number*) e foram identificadas as sondas de referência para garantir o sucesso da reacção. Em alguns casos não foi possível analisar um dos pares da amostra (DNA do tumor ou DNA normal – mucosa ou sangue), por não haver disponível.

Nesta análise, foi avaliada a região promotora do gene *TCF7L2*, tal como o exão 1, o intrão 2, 4 e 5, o exão 14 e o intrão 16. Os resultados obtidos por esta técnica encontram-se descritos na tabela 3.6.. Para alguns casos, em que foi feita a análise de LOH, não foi possível realizar a análise de *copy number*, uma vez que estas amostras de DNA são extraídas a partir de tecido incluído em parafina, por vezes o DNA está bastante fragmentado não sendo possível a realização desta técnica com êxito.

Tabela 3.6. Resultados da análise do *copy number* por MS-MLPA e indicação da resposta à QRT.

Amostra	P0,2	P1,1	P1,2	P1,3	Exão 1	Intrão 2,2	Intrão 4	Intrão 5	Exão 14	Intrão 16	Resposta à QRT
2241	0,84	0,91	1,11	0,83	0,91	0,55	0,94	0,75	0,76	0,99	(+))
A1414	1,08	1,2	1,25	1,14	1,11	1,17	1,12	1,12	1,11	1,14	
2269	1,03	1,39	1,14	1,35	1,7	1,4	1,37	1,28	1,04	1,19	(+))
2270s	1,18	1,29	1	1,27	1,31	1,12	1,2	1,25	1,17	1,12	
2525*											(+))
2526s	1,03	1,02	1,08	0,99	0,96	0,96	0,97	0,94	1,1	0,97	
2965	0,68	1,85	2,71	1,23	2,49	0,99	1,16	1,21	0,74	1,87	(+))
2966	0,59	0,94	1,43	1,45	1,01	0,84	1,26	0,69	0,97	0,66	
1923	1,15	1,01	1,42	0,8	1,96	1,05	1,61	1,16	1,15	1,79	(+))
1924s	1,31	2,61	2,33	1,83	2,26	1,82	1,38	1,59	0,82	1,25	
1937	1,02	0,97	0,92	1,52	1,57	0,64	0,99	0,84	0,76	0,68	(+))
1938	1,24	0,62	0,92	1,43	0,77	0,68	0,81	0,62	1,15	0,89	
E431T	0,97	0,61	0,54	0,61	0,41	0,13	1	0,93	0,86	0,94	(+/-)
E431*											
E436T	0,96	0,98	1,17	1,08	0,96	1,04	0,92	1,02	0,92	0,98	(+/-)
E436	0,93	1,09	1,59	0,97	0,75	0,77	0,79	0,93	0,86	0,79	
1927	1,37	1,35	1,09	1,56	1,44	1,26	1,25	1,15	1,27	1,15	(+/-)
1928	1,14	1,18	1,34	1,3	1,5	0,75	1	0,91	1,18	0,96	
1969	2,51	3,64	1,49	3,27	3,45	3,39	2,2	3,52	2,68	3,86	(+/-)
1970s	0,05	0,3	0,26	0,89	0,88	0,36	0,68	0,31	0,44	0,34	
2837	0,71	0,84	1,47	1,13	0,82	0,63	0,6	0,68	0,92	0,84	(+/-)
2838	0,77	0,87	1,84	1,5	1,29	0,92	0,42	0,91	1,2	0,49	
815T	0,82	0,84	0,99	0,95	0,89	0,99	0,86	0,88	0,85	0,91	(-))
816	0,66	1,04	1,13	0,81	0,93	0,89	0,85	0,95	0,73	0,93	
1913	0,62	2,85	3,15	1,86	1,82	1,92	0,87	2,03	0,4	1,06	(-))
1914s	1,01	1,02	0,97	0,94	0,93	0,88	1,03	0,97	0,92	0,96	
2817	0,97	1,09	0,95	1,23	1,05	0,88	0,98	0,92	1,01	0,64	(-))
2818s	1,09	1,21	1,09	1,14	1,14	1,08	1,18	1,11	1,13	1,04	
2845	1,33	2,15	2,35	2,34	2,3	1,11	1,02	0,99	0,84	0,91	(-))
2846	0,89	1,33	2,22	1,63	1,18	1,37	0,74	0,82	1,23	1,04	
2881	1,49	1,82	2,99	2,35	2,17	1,21	0,97	1,38	1,27	1,16	(-))
2882*											

Legenda: Cor azul – ganhos em relação as amostras de referência (a intensidade da cor é proporcional ao valor do ganho); Cor vermelha - perda em relação as amostras de referência (a intensidade da cor é proporcional ao valor da perda); P- região promotora; S- sangue; *não foi possível amplificar as sondas nesta amostra; QRT - quimiorradioterapia; (+) resposta completa; (+/-) resposta moderada; (-) sem resposta; s – DNA extraído a partir de sangue periférico; E - DNA extraído a partir de tecido congelado.

Em relação às amostras que foram estudadas quer pela análise de LOH quer pela análise de *copy number*, verificou-se que a maioria dos casos que apresentavam alterações de *copy number* (5/7 casos; ≈71%) tinha perda de heterozigotia. Em três casos, a perda de heterozigotia, está relacionado com o ganho da região dos promotores da mucosa tumoral (1913, 1969 e 2845), um dos casos com o ganho da região promotora da mucosa normal (1923) e o último tem vários tipos de alterações em ambas as mucosas (2837). Existem dois casos (815T e 2817) em que os resultados indicam a presença de LOH, contudo não se verifica alterações do *copy number*, assim sendo estas perdas poderão se localizar noutra região do gene não analisada por MS-MLPA ou então corresponder a LOH com *copy-neutral*

De acordo com os resultados, observa-se que existe uma associação significativa entre resposta do tumor do recto à quimiorradioterapia pré-operatória e a deleção do intrão 2 ou o ganho do intrão 16, sendo que 75% dos tumores que respondem total ou moderadamente à terapia apresentam uma destas alterações de *copy number* e nenhum tumor resistente à terapia apresenta estas alterações (6/8 tumores vs 0/6 tumores, $p=0,009$).

Por outro lado, observa-se que o ganho do promotor 1.3 e exão 1 é mais frequente em tumores que não responderam à terapia, sendo que, cerca de 66% dos tumores que não responderam à terapia apresentam ganho do promotor 1.3 e exão 1 (4/6 tumores); 25% dos tumores que responderam moderadamente têm ganho nesta região (1/4 tumores) e dos tumores que responderam completamente à terapia nenhum apresenta ganho do promotor 1.3 e exão 1 (0/4 tumores) [4/6 tumores resistentes vs. 1/8 tumores com boa resposta, $p=0,063$]. Uma vez que o ganho do intrão 16 parece associar-se a uma melhor resposta, se se incluir a deleção deste intrão ao marcador de pior resposta, isto é, ganho do promotor 1.3 e exão 1, verifica-se uma correlação significativa (1/8 tumores vs 5/6 tumores, $p=0,016$). Além disso, os resultados também demonstram que o único tumor resistente à terapia, cujo doente apresentou metástases síncronas, apresenta deleção no intrão 16. No entanto, é necessário analisar o *copy number* deste intrão em mais tumores do recto com metástases síncronas, de modo a avaliar se a deleção do intrão 16 também se associa à presença de metástases síncronas.

Além do *copy number*, também foi avaliada a metilação (Tabela 3.7). No entanto, não se observaram diferenças significativas entre as amostras normais correspondentes às referências e as amostras analisadas.

Tabela 3.7. Resultados da análise de metilação por MS-MLPA em tumores do recto.

Amostra	P 0,2	P1,1	P 1,2	P 1,3	Exão 1	Intrão 2,2	Intrão 4	Intrão 5	Exão 14	Intrão 16
2241	103%	16%	41%	7%	0	0	95%	5%	102%	105%
1414	97%	4%	25%	2%	10%	2%	95%	2%	79%	91%
2269	105%	16%	44%	5%	16%	0	84%	3%	88%	68%
2270	95%	4%	26%	0	10%	0	88%	2%	74%	94%
2525*										
2526	99%	5%	36%	0	10%	4%	88%	2%	85%	94%
2965	96%	37%	67%	13%	18%	18%	85%	12%	84%	92%
2966	197%	14%	118%	14%	0	0	46%	27%	68%	122%
1923	105%	29%	51%	10%	16%	6%	106%	7%	88%	97%
1924	81%	27%	42%	11%	11%	5%	103%	9%	56%	82%
1937	139%	48%	82%	15%	13%	0	131%	31%	94%	204%
1938	87%	58%	58%	0	0	0	101%	22%	63%	75%
E431T	83%	7%	65%	3%	8%	0	86%	3%	95%	86%
E431*										
E436T	88%	7%	32%	1%	10%	3%	95%	3%	93%	86%
E436	78%	4%	40%	0	10%	0	79%	3%	86%	82%
1927	92%	15%	40%	7%	12%	0	73%	5%	79%	77%
1928	83%	7%	48%	14%	0	0	96%	10%	74%	87%
1969	38%	3%	20%	1%	4%	1%	43%	1%	42%	30%
1970	118%	12%	30%	0	12%	0	116%	0	103%	97%
2837	101%	28%	51%	0	17%	0	0	17%	69%	0
2838	93%	33%	82%	24%	32%	0	149%	21%	62%	125%
815T	94%	6%	33%	0	7%	3%	96%	2%	100%	88%
816	97%	8%	36%	1%	10%	3%	90%	2%	90%	86%
1913	104%	66%	54%	54%	36%	39%	63%	32%	0	90%
1914	93%	14%	34%	2%	10%	3%	97%	4%	73%	99%
2817	94%	12%	40%	4%	12%	0	76%	0	70%	115%
2818	101%	5%	30%	0	11%	0	98%	0	79%	106%
2845	117%	16%	60%	6%	21%	15%	78%	8%	81%	99%
2846	89%	17%	48%	8%	31%	0	79%	0	75%	75%
2881	87%	29%	52%	5%	0	0	105%	8%	106%	116%
2882*										

Legenda: cor azul representa hipermetilação; cor vermelha representa hipometilação; * não foi possível amplificar as sondas desta amostra

De modo a se confirmar algumas das alterações de *copy-number* detectadas, foram efectuados estudos de *copy-number* por *real-time* PCR e estudos de expressão. No caso da deleção no intrão 4 detectada pela análise do *copy number* (1912), foi efectuada a análise do RNA para avaliar a existência de um transcrito adicional gerado pela deleção, eventualmente do exão 4 ou exões flanqueadores. No entanto, não foi possível amplificar o cDNA a partir do RNA da mucosa normal, uma vez que não havia RNA disponível, como alternativa recorreu-se ao RNA desta amostra extraído a partir de sangue periférico (1499).

Para esta análise, para além da amostra em análise (1499), utilizaram-se duas amostras de indivíduos sem CCR (controlos normais - N1 e N2), como também uma amostra de tumor sem deleção no intrão 4 (controlo normal - N3). O resultado da amplificação do cDNA entre os exões 2 e 7 da amostra 1499 e as amostras normais (sem deleção no intrão 4) (Figura 3.3) demonstra que independentemente da amostra existe a presença de vários transcritos. Isto deve-se aos vários locais de *splicing* que este gene contém, como explicado anteriormente. Contudo, em comparação com as amostras normais, não se observa a existência de uma banda extra, a qual corresponderia ao

transcrito aberrante. Cada um dos transcritos observados foram excisados, no entanto, não foi possível realizar a sequenciação dos transcritos em tempo útil de modo a confirmar este resultado.

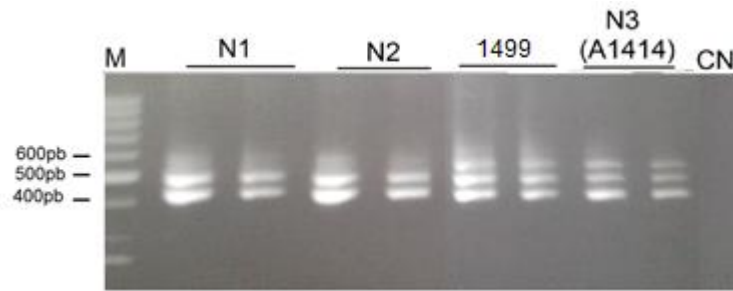


Figura 3.3. Resultado obtido através da amplificação do cDNA entre os exões 2 e 7 do gene *TCF7L2*, para a amostra 1499; Legenda: N1 – Controlo normal 1; N2 – Controlo normal 2; N3 – Controlo normal; CN – Controlo negativo; M – marcador *GeneRuler* 50pb (*ThermoScientific*).

3.4. Análise do *copy number* do gene *TCF7L2* por qPCR

De modo a confirmar as alterações de *copy number* indicadas pelo MS-MLPA, utilizou-se a técnica de qPCR, utilizando *primers* específicos (descritos nos anexos 8, tabela 2). Inicialmente realizou-se a análise para a região dos promotores, mas não foi possível obter resultados desta análise, pois a amplificação não era específica. Por isso, analisou-se as alterações do *copy number* no exão 1, como também, do exão e intrão 4 e do exão 14. Também se verificaram alterações de *copy number* no exão 16, mas não foi possível realizar este ensaio por não se dispor ainda de *primers*.

A análise do *copy number* do gene *TCF7L2* nas várias amostras foi normalizada em relação aos níveis de *copy number* do gene *housekeeping* *SDHA* (controlo endógeno).

Em relação à análise do exão 1 (Figura 3.4.), foram analisadas amostras com ganho neste exão indicado pela técnica MS-MLPA (1924, 1208 e 1660). Nesta análise foram incluídas duas amostras, que tinham sido analisadas num estudo anterior (1208 e 1660). Esta análise não foi conclusiva, uma vez que os valores foram bastante díspares. Nas duas amostras que apresentavam estar associadas a um ganho maior (1924 - 2,26; 1660 - 1,83) não revelaram nenhum aumento, sendo o seu valor semelhantes às amostras controlo (1914 e 2526) e a amostra 1208 que possivelmente o seu aumento era ligeiro, revelou aumento bastante elevado, por isso, será necessário repetir o ensaio. De facto, os valores dos triplicados neste ensaio foram bastante díspares, o que corrobora os resultados não concordantes.

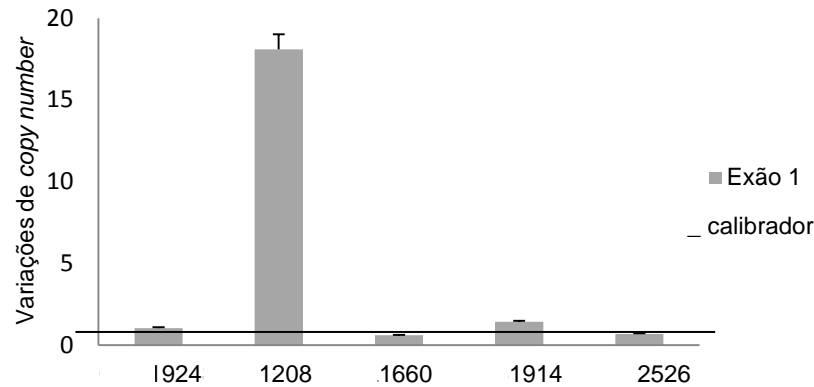


Figura 3.4. Resultados da análise do *copy number* do exão 1 do gene *TCF7L2* para as amostras, nas quais foram detectados ganhos do *copy number* pela técnica MS-MLPA. Os valores apresentados correspondem à média de triplicados. As amostras 1914 e 2526 são amostras controlo (sem ganhos neste exão em estudo).

Em seguida, analisaram-se as amostras com a perda no intrão e exão 4 (Figura 3.5). Neste conjunto, foi incluída uma amostra (704) que tinha sido analisada por MS-MLPA num estudo anterior. Em relação às amostras 1970 e 1924, a redução de *copy-number* não parece ser relevante, $\approx 0,7/ \approx 0,9$ (intrão e exão 4) e $\approx 0,9/\approx 0,8$ (intrão e exão 4), respectivamente, os quais estão muito próximos do valor das amostras normais. No entanto, estes resultados poderão indicar a presença de uma ligeira deleção, por isso é necessário repetir o ensaio para confirmar os resultados. O valor da amostra 1499, a qual corresponde ao DNA da mucosa tumoral 1911, não revela deleção nestas regiões, sendo o seu valor superior a 1. Uma vez que a deleção foi identificada na mucosa normal (0,48), isto pode indicar que é uma alteração somática e que não ocorre no sangue. Este resultado está de acordo com a amplificação de cDNA a partir de RNA extraído a partir de sangue deste doente, no qual não se observou a presença de um transcrito aberrante. Para confirmar esta deleção é necessário repetir o ensaio com DNA extraído da mucosa tumoral. O valor da amostra 704 não é concordante com o resultado do MS-MLPA, pois o seu valor ($\approx 0,8$ em ambos) é muito próximo do valor das amostras normais e o resultado do MS-MLPA indica que a deleção é significativa (0,38). Desta forma, é necessário repetir este ensaio, como também o MS-MLPA para confirmar os resultados.

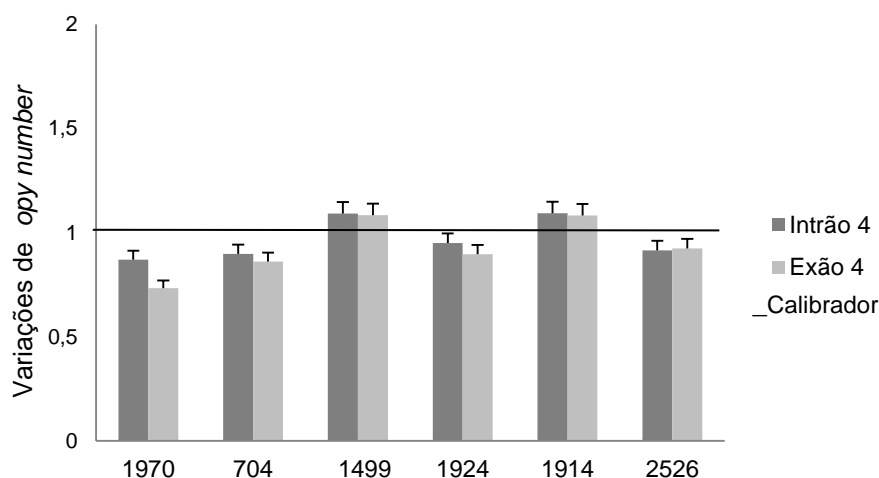


Figura 3.5. Resultados da análise do *copy number* do intrão e exão 4 do gene *TCF7L2* para as amostras, nas quais foram detectadas perdas do *copy number* pela técnica MS-MLPA. Os valores apresentados correspondem à média de triplicados. As amostras 1914 e 2526 são amostras controlo (sem alterações de *copy number*).

Por último, a análise do *copy number* do exão 14 na amostra 1970 (Figura 3.6.), parece confirmar a deleção deste exão ($\approx 0,6$), pois apresenta um valor de *copy number* inferior às amostras de controlo (sem alterações de *copy number*), o qual é compatível com o valor obtido no MS-MLPA.

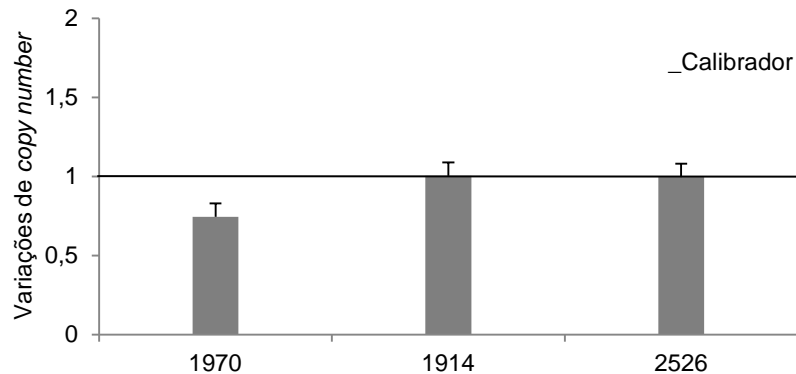


Figura 3.6. Resultados da análise do *copy number* exão 14 do gene *TCF7L2* para a amostra 1970, na qual foi detectada perda do *copy number* pela técnica MS-MLPA. Os valores apresentados correspondem à média de triplicados. As amostras 1914 e 2526 são amostras controlo (sem alterações de *copy number*).

3.5. Quantificação do nível de expressão do gene *TCF7L2* por qPCR

Nesta análise, pretendeu-se verificar se as alterações de *copy number* identificadas nas análises anteriores resultavam na alteração de expressão do gene *TCF7L2*. Para tal, fez-se dois tipos de ensaios, primeiramente quantificou-se os níveis de expressão entre os exões 3 e 4 para analisar as amostras que apresentam alterações de *copy number* no intrão 4 nos resultados de MS-MLPA e efectuou-se a quantificação dos níveis de expressão entre os exões 10 e 11, para quantificar a expressão total, considerando que à partida não existe *splicing* nestes exões. A análise da expressão do gene *TCF7L2* nas amostras foi normalizada em relação aos níveis de expressão do gene *housekeeping* (β -actina) (controlo endógeno).

Para esta análise, utilizou-se o cDNA das amostras 1924 e 1208, na primeira amostra foi detectado ganho de *copy number* na região promotora, no exão 1 e nos intrões 2 e 5 e na segunda amostra também se verificou o ganho na região promotora e nos intrões 4 e 5.

Os resultados da análise de expressão revelaram um aumento significativo dos níveis de expressão deste gene relativamente à expressão do gene *housekeeping* (β -actina) tanto no ensaio para os exões 3 e 4 (Figura 3.7.) como para o ensaio para os exões 10 e 11 (Figura 3.8.) em ambas as amostras. Porém, observa-se um aumento muito elevado dos níveis de expressão da amostra 1924, sendo necessário repetir o ensaio para confirmar o resultado.

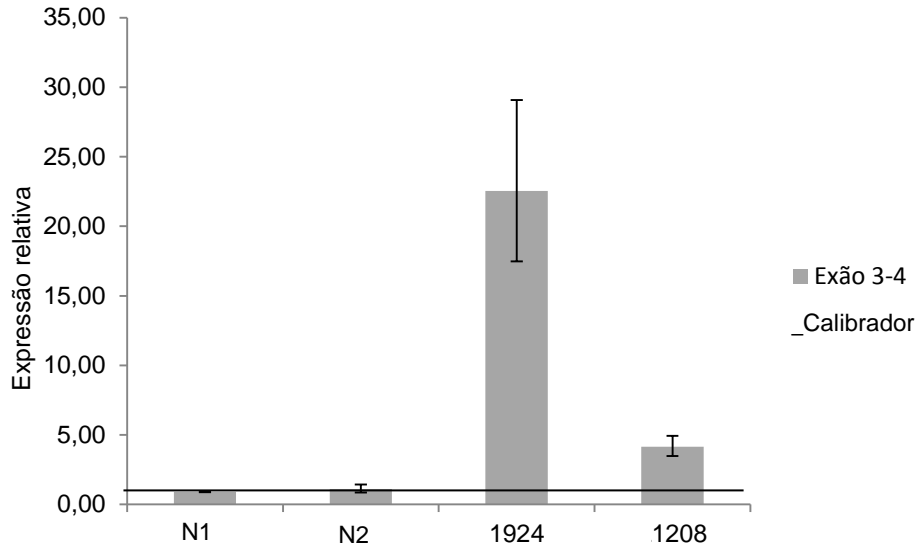


Figura 3.7. Resultados do ensaio de expressão relativa dos exões 3-4 do gene *TCF7L2* em amostras com alteração do *copy number* no intrão 4. Os valores apresentados correspondem à média de triplicados. As amostras N1 e N2 são amostras controle (indivíduos saudáveis).

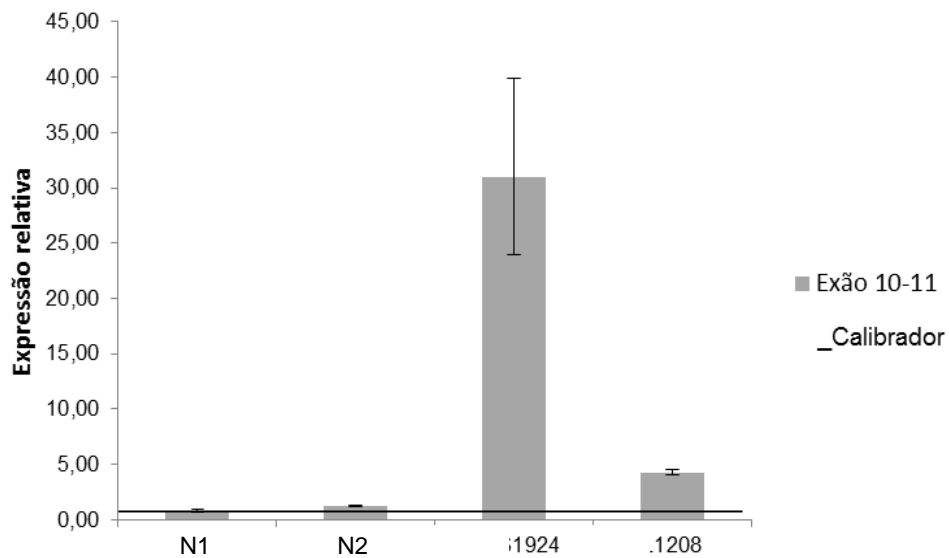


Figura 3.8. Resultados do ensaio de expressão relativa dos exões 10-11 do gene *TCF7L2* em amostras com alteração do *copy number*. Os valores apresentados correspondem à média de triplicados. As amostras N1 e N2 são amostras controle (indivíduos saudáveis).

4. Discussão

A via de sinalização WNT, também conhecida por via de sinalização APC/ β -catenina/TCF7L2, é uma das vias mais importantes para a regulação da proliferação, diferenciação e apoptose celular das criptas intestinais (Reya e Clevers, 2001). A perturbação do equilíbrio nesta via conduz à formação tumoral no cólon e recto. A maioria dos tumores colorectais (CCR) apresenta mutações nos genes que codificam para componentes desta via, tais como *APC*, *CTNNB1* e *AXIN2* (Fodde *et al.*, 2001; Reya e Clevers, 2001; Albuquerque *et al.*, 2011). Recentemente, alguns estudos têm associado o gene *TCF7L2*, o qual codifica para um factor de transcrição envolvido nesta via, a uma função de supressão tumoral no CCR.

4.1. Caracterização molecular do gene *TCF7L2* em tumores colorectais

Recentemente, o grupo de Gastrenterologia da UIMP, IPOLFG, EPE, observou que uma mutação *frameshift*, pela deleção de um nucleótido, no exão 17, em tumores colorectais com deficiência no sistema MMR, aumentava a sua frequência ao longo da sequência adenoma-carcinoma, sugerindo que esta mutação possa estar envolvida no desenvolvimento tumoral. Deste modo, no presente estudo foi realizada a análise mutacional em sete exões do gene *TCF7L2*, para averiguar se outras mutações neste gene poderiam contribuir para o desenvolvimento tumoral.

A análise mutacional permitiu a identificação de 7 mutações no gene *TCF7L2* num total de 68 amostras: duas mutações *frameshift* pela deleção de um nucleótido numa região repetitiva (região intrónica flanqueadora *upstream* do exão 13 e exão 17) e 5 substituições, das quais 2 são variantes *missense* (exão 15), 1 variante intrónica (região intrónica flanqueador *downstream* do exão 16) e 2 mutações silenciosas (exão 17). Apenas as mutações do tipo *frameshift* na região intrónica flanqueadora *upstream* do exão 13 e exão 17, assim como uma das mutações *missense* do exão 15 parecem ser patogénicas. A mutação *frameshift* no exão 17 referida anteriormente, correspondente à detectada anteriormente pelo grupo de Gastrenterologia, foi detectada em 8 amostras, encontrando-se já está descrita na literatura (Duval *et al.*, 1999). Esta mutação foi detectada apenas em tumores com instabilidade de microssatélites, o que está de acordo com o estudo realizado por Duval e seus colaboradores, que demonstra que a presença desta mutação é mais frequente em linhas celulares de CCR com instabilidade de alto grau. Esta deleção pode originar uma proteína truncada na região C-terminal, perdendo a capacidade de interagir com a proteína CtBP, a qual está envolvida na repressão transcricional mediada pelo domínio C-terminal. Assim de acordo com o previamente descrito, esta

mutação poderá ser um segundo evento mutacional seleccionado nas células tumorais para activar a via WNT, subsequentemente às mutações no *APC* e *CTNNB1* (β -catenina).

Para além desta *frameshift*, foi detectada uma deleção numa região repetitiva de 8 timinas na região intrónica flanqueadora do exão 13, apenas numa amostra com instabilidade de microsatélites (tumor 1911). Esta mutação também já tinha sido descrita por Duval e seus colaboradores, em 2000, considerando-a como uma alteração não codificante. No entanto, a análise *in silico* feita pelo VEP indica que poderá ser uma variante de *splicing* e o *software Mutation Taster* indica que é uma mutação patogénica por aumentar o local aceitador de *splicing*, aumentando a probabilidade da ocorrência deste processo. A análise de *copy number*, desta amostra, indica a deleção do exão 14, o que eventualmente poderá está associado com um eventual *splicing* do exão 13 e a uma hipotética selecção pela eliminação do exão 14, de modo a favorecer a tumorigénese. Nesta amostra também foi detectada a presença da mutação *frameshift* no exão 17. Observando que na análise de *copy number* da amostra 2229, na qual foi apenas detectada a *frameshift* no exão 17, não ocorre deleção do exão 14, coloca-se a hipótese da mutação no exão 13 poder estar relacionada com a deleção do exão 14. Um estudo realizado por Weise e seus colaboradores, em 2009, demonstrou que, na maioria dos transcritos, os exões 14 e 15 são mutuamente exclusivos e que a deleção de um deles origina a alteração da grelha de leitura, produzindo uma proteína truncada no exão 17 (Weise *et al.*, 2009). Desta forma, juntamente, com a presença da *frameshift* no exão 17 neste tumor, a qual origina também uma proteína truncada, esta mutação poderá estar relacionada com o desenvolvimento de CCR precoce neste doente (35 anos).

A avaliação da mutação do exão 15 (amostra 1977), na qual ocorre a substituição de uma timina por uma citosina, não é concordante entre os *softwares*. O SIFT considera que o seu efeito é tolerante, enquanto, o *Polyphen* considera benigna ou possivelmente patogénica dependendo do transcrito e o *software Mutation Taster* determina-a como patogénica, por levar à perda do local de *splicing*, sugerindo a deleção deste exão. Sabe-se que a ausência do exão 15, em células renais, determina a forma longa do transcrito, a qual contribui para a inibição da apoptose através da redução da expressão dos genes-alvo da via WNT (Shiina *et al.*, 2003). Assim, a mutação no exão 15, ao levar à deleção deste exão poderá originar a alteração da grelha de leitura, produzindo uma proteína truncada no exão 17 tal como descrito anteriormente (Weise *et al.*, 2009). Para além desta análise, verifica-se que não ocorre alteração do *copy number* do exão 14 nesta amostra, o que está de acordo o estudo de Weise, o qual refere que a maioria dos transcritos não contém simultaneamente os dois exões. Desta forma, esta alteração poderá contribuir para a progressão tumoral. A suportar este facto, esta amostra corresponde ao único tumor com estadio M1 (metástases síncronas).

4.2. Caracterização molecular do gene *TCF7L2* em tumores do recto e relação com a resposta à quimiorradioterapia pré-operatória

A resistência dos tumores do recto à quimiorradioterapia pré-operatória representa um dos principais problemas clínicos para o tratamento destes doentes. A redução de expressão do gene *TCF7L2* foi associada a uma maior sensibilidade destes tumores à quimiorradioterapia (Kendziorra, *et*

al., 2005). Desta forma, com este estudo, pretendeu-se a identificação de alterações (epi)genéticas em tumores do recto, as quais poderão se correlacionar com a resposta à terapia.

No estudo dos tumores do recto, observou-se que a deleção do intrão 2 e o ganho do intrão 16 correlaciona-se significativamente com a melhor resposta à quimiorradioterapia pré-operatória (tumores que responderam completa ou moderadamente à terapia). Está descrito que a inclusão do exão 16 é incompatível com a produção de uma isoforma longa, no entanto, também se sabe que a presença das isoformas com este exão é mais frequente em células estaminais embrionárias, as quais se caracterizam por apresentarem uma proliferação excessiva (Weise *et al.*, 2009). Sabendo, que a quimioterapia origina erros no DNA, quanto maior a actividade proliferativa maior a incorporação de erros, conduzindo à apoptose celular. Desta forma, se o ganho do intrão 16 reflectir um ganho do exão 16, estas células tumorais, ao expressarem esta isoforma têm maior capacidade proliferativa, podendo assim existir uma melhor resposta à terapia. Deste modo, estas alterações poderão funcionar como indicadores de resposta à quimiorradioterapia pré-operatória em tumores do recto.

Por outro lado, verifica-se que a presença de perda de heterozigotia do gene *TCF7L2* foi significativamente mais frequente em tumores resistentes ou que responderam moderadamente à quimiorradioterapia pré-operatória. Nestes casos, verifica-se que a perda de heterozigotia correspondia a ganhos de *copy number* nos promotores e no exão 1 pela análise de MS-MLPA. Observa-se que a frequência deste ganho aumenta gradualmente de acordo com o tipo de resposta à terapia (de resposta completa a ausência de resposta). Os resultados demonstraram que existe uma associação entre o ganho de *copy number* da sonda promotor 1.3 e do exão 1 do gene *TCF7L2* e a resistência dos tumores do recto.

O facto de se terem incluído na análise três tumores previamente submetidos a tratamento (2837, 2845 e 2881), reforça a necessidade de aumentar o número de tumores do recto pré-terapia, apesar do seu padrão revelar-se igual a outros casos que não foram submetidos a tratamento.

O estudo realizado por Ghadimi e os seus colaboradores, em 2005, demonstrou que o gene *TCF7L2* era significativamente sobre-expresso nos tumores resistentes à quimioterapia pré-operatória baseada em 5-FU (Ghadimi *et al.*, 2005). Assim, o ganho do promotor e exão 1 poderá ser uma das alterações responsáveis pela sobre-expressão deste factor, que leva a resistência à terapia. Desse modo, realizou-se a quantificação relativa do *copy number* do exão 1 das amostras que apresentaram este ganho, para confirmar o resultado. Relativamente a este ensaio, os resultados não foram conclusivos, pois os valores obtidos para os 3 ensaios do triplicado foram bastantes díspares. Porém, verificou-se pela quantificação da expressão do *TCF7L2* nestas amostras que a expressão deste gene aumenta em amostras com ganho. Assim, o ganho da região do promotor e exão 1 poderá ser um potencial marcador preditivo de resposta à terapia, uma vez que poderá aumentar a expressão do gene, o que está relacionado com a resistência dos tumores do recto à quimiorradioterapia pré-operatória

Para além destas alterações de *copy number*, verificou-se que esta associação aumenta com a deleção do intrão 16. Sabe-se que os transcritos com ausência do exão 16 são mais frequentes em células diferenciadas (Weise *et al.*, 2009). Assim, em células com menor proliferação a terapia não é tão eficaz na indução de erros no DNA, não conduzindo à apoptose celular e, por isso, os tumores são

mais resistentes à quimioterapia. Assim, estas alterações poderão ser indicadoras de resistência à quimiorradioterapia. Será, no entanto, necessário realizar a análise de *copy number* dos exões 16 e 17, pois a sonda do intrão 16 encontra-se bastante perto do exão 17, o que poderá implicar a sua deleção. Assim, será necessário determinar se o ganho do intrão 16 se reflecte também num ganho do exão 16.

No que diz respeito à LOH, em dois tumores que apresentaram resistência à terapia (815T e 2817) e LOH, não se verificaram os referidos ganhos de *copy number*, pelo que a LOH poderá ocorrer em outros locais do gene que não são abrangidos pelas sondas de MS-MLPA ou ser *copy-neutral*.

Para além disso, observou-se que alguns tumores que responderam à terapia apresentaram ganho de uma ou duas sondas do promotor. Tal como se pode observar no tumor 1923, verificou-se um ganho no exão 1, no entanto, os promotores apresentam perda em relação ao DNA normal. Deste modo, observa-se que a perda dos promotores poderá estar associada a uma melhor resposta à terapia. Em relação ao outro caso, pode ocorrer a mesma situação no entanto, não foi possível obter resultado para a amostra de mucosa normal.

No que diz respeito à deleção do exão 14 e do exão 4 será necessário analisar um maior número de amostras por MS-MLPA, assim como confirmar o resultado por qPCR, de modo a avaliar qual o seu impacto na resposta à quimiorradioterapia pré-operatória no recto.

Em relação à ausência de alterações de metilação significativas, uma vez que a expressão do gene *TCF7L2* é bastante regulada por *splicing* alternativo, a metilação poderá não ser um evento importante para a regulação deste gene, pelo menos do recto.

5. Conclusão

Os cancros colorectais de origem esporádica correspondem a 70% dos CCR e surgem pela acumulação de mutações somáticas, não existindo história família e desenvolve-se por volta dos 70 anos de idade. Existem poucos marcadores de prognóstico, principalmente de resposta à terapia (Jasperson *et al.*, 2010). Até ao momento, já tinha sido detectada uma associação entre uma alteração específica no gene *TCF7L2* e a metastização (Gonçalves, 2010) No entanto, pouco se sabe sobre alterações mutacionais neste gene, há excepção da mutação *frameshift* no exão 17 em tumores instáveis, descrita por Duval (Duval *et al.*, 1999).

A referida mutação no exão 17 verificou-se igualmente frequente neste estudo em tumores com instabilidade de microssatélites. No entanto, pela análise mutacional realizada, averiguou-se que as mutações neste gene não parecem ser frequentes em tumores estáveis, particularmente nos exões que foram analisados. No entanto, as mutações que foram detectadas parecem estar associadas com alterações de *splicing*, tendo sido detectadas em tumores diagnosticados numa idade jovem ou associados a metástases à distância. Será, no entanto, necessário analisar a expressão do RNA nestas amostras de modo a confirmar a referida alteração ao nível do *splicing*.

Em relação à análise realizada nos tumores do recto, os ganhos de *copy number* na região *upstream* e 5' do gene *TCF7L2* parecem estar associados à ausência de resposta do tumor à quimiorradioterapia pré-operatória, os quais originam um aumento dos níveis de expressão do gene. Para além desta alteração, foi ainda demonstrado que o ganho na região 3' do gene parece estar associado a uma melhor resposta.

Deste modo, as mutações e alterações de *copy number* poderão ser potenciais marcadores de prognóstico no CCR e de resposta à quimiorradioterapia no recto. No entanto, os resultados são ainda preliminares, sendo necessário a análise de um maior número de amostras.

6. 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. *Nature methods*, 7(4):248–249.

Albuquerque C., Bakker E. R., van Veelen W., Smits R. 2011. Colorectal cancers choosing sides. *Biochim Biophys Acta* 1816(2):219-31.

Albuquerque C., Baltazar C., Filipe B., Penha F., Pereira T., Smits R., Cravo M., Lage P., Fidalgo P., Claro I., Rodrigues P., Veiga I., Ramos J. S., Fonseca I., Leitão C. N., Fodde R. 2010. Colorectal cancers show distinct mutation spectra in members of the canonical WNT signaling pathway according to their anatomical location and type of genetic instability. *Genes Chromosomes Cancer* 49(8):746-59.

Albuquerque C., Breukel C., van der Looij R., Fidalgo P., Lage P., Slors F. J., Leitão C. N., Fodde R., 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):1549-60.

Angus-Hill M. L., Elbert K. M., Hidalgo J., Capecchi M.R. 2011. T-cell factor 4 functions as a tumor suppressor whose disruption modulates colon cell proliferation and tumorigenesis. *Proc Natl Acad Sci U S A* 108(12):4914-9.

Armaghany T., Wilson J. D., Chu Q., Mills G. 2012. Genetic alterations in colorectal cancer. *Gastrointest Cancer Res* 5(1):19-27.

Boland C. R. 2005. Evolution of the nomenclature for the hereditary colorectal cancer syndromes. *Fam Cancer* 4(3):211-8.

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

Boyle P., Langman J. S. 2000. ABC of colorectal cancer: Epidemiology. *BMJ* 321(7264):805-8.

Campbell, S.L., R. Khosravi-Far, K.L. Rossman, G.J. Clark, C.J. 1998 .Increasing complexity of Ras signaling. *Oncogene* 17(11 Reviews):1395–1413.

Carethers J. M. 2008. Systemic treatment of advanced colorectal cancer: tailoring therapy to the tumor. *Therap Adv Gastroenterol* 1(1):33-42.

Centelles, J. J. 2012. General aspects of colorectal cancer. *ISRN Oncol.* doi: 10.5402/2012/139268.

- Chu, E. 2008. New Treatment strategies for metastatic colorectal cancer. CMP Healthcare Media, Manhasset, NY.
- Diasio R. B., Harris B. E. 1989. Clinical pharmacology of 5-fluorouracil. Clin Pharmacokinet Apr;16(4):215-37.
- Duval A., Gayet J., Zhou X. P., Iacopetta B., Thomas G., Hamelin R. 1999. Frequent frameshift mutations of the *TCF-4* gene in colorectal cancers with microsatellite instability. Cancer Res 59(17):4213-5.
- Duval A., Rolland S., Tubacher E., Bui H., Thomas G., Hamelin R. 2000. The human T-cell transcription factor-4 gene: structure, extensive characterization of alternative splicings, and mutational analysis in colorectal cancer cell lines. Cancer Res 60(14):3872-9.
- Erich A. E. 1989. Polymerase Chain Reaction. Journal of Clinical Immunology 9(6):437.
- Fearon E. R., Vogelstein B. 1990. A genetic model for colorectal tumorigenesis. Cell 61(5):759-67.
- Fodde, R. e Losekoot, M. 1994. Mutation detection by denaturing gradient gel electrophoresis (DGGE). Hum Mutat 3: 83–94.
- Geiersbach K. B., Samowitz W. S. 2011. Microsatellite instability and colorectal cancer. Arch Pathol Lab Med 135 (10):1269-77.
- Ghadimi B. M., Grade M., Difilippantonio M. J., Varma S., Simon R., Montagna C., Füzési L., Langer C., Becker H., Liersch T., Ried T. 2005. Effectiveness of gene expression profiling for response prediction of rectal adenocarcinomas to preoperative chemoradiotherapy. J Clin Oncol 23(9):1826-38.
- Giles R. H., van Es J. H., Clevers H. 2003. Caught up in a Wnt storm: Wnt signaling in cancer. Biochim Biophys Acta 1653(1):1-24.
- GLOBOCAN, 2012 - http://globocan.iarc.fr/Pages/fact_sheets_cancer.aspx.
- Goncalvez, N. 2010. Terapêutica no Cancro do Cólon e Recto: contribuição da via de sinalização APC/ β -catenina/TCF7L2. Tese de Mestrado. Faculdade de Ciências, Universidade de Lisboa.
- Graham J., Muhsin M., Kirkpatrick P. 2004. Cetuximab. Nat Rev Drug Discov 3(7):549-50.
- Haggar F. A., Boushey R. P. 2009. Colorectal cancer epidemiology: incidence, mortality, survival, and risk factors. Clin Colon Rectal Surg 22 (4):191-7.
- Hecht A., Stemmler M. P. 2003. Identification of a promoter-specific transcriptional activation domain at the C terminus of the Wnt effector protein T-cell factor 4. J Biol Chem 278(6):3776-85.
- Heinen C. D. 2010. Genotype to phenotype: analyzing the effects of inherited mutations in colorectal cancer families. Mutat Res 693(1-2):32-45.

- Hurwitz H., Fehrenbacher L., Novotny W., Cartwright T., Hainsworth J., Heim W., Berlin J., Baron A., Griffing S., Holmgren E., Ferrara N., Fyfe G., Rogers B., Ross R., Kabbinavar F. 2004. Bevacizumab plus irinotecan, fluorouracil, and leucovorin for metastatic colorectal cancer. *N Engl J Med* 350(23):2335-42.
- Jasperson K. W., Tuohy T. M., Neklason D. W., Burt R. W. 2010. Hereditary and familial colon cancer. *Gastroenterology* 138(6):2044-58.
- Jemal A., Bray F., Center M. M., Ferlay J., Ward E., Forman D. 2011. Global cancer statistics. *CA Cancer J Clin* 61(2):69-90.
- Kashfi S. M., Golmohammadi M., Behboudi F., Nazemalhosseini-Mojarad E., Zali M. R. 2013. *MUTYH* the base excision repair gene family member associated with colorectal cancer polyposis. *Gastroenterol Hepatol Bed Bench* 6(Suppl 1):S1-S10.
- Kendziorra E., Ahlborn K., Spitzner M., Rave-Fränk M., Emons G., Gaedcke J., Kramer F., Wolff H. A., Becker H., Beissbarth T., Ebner R., Ghadimi B. M., Pukrop T., Ried T., Grade M. 2011. Silencing of the Wnt transcription factor TCF4 sensitizes colorectal cancer cells to (chemo-) radiotherapy. *Carcinogenesis* 32(12):1824-31.
- Kerr, D., 2003. Clinical development of gene therapy for colorectal cancer. *Nat Rev Cancer* 3(8): p. 615-22.
- Korinek V., Barker N., Moerer P., van Donselaar E., Huls G., Peters P. J., Clevers H. 1998. Depletion of epithelial stem-cell compartments in the small intestine of mice lacking Tcf-4. *Nat Genet* 19(4):379-83.
- Leggett B. e Whitehall V. 2010. Role of the serrated pathway in colorectal cancer pathogenesis. *Gastroenterology* 138(6):2088-100.
- Lenz H. J. 2007. Cetuximab in the management of colorectal cancer. *Biologics* 1(2):77-91.
- Ligtenberg M. J., Kuiper R. P., Chan T. L., Goossens M., Hebeda K. M., Voorendt M., Lee T. Y., Bodmer D., Hoenselaar E., Hendriks-Cornelissen S. J., Tsui W. Y., Kong C. K., Brunner H. G., van Kessel A. G., Yuen S. T., van Krieken J. H., Leung S. Y., Hoogerbrugge N. 2009. 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):112-7.
- Lindor N. M. 2009. 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., Jass J., LeMarchand L., Grove J., Potter J., Newcomb P. 2005. Lower cancer incidence in Amsterdam-I criteria families without mismatch repair deficiency: familial colorectal cancer type X. *JAMA* 293(16):1979-85.

- Lodish H., Berk A., Matsudaira P., Kaiser C. A., Krieger M., Scott M. P., Zipursky L., Darnell J. 2004. *Molecular Cell Biology*, 5th edn. NY, EUA: W. H. Freeman.
- Longley D. B., Harkin D. P., Johnston P. G. 2003. 5-fluorouracil: mechanisms of action and clinical strategies. *Nat Rev Cancer* 3(5):330-8.
- 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* (Oxford, England), 26(16):2069–2070.
- Narayan S. e Roy D. 2003. Role of *APC* and DNA mismatch repair genes in the development of colorectal cancers, *Mol Cancer* 2: 41. doi: 10.1186/1476-4598-2-41.
- Ng, P. C. and Henikoff, S. 2003. SIFT: predicting amino acid changes that affect protein function. *Nucleic Acids Research*, 31(13):3812–3814.
- Noordhuis P, Holwerda U, Van der Wilt CL, Van Groeningen CJ, Smid K, Meijer S, Pinedo HM, Peters GJ. 2004. 5-Fluorouracil incorporation into RNA and DNA in relation to thymidylate synthase inhibition of human colorectal cancers. *Ann Oncol Jul*;15(7):1025-32.
- 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 Research*, 33.
- Patel S. G., Ahnen D. J. 2012. Familial colon cancer syndromes: an update of a rapidly evolving field. *Curr Gastroenterol Rep* 14(5):428-38.
- Pino, M. S., e Chung, D. C. 2010. The chromosomal instability pathway in colon cancer. *Gastroenterology* 138(6): 2059–72.
- Polakis P. 2000. Wnt signaling and cancer. *Genes Dev* 14(15):1837-51.
- Reya T., Clevers H. 2005. Wnt signalling in stem cells and cancer. *Nature* 434(7035):843-50.
- Sambrook J, Russel DW. 2001. *Molecular Cloning: A Laboratory Manual*. Vol. 2, 3rd ed. NY, EUA: Cold Spring Harbor Laboratory Press.
- Sanger F., Nicklen S., Coulson A. R. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* 74: 5463-5467.
- 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.
- Tang W., Dodge M., Gundapaneni D., Michnoff C., Roth M., Lum L. 2008. A genome-wide RNAi screen for Wnt/beta-catenin pathway components identifies unexpected roles for TCF transcription factors in cancer. *Proc Natl Acad Sci U S A* 105(28):9697-702.

- Terdiman J. P., Conrad P., Moslein G., Goldberg R., Ziogas A., Anton-Culver H., de Andrade M., Siegmund K., Thibodeau S. N., Boardman L. A., Seminara D. 2005. Lower cancer incidence in Amsterdam-I criteria families without mismatch repair deficiency: familial colorectal cancer type X. *JAMA* 293(16):1979-85.
- Toyota M., Ahuja N., Ohe-Toyota M., Herman J. G., Baylin S. B., Issa J. P. 1999. CpG island methylator phenotype in colorectal cancer. *Proc Natl Acad Sci U S A* 96(15):8681-6.
- Travis A., Amsterdam A., Belanger C., Grosschedl R. 1991. LEF-1, a gene encoding a lymphoid-specific protein with an HMG domain, regulates T-cell receptor alpha enhancer function [corrected]. *Genes Dev* 5(5):880-94.
- Umar A., Boland C. R., Terdiman J. P., Syngal S., de la Chapelle A., Rüschoff J., Fishel R., Lindor N. M., Burgart L. J., Hamelin R., Hamilton S. R., Hiatt R. A., Jass J., Lindblom A., Lynch H. T., Peltomaki P., Ramsey S. D., Rodriguez-Bigas M. A., Vasen H. F., Hawk E. T., Barrett J. C., Freedman A. N., Srivastava S. 2004. Revised Bethesda Guidelines for hereditary nonpolyposis colorectal cancer (Lynch syndrome) and microsatellite instability. *J Natl Cancer Inst* 96(4):261-8.
- van de Wetering M., Oosterwegel M., Dooijes D., Clevers H. 1991. Identification and cloning of TCF-1, a T lymphocyte-specific transcription factor containing a sequence-specific HMG box. *EMBO J* 10(1):123-32.
- Varnat F., Siegl-Cachedenier I., Malerba M., Gervaz P., Ruiz i Altaba A. 2010. Loss of WNT-TCF addiction and enhancement of HH-GLI1 signalling define the metastatic transition of human colon carcinomas. *EMBO Mol Med* 2(11):440-57.
- Weise A., Bruser K., Elfert S., Wallmen B., Wittel Y., Wöhrle S., Hecht A. 2010. Alternative splicing of *Tcf7l2* transcripts generates protein variants with differential promoter-binding and transcriptional activation properties at Wnt/beta-catenin targets. *Nucleic Acids Res* 38(6):1964-81.
- Wilke H. J., Van Cutsem E. 2003. Current treatments and future perspectives in colorectal and gastric cancer. *Ann Oncol* 14 Suppl 2:ii49-55.
- World Health Organization - World Health Statistics 2012
http://www.who.int/gho/publications/world_health_statistics/2012/en/.
- Yalçın S. 2009. The increasing role of pharmacogenetics in the treatment of gastrointestinal cancers. *Gastrointest Cancer Res* 3(5):197-203.
- Zambirinis C. P., Theodoropoulos G., Gazouli M. 2009. Undefined familial colorectal cancer, *World J Gastrointest Oncol* 1(1):12-20.

7. Anexos

Anexo 1 – Protocolos para a extracção de ácidos nucleicos

1.1 Protocolo para extracção de DNA genómico a partir de tecido incluído em parafina

- 1) Adicionar 1ml de tampão de lise e 40µl de proteinase K (20mg/ml) (*Sigma*) a cada tubo de 2ml contendo os cortes de parafina, agitar em *vórtex* e incubar a 56°C *over-night* (o/n);
Nota: Se o tecido não estiver totalmente digerido após o período de incubação, adicionar novamente 40µl de proteinase K e voltar a incubar. Este passo repete-se no máximo 3 vezes.
- 2) Armazenar os tubos a -20°C durante, pelo menos 3 horas;
- 3) Descongelar e distribuir igual volume, cerca de 500µl, por dois tubos de 2ml;
- 4) Adicionar, igual volume de fenol/clorofórmio (1:1) a cada tubo, agitar em *vórtex* e centrifugar a 14000rpm durante 15 minutos a 4°C;
- 5) Pipetar a fase aquosa para novos tubos de 2ml;
- 6) Adicionar igual volume de clorofórmio/álcool isoamílico (24:1) a cada tubo, agitar em *vórtex* e centrifugar a 14000rpm durante 15 minutos a 4°C;
- 7) Transferir a fase aquosa para novos tubos de 2ml, adicionar o dobro do volume de álcool absoluto gelado, misturar por inversão e colocar os tubos a -20°C (o/n);
- 8) Centrifugar os tubos a 14000rpm durante 15 minutos a 4°C, descartar o sobrenadante e secar o *pellet* a 37°C num banho seco;
- 9) Ressuspender o *pellet* num volume de 20-60µl de TE pH8 (*Citogene*, Portugal), conforme o *pellet* obtido;
- 10) Juntar o volume dos dois tubos e guardar a 4°C até o DNA ser quantificado.

Tampão de lise

- 1) Pesar 5,85g de cloreto de sódio (*Merck*) e 5g de SDS (*Sigma*);
- 2) Adicionar 5ml de Tris-HCl 2M (pH8) e 50ml de EDTA 0,5M (pH8);
- 3) Agitar em placa de agitação e perfazer para um volume final de 1L.

1.2 Protocolo de extracção de DNA a partir de sangue periférico utilizando o *kit* comercial *Citogene® Blood Kit – Genomic DNA Purification Kit* (*Citomed*, Portugal).

- 1) Transferir 9ml de RBC (*Red-blood cells*) *Lysis* para um tubo tipo *falcon* de 15ml, adicionar cerca de 3ml de sangue periférico com o auxílio de uma pipeta de *Pasteur* e homogeneizar;
- 2) Incubar durante 5 minutos à temperatura ambiente e centrifugar a 2000g durante 2 minutos à temperatura ambiente;
- 3) Descartar o sobrenadante para a lixívia e soltar o *pellet*;
- 4) Adicionar 4ml de RBC *Lysis* e repetir os pontos 2 e 3;

- 5) Acrescentar 3ml de *Cell Lysis* e ressuspender;
Nota: Caso haja a formação de agregados celulares agitar em *vórtex* e colocar em banho-maria a 37°C durante algumas horas.
- 6) Acrescentar 15µl de RNAse a cada tubo, misturar por inversão e colocar a 37°C durante pelo menos 15 minutos, seguido de 2 a 3 minutos em gelo;
- 7) Acrescentar 1ml de *Protein Precipitation*, agitar em *vórtex* durante 20 a 30 segundos e centrifugar a 2000g durante 5 minutos;
- 8) Transferir o sobrenadante para novos tubos tipo *falcon* de 15ml e centrifugar a 2000g durante 5 minutos;
- 9) Transferir o sobrenadante para novos tubos tipo *falcon* de 15ml contendo 3ml de isopropanol 100% (*Merck*), agitar suavemente e recolher o DNA com o auxílio de uma pipeta para um tubo tipo *ependorf* de 1,5ml contendo 1ml de etanol a 70% (v/v);
- 10) Centrifugar o tubo a 14000rpm durante 7 minutos a 4°C, descartar o sobrenadante e secar o *pellet* de DNA num banho seco a 37°C;
- 11) Adicionar *DNA hidration* conforme o tamanho do *pellet* (50-200µl) e colocar os tubos num banho-maria a 65°C durante cerca de 30 minutos;
- 12) Deixar as amostras arrefecer à temperatura ambiente e em seguida guarda-las a 4°C até serem quantificadas.

Solução etanol 70% (v/v)

Numa proveta de 100ml medir 70ml de etanol (*Merck*) e prefazer para um volume final de 100ml com água bidestilada.

1.3 Protocolo de extracção de DNA a partir de tecido congelado utilizando o kit comercial *Citogene*[®] *Tissue Kit – Genomic DNA Purification Kit (Citomed, Portugal)*.

- 1) Retirar as amostras de tecido de azoto líquido, descongela-las em gelo e cortá-las com rapidez;
- 2) Colocar cerca de 150mg de tecido num tubo tipo *ependorf* de 1,5ml, adicionar 5ml de *Cell Lysis* e 25µl de Proteinase K (20mg/ml). Misturar por inversão (cerca de 25 vezes) e incubar a 55°C (o/n);
- 3) Adicionar 25µl de solução de RNAse ao lisado celular. Misturar, invertendo cerca de 25 vezes e incubar a 37°C durante 15-60 minutos.
- 4) Deixar as amostras arrefecer à temperatura ambiente, adicionar 1,5ml de solução *Protein Precipitation*, agitar em *vórtex* (cerca de 20 segundos) e centrifugar durante 10 minutos a 2000g;
- 5) Transferir o sobrenadante para um tubo de 15ml contendo 5ml de isopropanol 100% (*Merck*), agitar suavemente e recolher o DNA com o auxílio de uma pipeta para um tubo tipo *ependorf* de 1,5ml contendo 1ml de etanol a 70% (v/v);
- 6) Continuar o protocolo de acordo com o descrito no ponto 1.2 (passos 10 a 12).

1.4 Protocolo de extracção de RNA a partir de sangue periférico com Tri Reagent® (Ambion, Life Technologies).

- 1) Num tubo tipo *falcon* de 50ml, adicionar o volume de sangue de 3 tubos de hemograma, perfazer para um volume final de 45ml com solução de lise RBC 1x e agitar;
- 2) Colocar os tubos em gelo durante 30 minutos;
- 3) Centrifugar 10 minutos a 1200rpm;
- 4) Descartar o sobrenadante para um recipiente com lixívia e soltar o *pellet*;
- 5) Perfazer para um volume final de 45ml com PBS 1x, agitar e repetir os pontos 3 e 4;
- 6) De acordo com o tamanho do *pellet* obtido, adicionar 1 - 3ml de PBS 1x () e distribuir alíquotas de 1ml da solução por tubos tipo *ependorf* de 1,5ml;
- 7) Realizar uma centrifugação curta a 5000rpm e descartar o sobrenadante para um recipiente com lixívia;
- 8) Numa *hotte*, adicionar 1ml de *Tri Reagent*® a cada tubo e homogeneizar com auxílio de uma micropipeta;
Nota: As alíquotas que não são extraídas devem ser armazenadas a -80°C.
- 9) Adicionar 200µl de clorofórmio a cada tubo e agitar fortemente por 20 segundos (manualmente);
- 10) Deixar repousar à temperatura ambiente entre 2 a 3 minutos e centrifugar a 4°C a 14000rpm por 20 minutos;
- 11) Transferir a fase aquosa para um novo tubo, adicionar 500µl de isopropanol (*Merck*) e deixar incubar a -20°C (o/n);
- 12) Centrifugar a 4°C a 14000rpm por 20 minutos;
- 13) Desprezar o sobrenadante, adicionar ao *pellet* 1ml de etanol a 75% gelado (preparado com DEPC), agitar em *vórtex* e centrifugar 5 minutos a 8000rpm a 4°C;
- 14) Desprezar sobrenadante e secar o *pellet* ao ar;
- 15) Ressuspender cada *pellet* em 20µl de água com DEPC e deixar 30 minutos a solubilizar antes de quantificar;

Solução de lise – RBC 1x

Num proveta de 1L, adicionar 100ml de RBC 10x (*Biolegend*) e perfazer para um volume final de 1L com água bidestilada.

Solução PBS 1x

Num proveta de 1L, adicionar 100ml de PBS 10x (*Lonza*) e perfazer para um volume final de 1L com água bidestilada.

Anexo 2 Protocolos para efectuar as soluções utilizadas electroforese em gel de agarose.

Tampão de electroforese – TBE 1x

- 1) Adicionar 200ml de TBE 10x (0,89M Tris Borato pH 8.3 + 20mM Na₂ EDTA – *National Diagnostics*) e perfazer para um volume final de 2000ml, em água bidestilada.

Gel de agarose 0,8% (p/v)

- 1) Pesar 1,2g de agarose (*Seakem*[®] LE Agarose, Lonza, E.U.A.) num balão *erlenmeyer* de 500ml e adicionar 150ml de solução tamponada TBE 1x;
- 2) Dissolver totalmente no microondas durante 3 minutos à potência máxima;
- 3) Deixar a solução arrefecer ligeiramente e adicionar 7,5µl de brometo de etídeo (10mg/ml, *Invitrogen*), agitar até homogeneizar;
- 4) Verter a solução para um molde contendo os pentes necessários para o número de poços pretendidos e deixar o gel solidificar completamente durante, aproximadamente, 20 minutos.

Gel de agarose 2% (p/v)

- 1) Pesar 5g de agarose (*Seakem*[®] LE Agarose, Lonza, E.U.A.) num balão *erlenmeyer* de 500ml e adicionar 250ml de solução tamponada TBE 1x;
- 2) Dissolver totalmente no microondas durante 3 minutos à potência máxima;
- 3) Deixar a solução arrefecer ligeiramente e adicionar 12,5µl de brometo de etídeo (10mg/ml, *Invitrogen*) e agitar até homogeneizar;
- 4) Verter a solução para um molde contendo os pentes necessários para o número de poços pretendidos e deixar o gel solidificar completamente durante, aproximadamente, 20 minutos.

Orange G 5x

- 1) Pesar 0,125g de *Orange G* (*Sigma-Aldrich*) e 12,5g de *Ficoll* (*Sigma-Aldrich*), perfazer o volume para 50ml com água bidestilada;
- 2) Agitar em placa de agitação com temperatura para dissolver e distribuir a solução em tubos tipo *ependorf* de 2ml, armazenando a -20°C.

Orange G 1x

- 1) Num tubo de 1,5ml, adicionar 400µl de *Orange G 5x* a 1ml de água bidestilada, agitar em *vórtex* e armazenar a 4°C.

Marcador *GeneRuler 50 bp DNA Ladder* (*Thermo Scientific*) e marcado *Lambda/HindIII* (*Fermentas*)

- 1) Adicionar 0,5µl de marcador; 2,µl de *Orange G 5x* (*Sigma-Aldrich*) e 7µl de água bidestilada.
- 2) Agitar em *vórtex* e armazenar a 4°C.

Anexo 3 Sequências e características dos primers desenhados para a análise de mutações nos exões 1, 4, 13, 14, 15, 16 e 17 do gene *TCF7L2*.

Exão	Sequência	%GC	Tamanho do fragmento (pb)	Método *
1	F - CTTCCAAAATTGCTGCTGGTG	47,6	241	SA
	R - CCCGAGGGGCTTTTCCTA	61,1		
4	F -(a)GAACGCTTTGATTTGGTTTC	40	184	DGGE
	R - GCTTCAGAATCTCTTGCGT	47,4		
	F - GTAGTTTGTGACATAAGCAG	40		
	R - GACAATTCAAGCTTCAGAATC	40		
13	F-(a)GAAGGCTTTGTATAATTTGTTC	31,8	175	DGGE
	R - CTATTTTAATGACTCAAAATGAC	26,1		DGGE/SA
	F - GGCTTTTACAAACAAACAAAAC	31,8		SA
14	F - ACGTTCTCCTCCTCTGCT	55,6	198	DGGE/SA
	R - (b) ACCAGCACAGCGAATCTG	55,6		DGGE
	R - TAGACCAGCTGCAAACCAG	52,6		SA
15	R - GCTGAGATTTACATCCAACCTG	45,5	73	SA
	R - GCTTTGATTTCCCAAAGGC	45		
16	F - GCTCCTGTCTCATTTCCCTTC	47,6	146	SA
	R - GAGGCACTACAAACAGGGAAG	55		
17A	F - CTCTCTCCCTTGGCATCTGTGC	59,1	176	GS
	R - AGGGGAGCCTAGCAGGTTC	63,2		
17B	F - GCTTACTAGATTCGCCTCCCC	57,1	139	SA
	R - GGAGGCATCATGGACAGGTG	60		
17C	F - GTCACAGACTGAGCAGACCCAG	59,1	173	SA
	R - CAGGCTGCAAAGCGGCTG	66,7		
17D	F - TCCGCCCTCTGTCCCAAC	66,7	194	SA
	R - AAGCAGCGGGGTTACAGAC	63,2		

*Método utilizado para a análise mutacional: SA – Sequenciação Automática; DGGE – *Denaturing Gradient Gel Electrophoresis*; GS – Análise de fragmentos por electroforese capilar utilizando o software *GeneScan (Applied Biosystems)*. (a)C-clamp: CCCGCCCCGGCCCGCCCCGCCCCCGCC CCCCCTCCCGCCCCGCCCCCTGGCGCCCCGC (b)C-clamp: CCCACGCCACCCGACGCCCCAGC CCGACCCCCCGCGCCCCGGCGCCCCCGC

Anexo 4 Condições e programas utilizadas na amplificação dos vários exões do gene *TCF7L2*, analisados no presente estudo.

Tabela 3 – Condições utilizadas na amplificação dos vários exões em estudo do gene *TCF7L2*.

Exão	Temperatura de hibridação (°C)	[MgCl ₂] (mM)	[Primer] (pmol/μl)	[Polimerase] (U/μl)	Reagentes	Ciclos	Termociclador
1	64	4	0,12	0,4	AB	40	Veriti
4 ^(a)	59	5,5	0,12	0,4	AB	40	Veriti
4 ^(b)	59	5	0,12	0,4	AB	40	Veriti
13 ^{(a)(b)}	58,5	5,5	0,12	0,4	AB	40	Veriti
14	59	5,5	0,12	0,4	AB	40	Veriti
15	60	5	0,12	0,4	B	35	Biometra
16	62	2,5	0,12	0,4	AB	40	Veriti
17A	66	4	0,12	0,4	B	35	Veriti
17B	67	2	0,12	0,4	AB	40	Veriti
17C	69	2,5	0,12	0,4	AB	40	Veriti
17D	65	2,5	0,12	0,4	AB	40	Veriti

Legenda: (a) Condições utilizadas para a amplificação do exão para DGGE; (b) Condições utilizadas para a amplificação do exão para Sequenciação automática; AB – *Applied Biosystems, Life technologies*; B – *Bioline, Life Science*

Tabela 2 – Programa de amplificação por PCR utilizando o *kit Biotaq (Bioline)*.

Etapa	Temperatura (°C)	Tempo	Ciclos
Desnaturação inicial	95	5 minutos	1
Desnaturação	94	50 segundos	
Hibridação	Variável	30 segundos	35
Extensão	72	50 segundos	
Extensão final	72	7 minutos	1
Pausa	15	∞	-

Tabela 3 - Programa de amplificação por PCR utilizando o *kit Amplitaq[®] Gold (Applied Biosystems)*.

Etapa	Temperatura (°C)	Tempo	Ciclos
Desnaturação inicial	95	5 minutos	1
Desnaturação	94	50 segundos	
Hibridação	Variável	30 segundos	40
Extensão	70	50 segundos	
Extensão final	70	7 minutos	1
Pausa	15	∞	-

Tabela 4 - Programa utilizado na amplificação por PCR dos exões para DGGE. (*temperatura variável de acordo com a temperatura de hibridação).

Etapa	Temperatura (°C)	Tempo	Ciclos
Desnaturação inicial	95	5 minutos	1
Desnaturação	94	50 segundos	
Hibridação	Variável	30 segundos	40
Extensão	70	50 segundos	
Extensão final	70	7 minutos	1
	60*		
	50*		
	40	3 minutos cada	-
	30		
	20		
Pausa	15°C	∞	-

Anexo 5 – Protocolo de precipitação e purificação do DNA Etanol / EDTA / Acetato de sódio – *BigDye*[®] Terminator v1.1 Cycle Sequencing Kit (Applied Biosystems) após a reacção de sequenciação.

- 1) Preparar uma mistura reaccional num tubo tipo *ependorf* de 2ml com 2µl de EDTA (125mM), 2µl de acetato de sódio (3M) e 50µl de etanol absoluto por cada amostra;
- 2) Agitar em *vórtex* e realizar um *spin down* à velocidade máxima de 5000rpm;
- 3) Distribuir 54µl do sobrenadante da mistura por tubos de 1,5ml;
- 4) Adicionar o volume total da reacção de sequenciação e homogeneizar em *vórtex*;
- 5) Incubar durante 30 minutos à temperatura ambiente;
- 6) Centrifugar a 14000rpm durante 40 minutos a 4°C;
- 7) Retirar completamente o sobrenadante;
- 8) Adicionar 100µl de etanol 70% (v/v) e agitar ligeiramente em *vórtex*;
- 9) Centrifugar a 14000rpm durante 15 minutos a 4°C;
- 10) Retirar completamente o sobrenadante;
- 11) Secar o *pellet* a 37°C durante 10 minutos

Anexo 6 Protocolos para efectuar as soluções utilizadas na técnica de DGGE.

Solução de Acrilamida 40% (v/v)

- 1) Num gobelé de 2000ml coberto com papel de alumínio, pesar 380g de acrilamida (*Life Technologies*) e 20g de bis-acrilamida (*Sigma*), dissolver em água bidestilada e perfazer para um volume final de 1L.
- 2) Adicionar 50g de resina (*Bio-rad*) sob agitação e desionizar durante 1 hora;
- 3) Filtrar a solução com papel de filtro para um frasco escuro e armazenar a 4°C.

Formamida

- 1) Num gobelé de 2000ml, pesar 50g de resina (*Bio-rad*) e adicionar 1L de formamida (*Merck*) sob agitação e desionizar durante 1 hora;
- 2) Filtrar a solução com papel de filtro para um frasco escuro e armazenar a 4°C.

Tampão TAE 50x

- 1) Num gobelé de 2000ml, pesar 242,2g de tris-base (*Sigma*), 82,03g de acetato de sódio (*Merck*) e 18,61g de EDTA (*Life Technologies*);
- 2) Ajustar o pH a 7,5 com ácido acético (*Merck*) e perfazer para um volume final de 1L com água bidestilada;
- 3) Armazenar a solução num frasco escuro à temperatura ambiente.

Solução de Acrilamida 0% (v/v)

- 1) Adicionar 162,5ml de acrilamida 40% (v/v), 20ml de tampão TAE 50x e perfazer para um volume final de 1L com água bidestilada;
- 2) Armazenar num frasco escuro a 4°C.

Solução de Acrilamida 80% (v/v)

- 1) Num gobelé de 2000ml envolvido em papel de alumínio, pesar 336g de ureia (*Promega*), adicionar 162,5ml de acrilamida 40% (v/v), 20ml TAE 50x e 320ml de formamida desionizada;
- 2) Dissolver em água bidestilada sob agitação e temperatura e perfazer para o volume de 1L;
- 3) Filtrar a solução com papel de filtro para um frasco escuro e armazenar a 4°C.

APS

- 1) Num *gobelé* de 50ml, pesar 2,5g de persulfato de amónio (*Merck*) dissolver em água bidestilada sob agitação;
- 2) Perfazer para um volume final de 25ml com água bidestilada, aliquotar e armazenar a solução a -20°C.

Solução de Deposição DGGE

- 1) Num gobelé de 100ml, pesar 0,1g de azul bromofenol 0,2% (p/v), adicionar 25 ml de glicerol 99% (*Sigma*), 1ml de TAE 50x, perfazer a solução para um volume final de 50ml e homogeneizar sob agitação;
- 2) Armazenar num frasco escuro a 4°C.

Tampão TAE1x

Diluir 40ml de tampão TAE50x em 2000ml de água destilada.

Anexo 7 Características e condições de PCR dos marcadores de microssatélites

Tabela 1 - Sequência dos *primers* e tamanho do fragmento amplificado relativos aos marcadores de microssatélites: D10S1269, D10S168 e D10S562.

Marcador de Microssatélites	Sequência 5' – 3'	Tamanho (pb)
D10S1269	F – CTCAGATCGCGGAGTAGA * R – TGCTGTATTTATTACCAGTGCTTG	222-232
D10S168	F – CATGGCACTAATAGAGTTAAC ** R – TTCACTTGGGATGGAGGCA	159-165
D10S562	F – CCTGGCAGATGGAGGTT ** R – TCGGAGTGCTTCCTTAAAATAC	185-205

* *Primer forward* marcado com o fluorocromo FAM™ (6-Carboxifluorescein) – cor azul;

** *Primer forward* marcado com o fluorocromo TET™ (5'-Tetrachloro-Fluorescein Phosphoramidite) – cor verde.

Tabela 2 - Localização dos marcadores de microssatélites: D10S1269, D10S168 e D10S562.

Marcador de microssatélites	UniSTS	Posição (pb)
D10S1269	36608	114,521,464 - 114,521,689
D10S168	28320	114,745,515 - 114,745,679
D10S562	33530	116,640,361 - 116,640,545

Tabela 3 - Condições utilizadas na amplificação por PCR dos marcadores de microssatélites D10S1269, D10S168 e D10S562.

Marcador de Microssatélites	Temperatura de hibridação (°C)	[MgCl ₂] (mM)	[Primer] (pmol/μl)	[Polimerase] (U/μl)	Termociclador
D10S1269	59,5	5	0,12	0,4	Veriti
D10S168	61	7	0,12	0,4	Veriti
D10S562	61	4	0,12	0,4	Veriti

Tabela 4 - Programa utilizado para a amplificação por PCR dos marcadores de microssatélites utilizando o *kit* Biotaq (*Bioline*)

Etapa	Temperatura (°C)	Tempo	Ciclos
Desnaturação inicial	95	5 minutos	1
Desnaturação	94	1 minuto	
Hibridação	Variável	50 segundos	40
Extensão	72	1 minuto 30 segundos	
Extensão final	72	10 minutos	1
Pausa	15	∞	-

Anexo 8 Características e condições de qPCR

Tabela 1 Programa utilizado nas reacções de qPCR no aparelho *LightCycler®Plus 480 Multiwell Plates* (Roche).

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

Tabela 2 Sequência dos *primers* utilizados para análise de *copy-number* do gene *TCF7L2* por qPCR utilizando *SYBR Green*.

Região a amplificar	Sequência
Exão 1	F- CTTCAAAGACGAGGGCGAAC
	R - TAAATCCCTCTCTGCCGAGG
Exão 4	F - CTACAGCTCCATTTTCAGTCCG
	R - CCTGAACTGCAATCTGGTGTTT
Intrão 4	F- GTAAAACGGGGTGGCACAAT
	R- CAGAAACAACCCTCTGGCTGT
Exão 14	F - TCCTAAGAAATGCCGAGCGC
	R - GCAAACCAGCACAGCGAATC

Tabela 3 Programa utilizado nas reacções de q PCR no aparelho *7900HT Sequence Detection System* (Applied Biosystems).

Etapa	Temperatura (°C)	Tempo	Ciclos
Incubação	50	2 minutos	-
Activação da polimerase	95	10 minutos	-
Desnaturação	95	15 segundos	40
Hibridação e extensão	60	1 minuto	

Anexo 9 Resultados da análise de mutações nos exões 1, 4, 13, 14, 15, 16 e 17 do gene TCF7L2.

DNA	Exão 1	Exão 4	Exão 13	Exão 14	Exão 15	Exão 16	Exão 17			
							Frag 17A	Frag 17B	Frag 17C	Frag 17D
155	-	-	N	N	-	-	-	-	-	-
249	-	N	N	-	-	N	-	N	*	*
253	-	N	N	N	-	N	-	N	*	*
295	-	N	N	-	-	-	-	-	-	-
297	-	N	N	N	-	N	-	-	-	-
301	*	-	N	N	-	N	-	N	*	*
305	-	N	N	N	-	N	-	-	-	-
319	-		N	N	-	N	-	N	-	-
337	-		N	-	-	*	-	*	*	*
383	-	N	-	N	-	N	-	N	-	-
387	-	N	-	N	-	-	-	-	-	-
389	-	N	N	N	-	N	-	N	-	-
489	-	N	N	-	-	-	-	-	-	
525	*	N	N	N	-	N	-	N	*	N
537	*	N	N	-	-	N	-	N	*	N
585	-	N	N	N	-	N	-	N	*	N
593	-	N	N	-	-	N	-	N	-	-
595	-	N	N	-	-	mut	-	N	*	mut
603	*	N	N	N	N	N	N	N	N	*
611	*	N	N	N	mut	N	N	N	N	N
645	-	N	N	-	-	N	-	N	N	N
785	-	N	N	-	N	N	N	N	N	N
815	-	N	N	-	-	N	-	-	-	-
817	-	N	N	-	-	N	-	N	N	N
1083	-	N	N	-	-	N	-	N	N	N
1233	-	N	N	-	-	N	-	N	N	N
1421	-	N	N	-	-	N	-	N	N	N
1529	-	N	-	-	-	N	-	N	N	N
1639	-	N	N	N	N	N	N	N	N	N
1653	N	N	N	N	-	N	-	N	N	N
1705	*	N	N	N	N	N	N	N	N	N
1707	*	N	N	N	N	N	N	N	N	N
1877	*	N	-	-	N	N	N	N	*	*
1899	-	-	-	-	N	N	N	-	-	-
1915	N	N	N	*	N	N	N	N	N	N
1969	N	-	N	-	N	N	N	N	*	N
1971	N	-	N	-	N	N	N	N	*	*
1977	N	N	N	-	mut	N	N	-	N	N
1983	*	-	N	-	-	N	-	-	-	N
1991	N	N	N	-	N	N	N	N	N	N
2011	*	N	N	-	N	N	-	N	N	N
2031	N	N	N	N	N	N	N	-	N	N
1929	N	N	N	N	*	N	N	N	N	*

Continuação do anexo 9

DNA	Exão 1	Exão 4	Exão 13	Exão 14	Exão 15	Exão 16	Exão 17			
							Frag 17A	Frag 17B	Frag 17C	Frag 17D
1935	N	N	-	-	N	N	N	N	N	N
1937	*	-	-	-	-	N	*	N	N	N
1953	*	N	N	N	-	N	N	N	N	N
1987	N	N	N	-		N	N	N	mut	N
2043	N	N	N	-	-	N	N	N	N	mut
2059	*	N	N	-	N	N	N	N	N	N
2115	N	N	N	N	N	N	N	N	N	N
2155	N	N	N	N	-	N	-	N	N	N
2159	N	N	N	N	-	N	-	N	N	N
2187	N	-	-	N	-	N	-	N	N	N
2197	N	N	N	N	N	N	N	N	N	N
2209	N	N	N	N	-	N	N	N	N	N
2241	N	N	N	N	-	N	N	N	N	N
353	-	-	-	-	-	-	N	-	N	-
455	-	-	-	-	-	-	N	N	N	-
579	-	-	-	-	-	-	mut	N	N	-
621	-	-	-	-	-	-	N	N	N	-
725	-	-	-	-	-	-	mut	-	N	-
1019	-	-	-	-	-	-	mut	N	N	-
1027	-	-	-	-	-	-	mut	N	N	-
1265	-	-	-	-	-	-	N	-	N	-
1911	N	N	mut	N	-	N	mut	N	N	N
1923	-	-	-	-	-	N	mut	N	N	N
1973	N	N	N	N	-	N	mut	N	N	N
2229	N	N	N	-	N	N	mut	*	N	*

Legenda: Frag: fragmento; N: normal (sem mutação); mut: com mutação; * não foi possível amplificar por PCR; - não foi feita análise.