



**Ministério da Saúde
Instituto Nacional de Câncer
Coordenação de Pós-graduação**

**INSTITUTO NACIONAL DE CÂNCER
Pós-Graduação em Oncologia**

Bárbara Luísa Soares

**Detecção de Variantes nos genes *MSH2*, *MSH6*,
MLH1 e *PMS2* utilizando Sequenciamento de Nova
Geração**

Orientador: Prof. Dr. Miguel Ângelo Martins Moreira

**RIO DE JANEIRO
Fevereiro de 2017**



**Ministério da Saúde
Instituto Nacional de Câncer
Coordenação de Pós-graduação**

**INSTITUTO NACIONAL DE CÂNCER
Pós-Graduação em Oncologia**

Bárbara Luísa Soares

**Detecção de Variantes nos genes *MSH2*, *MSH6*,
MLH1 e *PMS2* utilizando Sequenciamento de Nova
Geração**

Dissertação de mestrado apresentada ao
Programa de Pós-Graduação em Oncologia do
Instituto Nacional de Câncer José de
Alencar Gomes da Silva como parte
dos requisitos para obtenção do grau
de Mestre em Oncologia

Orientador: Prof. Dr. Miguel Ângelo Martins Moreira

**RIO DE JANEIRO
Fevereiro de 2017**

S676d

Soares, Bárbara Luísa.

Detecção de variantes nos genes MSH2, MSH6, MLH1 e PMS2 utilizando Sequenciamento de nova geração. / Bárbara Luísa Soares. - Rio de Janeiro, 2017.

133 f.

Orientador: Miguel Ângelo Martins Moreira.

Dissertação (Mestrado em Oncologia)– Instituto Nacional de Câncer José Alencar Gomes da Silva, Programa de Pós-Graduação em Oncologia, 2016.

1. Neoplasias Colorretais Hereditárias sem Polipose. 2.Reparo do DNA. 3. Biomarcadores Tumorais. I. Moreira, Miguel Ângelo Martins. II.Título.

CDD 572.865



**Ministério da Saúde
Instituto Nacional de Câncer
Coordenação de Pós-graduação**

INSTITUTO NACIONAL DE CÂNCER

Detecção de Variantes nos genes *MSH2*, *MSH6*, *MLH1* e *PMS2* utilizando Sequenciamento de Nova Geração

RESUMO

Dissertação de mestrado

Bárbara Luísa Soares

A Síndrome de Lynch (SL) é uma herança autossômica dominante de alta penetrância, que atinge aproximadamente 3% dos pacientes com câncer colorretal. Os indivíduos acometidos apresentam mutações germinativas em um dos genes responsáveis pelo reparo de DNA por mal pareamento (Mismatch Repair- MMR): *MSH2*, *MLH1*, *MSH6* ou *PMS2*. Por serem genes longos, a identificação de mutações nestes genes se torna demorada e com custo elevado. Desta maneira, o projeto teve como objetivo padronizar uma estratégia de identificação de variantes que permitisse reduzir os custos e tempo de rastreamento molecular dos genes MMR. Para isso, foram padronizadas 16 reações de PCRs-*multiplex* para os genes *MSH2*, *MLH1* e *MSH6* e 5 reações de PCRs de longo alcance para o gene *PMS2*. Estes produtos foram sequenciados utilizando a tecnologia de Nova Geração (NGS), através do equipamento *HiSeq2500*. A estratégia foi validada através do sequenciamento pelo método de Sanger dos genes de MMR em 66 pacientes, provenientes de quatro centros distintos do Brasil (INCA-RJ, HCPA- RS, HJUBB- PA e ACCAM- SP), e que preencheram os critérios de Bethesda para SL. As profundidades médias de cobertura obtidas para os genes *MSH2*, *MSH6*, *MLH1* e *PMS2* foram de 7.988, 36.313, 11.899 e 4.772 vezes, respectivamente. Foram identificadas 98 alterações em éxons e íntrons dos quatro genes, sendo que 25 eram patogênicas ou VUS (7 em *MSH2*, 5 em *MSH6*, 12 em *MLH1* e 1 em *PMS2*) e foram encontradas em 32 pacientes. A estratégia padronizada foi eficaz na identificação de variantes dos genes MMR, permitiu reduzir em três vezes o número de reações de PCR realizadas por amostra e em 2,15 vezes o custo de rastreamento molecular para os quatro genes MMR. Dos fragmentos sequenciados, 3,1% apresentaram profundidade média de cobertura <30X e não foram eficientes na detecção de sítios variáveis nos genes MMR.



Ministério da Saúde
Instituto Nacional de Câncer
Coordenação de Pós-graduação
INSTITUTO NACIONAL DE CÂNCER

Deteção de Variantes nos genes *MSH2*, *MSH6*, *MLH1* e *PMS2* utilizando Sequenciamento de Nova Geração

ABSTRACT

Dissertação de mestrado

Bárbara Luísa Soares

Lynch Syndrome (LS) is an autosomal dominant inheritance of high penetrance that affects approximately 3% of the cases of colorectal cancer. Individuals affected with this syndrome inherit germline mutations in one of the genes responsible for the Mismatch Repair: *MSH2*, *MLH1*, *MSH6* or *PMS2*. These large numbers of genes turns mutations' identification time consuming and costly. The project aimed to establish a molecular screening method in order to optimize the cost-effectiveness for screening these genes. We performed 16 Multiplex PCRs for *MSH2*, *MSH6* and *MLH1* genes and 5 Long-Range PCRs for *PMS2* gene, coupled with Next Generation Sequencing (NGS) technologies. Our strategy was validated by the screening of sixty-six patients who filled Bethesda criteria for LS from different hospitals of Brazil (INCA-RJ, HCPA-RS, HJUBB-PA and ACCAM-SP). The depth of coverage for *MSH2*, *MSH6*, *MLH1* and *PMS2* genes was 7.988, 36.313, 11.899 and 4.772 times, respectively. Ninety-eight alterations were found in exons and introns regions for the four MMR genes. From this, 25 were pathogenic or VUS found in 32 patients (7 in *MSH2*, 5 in *MSH6*, 12 in *MLH1* e 1 in *PMS2*). The strategy was efficient to identify genetic changes at the MMR genes, since allowed to reduce to 3 times the number of PCR reactions performed per patient and to reduce in 2,15 times the costs to molecular screening of the four MMR genes. Approximately 3% of the amplicons sequenced had a depth of coverage <30X and were not efficient in variation detection at MMR genes.

Este trabalho foi realizado na Divisão de Genética da
Coordenação de Pesquisa do Instituto Nacional de
Câncer José de Alencar Gomes da Silva, sob
Orientação do Dr. Miguel Ângelo, e contou com apoio
Financeiro da FAPERJ, CNPq e Ministério da Saúde.

AGRADECIMENTOS

Gostaria de começar agradecendo às primeiras pessoas que me deram oportunidade de ingressar neste curso de Mestrado: minha família. Aos meus pais, pelo apoio incondicional, por acreditarem na minha palavra e no meu potencial, e por colocarem à minha formação em primeiro lugar, e ao meu irmão Eduardo, pela amizade e apoio em todos os momentos desafiadores que vivi durante esta Jornada.

Ao meu orientador Miguel Moreira, pelos ensinamentos e pela presença constante durante estes dois anos. Foram tempos de muito aprendizado, que contribuíram de maneira muito positiva para o meu crescimento pessoal, emocional e profissional. Sua tranquilidade e paciência para resolver os problemas que surgiram, serviram de exemplo para mim. Agradeço pela confiança depositada na minha palavra.

Não poderia deixar de agradecer às pessoas que estiveram comigo todos os dias durante estes dois anos, meus grandes amigos do Laboratório de Genética. Ao Renan, pela amizade e colaboração nos projetos. Ao Ayslan, por ser meu colaborador principal e por me ensinar tantas coisas no maluco mundo da bioinformática. Às grandes amigas que ganhei para a vida, Carolzinha e Mirela, por compartilharem momentos difíceis, felizes e desafiadores, que vão além da bancada. À Hanninha e Mariosinha, por me inserirem no laboratório e por todo o carinho que têm comigo. À Shay, pelos almoços na copa e pela amizade que construímos. À Pricilinha e Diogo, por serem meu casal preferido no laboratório e pelas risadas diárias. À Cintinha, pelo carinho. À Bruna e Val, pela agilidade dos resultados das corridas e pelos momentos de risadas homéricas. À Morganinha, pelos bares de sexta à noite. Ao Sérgio, pelas conversas produtivas. Ao Mateuzinho, que mesmo longe, continua ativamente com a gente no grupo da Genética. Ao Régis, pela implicância, Tia Adriana, e à Mari, por serem “bichas”. E a todos os funcionários e bolsistas que contribuem para que o funcionamento do laboratório ocorra de maneira eficaz. Neste momento, também não poderia deixar de agradecer aos outros chefes da Genética, Héctor, Cibele e Marcelo, pela oportunidade que me ofereceram de participar do Laboratório e por estarem sempre atentos às nossas necessidades.

Gostaria também de agradecer aos meus avôs, tios, tias, primos e primas, que mesmo de longe, acompanharam a minha batalha diária e me aconchegaram nos momentos mais difíceis durante estes anos. Vovó Joana, Vô Raimundo, Tia Tete e Dindinha vocês foram essenciais. Ao Felipe, pela paciência, carinho e amor diários e por ser parte da minha família no Rio de Janeiro. Às amigas que construí no Rio.

Aos meus amigos de Colégio, faculdade e da vida, por me proporcionarem momentos tão alegres e que contribuíram para o meu descanso mental a cada encontro que estivemos juntos, especialmente às Azamigas, Lohzinha, Chris, Naninha, Nath, Didi Arrizia e Juhzinho.

À Pós-Graduação do INCA, composta por funcionários e professores, que contribuíram diretamente para o meu desenvolvimento profissional. E à minha banca examinadora, que contribuirá de maneira decisiva para a minha formação como Mestre em Oncologia pelo Instituto Nacional de Câncer. Meu muito obrigada a todos que torceram por mim!

LISTA DE ABREVIATURAS

ACCAM: AC. Camargos Cancer Center

APC: *Adenomatous Polyposis Coli*

ATP: *Adenosine Triphosphate*

BMPR1A: *Bone Morphogenetic Protein Receptor Type 1A*

BRAF: *B-Raf proto-oncogene, serine/threonine kinase*

CDKN2A: *Cyclin Dependent Kinase Inhibitor 2A*

CFS: Síndrome Familiar de Câncer (do inglês, *Cancer Family Syndrome*)

CIMP: metilação das ilhas CpG (do inglês, *CpG Island Methylated Phenotype*)

CIN: instabilidade cromossômica (do inglês, *chromosomal instability*)

CMMR-D: Deficiência Constitucional de Mismatch Repair (do inglês, *Constitutional Mismatch Repair- Deficiency*)

DCC: *DCC Netrin 1 Receptor*

EPCAM: *Epithelial Cell Adhesion Molecule*

EXO1: *Exonuclease 1*

FAP: Síndrome Adenomatosa Familiar (do inglês, *Familial Adenomatous polyposis*)

FCCTX: Câncer Colorretal Familiar do Tipo X (do inglês, *Familial Colorectal Cancer Type X*)

HCPA: Hospital das Clínicas de Porto Alegre

HJUBB: Hospital Universitário João de Barros Barretos

HNPCC: Síndrome Hereditária de Câncer Colorretal Não Poliposo (do inglês, *Hereditary non-Polyposis Colorectal cancer*)

HTX: Domínio Hélice- Volta- Hélice (do inglês, *Helix- Turn- Helix Domain*)

IARC: Agência Internacional para Pesquisa em Câncer (do inglês, *International Agency for Research on Cancer*)

IDLs: inserção e deleção de loops

IHC: Imunohistoquímica

INCA: Instituto Nacional de Câncer

InSiGHT: Group Internacional Colaborativo em HNPCC (do inglês, *International Collaborative Group on HNPCC*)

KRAS: *K-Ras proto- oncogene*

MAP: Polipose associada a MUTYH (do inglês, *MYH-associated polyposis*)

MINT1: Amyloid Beta Precursor Protein Binding Family A Member 1

MINT2: Amyloid Beta Precursor Protein Binding Family A Member 2

MINT3: Amyloid Beta Precursor Protein Binding Family A Member 3

MLH1: *MutL Homolog 1*

MLPA: *Multiplex ligation-dependent probe amplification*

MMR: reparo de bases por mal pareamento (do inglês, *Mismatch Repair*)

MSH2: *MutS Homolog 2*

MSH3: *MutS Homolog 3*

MSH6: *MutS Homolog 6*

MSI: instabilidade de microssatélites (do inglês, *microsatellite instability*)

MSI-H: Alta Instabilidade de Microssatélites (do inglês, *High microsatellite instability*)

MSI-L: Baixa Instabilidade de Microssatélites (do inglês, *Low microsatellite instability*)

MSS: Microssatélite Estável (do inglês, *microsatellite stable*)

MUTYH: *MutY DNA glicosilase*

NGS: Sequenciamento de Nova Geração (do inglês, *Next Generation Sequencing*)

OMS: Organização Mundial da Saúde

PCNA: *Proliferating Cell Nuclear Antigen*

PIK3CA: *Phosphatidylinositol-4,5-Bisphosphate 3-Kinase Catalytic Subunit Alpha*

PJS: Síndrome *Peutz-Jeghers* (do inglês, *Peutz-Jeghers Syndrome*)

PMS2: *PMS1 homolog 2*

PMS2CL: *PMS2 C-terminal like pseudogene*

PREMM_{1,2,6}: Modelo de Predição para Mutações nos genes *MLH1*, *MSH2*, *MSH6* (do inglês, Prediction Model for *MLH1*, *MSH2*, *MSH6* Gene Mutations)

RB1: *Retinoblastoma transcriptional corepressor 1*

RPA: *Replication Protein A*

SL: Síndrome de Lynch

SMAD4: *SMAD Family member 4*

SNV: Variantes de Nucleotídeos Únicos (do inglês, *Single Nucleotide Variants*)

STK11: *Serine/Threonine Kinase 11*

TP53: *Tumor Protein p53*

VUS: Variantes de Significado desconhecido (do inglês, *Variants of Unknown Significance*)

WNT: Via de Sinalização WNT (do inglês, *Wingless- type protein*)

ÍNDICE

1) INTRODUÇÃO	1
1.1) Oncogenes, supressores tumorais e câncer	1
1.2) Incidência do câncer colorretal	2
1.3) Carcinogênese do câncer colorretal	3
1.4) Variantes do câncer colorretal	7
1.5) Síndrome de Lynch (SL)	9
1.5.1) Características clínicas	10
1.5.2) Características moleculares	11
1.5.3) Diagnóstico clínico e Aconselhamento Genético	11
1.5.4) Diagnóstico molecular	15
1.6) Sistema de reparo por mal pareamento (MMR)	19
1.6.1) Principais genes envolvidos com o MMR	21
1.7) Pseudogenes de PMS2	24
1.8) Sequenciamento de Nova geração (NGS)	26
2) JUSTIFICATIVAS	29
3) OBJETIVOS	29
3.1) Objetivos Gerais	29
3.2) Objetivos Específicos	29
4) MATERIAL E MÉTODOS	30
4.1) Descrição das amostras	30
4.2) Reações de PCR multiplex e LR-PCRs	32
4.3) Preparação das bibliotecas e NGS (HiSeq2500)	33
4.4) Análise dos dados obtidos por bioinformática	33
4.5) Modelos de predição de sítios variáveis <i>in-silico</i>	34
4.6) Sequenciamento de Sanger	35
4.7) Análise de custos	36
5) RESULTADOS	36
5.1) Reações de PCR multiplex e LR-PCRs	36
5.2) Corrida de sequenciamento no HiSeq2500	39
5.3) Resultados de Sequenciamento de Nova Geração	44
5.3.1) Variantes encontradas no gene <i>MSH2</i>	46
5.3.2) Variantes encontradas no gene <i>MSH6</i>	50
5.3.3) Variantes encontradas no gene <i>MLH1</i>	54
5.3.4) Variantes encontradas no gene <i>PMS2</i>	58
5.3.5) Análise do pseudogene <i>PMS2CL</i>	60
5.4) Rastreamento molecular das famílias com SL	60
5.5) Validação das alterações patogênicas/VUS pela metodologia de Sanger	62
5.6) Análise de custo	63
6) DISCUSSÃO	64
6.1) Metodologias de PCR multiplex e LR-PCR acopladas ao NGS	64
6.2) Limitações das metodologias de PCRs multiplexes e LR-PCRs acopladas ao NGS	70
6.3) Padronização das análises de bioinformática	71
6.4) Validação da metodologia padronizada para rastreamento dos genes MMR	72
6.4.1) Alterações genéticas identificadas no gene <i>MSH2</i>	73
6.4.2) Alterações genéticas identificadas no gene <i>MSH6</i>	75
6.4.3) Alterações genéticas identificadas no gene <i>MLH1</i>	76
6.4.4) Alterações genéticas identificadas no gene <i>PMS2</i> e análise do pseudogene <i>PMS2CL</i>	79
6.5) Novas variantes depositadas nos bancos de dados	80
7) CONCLUSÕES	81

8) REFERÊNCIAS BIBLIOGRÁFICAS	82
9) ANEXOS	98
ANEXO 1	98
ANEXO 2	102
ANEXO 3	103
ANEXO 4	105
ANEXO 5	106
ANEXO 6	107
ANEXO 7	108
ANEXO 8	109
ANEXO 9	110
ANEXO 10	111
ANEXO 11	112
ANEXO 12	113
ANEXO 13	114
ANEXO 14	116
ANEXO 15	117
ANEXO 16	119
ANEXO 17	120
ANEXO 18	121
ANEXO 19	122
ANEXO 20	123
ANEXO 21	124
ANEXO 22	125
ANEXO 23	126
ANEXO 24	127
ANEXO 25	128
ANEXO 26	129
ANEXO 27	130
ANEXO 28	132
ANEXO 29	131
ANEXO 30	133

LISTA DE FIGURAS

Figura 1.1: Estimativa da incidência dos dez tipos de câncer mais frequentes para o Brasil, no triênio 2015 – 2017, excluindo-se os casos de câncer de pele não-melanoma.....	3
Figura 1.2: Processo de carcinogênese colorretal.....	4
Figura 1.3: O fenômeno de escorregamento da DNA polimerase.	7
Figura 1.4: Incidência dos diferentes tipos de câncer colorretal.	9
Figura 1.5: Tipos histológicos de carcinoma colorretal mucinoso e células em formato de anel de sinete.	10
Figura 1.6: Perfil imunohistoquímico de proteínas MMR.	16
Figura 1.7: Estratégia de teste genético quando um paciente é clinicamente afetado e a mutação de sua família é desconhecida.	18
Figura 1.8: Mecanismo de reparo de DNA por Mismatch Repair em mamíferos.	21
Figura 1.9: Domínios presentes nas proteínas dos complexos de MutS em E.coli e em eucariotos (MSH2, MSH6, MSH3).	23
Figura 1.10: Estrutura ortogonal do dímero formado pelo complexo MutL.....	24
Figura 1.11: Localização do gene PMS2 e seus pseudogenes no cromossomo 7 do genoma.....	24
Figura 1.12: Etapas do processo de sequenciamento por síntese utilizando as plataformas da Illumina.....	28
Figura 5.1: Reações de PCRs multiplex e de Longo Alcance padronizadas para os genes (a) <i>MSH2</i> , (b) <i>MSH6</i> , (c) <i>MLH1</i> e (d) <i>PMS2</i>	38
Figura 5.2: Profundidade média de cobertura observada por éxon, profundidades mínimas e máximas de cobertura para cada um dos éxons e pacientes que apresentam cobertura média por éxon igual a 0 para o gene <i>MSH2</i>	40
Figura 5.3: Profundidade média de cobertura observada por éxon, profundidades mínimas e máximas de cobertura para cada um dos éxons e pacientes que apresentam cobertura média por éxon igual a 0 para o gene <i>MSH6</i>	41
Figura 5.4: Profundidade média de cobertura observada por éxon, profundidades mínimas e máximas de cobertura para cada um dos éxons e pacientes que apresentam cobertura média por éxon igual a 0 para o gene <i>MLH1</i>	42
Figura 5.5: Profundidade média de cobertura observada por éxon, profundidades mínimas e máximas de cobertura para cada um dos éxons e pacientes que apresentam cobertura média por éxon igual a 0 para o gene <i>PMS2</i>	43
Figura 5.6: Frequências das variantes identificadas nos genes <i>MSH2</i> , <i>MSH6</i> , <i>MLH1</i> e <i>PMS2</i> por NGS.....	45

Figura 5.7: Mapa das leituras da região cromossômica L3: 37061803-37061806 do paciente 872, que possui duas alterações em cis no éxon 11 do gene MLH1 (c.888_889delAG e c.890A>C).58

LISTA DE QUADROS

Quadro I.I: Critérios clínicos de Amsterdã I para o diagnóstico da Síndrome de Lynch.	12
Quadro I.II: Critérios clínicos de Amsterdã II para o diagnóstico de Síndrome de Lynch. ...	13
Quadro I. III: Critérios clínicos de Bethesda para o diagnóstico de Síndrome de Lynch.	13
Quadro I. IV: Critérios clínicos de Bethesda Revisados para o diagnóstico de Síndrome de Lynch, visando à identificação de indivíduos candidatos aos testes de rastreamento para a SL.	14

LISTA DE TABELAS

Tabela 1.1: Risco cumulativo de desenvolver câncer colorretal, endométrio, ovário, trato urinário e estômago aos 70 anos de idade para SL.....	11
Tabela 1.2: Identidade das sequências entre os éxons 1 a 15 de PMS2 e seus pseudogenes..	25
Tabela 4.1: Informações clínicas, imunohistoquímica, dados de MLPA e instabilidade de microssatélites dos pacientes incluídos.	31
Tabela 5.1: Reações de PCRs multiplexes e de longo alcance padronizadas para os genes de MMR.	37
Tabela 5.2: Alterações benignas e VUS encontradas nas regiões codificantes e flanqueadoras de <i>MSH2</i>	48
Tabela 5.3: Alterações patogênicas ou de significado incerto encontradas em regiões codificantes e em sítios de splicing do gene <i>MSH2</i>	49
Tabela 5.4: Predição in silico para a variante c.2078G>A presente no éxon 13 do gene <i>MSH2</i>	50
Tabela 5.5: Alterações benignas e de significado incerto (VUS) encontradas nas regiões codificantes e flanqueadoras de <i>MSH6</i>	52
Tabela 5.6: Alterações patogênicas ou de significado incerto encontradas em regiões codificantes do gene <i>MSH6</i>	53
Tabela 5.7: Predições in silico para as variantes de significado incerto encontradas nas regiões codificantes do gene <i>MSH6</i>	53
Tabela 5.8: Alterações benignas e de significado incerto (VUS) encontradas nas regiões codificantes e flanqueadoras de <i>MLH1</i>	55
Tabela 5.9: Alterações patogênicas ou de significado incerto (VUS) encontradas em regiões codificantes e em sítios de splicing do gene <i>MLH1</i>	56
Tabela 5.10: Predições in silico para as variantes missense de significado incerto encontradas nas regiões codificantes do gene <i>MLH1</i>	57
Tabela 5.11: Variantes benignas e de significado incerto (VUS) encontradas nos pacientes rastreados para o gene <i>PMS2</i>	59
Tabela 5.12: Comparação entre as profundidades médias de coberturas para os éxons 9, 11-15 de PMS2 e as regiões correspondentes do pseudogene <i>PMSCL</i>	60
Tabela 5.13: Rastreamento molecular de familiares dos pacientes que apresentaram alterações patogênicas em <i>MSH2</i> e <i>MLH1</i>	61
Tabela 5.14: Análise do custo efetividade para o rastreamento molecular dos genes MMR utilizando as metodologias de Sanger e NGS, em Reais, por Paciente.	63

Tabela 6.1: Resumo dos principais trabalhos que utilizaram o NGS para rastreamento molecular dos genes envolvidos com a SL ou síndromes hereditárias relacionadas, ressaltando os tipos de amostras utilizadas, o modo de preparação das bibliotecas, as plataformas utilizadas, o modo de preparação das bibliotecas, a plataforma utilizada para o sequenciamento, os genes rastreados, a porcentagem de leituras mapeadas e a cobertura média encontrada.....67

Tabela 6.2: Alterações encontradas neste estudo que ainda não descritas pelos bancos de dados LOVD e ClinVar.81

1) INTRODUÇÃO

1.1) Oncogenes, supressores tumorais e câncer

O câncer representa um conjunto de mais de 100 doenças que têm em comum a capacidade de crescimento das células de maneira desordenada (MINISTÉRIO DA SAÚDE, 2012). Para o ano de 2012, foram estimados 14,1 milhões de novos casos de câncer no mundo, de acordo com a *International Agency for Research on Cancer* (IARC), culminando com mais de 8,2 milhões de mortes (FERLAY *et al.*, 2015).

O câncer é uma doença provocada por alterações do padrão de expressão de proteínas relacionadas com o controle do ciclo celular. Estas proteínas são geralmente codificadas por oncogenes e genes supressores tumorais (BERGER; KNUDSON; PANDOLFI, 2015). Quando os tecidos são submetidos a ação de agentes carcinogênicos, podem vir a sofrer mutações somáticas em oncogenes e genes supressores tumorais, que fazem com que as células adquiram vantagens seletivas. Essas vantagens promovem a ativação de vias de sinalização que estimulam o crescimento descontrolado das células (FEITELSON *et al.*, 2015), culminando com a transformação progressiva de células normais em células neoplásicas malignas (HANAHAN; WEINBERG, 2011). O processo de tumorigênese é sustentado pela autossuficiência em sinais de crescimento, capacidade replicativa ilimitada, estímulo à formação de novos vasos (angiogênese), evasão dos mecanismos de morte celular programada e capacidade de invasão e metástase para outros tecidos (HANAHAN AND WEINBERG, 2010).

Os oncogenes são genes que codificam para fatores de crescimento, receptores transmembrânicos e transdutores de cascatas de sinalização, capazes de controlar a proliferação e diferenciação celular (BOLAND; GOEL, 2011). Uma vez que superativados, os produtos gênicos dos oncogenes levam a um estresse replicativo nas células, que culminam com a perda de heterozigosidade e instabilidade genômica (BARTKOVA *et al.*, 2006), principalmente nos sítios frágeis comuns de DNA (GORGOULIS *et al.*, 2005), promovendo a transformação neoplásica em células humanas (HALAZONETIS; GORGOULIS; BARTEK, 2008).

Os genes supressores tumorais têm como função impedir que haja a ativação de oncogenes responsáveis pela progressão tumoral (antioncogenes), controlar o dano de DNA nos *checkpoints* do ciclo celular e/ou realizar o reparo de DNA em situações de dano (NEGRINI; GORGOULIS; HALAZONETIS, 2010). Quando o dano de DNA não é reparado, as células são induzidas à apoptose e são impedidas de participarem dos ciclos celulares.

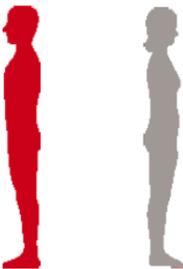
Quando os genes supressores tumorais apresentam mutações pontuais, deleções, inserções ou metilação em suas regiões promotoras, suas funções são perdidas, contribuindo para o processo de tumorigênese (BOLAND; GOEL, 2011).

Em 1971, Knudson criou um modelo de herança recessiva para o processo de carcinogênese mediado pelo gene supressor de tumor *RBI*. Este modelo ficou conhecido como modelo dos dois *hits* e foi aplicado, posteriormente, para outros genes supressores de tumor. O modelo descreve que a perda da função total de um supressor tumoral ocorre quando há inativação da cópia do gene materno e da cópia do gene paterno (KNUDSON, 1971). Em cânceres esporádicos, o primeiro evento mutacional inativa um alelo do gene supressor tumoral e o segundo evento mutacional inativa o outro alelo, por meio de recombinação, perda cromossômica ou mutação pontual, e ambos os eventos mutacionais são somáticos. Em cânceres hereditários, uma cópia do alelo mutado é herdada de um dos pais e, a outra cópia é inativada por meio de mutações somáticas nas células (KNUDSON, 1985).

1.2) Incidência do câncer colorretal

Diversos tecidos do organismo podem ser acometidos com o processo de formação de tumores, entre eles, o colón e o reto do intestino. O câncer colorretal é o terceiro tipo de câncer mais frequente no mundo e a quarta causa de mortes relacionadas ao câncer. Acomete mais de 1,3 milhões de novos casos e mata mais de 690 mil indivíduos anualmente (FERLAY *et al.*, 2015).

No Brasil, para o biênio 2016/2017, são esperados mais de 34 mil novos casos de câncer colorretal por ano. Este tipo de câncer representa o segundo mais frequente entre mulheres e o terceiro entre homens, excluindo-se os casos de câncer de pele não melanoma (Figura 1.1) (MINISTÉRIO DA SAÚDE, 2015). De acordo com dados do Instituto Nacional de Câncer (INCA), a incidência do câncer colorretal varia entre as regiões do Brasil, ocupando a segunda posição na região sudeste (24,27 mil), a terceira posição nas regiões Sul (22,35 mil) e Centro-Oeste (14,16 mil) e a quarta nas regiões Nordeste (7,05 mil) e Norte (5,34 mil).

Localização Primária	Casos	%			Localização Primária	Casos	%
Próstata	61.200	28,6%	Homens 		Mama feminina	57.960	28,1%
Traqueia, Brônquio e Pulmão	17.330	8,1%			Cólon e Reto	17.620	8,6%
Cólon e Reto	16.660	7,8%			Colo do útero	16.340	7,9%
Estômago	12.920	6,0%			Traqueia, Brônquio e Pulmão	10.890	5,3%
Cavidade Oral	11.140	5,2%			Estômago	7.600	3,7%
Esôfago	7.950	3,7%			Corpo do útero	6.950	3,4%
Bexiga	7.200	3,4%			Ovário	6.150	3,0%
Laringe	6.360	3,0%			Glândula Tireoide	5.870	2,9%
Leucemias	5.540	2,6%			Linfoma não Hodgkin	5.030	2,4%
Sistema Nervoso Central	5.440	2,5%			Sistema Nervoso Central	4.830	2,3%

*Números arredondados para múltiplos de 10.

Figura 1.1: Estimativa da incidência dos dez tipos de câncer mais frequentes para o Brasil, no triênio 2015 – 2017, excluindo-se os casos de câncer de pele não-melanoma (Fonte: MINISTÉRIO DA SAÚDE, 2015).

1.3) Carcinogênese do câncer colorretal

O câncer colorretal se desenvolve com a aquisição de alterações genéticas provocadas por fatores extrínsecos e intrínsecos. Entre os fatores extrínsecos associados, encontram-se a dieta rica em carnes vermelhas e gorduras, pobre em fibras, consumo de bebidas alcoólicas, tabagismo, obesidade e idade avançada (GILSING *et al.*, 2015) (VAN DEN BRANDT; GOLDBOHRM, 2006). Os fatores intrínsecos correspondem a herança de mutações germinativas em supressores tumorais, frequentemente encontradas em famílias que preenchem os critérios clínicos para as síndromes hereditárias de câncer colorretal. Nestes casos, o paciente herda um alelo do gene supressor já mutado, e a inativação do segundo alelo ocorra pela aquisição de mutações somáticas (SCHLUSSEL *et al.*, 2014). Existem ainda fatores epigenéticos, fatores relacionados com o microambiente do cólon e reto e fatores associados ao sistema imune, que são determinantes para o aumento de risco (HERCEG, 2007; LI; MARTIN, 2016).

Em 1989, FEARON & VOLGESTEIN, apresentaram uma das primeiras evidências sobre o processo de tumorigênese no cólon e reto do intestino. Foi sugerida que a carcinogênese colorretal se dava pela aquisição de alterações genéticas de maneira sequencial em oncogenes e genes supressores tumorais (FEARON; VOGELSTEIN, 1989). O processo de carcinogênese colorretal envolve a transformação de tecidos epiteliais hiperproliferativos em displasias nas criptas intestinais, com posterior desenvolvimento de adenomas tubulares, progressivos e vilosos e formação de carcinoma invasivo. Após a formação dos carcinomas, os tumores continuam a progredir, dando início ao processo metastatização para outros tecidos do organismo (HILL; MORSON; BUSSEY, 1978). FEARON & VOLGESTEIN destacaram mutações no gene *adenomatous polyposis coli* (APC) (com perda da região cromossômica 5q, onde está localizado o gene), a hipometilação do DNA, mutações nos

genes *K-Ras Proto-Oncogene (KRAS)*, *DCC Netrin 1 Receptor (DCC)* (perda da região cromossômica 18q) e no gene *TP53* (perda da região cromossômica 17p), como causas do desenvolvimento do câncer colorretal (FEARON; VOGELSTEIN, 1989).

Com o passar dos anos, outras alterações somáticas e aberrações cromossômicas envolvidas com a tumorigênese colorretal foram descritas, como pode ser visto na Figura 1.2. Atualmente, já se sabe que existem três mecanismos principais que levam a carcinogênese colorretal: instabilidade cromossômica (CIN), metilação das ilhas CpGs (CIMP) e instabilidade de regiões microssatélites (MSI).

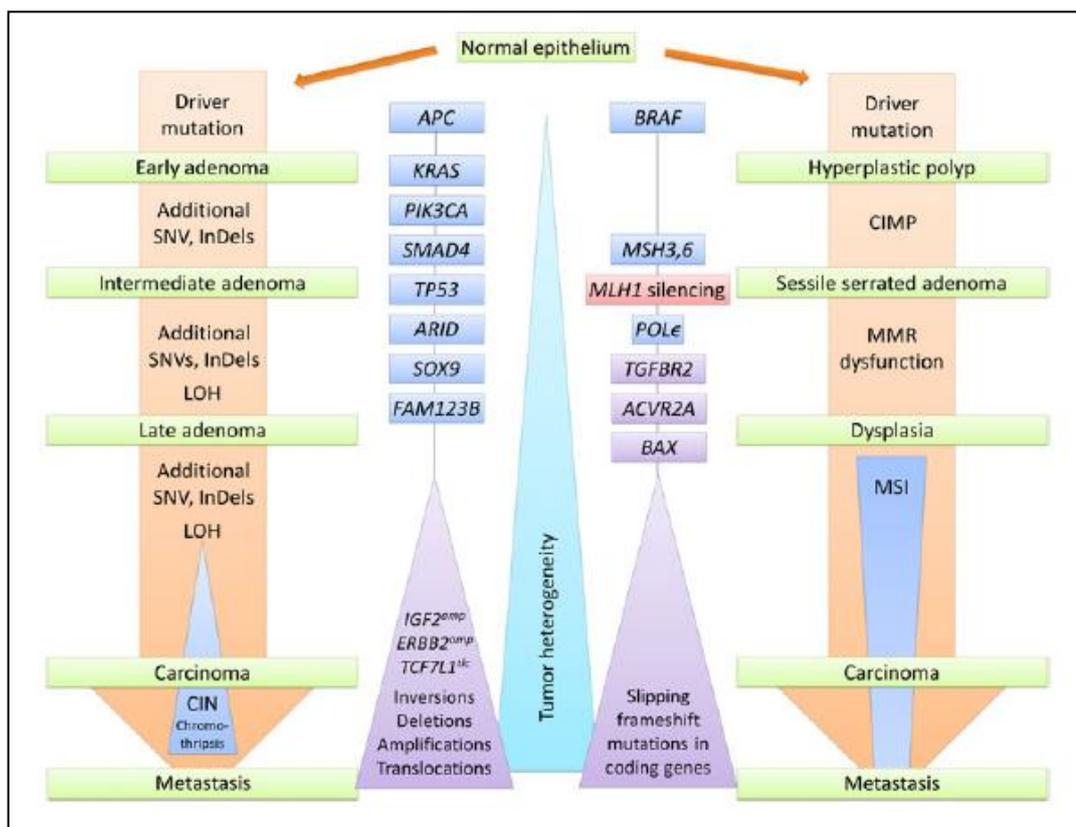


Figura 1.2: Processo de carcinogênese colorretal. O epitélio normal é acometido com mutações que dão início ao processo de tumorigênese propriamente dito, como por exemplo, mutações nos genes *APC*, *KRAS* ou *BRAF*. Mutações nestes genes provocam a formação de adenomas, que acumulam mutações pontuais e inserção ou deleção em outros genes, levando ao desenvolvimento dos carcinomas. Os carcinomas adquirem instabilidade cromossômica ou instabilidade de regiões microssatélites, sendo capazes de proliferar de maneira descontrolada e promover a metástase do tumor para outros tecidos e órgãos (Fonte: AMARO; CHIARA; PFEFFER, 2016).

A via clássica de formação de carcinomas gera instabilidade cromossômica e está associada com 65-70% dos casos de câncer colorretal esporádicos. Esta via ocorre como consequência do surgimento de mutações nos genes *APC*, *KRAS*, *Phosphatidylinositol-4,5-Bisphosphate 3-Kinase Catalytic Subunit Alpha (PIK3CA)*, *SMAD Family Member 4 (SMAD4)* e *TP53*, e posterior perda de heterozigotidade nos braços longo e curto dos cromossomos 18 e 17, respectivamente (AMARO; CHIARA; PFEFFER, 2016). Mutações nestes genes provocam defeitos na segregação cromossômica (WANG *et al.*, 2004), perda dos

telômeros, transformação oncogênica de adenomas em carcinomas (PLENTZ *et al.*, 2003) e comprometimento dos diferentes mecanismos de reparo de DNA (PRITCHARD; STEPHENS; DONNELLY, 2000; SMITH *et al.*, 2010; TARIQ; GHAS, 2016).

Além da via de Instabilidade Cromossômica, existe uma a via alternativa, que gera outros dois outros fenótipos observados na carcinogênese em câncer colorretal. Esta via é conhecida como *CpG Island Methylator Phenotype* (CIMP) e se inicia com a presença de mutações no oncogene *B-Raf Proto-Oncogene (BRAF)*, seguida da metilação das ilhas CpGs de genes envolvidos com o reparo de DNA, originando adenomas serrados sésseis, que são caracterizados pela presença de criptas em formas de dente. CIMP provoca ainda a metilação do gene *MutL Homolog 1 (MLH1)*, provocando displasia nas células e levando a formação de carcinomas com instabilidade nas regiões microssatélites (MSI) (WALTHER *et al.*, 2009).

O processo de metilação é caracterizado pela adição de grupamentos metil (CH₃) na posição 5' de anéis pirimidínicos (Guaninas e Citosinas) de regiões promotoras. A hipermetilação suprime a expressão de genes supressores de tumor, como *MLH1* e *MSH2*, bem como genes envolvidos na via de sinalização WNT, como é o caso do *APC* (YIU; YIU, 2016). A metilação de *MLH1* tem sido frequentemente associada com a mutação V600E no gene *BRAF*, principalmente em casos de câncer colorretal esporádicos. A hipermetilação é ainda mais frequente em casos esporádicos com alta instabilidade de microssatélites (POYNTER *et al.*, 2009). Além do *MLH1*, outros genes podem se encontrar hipermetilados na região promotora, como os genes *CDKN2A*, *MINT1*, *MINT2* e *MINT3* (TARIQ; GHAS, 2016), que têm sido amplamente empregados na confecção de painéis de sequenciamento de nova geração (NGS) utilizados para avaliar hipermetilação das promotoras de DNA (SEPPÄLÄ *et al.*, 2015).

A instabilidade das regiões de microssatélites (MSI), o terceiro mecanismo de carcinogênese do câncer colorretal, é provocado pela inativação de genes de reparo de DNA que participam do *Mismatch (MSH2, MSH3, MLH1, MSH6, PMS2)*. A principal função destes genes é corrigir nucleotídeos de DNA que foram mal pareados durante o processo de replicação, por meio do reconhecimento dos complexos protéicos *MutS* (MSH2/MSH6 ou MSH2/MSH3) e *MutL* (MLH1/PMS2) (IYER *et al.*, 2006). Quando estes genes estão inativados, a adição de nucleotídeos pela DNA polimerase passa a não ser conferida e um grande número de mutações começa a se manifestar nas células. A própria incorporação de nucleotídeos pela polimerase pode ser responsável pela deleção ou inserção de nucleotídeos à fita recém-sintetizada, através do mecanismo de escorregamento da DNA polimerase (*slippage event*) (KLOOR *et al.*, 2014).

As regiões microssatélites são sequências que consistem em unidades repetitivas de DNA de 2-6 pares de base e estão presentes em grande quantidade no genoma humano. Durante o processo de replicação, é possível que ocorra uma dissociação e reanelamento ocasional entre as fitas molde e complementar da molécula de DNA. Caso esse reanelamento ocorra de maneira incorreta, as unidades de repetições da fita molde ficam diferentes das unidades de repetições da fita complementar, e isto dá origem às chamadas regiões de inserção e deleção de loops (IDLs).

Tanto as regiões de mal pareamento de bases, quanto as regiões IDLs podem ser reparadas pelo complexo de proteínas do MMR. Entretanto, caso a função destas proteínas esteja comprometida, os erros inerentes ao processo de replicação não serão reparados, originando instabilidade nas regiões microssatélites, e eventualmente, contribuindo para o desenvolvimento do câncer colorretal (JIRICNY, 2006). A Figura 1.3 ilustra todo o mecanismo descrito acima.

A MSI atinge 15% de todos os casos de câncer colorretal, sendo 3% referentes aos casos de câncer colorretal hereditário não poliposo (Síndrome de Lynch) e os 12% restantes referentes aos casos de câncer colorretal esporádico. Dos casos hereditários, aproximadamente 90% possuem instabilidade nas regiões de microssatélites. Os casos de câncer colorretal esporádicos com MSI, são provocados pela hipermetilação da região promotora do gene *MLH1* (envolvidos também com CIMP) (DE LA CHAPELLE; HAMPEL, 2010; SUN *et al.*, 2014).

Pacientes com câncer colorretal e que são MSI positivos, apresentam melhor prognóstico quando comparado aos casos de câncer colorretal MSS, uma vez que a via MSI não se encontra envolvida com a perda de heterozigosidade (THIBODEAU; BREN; SCHAID, 1993).

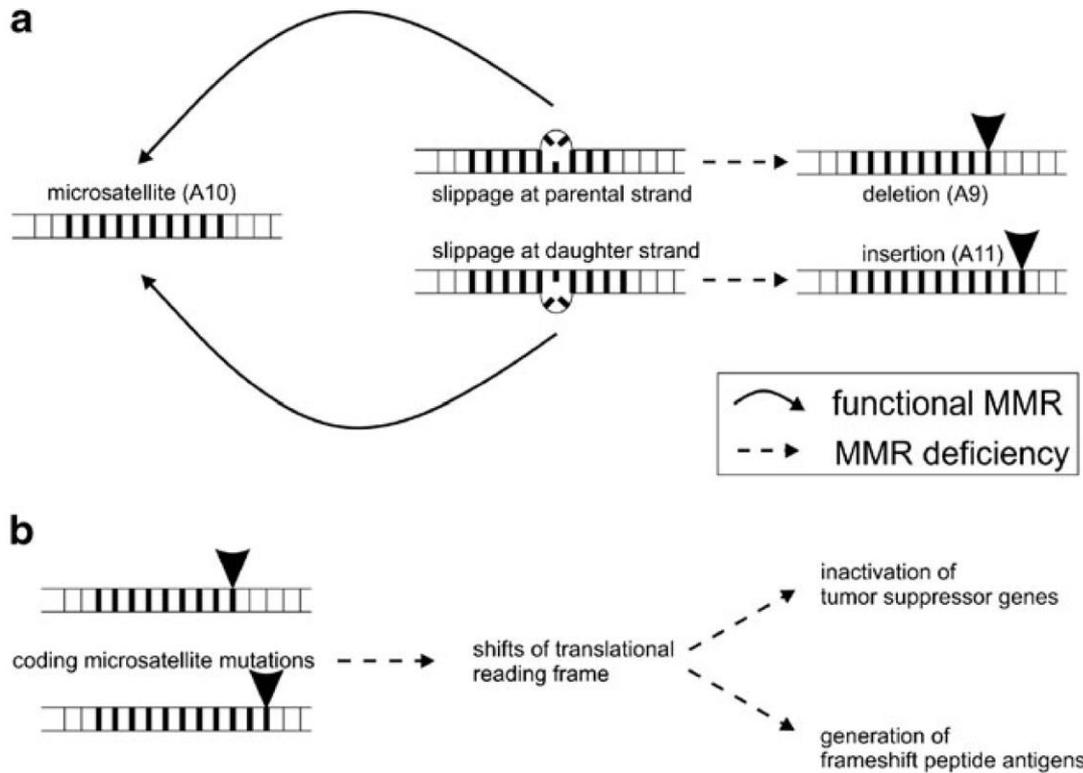


Figura 1.3: O fenômeno de escorregamento da DNA polimerase. As bases mal pareadas devido ao fenômeno de escorregamento da DNA polimerase são corrigidas pelo mecanismo de MMR. (a) Células que apresentam o mecanismo de MMR defeituoso, apresentam mutações de inserções ou deleções na sequência de DNA, como é indicado pelas setas em preto. (b) Regiões de microssatélites que são afetadas por mutações de inserção ou deleção, podem resultar na inativação funcional da proteína codificada por este gene (Fonte: KLOOR *et al.*, 2014).

1.4) Variantes do câncer colorretal

Existem três tipos principais de câncer colorretal: o esporádico, o hereditário e o familiar. O câncer colorretal esporádico acomete a maior parte dos indivíduos, correspondendo de 75-85% dos casos (YAMAGISHI *et al.*, 2016). Ocorre como consequência de mutações somáticas nas células (VATANDOOST; GHANBARI, 2015), associadas a fatores ambientais como hábitos alimentares, tabagismo, obesidade e idade avançada (GILSING *et al.*, 2015).

O câncer familiar acomete aproximadamente 20-30% dos casos e se desenvolve pela presença de mutações herdadas pelas células. De todos os casos de cânceres familiares, 3 a 5% são conhecidos como síndromes hereditárias e encontram-se associados à presença de mutações germinativas em genes conhecidos, como os genes de MMR. Os outros 25% restantes apresentam uma etiologia pouco conhecida e podem ser causados por mutações em genes ainda não identificados (JASPERSON *et al.*, 2010).

Existem diversas síndromes hereditárias relacionadas com o câncer colorretal: Síndrome Adenomatosa Familiar (FAP), Polipose associada a MUTYH (MAP), Polipose Juvenil e Síndrome *Peutz-Jeghers* (PJS) e Síndrome Hereditária de Câncer Colorretal Não Poliposo (HNPCC) ou Síndrome de Lynch (SL) (JASPERSON; BURT, 2015), que é a mais frequente (NOLL; PAREKH; KARLITZ, 2016).

A FAP é uma herança autossômica dominante caracterizada pela formação de múltiplos pólipos colônicos. Acomete menos de 1% dos casos de câncer colorretal e surge como consequência de mutações germinativas no gene *APC*. É uma síndrome de penetrância completa, em que 100% dos indivíduos que apresentam mutações são acometidos com câncer colorretal (JUNG *et al.*, 2016).

A MAP é uma herança autossômica recessiva causada por mutações bialélicas no gene *MUTYH*. Acomete menos de 1% dos indivíduos com câncer colorretal e assim como a FAP, é caracterizada por formação de múltiplos pólipos adenomatosos. Indivíduos acometidos com MAP apresentam risco de 80% de desenvolverem câncer até os 80 anos de idade (LUBBE *et al.*, 2009).

A Polipose Juvenil e a PJS são heranças autossômicas dominantes muito raras provocadas por mutações em *SMAD4* ou *BMPRIA* e *STK11*, respectivamente. A Polipose Juvenil acomete 1 em cada 100.000-160.000 nascimentos, enquanto PJS acomete 1 a cada 50.000- 200.000 nascimentos. São caracterizadas pela formação de pólipos gastrointestinais no estômago, intestino delgado, cólon e reto (LATCHFORD; PHILLIPS, 2011; CAMPOS; FIGUEIREDO; REAL MARTINEZ, 2015;). A Figura 1.4 é mostra a incidência dos diferentes tipos de câncer colorretal.

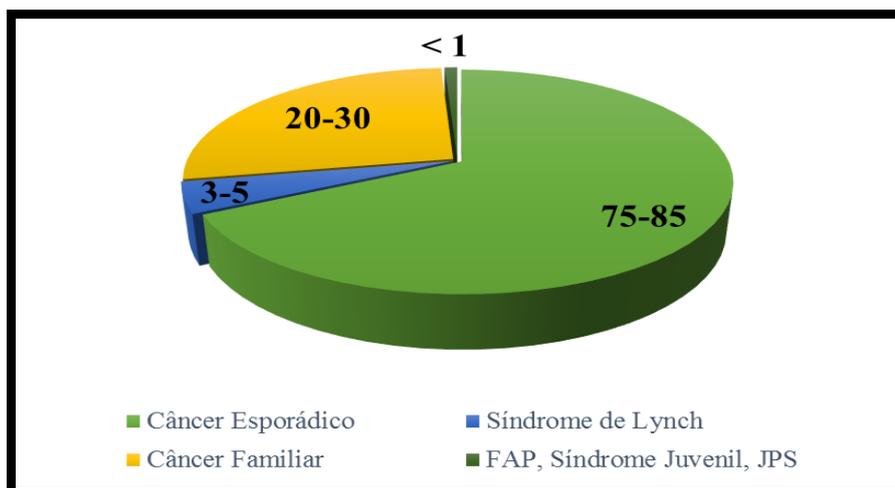


Figura 1.4: Incidência dos diferentes tipos de câncer colorretal. Em verde, casos de câncer colorretal esporádicos; em amarelo, casos de câncer familiar; em azul, casos de Síndrome de Lynch; em verde escuro, casos de FAP, Síndrome Juvenil e JPS.

1.5) Síndrome de Lynch (SL)

A SL começou a ser estudada em 1895, por *Aldred Scott Wartin*, um patologista da *Michigan University* (Estados Unidos). Wartin deu início aos estudos ao avaliar uma família de origem alemã (Família G) com um grande número de membros (146), que apresentava uma frequência elevada de tumores no intestino (cólon e reto), estômago e útero (WARTHIN, 1913). Além da família G, *Wartin* estudou a transmissão fenotípica de cânceres em outras famílias, permitindo com que ele demonstrasse que a transmissão dos cânceres entre indivíduos relacionados se dava de maneira autossômica dominante (WARTHIN, 1913).

Alguns anos depois, na década de 1960, o pesquisador *Henry Lynch*, deu início ao estudo da família N (de Nebraska), que apresentava um grande número de indivíduos acometidos com cânceres no intestino, sobretudo no cólon e no reto. Inicialmente, as famílias G e N foram associadas à Polipose adenomatosa Familiar, embora não apresentassem as características clínicas comumente encontradas em pacientes FAP positivos, a exemplo dos múltiplos pólipos adenomatosos (LYNCH *et al.*, 1966). Em 1971, essa síndrome ficou conhecida como *Cancer Family Syndrome (CFS)* (LYNCH; KRUSH, 1971) e em 1984 foi renomeada para Síndrome de *Lynch* ou *Hereditary Non-poliposis colorectal cancer (HNPCC)* (BOLAND; TRONCALE, 1984).

A SL é uma herança autossômica dominante de alta penetrância, em que os pacientes acometidos herdam, de um dos pais, um alelo dos genes MMR mutado. Os portadores de mutações apresentam risco aumentado de desenvolver diversos tumores, principalmente tumores no cólon e reto do intestino (risco aumentado em 80%), endométrio (risco aumentado em 65%) e ovário (risco aumentado em 6-21%) (GOODENBERGER *et al.*, 2015; HAMPEL,

2016). O efeito patogênico da mutação herdada é observado após o evento de segundo *hit*, que leva a inativação somática do segundo alelo funcional (AHADOVA *et al.*, 2016).

Aproximadamente 3% de todos os casos de câncer colorretal (WANG *et al.*, 2016), 5% de todos os casos de câncer de endométrio (DASHTI *et al.*, 2015) e 10 a 15 % de todos os casos de câncer de ovário hereditários (TOSS *et al.*, 2015) estão associados com a SL.

1.5.1) Características clínicas

Além dos tumores colorretais, endométrio e ovário, os pacientes com SL apresentam um risco aumentado de desenvolverem tumores extracolônicos em diferentes tecidos, como: estômago, bexiga, intestino delgado, sistema hepatobiliar (WATSON *et al.*, 2009), pâncreas (KASTRINOS *et al.*, 2009), mama (WATSON *et al.*, 2009), próstata (BAUER *et al.*, 2011) e nas glândulas adrenocorticais (RAYMOND *et al.*, 2013), embora em menor proporção. É comum o desenvolvimento de tumores sincrônicos (detectados no mesmo momento do tumor primário) e metacrônicos (detectados após a ressecção ou diagnóstico do tumor primário) (PAJARES; PEREA, 2015). Além disso, o risco de desenvolvimento de câncer é aumentado em 17% para pacientes jovens que preenchem os critérios clínicos para SL (HAMPEL *et al.*, 2005).

Os tumores colorretais em SL tendem a ser não- poliposos e a maior parte encontra-se localizada no cólon proximal (direito) (LYNCH; LYNCH; HARRIS, 1977). Além disso, é comum a presença de infiltrado linfocitário e tumores pouco diferenciados mucinosos ou células no formato de anel de sinete (ver Figura 1.5).

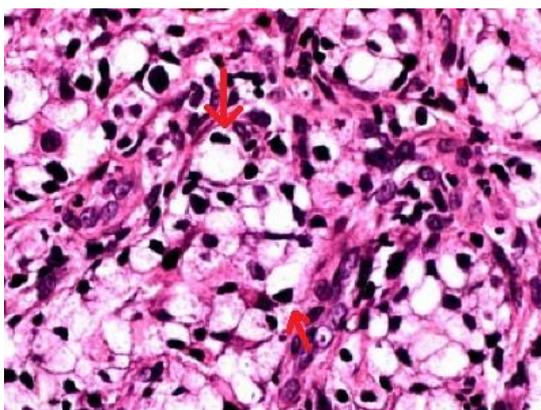


Figura 1.5: Tipos histológicos de carcinoma colorretal mucinoso e células em formato de anel de sinete. As células em formato de anel de sinete (indicado pelas setas em vermelho) são abundantes em mucinas no citoplasma e no núcleo (Fonte: ADAPTADO de [http://www.gastrointestinalatlas.com/english/colon carcinoma ii.html](http://www.gastrointestinalatlas.com/english/colon_carcinoma_ii.html)).

1.5.2) Características moleculares

A Síndrome de Lynch surge como consequência da presença de mutações germinativas em um dos quatro principais genes de reparo de DNA envolvidos com o *Mismatch Repair*: *MutS homolog 2 (MSH2)*, *MutL homolog 1 (MLH1)*, *mutS homolog 6 (MSH6)* e *Pms1 homolog 2 (PMS2)* (SIRAJ *et al.*, 2015). Além de variações nestes genes, pode ser também observada a deleção germinativa do gene *Epithelial Cell Adhesion Molecule (EPCAM)*, que se posiciona ao lado do gene de *MSH2*. A perda do códon de parada de *EPCAM*, é capaz de promover a hipermetilação da região promotora de *MSH2*, culminando com o seu silenciamento (LIGTENBERG *et al.*, 2009).

O MMR é um dos mecanismos responsáveis por assegurar que o processo de replicação ocorra de maneira correta, sem que haja a incorporação ou deleção de nucleotídeos na sequência de DNA. Alterações moleculares nos genes envolvidos com o MMR, levam a perda da função das proteínas codificadas por eles (SIRAJ *et al.*, 2015), deixando as células mais susceptíveis à instabilidade das regiões de microssatélites e contribuindo para o início do processo de carcinogênese colorretal (NOLL; PAREKH; KARLITZ, 2016).

De acordo com a *International Society for Gastrointestinal and Hereditary Tumors (InSIGHT)*, mutações nos genes *MLH1*, *MSH2*, *MSH6* e *PMS2*, contribuem para 42%, 32%, 18% e 8% dos casos de Síndrome de Lynch, respectivamente (PLAZZER *et al.*, 2013). Além disso, os portadores de mutações nos genes MMR apresentam riscos cumulativos diferentes de desenvolver câncer dependendo do gene que se encontra mutado, conforme é demonstrado na Tabela 1.1 (GOODENBERGER *et al.*, 2015).

Tabela 1.1: Risco cumulativo de desenvolver câncer colorretal, endométrio, ovário, trato urinário e estômago aos 70 anos de idade para SL (Fonte: ADAPTADO de GOODENBERGER *et al.*, 2015).

Tipo de câncer	Risco câncer <i>MLH1</i> e <i>MSH2</i> (%)	Risco câncer <i>MSH6</i> (%)	Risco câncer <i>PMS2</i> (%)
Colorretal	28-75	10-70	15-20
Endométrio	27-60	15-71	15
Ovário	6-21	1	Risco aumentado
Trato urinário	8-9	<1	Risco aumentado
Estômago	5-20	<1	Risco aumentado

1.5.3) Diagnóstico clínico e Aconselhamento Genético

De acordo com a Organização Mundial da Saúde (OMS), o aconselhamento genético é uma prática da medicina que tem como objetivo compartilhar informações a respeito dos aspectos clínicos e moleculares de uma determinada doença genética. A prática de

aconselhamento é realizada por profissionais qualificados (médicos geneticistas, enfermeiros e psicólogos) e possibilita a detecção precoce de uma doença genética em determinadas famílias. Através do aconselhamento é possível planejar medidas de prevenção da manifestação da doença em familiares que ainda não foram acometidos. O aconselhamento genético pode ser utilizado para várias doenças hereditárias, inclusive para cânceres hereditários, como é o caso da Síndrome de Lynch (<http://www.who.int/genomics/professionals/counselling/en/>).

A identificação de pacientes suspeitos com SL ocorre de acordo com as características clinicopatológicas e moleculares que este grupo de indivíduos apresenta em comum. O aconselhamento genético em pacientes com câncer colorretal ganhou maior importância a partir do desenvolvimento dos critérios de Amsterdã para o diagnóstico clínico da síndrome na década de 1990 (HAMPEL, 2016).

Os critérios de Amsterdã foram elaborados pelo InSiGHT, em uma conferência realizada na cidade de Amsterdã na Holanda (VASEN *et al.*, 1991). A forte história familiar de câncer colorretal presente em pacientes jovens que não apresentam FAP é um fator considerado na detecção de indivíduos com SL, conforme mostrado no Quadro I.I.

Quadro I.I: Critérios clínicos de Amsterdã I para o diagnóstico da Síndrome de Lynch.

Critérios de Amsterdã para o diagnóstico clínico da SL
Para o diagnóstico clínico de SL pelos Critérios de Amsterdã, é necessária a identificação de pelo 3 parentes com histologia sugerindo câncer colorretal:
• Um dos membros ser parente em 1º grau dos outros dois
• Pelo menos duas gerações acometidas
• Pelo menos um dos membros com CCR deve diagnosticado < 50 anos
• Exclusão de Polipose Adenomatosa Familiar (FAP)

Fonte: ADAPTADO de ROSSI, 2011.

Em 1999, durante uma nova conferência realizada na cidade de Amsterdã, pesquisadores do mesmo grupo revisaram os critérios de Amsterdã e incluíram a formação de tumores extracolônicos (Ver Quadro I.II). Este critérios ficaram conhecidos como Critérios de Amsterdã II (VASEN *et al.*, 1999).

Quadro I.II: Critérios clínicos de Amsterdã II para o diagnóstico de Síndrome de Lynch.

Critérios de Amsterdã II para o diagnóstico clínico de SL
Para o diagnóstico clínico de SL são necessários pelo menos três membros de uma mesma família com CCR, ou adenocarcinoma de endométrio, ou carcinoma de células transicionais de vias excretoras renais (pelve renal ou ureter), ou adenocarcinoma de intestino delgado:
• Um dos membros ser parente em 1º grau dos outros dois
• Pelo menos duas gerações acometidas
• Pelo menos um dos membros com CCR e idade < 50 anos
• Exclusão de Polipose Adenomatosa Familiar (FAP)

Fonte: ADAPTADO de ROSSI, 2011.

Com a descoberta da MSI e sua associação com a síndrome HNPCC, os critérios de Amsterdã I e II tornaram-se restritos. Fez-se necessária a elaboração de critérios clínicos mais abrangentes para o diagnóstico de SL, que incluíssem a detecção de pacientes que não preencheram os critérios de Amsterdã I e II. Durante uma conferência realizada na cidade de Bethesda, nos Estados Unidos, no ano de 1996, o Instituto Nacional de Câncer dos Estados Unidos (do inglês, *US National Cancer Institute*) elaborou um novo conjunto de critérios clínicos para seleção de casos de câncer colorretal com MSI, que ficaram conhecidos como Critérios de Bethesda (Ver Quadro I.III) (RODRIGUEZ-BIGAS *et al.*, 1997). Em 2004, estes critérios foram revisados e houve a inclusão da avaliação molecular dos tumores e o rastreamento de mutações germinativas nos genes de MMR (Ver Quadro I.IV) (UMAR *et al.*, 2004).

Quadro I. III: Critérios clínicos de Bethesda para o diagnóstico de Síndrome de Lynch.

Critérios de Bethesda para o diagnóstico de SL
• Indivíduos que preenchem os Critérios de Amsterdã
• Indivíduos com dois tumores relacionados à SL, incluindo CCR sincrônico ou metacrônico, ou tumores extracolônicos
• Indivíduos com CCR, e um parente em 1º grau com CCR, e/ou tumor extracolônico relacionado à síndrome, e/ou adenoma colorretal, um dos tumores diagnosticados < 45 anos, e o adenoma diagnosticado < 40 anos
• Indivíduos com CCR, ou câncer endometrial, diagnosticado < 45 anos
• Indivíduos com câncer no cólon direito, com padrão histológico indiferenciado (sólido/cribriforme), antes dos 45 anos
• Indivíduos com CCR, com células em anel de sinete, diagnosticado < 45 anos
• Indivíduos com adenomas, diagnosticados < 40 anos

Fonte: ADAPTADO de ROSSI, 2011.

Quadro I. IV: Critérios clínicos de Bethesda Revisados para o diagnóstico de Síndrome de Lynch, visando à identificação de indivíduos candidatos aos testes de rastreamento para a SL.

Critérios de Bethesda revisados para o diagnóstico de SL
Para justificar a procura por MSI, os critérios de Bethesda revisados recomendam:
• CCR diagnosticado em paciente com menos de 50 anos
• Presença de CCR sincrônico ou metacrônico, ou outro tumor extracolônico, associado à síndrome, independentemente da idade
• CCR com histologia sugerindo MSI*, diagnosticado em paciente com menos de 60 anos
• CCR diagnosticado em um ou mais parentes em 1º grau, com tumor relacionado à síndrome, com um dos tumores tendo sido diagnosticado antes dos 50 anos
• CCR diagnosticado em um ou mais parentes de 1º ou 2º graus, com tumores relacionados à síndrome, independentemente da idade

Fonte: ADAPTADO de ROSSI, 2011.

Atualmente, tanto os critérios clínicos de Amsterdã, quanto os critérios clínicos de Bethesda são utilizados para o diagnóstico clínico da SL em pacientes diagnosticados com menos de 50 anos. Entretanto, a organização *National Comprehensive Cancer Network* recomenda o rastreamento da SL utilizando os testes de imunohistoquímica e instabilidade de microssatélites para todos os pacientes que preencheram os critérios clínicos de Bethesda (SIRAJ *et al.*, 2015).

Na prática, muitos indivíduos portadores de mutações em um dos genes MMR, não preenchem os critérios de Amsterdã II e os critérios de Bethesda Revisados, o que muitas vezes dificulta a identificação clínica destes pacientes (VASEN *et al.*, 2013).

Além dos critérios clínicos, os participantes do aconselhamento genético podem ser avaliados por modelos de predição, que estimam o risco de se encontrar uma mutação deletéria em um dos genes de MMR. A predição é feita antes de se realizar o rastreamento molecular para SL e os principais modelos utilizados são: *Prediction Model for MLH1, MSH2, MSH6 Gene Mutations* (PREMM_{1,2,6}), MMRpro e MMRpredict. A história familiar, tipos e número de tumores frequentemente encontrados, idade do diagnóstico e outros fatores são avaliados por esses programas (CHEN *et al.*, 2006).

Os indivíduos que apresentam história familiar são encaminhados para diferentes acompanhamentos clínicos dependendo da localização do tumor do gene de MMR acometido. A colonoscopia é recomendada a cada 1-2 anos entre as idades de 20-30 anos, esofagogastroduodenoscopia a cada 3-5 anos entre as idades de 20-55 anos, exame neurológico anual a partir de 25-30 anos (JASPERSON; BURT, 2015).

A identificação de pacientes que apresentam SL é importante, uma vez que estes pacientes apresentam um risco cumulativo bem superior em comparação com o restante da população de desenvolver câncer colorretal e outros cânceres (ZUMSTEIN *et al.*, 2016). Os pacientes identificados precocemente podem ser beneficiar de estratégias de prevenção, como a remoção de pólipos adenomatosos (RICCIARDIELLO; AHNEN; LYNCH, 2016). E, em alguns casos mais extremos, cirurgias profiláticas de remoção do útero e ovário (histerectomia com salpingooforectomia bilateral) (JASPERSON; BURT, 2015). Além disso, acredita-se que fatores como modificações na dieta, podem reduzir levar a redução do risco de câncer colorretal (SONG; GARRETT; CHAN, 2015).

A quimioprevenção é uma estratégia bastante discutida em câncer colorretal e utiliza de compostos naturais ou sintéticos para a prevenção do desenvolvimento do câncer. Vários compostos já foram testados como agentes de quimioprevenção, entretanto a aspirina é o agente mais bem estudado para SL. Em 2011, *Burn* e colaboradores acompanharam 937 pacientes com diagnóstico clínico ou genético de SL e observaram que os pacientes que tomaram aspirina por ≥ 2 anos tiveram uma menor incidência de câncer colorretal e outros cânceres relacionados a SL, em comparação com os pacientes que receberam o placebo. Este estudo recebeu o nome de CAPP2 (BURN *et al.*, 2011).

Apesar dos resultados promissores, o estudo de CAPP2 não permitiu definir a dose correta e os efeitos adversos da aspirina. Portanto, novos estudos de triagem randomizados estão sendo realizados (ensaio CAPP3), com o objetivo de definir a dose correta do medicamento(<http://www.isrctn.com/ISRCTN16261285?q=capp3&filters=&sort=&offset=1&totalResults=1&page=1&pageSize=10&searchType=basic-search>). A síndrome de Lynch está associada com a rápida progressão de pólipos para os tumores malignos, portanto, estratégias de quimioprevenção são interessantes para estes tipos de tumores (WU; PATTERSON; HAWK, 2011).

1.5.4) Diagnóstico molecular

As famílias que preenchem os critérios clínicos de SL são encaminhadas para o rastreamento molecular, que se inicia pela procura da instabilidade de regiões microssatélites (MSI) e avaliação da expressão das proteínas de MMR por meio de Imunohistoquímica (IHC) (JULIÉ *et al.*, 2008).

Cinco marcadores moleculares, presentes em regiões de microssatélites e que se encontram frequentemente alterados em câncer colorretal hereditário, são convencionalmente utilizados na procura por MSI: BAT25 e BAT26 (monocleotídicos) e D5S446, D2S123 e D17S250 (dinucleotídicos). Os tumores rastreados que apresentam mais de dois marcadores

alterados são classificados como tumores com alta instabilidade nas regiões de microssatélites (MSI- *high* ou MSI-H), os que apresentam ao menos um dos marcadores alterados são classificados como tumores de baixa instabilidade de microssatélites (MSI-*low* ou MSI-L) e os tumores não alterados para nenhum dos marcadores testados são considerados estáveis (MSI *stable* ou MSS) (BOLAND *et al.*, 1999). Além dos marcadores MSI inicialmente sugeridos por Boland *et al.*, outros marcadores moleculares são também empregados para avaliar a MSI em pacientes com SL, a exemplo dos marcadores monocleotídicos MONO-27, NR-21, NR-24 e CAT-25 (SURAWEERA *et al.*, 2002; BACHER *et al.*, 2004; BIANCHI *et al.*, 2009).

O padrão de expressão das proteínas de MMR também é investigado. A IHC, que utiliza de anticorpos contra cada uma das proteínas de MMR: MLH1, MSH2, MSH6, PMS2, auxilia na procura da região do DNA que provavelmente se encontra mutado em pacientes com SL. A perda de expressão da proteína codificada pelo gene investigado pode ser um indício da presença de mutações neste gene (ZUMSTEIN *et al.*, 2016). A Figura 1.6 é um exemplo do teste de IHC, para a perda de expressão da proteína MLH1.

Se o teste de MSI indica a presença de instabilidade de microssatélites (MSI-H ou MSI-L) e/ou a IHC revela a ausência de expressão de uma ou mais proteínas de MMR, recomenda-se a realização do teste genético para identificar se os genes de MMR e EPCAM se encontram mutados. O sequenciamento de *Sanger* é o teste genético padrão ouro utilizado para o rastreamento dos genes MMR (BOISSIÈRE-MICHOT *et al.*, 2016). A Figura 1.7 fornece um esquema detalhado a respeito das etapas de rastreamento molecular para pacientes que preencheram os critérios clínicos para SL.

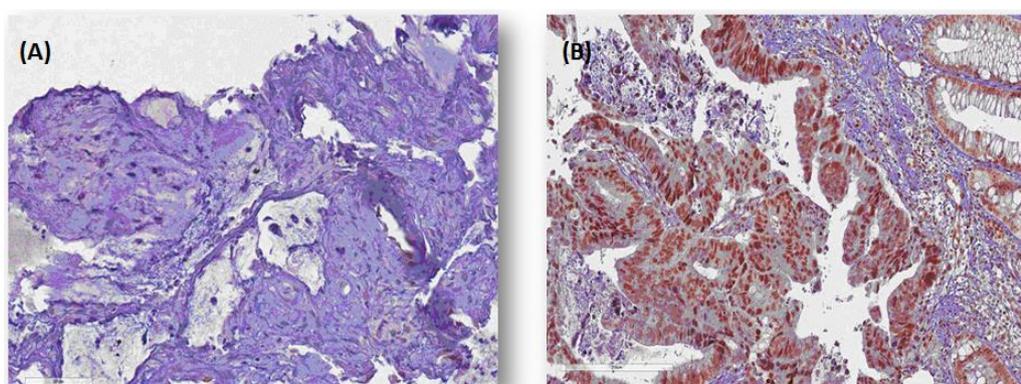


Figura 1.6: Perfil imunohistoquímico de proteínas MMR: (A) Expressão patológica da proteína de reparo MLH1; (B) Expressão normal da proteína MLH1 (presença de coloração). Paciente de 64 anos sem instabilidade nas regiões microssatélites (MSS) e que apresenta célula em formato de anel de sinete com mucina (Fonte: Aluna de Doutorado do Departamento de Genética do INCA, Marion Ama).

Quando há perda da expressão das proteínas MLH1 e PMS2 é recomendado que se realize o sequenciamento do tumor para verificar a presença da mutação somática *BRAF* V600E, que leva a hipermetilação da região promotora de *MLH1* e seu consequente silenciamento (DENG *et al.*, 2004). A hipermetilação de regiões promotoras também pode ser encontrada em casos de câncer colorretal hereditários, pela presença de mutações germinativas no gene *EPCAM* que levam ao silenciamento do gene *MSH2* (KWOK *et al.*, 2014).

Pacientes que preenchem os critérios de Amsterdã, entretanto são MSS e mostram um padrão de expressão normal para as quatro proteínas do MMR, normalmente são classificados como portadores da Síndrome de Câncer Colorretal Familiar do Tipo X (FCCTX). Estes indivíduos normalmente não possuem mutações nos genes MMR, podendo apresentar alterações em genes relacionados a outras síndromes genéticas (polipose associada a *MUTYH*, polipose serrada, polipose associada a *NTHL1* e polipose associada aos erros da polimerase) ou em genes de menor penetrância ainda não identificados (KRAVOCHUCK; CHURCH, 2016).

Existem ainda situações em que o indivíduo possui história familiar para SL, é instável para as regiões microssatélites, apresenta perda de expressão de pelo menos uma das proteínas MMR, não é portador de mutações nos genes de MMR ou *EPCAM* e não apresenta o gene *MLH1* hipermetilado. A etiologia e a significância clínica destes casos ainda é pouco conhecida e estes casos são denominados como Síndrome de Lynch-like (KANG *et al.*, 2015).

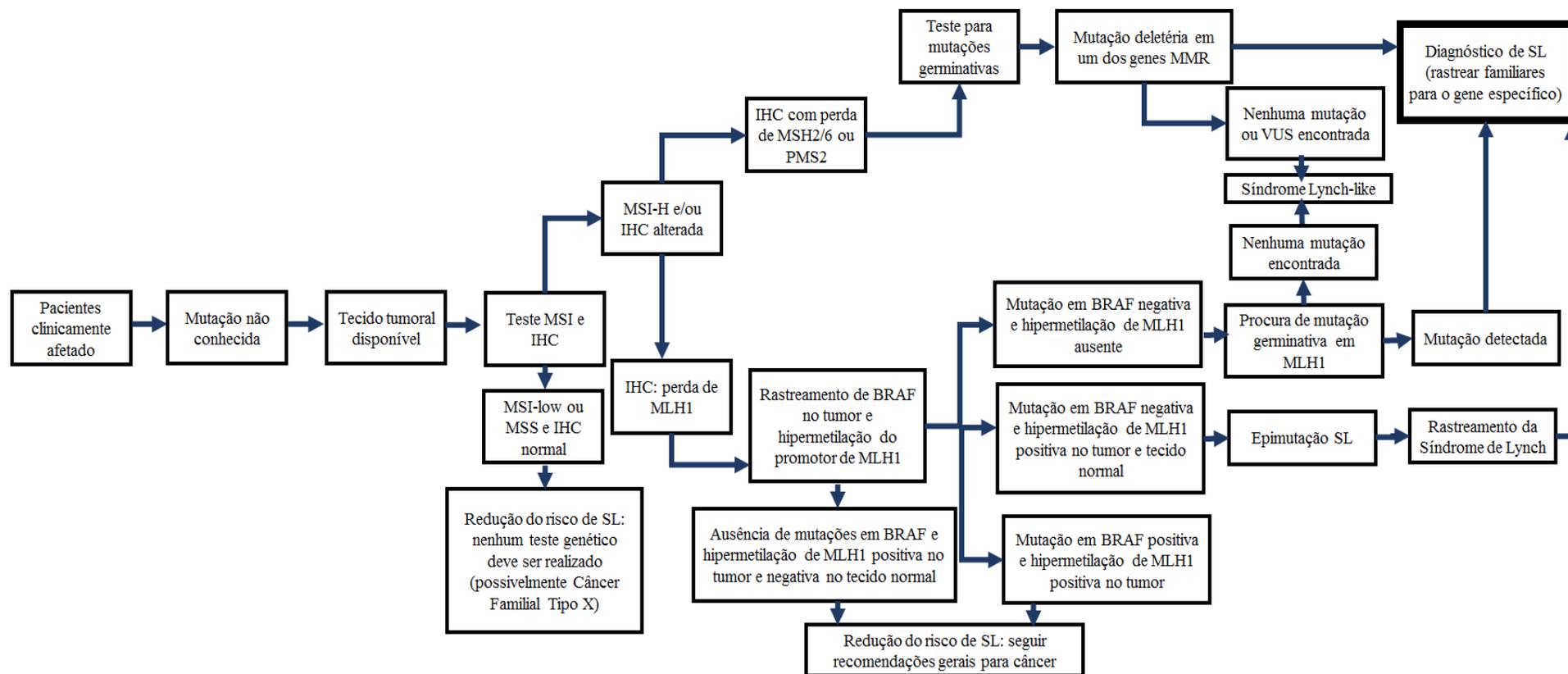


Figura 1.7: Estratégia de teste genético quando um paciente é clinicamente afetado e a mutação de sua família é desconhecida (Fonte: ADAPTADO de GIARDIELLO *et al.*, 2014).

1.6) Sistema de reparo por mal pareamento (MMR)

As proteínas MSH2, MSH6, MLH1 e PMS2 participam do reparo de DNA por mal pareamento ou *Mismatch Repair* (MMR). Elas são capazes de detectar erros de mal pareamento de nucleotídeos durante a replicação do DNA (fase S) e estimular a adição correta de nucleotídeos pela DNA polimerase. Caso o DNA mal pareado não seja reparado, são originadas mutações pontuais, deleções ou inserções, formando proteínas truncadas e não funcionais (BOLAND; GOEL, 2011).

O mecanismo de reparo por mal pareamento foi primeiramente descrito em bactérias ao final da década de 1970, mas só foi melhor caracterizado na década de 1990, por *Grilley* e colaboradores (GRILLEY *et al.*, 1990). Diferente do mecanismo que ocorre em *E.coli*, em organismos eucariotos, as proteínas envolvidas no reparo MMR formam heterodímeros constituídos por dois genes. A proteína MSH2 forma um heterodímero com MSH6 ou MSH3, dando origem aos complexos *MutSa* ou *MutSβ*, respectivamente e proteína MLH1 forma heterodímeros com PMS2 ou PMS1, originando os complexos *MutLa* e *MutLβ*, respectivamente (ACHARYA *et al.*, 1996).

O reparo por MMR se inicia com o reconhecimento do nucleotídeo mal pareado pelo complexo *MutSa* ou *MutSβ*. *MutSa* é capaz de reconhecer mal pareamentos de 1-2 pares de bases, enquanto *MutSβ* reconhece mal pareamentos de mais de 2 pares de bases (regiões de inserção-deleção loops - IDLs). Uma vez que *MutS* encontra-se ligado no sítio com dano de DNA, há o recrutamento do complexo *MutLa*. A formação do complexo de reparo a partir da junção de dois heterodímeros, recruta a proteína *proliferating cell nuclear antigen* (PCNA) à fita que deve ser reparada. PCNA é posicionada na fita com o auxílio da proteína *replication factor C* (RFC) e interage com as proteínas do complexo *MutLa*, estimulando a atividade endonucleásica de PMS2, que cliva os nucleotídeos mal pareados. A excisão destes nucleotídeos ocorre com a ação da enzima *Exonuclease 1* (EXO1). A proteína *replication protein A* (RPA), liga-se aos fragmentos de fita simples gerados após a ação da EXO1, até que haja uma reconstituição da fita pela ação da DNA polimerase δ . A nova fita reparada é ligada com o auxílio da DNA Ligase I (LI; MARTIN, 2016). A Figura 1.8 ilustra este mecanismo.

A identificação da base mal pareada em eucariotos ocorre pelo reconhecimento da descontinuidade da fita recém-sintetizada em comparação com a fita molde, e não pela metilação de sítios GATC, como é observado em *E.coli*. As terminações 5' ou 3' de fragmentos de *Okasaki* na fita retardada (*lagging strand*) ou as terminações 3' da fita líder, podem apresentar quebra de fitas em situações de mal pareamento de bases, funcionando como um sinal para que o mecanismo MMR seja ativado (FANG; MODRICH, 1993).

A atividade dos complexos *MutS* e *MutL* é regulada pela presença da molécula de *Adenosine-Triphosphate* (ATP). Uma vez que o mal pareamento de DNA é reconhecido pelo complexo *MutS*, o ADP ligado a cada uma das proteínas deste complexo é liberado e trocado por uma molécula de ATP, induzindo uma mudança conformacional na estrutura do complexo que vai estimular o reparo de DNA. A conversão de ADP para ATP, permite que o complexo *MutS* se deslize sobre a molécula de DNA e recrute o complexo *MutL*. A capacidade de escorregamento de *MutS* sobre a fita de DNA é conhecido como *sliding clamp state* e é essencial para que ocorra ativação de *MutL*. Uma vez *MutL* é ativado, *MutS* é liberado da estrutura formada sobre a base mal pareada, ocorre a hidrólise de ATP e o complexo torna-se disponível para novos ciclos de reparo. Caso não haja mal pareamento de bases, o ATP que se encontra ligado a *MutS* é desprendido do DNA e o estado de *sliding clamp* não é atingido (GROOTHUIZEN; SIXMA, 2016).

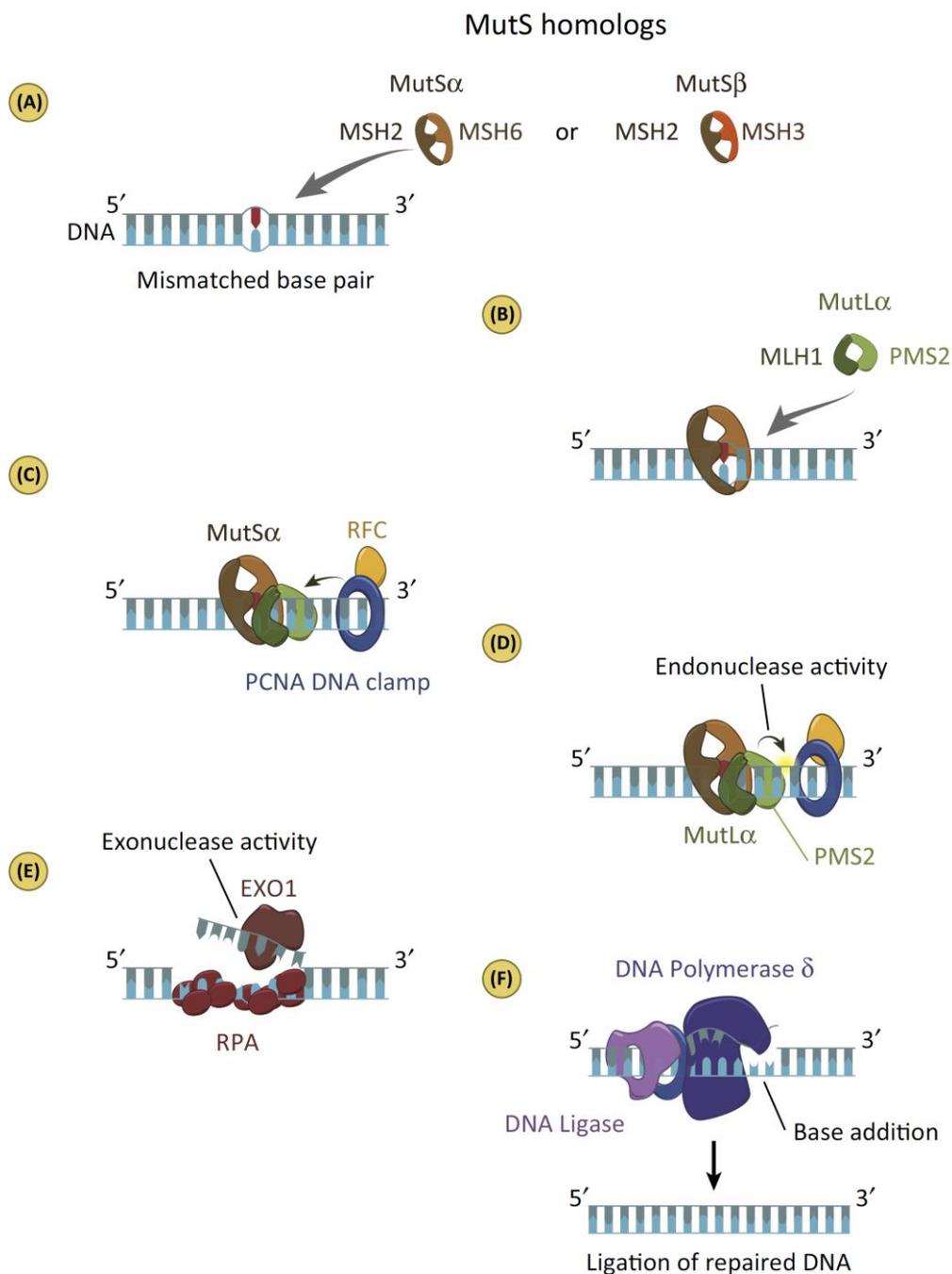


Figura 1.8: Mecanismo de reparo de DNA por Mismatch Repair em mamíferos. (A) O DNA mal pareado é reconhecido pelos complexos MutS α ou MutS β . (B) Posteriormente, MutL α é recrutado para o local do dano e (C) PCNA interage com este complexo, (D) ativando a proteína PMS2 que apresenta atividade endonucleásica. (E) A excisão dos nucleotídeos mal pareados ocorre pela ação da Exo1 e (F) os novos nucleotídeos são incorporados pela polimerase δ . A DNA Ligase faz a ligação da nova fita reparada. (Fonte: LI; MARTIN, 2016).

1.6.1) Principais genes envolvidos com o MMR

Como já foi discutido no tópico anterior, existem 6 principais genes envolvidos com o MMR de eucariotos. Entretanto, em pacientes com síndrome de Lynch, o rastreamento genético ocorre principalmente nos genes *MLH1*, *MSH2*, *MSH6* e *PMS2*, uma vez que *MSH3*

e *PMS1* têm funções pouco conhecidas na carcinogênese colorretal. Estudos demonstraram que a deleção completa do gene *MSH3* não propicia desenvolvimento de tumores em camundongos, o que permite prever que em situações que se tem dano de DNA, o complexo *MutS α* , por si só, é capaz de iniciar o reparo da maioria das sequências mal pareadas, mesmo que elas apresentem mais de 2 nucleotídeos mal pareados. Desta maneira, o rastreamento genético de *MSH3* é pouco usual para pacientes com SL (EDELHANN *et al.*, 2000). A proteína MLH1 pode formar um heterodímero com PMS1, dando origem ao complexo *MutL β* . Entretanto, a supressão da mutagênese mediada por *MutL β* é mais conhecida em leveduras e em mamíferos sua função foi pouco estudada, por isso o rastreamento do gene *PMS1* é raramente realizado (BOLAND; GOEL, 2011).

Os genes *MSH2* e *MSH6* juntos formam o heterodímero que constitui o complexo *MutS*. Eles são estruturalmente similares e contêm 5 domínios conservados: (1) *mismatch domain* (ligação da base mal pareada), (2) *connector domain* (domínio de ligação a *MutL*), (3) *lever domain* (α hélice que envolve e protege a molécula de DNA), (4) *clamp domain* (interage com a estrutura do DNA para estabilizar sua conformação durante o reparo) (5) *ATPase domain* (motivo Walker de ligação do ATP) e (6) *Helix-Turn-Helix domain* (HTX, envolvido com a ligação dos dímeros), como pode ser visto pela Figura 1.9 (LAMERS *et al.*, 2000; MENDILLO *et al.*, 2009; MUKHERJEE; FEIG, 2009).

O gene *MSH2* encontra-se localizado no braço curto do cromossomo 2 (2p21), na posição 2: 47.630.108-47.789.450 (GRCh37), apresenta 16 éxons codificantes e codifica para uma proteína de 943 aminoácidos (<http://www.ncbi.nlm.nih.gov/gene/4436>). Foi associado pela primeira vez com a Síndrome de Lynch em 1993 (FISHEL *et al.*, 1993). O gene *MSH6* está localizado no braço curto do cromossomo 2 (2p16), na posição 2: 47.922.669-48.037.240 (GRCh37). Apresenta 10 éxons codificantes que geram uma proteína de 1360 aminoácidos (<http://www.ncbi.nlm.nih.gov/gene/2956>). Foi identificado pela primeira vez identificado como um gene que participa do MMR em 1997 (MIYAKI *et al.*, 1997).

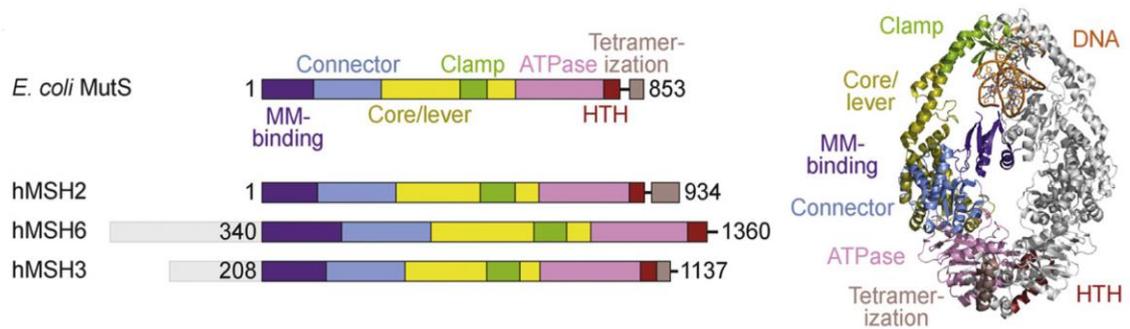


Figura 1.9: Domínios presentes nas proteínas dos complexos de MutS em *E. coli* e em eucariotos (MSH2, MSH6, MSH3). Em azul escuro, o domínio de ligação ao mismatch. Em azul claro, domínio conector. Em amarelo, domínio lever (ou core). Em rosa, domínio ATPásico. Em cinza, domínio HTH (Fonte: ADAPTADO de GROOTHUIZEN; SIXMA, 2016).

Os genes *MLH1* e *PMS2* participam da formação do complexo *MutL*, que é responsável por clivar os *mismatches* na molécula de DNA. As proteínas de *MutL* formam um complexo com dois domínios principais: domínio N-terminal, que contém um sítio ativo com atividade ATPásica, capaz de se ligar ao DNA e ao complexo *MutS* e; domínio C-terminal, que apresenta atividade endonucleásica. Os dois domínios são ligados por um motivo denominado *linker* (Ver Figura 1.10) (HINGORANI, 2016).

O gene *MLH1* está localizado no braço curto do cromossomo 3 (3p21.3), na posição 3: 37.034.823-37.107.380 (GRCh37). Este gene apresenta 19 éxons codificantes e codifica para uma proteína de 756 aminoácidos (<http://www.ncbi.nlm.nih.gov/gene/4292>). *PMS2* está localizado no braço curto do cromossomo 7 (7p22.2), na posição 7: 6.012.870-6.048.756 (GRCh37) e possui 15 éxons que codificam para uma proteína de 862 aminoácidos (http://www.ncbi.nlm.nih.gov/gene/?Term=related_functional_gene_5395%5Bgroup%5D). A primeira associação dos genes *MLH1* (BRONNER *et al.*, 1994) e *PMS2* (NICOLAIDES *et al.*, 1994) com a SL foi realizada em 1994.

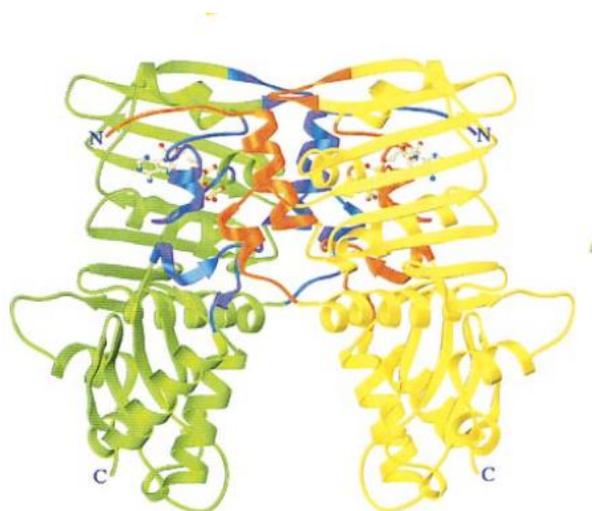


Figura 1.10: Estrutura ortogonal do dímero formado pelo complexo MutL. Cada subunidade é indicada com cores diferentes (verde e amarelo). O domínio ATPásico pode ser visualizado próximo ao domínio N- de cada subunidade, indicado por meio dos átomos nas cores azul e e vermelha (BAN; JUNOP; YANG, 1999).

1.7) Pseudogenes de PMS2

O gene *PMS2* apresenta 14 sequências de pseudogenes de acordo com a referência do genoma humano GRCh37. Nenhum deles é processado e todos estão localizados no mesmo cromossomo do gene funcional (cromossomo 7), como é demonstrado na Figura 1.11 (DEVOS *et al.*, 2004). A versão mais recente do genoma (GRCh38) inclui 16 regiões que apresentam mais de 90% de identidade com o gene funcional *PMS2* (VAN DER KLIFT *et al.*, 2016).

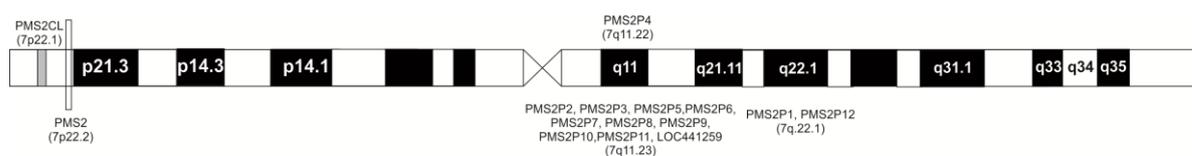


Figura 1.11: Localização do gene PMS2 e seus pseudogenes no cromossomo 7 do genoma.

A maior parte dos pseudogenes apresenta identidade com os éxons 1 a 5 de *PMS2*, com um grau de homologia que varia de entre 89 e 96,4%, como demonstrado na Tabela 1.2. Um único pseudogene, *PMS2CL*, apresenta homologia com a região 3' de *PMS2* (éxons 9 e 11 ao 15), com um grau de identidade das sequências maior do que 96,5%. Ainda não foi identificado nenhum pseudogene com sequência homóloga aos éxons 6, 7, 8 e 10 de *PMS2* (HAYWARD *et al.*, 2007; LI *et al.*, 2015).

PMS2CL é um segmento invertido de 100kb do gene funcional que foi duplicado e se encontra localizado a 700kb da região centrométrica de *PMS2*. Devido à alta homologia que

apresenta com o gene funcional, é comum a ocorrência de eventos de recombinação interparálogos entre as sequências de *PMS2* e *PMS2CL* durante o *crossing over*, gerando alelos híbridos que dificultam a diferenciação entre as duas sequências na população (HAYWARD *et al.*, 2007; VOGT *et al.*, 2016).

O grande número de pseudogenes encontrados em *PMS2* dificulta o rastreamento de cada um de seus éxons por PCR convencional. Para evitar esta interferência, alguns iniciadores e sondas de MLPA (MRC-Holland, P008 versão A1) passaram a ser desenhados em nucleotídeos que se diferiam entre os loci duplicados, mas não variavam na sequência do gene, as chamadas de variantes de sequências parálogas (PSVs) (VOS *et al.*, 2004). Esta estratégia mostrou-se adequada para a amplificação dos éxons 1-5 de *PMS2*, entretanto não evitou a amplificação de sequências homólogas a região 3' de *PMS2*, gerando resultados falso-positivos e negativos (GROOTHUIZEN; SIXMA, 2016).

Tabela 1.2: Identidade das sequências entre os éxons 1 a 15 de *PMS2* e seus pseudogenes (Fonte: Adaptado de LI *et al.*, 2015).

Éxons	Número de pseudogenes	Identidade das sequências (%)
E1	5	89.0~90.2
E2	13	96.3~97.8
E3	13	93.1~95.1
E4	13	93.1~95.1
E5	13	94.5~96.4
E6	0	-
E7	0	-
E8	0	-
E9	1	96.5
E10	0	-
E11	1	96.9
E12	1	100
E13	1	99.1
E14	1	99.5
E15	1	100

Legenda: E: éxon

Com a finalidade de evitar o problema que envolve o crossing-over desigual da região 3' de *PMS2*, estratégias de PCR de longo alcance aplicadas ao DNA (VAUGHN *et al.*, 2010; BORRAS *et al.*, 2013; TOMSIC *et al.*, 2013;) ou ao cDNA (WERNSTEDT *et al.*, 2012; TEN BROEKE *et al.*, 2015) foram desenvolvidas. Além disso, as sondas para o *Multiplex Ligation Probe Amplification* (MLPA) foram redesenhadas em regiões específicas ricas em SNPs (PSVs que se tornaram SNPs com a frequente conversão gênica que ocorre entre *PMS2* e

PMS2CL) (MRC-Holland, P008 versão B1) e têm sido úteis para a detecção de grandes deleções e duplicações nas regiões de sequências parálogas (VAUGHN *et al.*, 2011).

1.8) Sequenciamento de Nova geração (NGS)

A técnica padrão ouro para o rastreamento de genes MMR de pacientes com SL é o Sequenciamento de *Sanger*. Entretanto, *Sanger* é uma técnica altamente laboriosa quando se deseja obter sequências de um grande número de genes, como é o caso da SL. Com o advento das metodologias de Sequenciamento de Nova Geração (NGS) e a demanda de sequenciamento para múltiplos genes, à amplificação isolada de cada um dos éxons de um gene tornou-se um processo ultrapassado (HOPPMAN-CHANEY *et al.*, 2010). Atualmente, existem diversos estudos com doenças genéticas complexas utilizando NGS, considerando que o custo-benefício utilizado para rastreamento genético por meio desta metodologia tem se tornado cada vez melhor (HANSEN *et al.*, 2014).

Desde que o primeiro sequenciador de nova geração (454 Life Sciences, *Roche*) foi lançado em 2005 (MARGULIES *et al.*, 2005), novas plataformas de NGS foram sendo criadas com diferentes princípios de detecção diferente. As plataformas da empresa *Roche* utilizam PCR em emulsão para a preparação de bibliotecas genômicas e podem gerar 100Mb de DNA em 8 horas (250pb tamanho fragmento), enquanto a plataforma da *Helicos Heliscope* (*Cambridge, MA*) utiliza de fragmentos não amplificados de DNA e é capaz de gerar até 2000Mb de sequência por dia (TEN BOSCH; GRODY, 2008).

Além das plataformas destacadas acima, existem as plataformas da empresa *Illumina*, que vêm se destacando pelo menor custo de sequenciamento por amostra quando comparado aos outros sequenciadores (CAPORASO *et al.*, 2012). As novas plataformas da *Illumina* utilizam de sequenciamento por síntese e são capazes de gerar > 40Gb de sequências por corrida de sequenciamento (PRITCHARD *et al.*, 2012).

A metodologia de sequenciamento por síntese envolve a fragmentação enzimática do DNA, com posterior ligação de adaptadores nas extremidades 5' (P5) e 3' (P7) da sequência e desnaturação do DNA. As extremidades 5' das moléculas de DNA são aderidas à superfície sólida de clonagem (*flow cells*) por complementariedade. Durante a etapa de pareamento, o adaptador presente na extremidade livre 3' da molécula encontra seu oligonucleotídeo complementar no suporte sólido, formando pontes e permitindo à amplificação das sequências de DNA. Posteriormente, há a separação e linearização das fitas, formando *clusters* de moléculas idênticas ligadas ao suporte. Com a incorporação de nucleotídeos marcados com fluorocromos e excitação a laser é gerado um sinal que é capturado pelo dispositivo de leitura. A leitura é feita de forma sequencial, o que permite a montagem da sequência completa de

cada *cluster*. O processo de amplificação por pontes é ilustrado pela Figura 1.12 (CARVALHO; SILVA, 2010).

Para que seja possível sequenciar amostras distintas dentro de uma mesma reação, sequências específicas de nucleotídeos (indexes) são adicionadas nas extremidades 5' e 3' da molécula de DNA (junto com os adaptadores P5 e P7). A adição dos indexes permite que até 96 amostras sejam corridas simultaneamente em uma mesma linha (*lane*), uma vez que existem 12 e 8 opções de indexes ligados a P7 e P5, respectivamente (ILLUMINA, 2012).

Além do custo benefício por amostra ser melhor, as metodologias de NGS apresentam uma sensibilidade muito maior do que o sequenciamento de *Sanger*, uma vez uma única sequência de DNA é lida em um número muito superior de vezes. Isso possibilita que erros introduzidos pelas enzimas *Taq DNA polimerases* sejam mais facilmente detectados, reduzindo o número de resultados falso-positivos. As metodologias de NGS tem-se tornado cada vez mais acessíveis e utilizadas, especialmente para o rastreamento de genes ligados síndromes hereditárias (DE LEENEER *et al.*, 2011). Alguns trabalhos vem empregando plataformas de NGS para a detecção de variantes em pacientes com SL, através da utilização de painéis de câncer que utilizam o produto de PCR ou métodos de captura para a preparação das bibliotecas de DNA (HOPPMAN-CHANEY *et al.*, 2010; PRITCHARD *et al.*, 2012; HANSEN *et al.*, 2014; TALSETH-PALMER *et al.*, 2016).

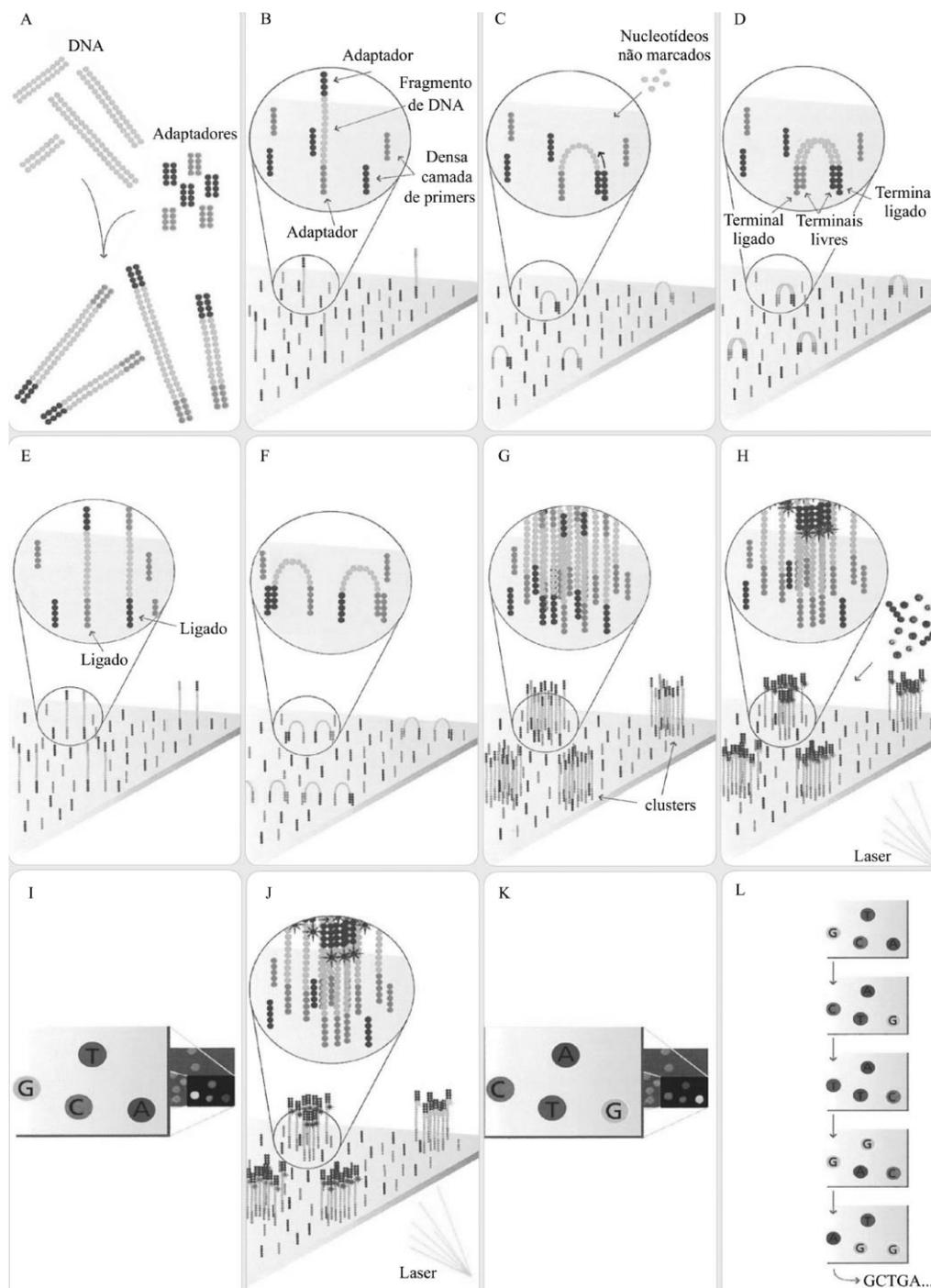


Figura 1.12: Etapas do processo de sequenciamento por síntese utilizando as plataformas da Illumina. (A) Moléculas de DNA fragmentadas são aderidas a superfície sólida por complementariedade. (B) Durante a etapa de anelamento, o adaptador presente na extremidade livre da molécula, encontra sua sequência complementar na *flow cell*, formando uma estrutura em forma de ponte. (C e D) A PCR é iniciada utilizando a extremidade 3' do nucleotídeo como primer. (E) Durante a etapa de desnaturação, ocorre a linearização das fitas e a (F) etapa de anelamento é repetida, formando (G) grupos de moléculas idênticas ligadas ao suporte. (H) Ocorre a incorporação de nucleotídeos marcados, a reação de PCR é finalizada e a excitação a laser gera um sinal correspondente a cada um dos nucleotídeos. (I, J, K) Este processo é repetido várias vezes de forma sequencial, permitindo a montagem do genoma (CARVALHO; SILVA, 2010).

2) JUSTIFICATIVAS

A Síndrome de Lynch se desenvolve como consequência de mutações germinativas em pelo menos um dos quatro genes responsáveis pelo reparo de DNA por mal pareamento. Os genes *MSH2*, *MSH6*, *MLH1* e *PMS2* são genes longos, que apresentam 16, 10, 19 e 15 éxons, respectivamente. Atualmente, a técnica padrão ouro utilizada para a identificação de mutações nestas regiões é o Sequenciamento de Sanger, entretanto, devido ao grande número de genes envolvidos com a SL, a metodologia muitas vezes se torna laboriosa e com custo elevado. Desta maneira, a padronização de metodologias mais específicas e que reduzam o custo e tempo de rastreamento molecular para SL é importante, principalmente ao se considerar que os indivíduos portadores da Síndrome possuem um risco cumulativo muito superior, em relação ao restante da população, de desenvolverem cânceres.

3) OBJETIVOS

3.1) Objetivos Gerais

O trabalho tem como objetivo padronizar PCRs *multiplexes* (para os genes *MSH2*, *MLH1*, *MSH6*) e PCRs de Longo Alcance (para o gene *PMS2*), acopladas a metodologia de Sequenciamento de Nova Geração por síntese, para a identificação de variantes polimórficos em pacientes que preencherem os critérios clínicos de Amsterdã e/ou Bethesda para HNPCC/Síndrome de Lynch.

3.2) Objetivos Específicos

- Padronização de PCRs multiplexes para os genes *MSH2/MSH6/MLH1* para identificar alterações nos éxons e regiões intrônicas flanqueadoras destes genes, em pacientes que preencheram os critérios clínicos para HNPCC;
- Padronização de PCR de Longo Alcance para o gene *PMS2* com o objetivo de identificar alterações nas regiões exônicas e flanqueadoras deste gene, em pacientes que preencheram os critérios clínicos para HNPCC;
- Padronização de análises de bioinformática para os resultados obtidos pelo Sequenciamento de Nova Geração, como o intuito de identificar variantes nos genes MMR em pacientes que preencheram os critérios clínicos para HNPCC.
- Validar a metodologia empregada no estudo (PCR multiplex/ LR-PCR acoplada ao Sequenciamento de Nova Geração) para a identificação de alterações nos genes *MSH2*, *MLH1*, *MSH6* e *PMS2*, por meio de Sequenciamento de *Sanger*, que é a técnica padrão ouro para rastreamento destes genes.

- Realizar estudo *in silico* do efeito funcional que a troca de um aminoácido pode provocar na estrutura da proteína, visando prever o significado clínico de alterações não descritas pelos bancos de dados LOVD e dbSNP.

4) MATERIAL E MÉTODOS

4.1) Descrição das amostras

Fizeram parte do projeto 63 pacientes diagnosticados com câncer colorretal, maiores de 18 anos, que possuíam história familiar para Síndrome de Lynch, de acordo com os critérios clínicos de Bethesda. Três outras amostras de pacientes com câncer endometrial foram incluídas nas análises por preencherem o critério de Bethesda. Destes 66 pacientes, 24 (36,4%) também preencheram os critérios clínicos de Amsterdam. A maior parte dos pacientes eram do sexo feminino (55%) e a média de diagnóstico do câncer foi de aproximadamente 42 anos de idade (19-81). Os pacientes desenvolveram tumores no cólon ascendente (35%), reto (17%), cólon sigmóide (9%), transverso (8%) e descendente (8%). Casos de câncer de endométrio (5%), colón e estômago (2%) e cólon e intestino delgado (2%) foram menos frequentes. A Tabela 4.1 descreve as características clínicas das amostras.

As amostras dos 66 pacientes foram coletadas em diferentes regiões do Brasil, sendo provenientes de quatro centros distintos: Hospital A.C. Camargos (ACCAM) (São Paulo), Hospital de Clínicas de Porto Alegre (HCPA) (Rio Grande do Sul), Hospital Universitário João de Barros Barreto (HUJJB) (Belém do Pará) e Instituto Nacional de Câncer (INCA) (Rio de Janeiro). Todos os pacientes aceitaram participar voluntariamente do projeto mediante assinatura de Termo de Consentimento Livre Esclarecido (Anexo 1). O projeto foi aprovado pelo CEP do INCA e pelas demais instituições participantes (CAAE-0254.1.001.007-11) (Anexo 2).

DNA genômico foi isolado a partir do sangue periférico dos pacientes, utilizando o Puregene – Blood Core Kit B (Qiagen, Valencia, CA, Estados Unidos), de acordo com o protocolo do fabricante. A concentração e pureza dos DNAs foram estimadas por meio do espectrofotômetro Nanodrop 2000 (Thermo Scientific, Wilmington, DE, Estados Unidos).

A análise de Imunohistoquímica foi realizada para a detecção da expressão das proteínas MSH2, MSH6, MLH1 e PMS2, utilizando tecido tumoral e tecido normal adjacente de 60 pacientes. A MSI foi avaliada em 21 pacientes, a partir de amostras do tecido tumoral e tecido normal adjacente, para os cinco marcadores monocleotídicos: BAT-25, BAT-26, MONO-27, NR-21 e NR-24. A presença de grandes rearranjos gênicos foi avaliada para 51 pacientes por *Multiplex Ligation and Probe Amplification* (MLPA) para os genes *MLH1* e

MSH2 (sonda P072) e *EPCAM* (sonda P003). A Tabela 4.1 e o Anexo 3 descrevem os dados obtidos para estas análises.

Tabela 4.1: Informações clínicas, imunohistoquímica, dados de MLPA e instabilidade de microssatélites dos pacientes incluídos.

Informações clínicas		
Número de pacientes por centro	N	Percentual
HCPA	35	53%
INCA	9	14%
HUJJB	7	11%
ACCAM	15	23%
Total	66	-
Sexo		
Feminino	36	55%
Masculino	30	45%
Idade do diagnóstico		
Média	42,32	-
Mínimo	19	-
Máximo	81	-
Local do primeiro tumor		
Cólon ascendente	23	35%
Cólon transverso	5	8%
Cólon descendente	5	8%
Cólon sigmóide	6	9%
Reto	11	17%
Cólon e estômago	1	2%
Cólon e intestino delgado	1	2%
Endométrio	3	5%
Não especificado (cólon)	11	17%
Informações moleculares		
IHC		
Positiva	31	52%
Negativa	27	45%
Inconclusivo	2	3%
Total	60	
MSI		
MSI-H	11	52%
MSI-L	0	-
MSS	8	38%
Inconclusivo	2	10%
Total	21	-
MLPA		
Normal	43	84%
Del E17-19 <i>MLH1</i>	3	6%
Del E3,8, 9 <i>EPCAM</i> + Del E1-4 <i>MSH2</i>	1	2%
Del E3,8, 9 <i>EPCAM</i> + Del E1-6 <i>MSH2</i>	1	2%
Del E3,8, 9 <i>EPCAM</i> + Del E1-8 <i>MSH2</i>	1	2%
Del E16 <i>MSH2</i>	1	2%
Del <i>MUTHY</i>	1	2%
Total	51	

4.2) Reações de PCR multiplex e LR-PCRs

As sequências de iniciadores utilizadas para cada um dos quatro genes foram adaptadas de trabalhos anteriormente publicados para os genes *MSH2* (BECK *et al.*, 1997; ZAHARY *et al.*, 2012), *MSH6* (CHADWICK *et al.*, 2001), *MLH1* (MENSENKAMP *et al.*, 2014) e *PMS2* (VAUGHN *et al.*, 2010) e são mostrados nos Anexos 4 a 7.

Para os genes *MSH2* (16 éxons), *MLH1* (19 éxons) e *MSH6* (10 éxons) foram utilizados pares de iniciadores que amplificavam os éxons e suas regiões intrônicas flanqueadoras e para o gene *PMS2* (15 éxons) foram utilizados iniciadores capazes de amplificar quase todo o gene.

PCRs *multiplexes* foram padronizadas para os genes *MLH1*, *MSH2* e *MSH6*, levando-se em consideração o tamanho dos *amplicons* gerados e a temperatura de anelamento para cada um dos iniciadores. Regiões que apresentavam temperaturas de pareamento similares e fragmentos com tamanhos distintos, dentro de um mesmo gene, foram alocadas em uma mesma reação (ou mix) de PCR. O gene *PMS2* foi amplificado por meio de 3 reações de PCR de longo alcance: LR-PCR1 (amplificando do éxon 1 ao 5), LR-PCR2 (amplificando do éxon 7 ao 9) e LR-PCR3 (amplificando do éxon 11 ao 15), sendo que os éxons 6 e 10 deste gene foram amplificados por PCR convencional.

As PCR multiplexes foram realizadas para volumes finais de 25 ou 50 µL a partir de 50 ng/µL de DNA; 0,4, 0,8 ou 2 pmol de cada iniciador; 1,0 U de *Taq DNA Platinum* (Invitrogen, São Paulo, SP, Brasil); Tampão de PCR 1X e 0,2mM de cada dNTP. As concentrações de MgCl₂ (1,5 a 2,5Mm) e a ciclagem dos PCRs variaram para cada conjunto de iniciadores, como pode ser visualizado nos Anexos 8, 9, 10 e 11.

As LR-PCRs partiram de uma concentração inicial de DNA a 100ng, utilizando 0,4pmol (LR-1/3) ou 0,8pmol de cada primer (LR-2); 1,25 U de Takara La Taq polimerase (Takara, Mountain View, CA, Estados Unidos), tampão de PCR II 1X e 0,40mM de cada dNTP, para um volume final de reação de 25 µL. As ciclagens consistem em uma desnaturação inicial de 94°C por 1 minuto (min), acompanhada por 35 ciclos de 30 segundos (seg) a 94°C, 30seg a 65°C (LR-1/3) ou 60°C (LR-2) e 15min a 68°C, extensão final de 72°C por 10 minutos. As reações dos éxons 6 e 10 eram realizadas a partir de 50ng de DNA, conforme descrito no protocolo do Anexo 12.

As reações foram realizadas nos termocicladores Veriti 96-well Fast Thermal Cycler, Proflex 3X 32-well PCR System (Life Technologies, Clatsop, CA, Estados Unidos) e C100 Touch Thermal Cycler (Bio-Rad Laboratories, Philadelphia, PA, Estados Unidos). Os

fragmentos provenientes dos PCRs foram separados e visualizados em gel de agarose 0,8% ou 3% ou em poliacrilamida não desnaturante 6% (para os mixes 1.1/1.2/1.3 do gene *MSH2*).

Após amplificados, os fragmentos foram purificados por meio do kit *Illustra GFX PCR DNA and Gel Band Purification* (GE HealthCare Life Sciences, Pittsburg, PA, Estados Unidos) e submetidos ao processo de preparação de bibliotecas.

4.3) Preparação das bibliotecas e NGS (HiSeq2500)

Os produtos de PCR purificados foram quantificados no espectrofotômetro Nanodrop 2000 e quantidades equimolares de cada produto de PCR correspondentes aos quatro genes do mesmo paciente foram misturadas. Esta mistura foi quantificada fluorimetricamente por meio do equipamento Qubit® 3.0 Fluorometer (Life Technologies, Clasbad, CA, Estados Unidos), utilizando o kit *Qubit® dsDNA HS Assay*, de acordo com o protocolo do fabricante. Todas as amostras foram diluídas para uma concentração de 0,2 ng/μL e 5 μL desta diluição (1ng de DNA) foram utilizados para a preparação de bibliotecas por meio do Illumina Nextera®XT DNA Sample Preparation Kit (Illumina, San Diego, Estados Unidos), de acordo com o protocolo do fabricante. Foram utilizadas combinações de pares de indexes diferentes para cada biblioteca, sendo pareados os indexes: N501, 502, 503, 504, 505, 506 com N701, 702, 703, 704, 705, 706, 707, 708, 709, 710, 711, 712, respectivamente.

As bibliotecas foram quantificadas por PCR em tempo Real no equipamento ECO Real Time PCR, utilizando o Kapa Quantification kit (Illumina, San Diego, Estados Unidos) e diluídas para 2nM. Após isso, foram clusterizadas em flow cell através do equipamento cBot 2 System (Illumina, San Diego, Estados Unidos) e sequenciadas na plataforma HiSeq 2500 (Illumina, San Diego, Estados Unidos), utilizando a metodologia de *paired-end* 100X100.

4.4) Análise dos dados obtidos por bioinformática

Os dados gerados da corrida foram convertidos do formato BCL para o formato FASTQ pelo programa CASAVA (Illumina). A qualidade de leituras (*reads*) obtidas foi verificada por meio do software *Prinseq*, de maneira que amostras com score $Q < 20$ foram excluídas da análise. As reads foram mapeadas de acordo com o genoma de referência humano GRCh37, por meio do programa BWA (para os genes *MSH2*, *MSH6* e *MLH1*), e de acordo com a sequência de referência do gene *PMS2* (NC_000007.13: 6048737-6012870) e do pseudogene *PMS2CL* (NC_000007.13: 6774936- 6791232). Houve a remoção de leituras duplicadas por meio do *Prinseq*. O remapeamento e recalibração das bases foi realizado com o auxílio do *Genome Analysis Toolkit* (GATK). Listas de sítios variáveis e índices de cobertura também foram obtidos pelo GATK, delimitando-se as regiões sequenciadas para

cada um dos genes: *MSH2* (NM_000251.2), *MLH1* (NM_000249.3), *MSH6* (NM_000179.2), *PMS2* (NM_000535.5) e *PMS2CL* (NR_002217.1). Variantes com uma frequência do alelo alternativo <0,2 (para SNVs) ou <0,25 (para INDELS) e com o número total de leitura <30X foram removidas das análises. Os scripts para todos os passos realizados se encontram disponíveis no Anexo 13.

Sítios variáveis foram descritos de acordo com os bancos de dados dbSNP (<http://www.ncbi.nlm.nih.gov/SNP/>) e *Leiden Open Variation Database* v 3.0 (LOVD) (http://chromium.lovd.nl/LOVD2/colon_cancer/home.php?used_old_url), seguindo as instruções da *Human Genome Variation Society* (<http://www.hgvs.org/>). A classificação das variantes foi realizada de acordo com as recomendações do Insight (<http://insight-group.org/variants/classifications/>), em: (1) Não patogênicas, (2) Possivelmente não patogênicas ou com pouca significância clínica, (3) Não classificadas ou de significado clínico desconhecido (VUS), (4) Possivelmente Patogênicas e (5) Patogênicas.

4.5) Modelos de predição de sítios variáveis *in-silico*

Para variantes *missense* de significado clínico não conhecido, foram utilizados três softwares de predição do efeito funcional que a substituição de um aminoácido poderia causar na estrutura da proteína: Polyphen-2 (<http://genetics.bwh.harvard.edu/pph2/>), *SIFT Human Protein* (http://sift.jcvi.org/www/SIFT_enst_submit.html) e *Protein Variation Effect Analyzer* (PROVEAN) (<http://provean.jcvi.org/index.php>).

O programa Polyphen calcula o efeito funcional de uma variante *missense* com base em dois modelos de predição: *HumDiv* e *HumVar*. O *HumDiv* é utilizado para a predição de mutações causadas em proteínas descritas pelo banco de dados *UniProtKB* (<http://www.ebi.ac.uk/uniprot>), que estão associadas com doenças mendelianas. O *HumVar* prediz alelos raros, regiões densas de mapeamento e análises de seleção natural para todas as doenças humanas (ADZHUBEI *et al.*, 2013). As alterações podem ser classificadas como: benignas (não causam nenhum efeito na proteína), possivelmente danosas (possivelmente afetam a estrutura e/ou função da proteína), provavelmente danosas (a probabilidade de que as alterações afetem a estrutura e/ou função da proteína é alta) ou não conhecidas (não existem dados suficientes para prever o efeito funcional na proteína, também conhecidas como VUS: *Variants of Uncertain Significance*). A classificação é realizada com base nos *scores* gerados de *HumVar* e *HumDiv*, de maneira que quanto mais próximo de 0, maior a probabilidade de que a variante seja benigna e quanto mais próximo de 1, maior a probabilidade de que seja patogênica.

O programa SIFT é capaz de distinguir variantes envolvidas com algumas doenças de alterações neutras, com base no tipo de aminoácido que está sendo trocado e também na posição em que a substituição ocorre (características físico-químicas). As alterações foram classificadas como deletérias, capazes de alterar a função da proteína ($Score \leq 0.05$) ou toleradas ($Score > 0.05$), que não comprometem a função da proteína (NG; HENIKOFF, 2003). Os transcritos utilizados como referência para este programa foram disponíveis pelo banco de dados Ensembl (<http://www.ensembl.org/index.html>).

O programa PROVEAN é utilizado para prever se substituições de aminoácidos que são capazes de alterar a função de uma proteína. O *software* utiliza o BLAST para alinhar a variante de interesse e seleciona até 30 grupos que possuem identidade de pelo 75% com a sequência de referência relacionada à esta variante. A partir deste alinhamento, são gerados scores que predizem o efeito deletério de uma determinada alteração no DNA. Scores PROVEAN $\leq -2,5$ indicam que a variante tem um efeito deletério e é capaz de alterar a função da proteína, enquanto $scores > -2,5$ indicam que a variante é neutra (CHOI, 2012).

4.6) Sequenciamento de Sanger

A metodologia de Sequenciamento de Sanger foi utilizada para a validação das alterações patogênicas encontradas nos genes de MMR por NGS e também para o rastreamento molecular dos familiares de probandos alterados para um dos genes de interesse.

As amostras foram sequenciadas no sequenciador automático *ABI PRISM 3700 Genetic Analyzer* (Applied Biosystems, Clusbad, CA, Estados Unidos), utilizando-se de 3,2 pmol dos iniciadores direto ou reverso e por meio do *Bigdye Terminator Cycle Sequencing kit*, de acordo com o protocolo recomendado pelo fabricante. Cada um dos centros participantes do projeto ficou responsável pelo sequenciamento de pelo menos um dos genes. Para o gene *PMS2*, os produtos de LR-PCR foram previamente submetidos às reações de *Nested-PCR* com os iniciadores disponíveis no Anexo 7.

Os eletroferogramas gerados foram visualizados no *ChromasPro DNA Sequencing Software v 2.0* (Technelysium) e os contigs gerados foram alinhados contra as sequências de referência dos genes *MSH2* (NG_007110.1), *MSH6* (NG_007111.1), *MLH1* (NG_007109.2) e *PMS2* (NG_008466.1) no programa *Molecular Evolutionary Genetics Analysis (MEGA)* v.6. Foram consideradas validadas as alterações que foram encontradas por meio das metodologias de *Sanger* e NGS.

4.7) Análise de custos

O custo para o sequenciamento dos quatro genes MMR por meio das metodologias de Sanger e NGS foi avaliado. O levantamento dos preços foi realizado por meio de cotações ou consultando as páginas na internet de cada uma das empresas fornecedoras, entre o período de agosto e setembro/2016. Apenas os preços de consumíveis foram incluídos, sendo que fatores como energia elétrica, manutenção dos equipamentos e mão de obra não foram contabilizados.

5) RESULTADOS

5.1) Reações de PCR multiplex e LR-PCRs

Foram padronizadas PCRs multiplexes para os genes *MSH2* (seis reações), *MSH6* (quatro reações) e *MLH1* (seis reações), conforme demonstrado na Tabela 5.1 e na Figura 5.1. A amplificação do gene *PMS2* foi realizada por meio de 3 reações de longo alcance e 2 PCRs convencionais para eliminar a interferência de pseudogenes similares aos éxons 1-5, 9 e 11-15 do gene funcional (TOMSIC *et al.*, 2013).

A amplificação éxon por éxon dos genes de MMR necessitaria de no mínimo 63 reações de PCR por paciente para os quatro genes MMR: 16 para *MSH2*, 13 para *MSH6*, 19 para *MLH1* e 15 para *PMS2*. Com a padronização de multiplexes e LR-PCRs, foi possível reduzir de 63 para 21 o número de reações realizadas por pacientes para amplificação dos quatro genes de MMR.

Tabela 5.1: Reações de PCRs multiplexes e de longo alcance padronizadas para os genes de MMR.

Gene	Mixes	Éxons
<i>MSH2</i>	1.1	9, 12, 15
	1.2	4, 7
	1.3	6, 13
	1.4	5, 8
	2	3, 10, 11, 14, 16
	3	1, 2
<i>MLH1</i>	4	5, 9, 19
	5	2, 4, 7, 10, 15, 18
	6	1, 6, 11
	7	14, 17
	7.2	13, 16
	7.6	3, 8, 12
<i>MSH6</i>	8	3, 6, 7
	9	2, 4, 5
	10	8, 9, 10
	11	1
<i>PMS2</i>	LR-1	1 ao 5
	E6	6
	LR-2	7 ao 9
	E10	10
	LR-3	11 ao 15

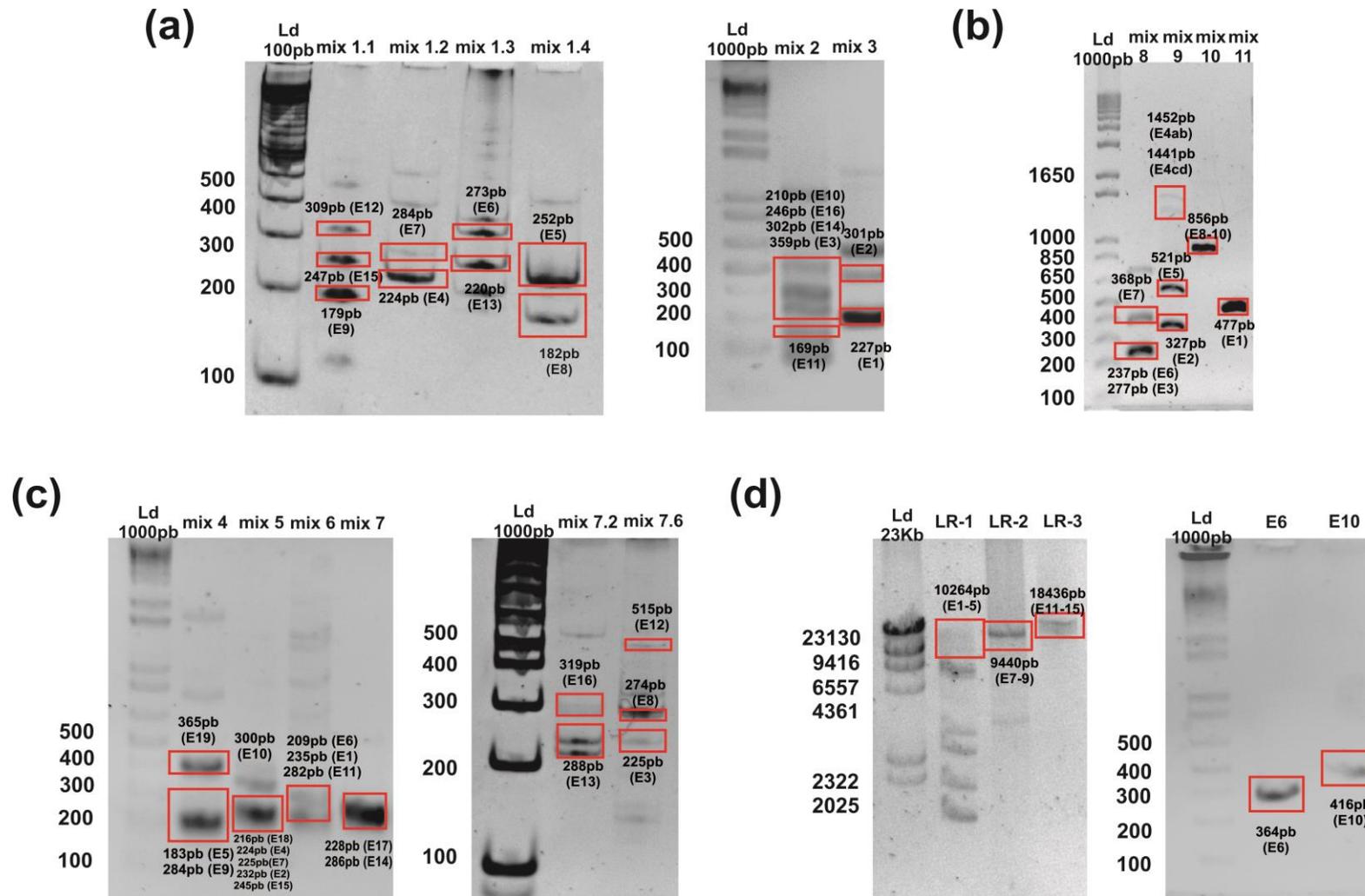


Figura 5.1: Reações de PCRs multiplex e de Longo Alcance padronizadas para os genes (a) *MSH2*, (b) *MSH6*, (c) *MLH1* e (d) *PMS2*. Os quadrados vermelhos indicam os amplicons esperados para cada PCR multiplex, acompanhado do tamanho de cada um dos fragmentos em pares de base (pb).

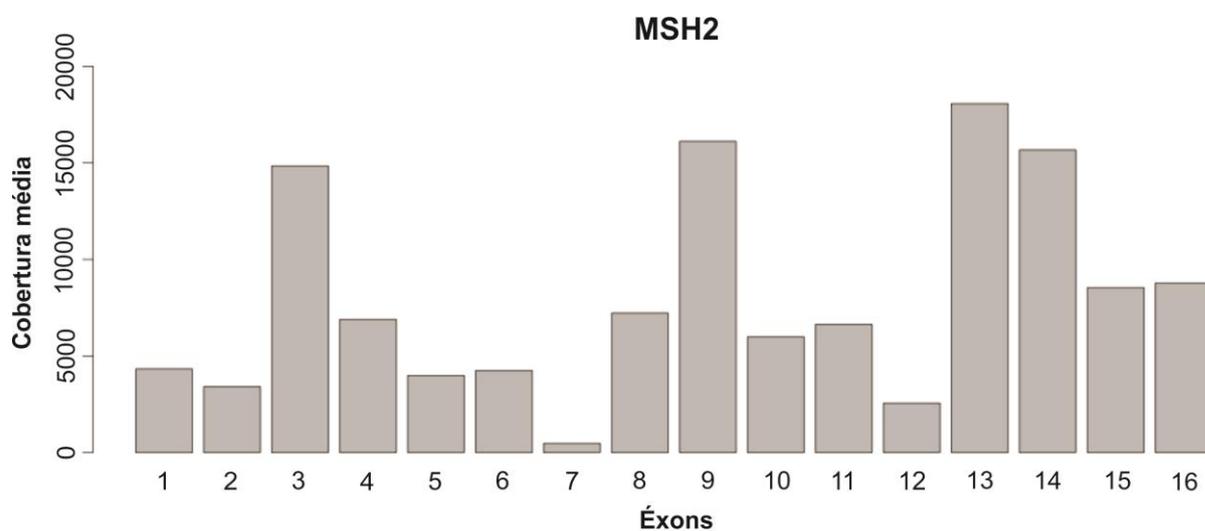
5.2) Corrida de sequenciamento no HiSeq2500

Uma única corrida foi realizada e todas bibliotecas de DNA foram clusterizadas em *flow cell* com uma densidade de 1003 k/mm³. No total, 369,98 milhões de clusters foram gerados. Destes clusters, 91,52% foram lidos em dois sentidos (2 X 169,03) pelo *HiSeq2500*. Em média 89,2% das leituras apresentaram *score* de qualidade superior a 30, sendo 91,70% no sentido senso e 85,90% no sentido anti-senso (Anexo 14). Em média, 4.385.195 sequências de leituras foram obtidas por amostra (variando de 3150 a 236.794.640) e aproximadamente 97% foram mapeadas nas regiões de interesse do genoma humano GRCh37.

A profundidade média de cobertura por éxon para os pacientes foi de 13.538 vezes, com um valor mínimo de 165 e máximo de 138.617 vezes. Os amplicons sequenciados para *MSH2*, *MSH6*, *MLH1* e *PMS2* tiveram coberturas médias de 7.988, 36.313, 11.899 e 4.772 vezes, respectivamente. As Figuras 5.2, 5.3, 5.4 e 5.5 ilustram as profundidades médias de cobertura obtidas por éxon e regiões flanqueadoras para cada um dos genes sequenciados e o Anexo 15 demonstra a cobertura média por gene para cada um dos pacientes.

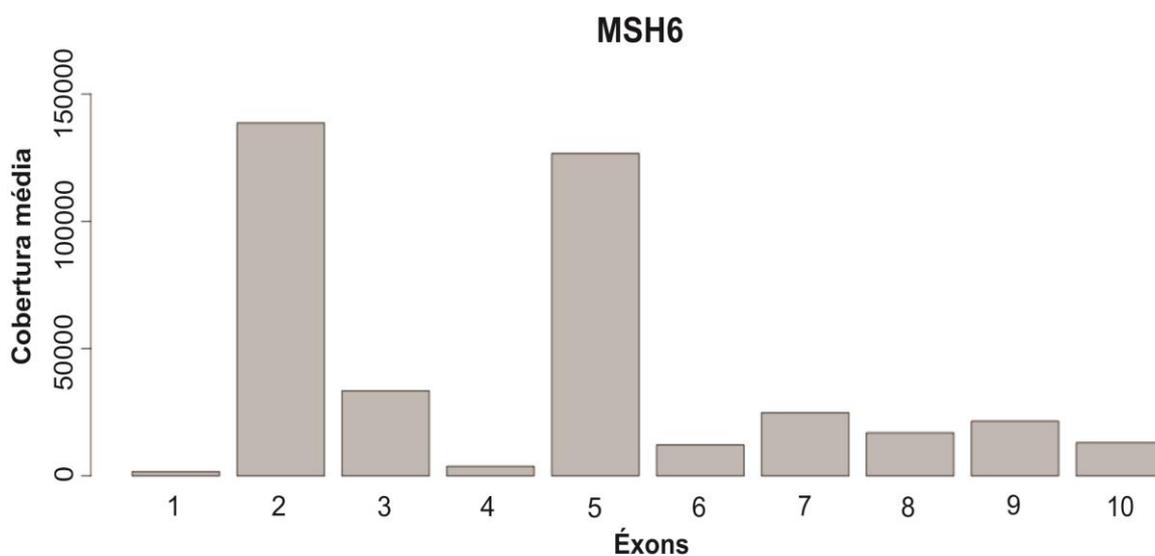
Os éxons 1, 2, 7 e 10 do gene *MSH2* apresentaram baixa profundidade média de cobertura (<30X) para 33% (22/65), 18% (12/54), 7,6% (5/65) e 4,6% (3/65) dos pacientes, respectivamente. O éxon 1 de *MSH6* apresentou baixa profundidade média de cobertura (<30X) para 8,0% (5/52) dos pacientes sequenciados para esta região. No total, 3,1% dos fragmentos que foram sequenciados apresentaram profundidade média de cobertura <30X e não foram eficientes na detecção de sítios variáveis nos genes MMR. O número de pacientes que apresentaram cobertura média igual a 0 está demonstrada nas Figura 5.2; 5.3; 5.4 e 5.5.

Um paciente (SDL) apresentou profundidade média de cobertura igual a 0 para todas os genes. Acredita-se que a ausência de cobertura para este paciente foi um erro metodológico e por isso este motivo ele foi eliminado das análises de NGS. O Sequenciamento de Sanger foi realizado para este paciente, que demonstrou ausência de alterações patogênicas nos genes *MSH2*, *MSH6* e *MLH1*.



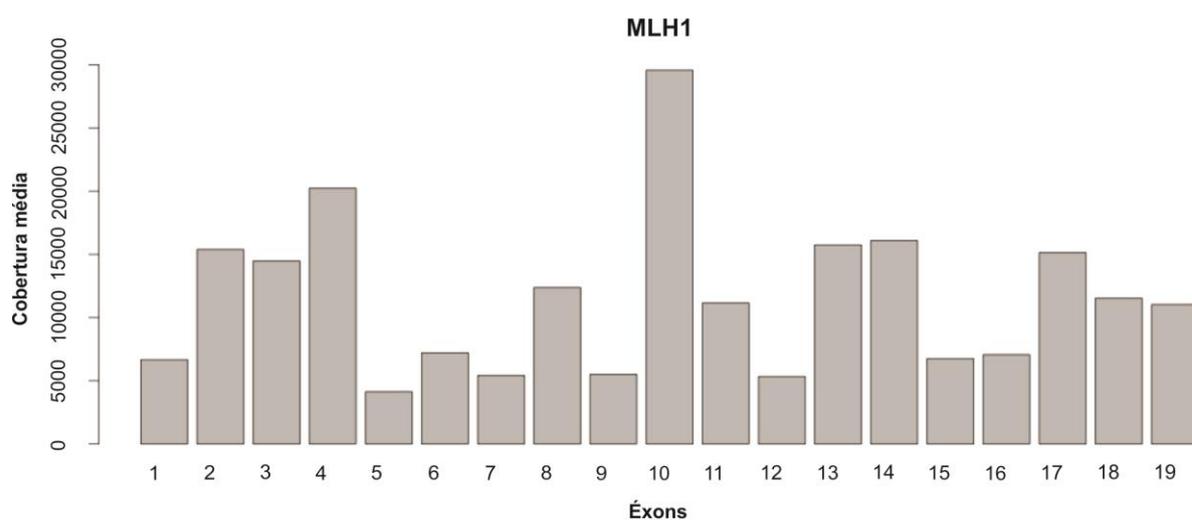
Gene	Éxons	Média cobertura éxon	Desvio Padrão	Min	Max	N pacientes DP=0
MSH2	E1	4336,71	6690,15	0	30291	2
	E2	3412,65	4613,97	1	23513	0
	E3	14838,37	18024,80	5	90304	0
	E4	6893,67	8446,72	438	40170	0
	E5	3989,52	4778,73	111	29616	0
	E6	4248,49	4477,03	197	24892	0
	E7	476,70	906,68	5	5125	0
	E8	7231,47	8043,76	34	49459	0
	E9	16119,50	15628,95	937	85817	0
	E10	5992,36	7431,98	1	39837	0
	E11	6639,49	8154,90	0	39470	0
	E12	2560,44	3440,24	34	16015	0
	E13	18066,30	21728,66	271	106146	0
	E14	15675,59	18820,67	3	87981	0
	E15	8539,36	9700,61	350	50978	0
	E16	8781,19	13243,52	0	88195	1

Figura 5.2: Profundidade média de cobertura observada por éxon, profundidades mínimas e máximas de cobertura para cada um dos éxons e pacientes que apresentam cobertura média por éxon igual a 0 para o gene *MSH2*. Legenda: N pacientes DP=0: Número de pacientes com profundidade média de cobertura igual a 0; as células marcadas em cinza representam as profundidades médias mínimas e máximas de cobertura para cada um dos éxons.



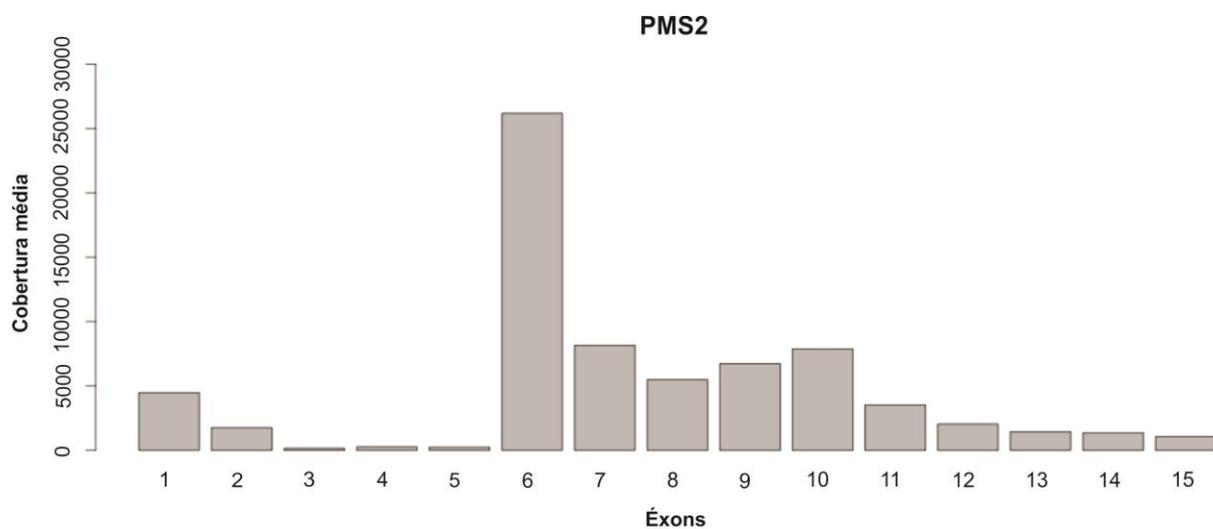
Gene	Éxons	Média cobertura éxon	Desvio Padrão	Min	Max	N pacientes DP=0
MSH6	E1	1695,97	2649,17	0	14492	2
	E2	138617,00	115298,06	8708	415234	0
	E3	33410,30	34453,89	18	126148	0
	E4	3801,57	3672,66	94	15333	0
	E5	126620,88	111889,22	6241	383255	0
	E6	12184,21	16559,22	1	106674	0
	E7	24801,10	27378,85	12	112860	0
	E8	16986,73	30089,92	4	197973	0
	E9	21561,48	38670,79	3	256424	0
	E10	13097,77	22890,04	3	148170	0

Figura 5.3: Profundidade média de cobertura observada por éxon, profundidades mínimas e máximas de cobertura para cada um dos éxons e pacientes que apresentam cobertura média por éxon igual a 0 para o gene *MSH6*. Legenda: N pacientes DP=0: Número de pacientes com profundidade média de cobertura igual a 0; as células marcadas em cinza representam as profundidades médias mínimas e máximas de cobertura para cada um dos éxons.



Gene	Éxons	Média cobertura éxon	Desvio Padrão	Min	Max	N pacientes DP=0
MLH1	E1	6664,90	6852,07	41	27483	0
	E2	15393,76	19502,97	155	86445	0
	E3	14473,56	17508,69	138	109212	0
	E4	20233,29	24835,45	413	118288	0
	E5	4142,26	4691,73	3	20986	0
	E6	7206,31	7967,71	55	33556	0
	E7	5423,40	6889,66	26	38507	0
	E8	12378,68	14888,92	84	96908	0
	E9	5504,00	7363,79	8	46793	0
	E10	29569,59	37431,20	133	163293	0
	E11	11158,07	13396,65	70	71244	0
	E12	5338,40	7000,93	158	31330	0
	E13	15743,33	15890,99	392	90042	0
	E14	16089,38	20589,52	567	103020	0
	E15	6751,73	8511,25	265	39347	0
	E16	7071,62	9388,16	237	41351	0
	E17	15139,04	17713,12	819	76832	0
	E18	11532,23	14690,90	170	80593	0
	E19	11041,22	16621,91	7	103193	0

Figura 5.4: Profundidade média de cobertura observada por éxon, profundidades mínimas e máximas de cobertura para cada um dos éxons e pacientes que apresentam cobertura média por éxon igual a 0 para o gene *MLH1*. Legenda: N pacientes DP=0: Número de pacientes com pr profundidade média de cobertura igual a 0; as células marcadas em cinza representam as profundidades médias mínimas e máximas de cobertura para cada um dos éxons.



Gene	Éxons	Média cobertura éxon	Desvio Padrão	Min	Max	N pacientes DP=0
PMS2	E1	4477,81	6614,08	0	29956	1
	E2	1766,71	2559,39	0	13420	2
	E3	165,28	422,06	0	2607	3
	E4	274,52	916,15	0	5734	3
	E5	249,50	974,60	0	6116	4
	E6	26177,87	32273,18	0	136356	1
	E7	8153,60	12929,83	0	71658	1
	E8	5491,87	8066,17	0	43352	1
	E9	6736,21	10173,07	0	52500	2
	E10	7873,16	7859,28	0	31051	1
	E11	3525,20	5100,85	0	27886	2
	E12	2055,77	4645,28	0	28033	2
	E13	1447,08	1879,91	0	7265	2
	E14	1362,59	1748,43	0	7706	2
	E15	1072,75	1507,75	0	6401	1

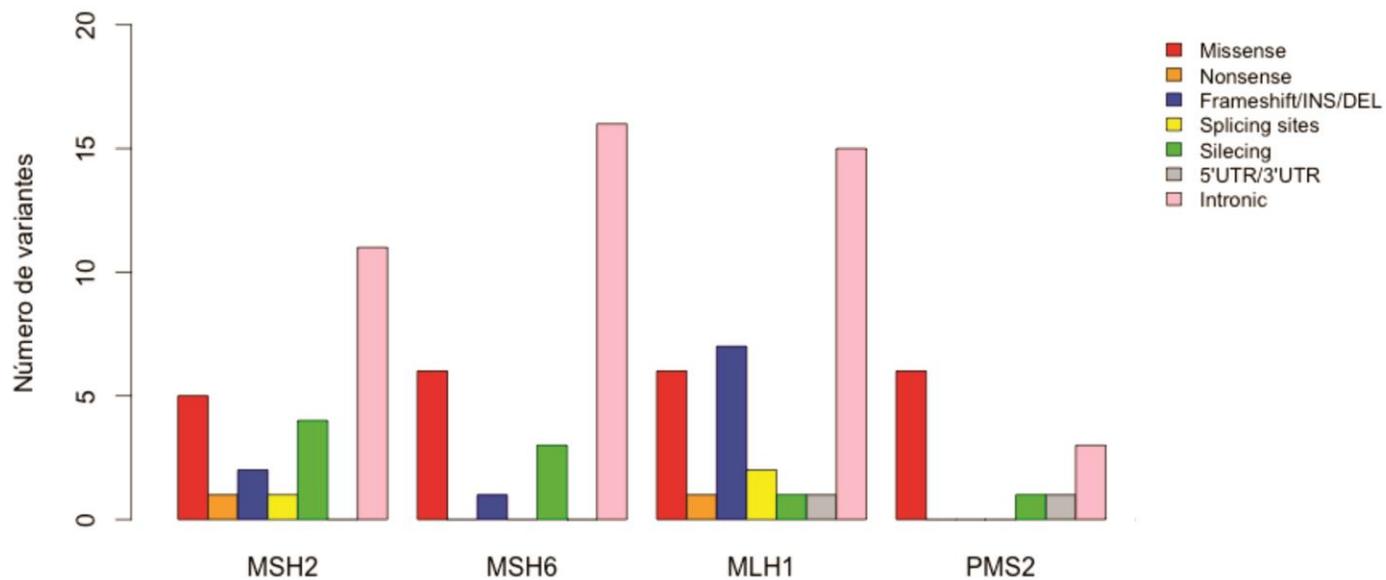
Figura 5.5: Profundidade média de cobertura observada por éxon, profundidades mínimas e máximas de cobertura para cada um dos éxons e pacientes que apresentam cobertura média por éxon igual a 0 para o gene *PMS2*. Legenda: N pacientes DP=0: Número de pacientes com profundidade média de cobertura igual a 0; as células marcadas em cinza representam as profundidades médias mínimas e máximas de cobertura para cada um dos éxons.

5.3) Resultados de Sequenciamento de Nova Geração

Noventa e seis sítios variáveis foram encontradas entre as regiões codificantes e flanqueadoras dos genes de MMR, sendo 24, 28, 32 e 12 alterações para os genes *MSH2*, *MSH6*, *MLH1* e *PMS2*, respectivamente. Do número total de variantes encontradas, vinte e cinco (25%) foram classificadas como mutações patogênicas ou de significado clínico desconhecido em regiões codificantes ou em sítios de *splicing*: 7 em *MSH2* (7,3% do total), 5 em *MSH6* (5,2% do total), 12 em *MLH1* (12,5% do total) e 1 em *PMS2* (1%). No total, trinta e dois pacientes foram afetados com variantes patogênicas e VUS nos genes MMR. As variantes encontradas em cada um dos genes serão descritas nas seções abaixo.

De todas as alterações encontradas, as variantes intrônicas (41,7%), de troca de sentido ou *missense* (23,3%) e silenciosas (12,6%) foram os tipos mais frequentes. Entre as alterações patogênicas, mutações *missense* (9,7%) e *frameshift* (7,7%) foram as mais comuns, seguidas de variantes em sítio de *splicing* (2,9%) e sem sentido (1,9%). A Figura 5.6 ilustra estes resultados.

Dos 35 indivíduos que apresentaram mutações patogênicas, incluindo as variantes encontradas por sequenciamento e por MLPA (metodologia realizada no Hospital de Clínicas de Porto Alegre), em *MSH2*, *EPCAM*, *MLH1* ou *MSH6*, vinte (57,1%) tiveram perda de expressão de pelo menos uma das proteínas de MMR e 13 (37,2%) apresentaram expressão normal, sendo dois resultados inconclusivos. Além disso, doze pacientes que apresentaram perda de expressão de pelo menos uma das quatro proteínas de MMR, não apresentaram alteração gênica quando sequenciados por NGS ou MLPA (ver Anexo 3).



Mutação	MSH2	MSH6	MLH1	PMS2
Missense	5	7	6	6
Nonsense	1	0	1	0
Frameshift/INS/DEL	2	1	5	0
Splicing	1	0	2	0
Silenciosas	4	4	1	2
5'UTR/3'UTR	0	0	1	1
Intrônicas	11	16	16	3
Total	24	28	32	12

Figura 5.6: Frequências das variantes identificadas nos genes *MSH2*, *MSH6*, *MLH1* e *PMS2* por NGS. As variantes foram subdivididas em *missense*, *nonsense*, *frameshift*, alterações em sítios de *splicing*, silenciosas e variantes encontradas na 5'/3' UTR. As variantes intrônicas incluem variantes de substituição de aminoácidos ou INDELS.

5.5.1) Variantes encontradas no gene *MSH2*

No total, foram identificadas 24 alterações no gene *MSH2*, sendo: 9 classificadas como benignas, 8 VUS presentes em regiões intrônicas, 6 classificadas como patogênicas e 1 VUS em regiões exônica. Os pacientes com variantes patogênicas e VUS em regiões codificantes tinham alterações nos éxons 3, 5, 6, 11 e 13 de *MSH2*. Destes pacientes, nove demonstraram perda de expressão de pelo menos uma das proteínas de MMR por IHQ e 3/4 avaliados por MSI apresentaram alta instabilidade em 4 ou 5 dos marcadores moleculares investigados. O paciente 2318 não foi rastreado por meio de IHQ. As Tabelas 5.2 e 5.3 fornecem informações a respeito das variantes benignas, patogênicas e VUS encontradas no estudo. Resultados sobre as predições *in silico* realizadas são descritos na Tabela 5.4.

Entre as variantes benignas encontradas nas regiões codificantes e íntrons de *MSH2*, sete eram mutações silenciosas e duas foram alterações *missense*. Onze VUS foram INDELS localizadas em íntrons com regiões microssatélites.

Entre as variantes patogênicas e VUS identificadas em regiões codificantes, foram encontradas: uma variante nonsense no éxon 13 c.2152C>T (p.Q718X), três variantes *missenses* nos éxons 6 (c.1046C>G - p.P349R) e 13 (c.2078G>A - p.C693Y e c.2021G>A - p.G674D), duas alterações *frameshift* nos éxons 3 (c.388_389delCA - p. Q130Vfs*2) e 11 (c.1738_1741delGAA - p.E580Lfs) e uma variante em sítio de reconhecimento de *splicing* no éxon 5 (c.942+3A>T - p.V265_Gln314del).

Dos doze pacientes que apresentaram alterações patogênicas ou VUS em regiões codificantes de *MSH2*, seis (50%) herdaram a mutação c.2152C>T - p.Q718X (pacientes 874/1112/2106/871/2309/2318), encontrada em heterozigose para todos os pacientes, com cobertura média para o alelo alterado de 1276,66 vezes.

Das variantes *missense* encontradas, uma era VUS c.2078G>A (p.C693Y), presente em heterozigose no paciente 1108. Análises de predição *in silico* (Tabela 5.4) demonstram que a variante é capaz de promover a troca de uma Cisteína na posição 693 da proteína por uma Tirosina (C693Y) no domínio *ATPásico* de *MSH2*, podendo ser considerada danosa/provavelmente danosa pelos três programas de predição utilizados.

A única alteração identificada em sítio de reconhecimento *splicing* foi a alteração c.942+3A>T (p.V265_Gln314del), capaz de provocar deleção completa do éxon 5. A alteração foi encontrada em homozigose e alelo alterado estava presente em 93,8% das leituras mapeadas nesta região. As variantes encontradas em regiões codificantes do gene *MSH2* tiveram uma cobertura média de 1052,75 vezes para o alelo alterado, com um mínimo

de 90 e um máximo de 2171 vezes. Os mapas e eletroferogramas para cada uma das variantes patogênicas encontradas para *MSH2* são apresentados nos Anexos 16,17 e 18.

Tabela 5.2: Alterações benignas e VUS encontradas nas regiões codificantes e flanqueadoras de *MSH2*.

Região	Het/Hom	ID(dbDSNP)	cDNA	Proteína	Domínio	Classificação	Número de pacientes alterados
Íntron 1	Het	rs2303426	c.211+9C>G	-	-	Benigna	15
Éxon 1	Het	-	c.193C>T	p.(=)	<i>Mismatch</i>	Benigna	1
Éxon 2	Het	rs35898375	c.339G>A	p.(=)	<i>Mismatch</i>	Benigna	1
Éxon 3	Het	rs17217772	c.380A>G	p.N127S	<i>Mismatch</i>	Benigna	2
Éxon 3	Het	rs1800151	c.573C>T	p.(=)	<i>Mismatch</i>	Benigna	1
Éxon 6	Het	rs4987188	c.965G>A	p.G322D	<i>Lever</i>	Benigna	4
Íntron 6	Het	rs17224360	1077-10T>C	p.(=)	-	Benigna	3
Éxon 11	Het	rs200056411	c.1680T>C	p.(=)	<i>Lever</i>	Benigna	1
Íntron 2	Het	rs201372136	c.212-16delT	-	-	VUS	49
Íntron 5	Het	rs63750704	c.942+3_942+5delAAA	-	-	VUS	5
Íntron 5	Het	rs780636964	c.942+3delA	-	-	Benigna	1
Íntron 5	Het	-	c.942+3_942+4delAA	-	-	VUS	2
Íntron 5	Het	-	c.942+3_942+10delAAAAAAAA	-	-	VUS	2
Íntron 5	Het	-	c.942+7_942+10delAAA	-	-	VUS	4
Íntron 5	Het	-	c.942+3_942+7delAAAA	-	-	VUS	3
Íntron 5	Het	rs768805063	c.942+4_942+7delAAAA	-	-	VUS	1
Íntron 5	Het	-	c.942+6delA	-	-	VUS	1

Legenda: Het/Hom: alteração presente em heterozigose (Het) ou homozigose (Hom).

Tabela 5.3: Alterações patogênicas ou de significado incerto encontradas em regiões codificantes e em sítios de splicing do gene *MSH2*.

Caso	Centro	Tumor primário (idade)	Éxon	Het/Hom	ID(dbDSNP)	cDNA	Proteína	Domínio	Classificação	DP	Cobertura alelo alterado (%)
874	HCPA	Cólon sigmóide (29)	13	Het	rs587779139	c.2152C>T	p.Q718X	<i>ATPase</i>	Patogênica	818	422 (51,6)
1108	HCPA	Cólon ascendente (42)	13	Het	-	c.2078G>A	p.C693Y	<i>ATPase</i>	VUS	2994	1563 (52,2)
1109	HCPA	Retossigmóide (30)	6	Het	rs587779067	c.1046C>G	p.P349R	<i>Lever</i>	Patogênica	973	475 (48,8)
1112	HCPA	Reto (34)	13	Het	rs587779139	c.2152C>T	p.Q718X	<i>ATPase</i>	Patogênica	4096	2010 (49,1)
2114	HCPA	Cólon ascendente (29)	11	Het	-	c.1738_1741delGAAA	p.E580Lfs	<i>Lever</i>	Patogênica	1388	364 (26,2)
2106	HCPA	Retossigmóide (50)	13	Het	rs587779139	c.2152C>T	p.Q718X	<i>ATPase</i>	Patogênica	4433	2171 (49,0)
871	HCPA	Cólon ascendente (30)	13	Het	rs587779139	c.2152C>T	p.Q718X	<i>ATPase</i>	Patogênica	342	152 (44,4)
2309	HCPA	Cólon descendente e sigmóide (44)	13	Het	rs587779139	c.2152C>T	p.Q718X	<i>ATPase</i>	Patogênica	2250	1089 (48,4)
2318	HCPA	Cólon e estômago (66)	13	Het	rs587779139	c.2152C>T	p.Q718X	<i>ATPase</i>	Patogênica	3633	1816 (50,0)
8	ACCA M	Cólon sigmóide (44)	13	Het	rs267607996	c.2021G>A	p.G674D	<i>ATPase</i>	Patogênica	2866	1252 (43,7)
MRV	HUJJB	Cólon descendente (21)	5	Hom	rs193922376	c.942+3A>T	p.V265_Gln314del	<i>Connector</i>	Patogênica	96	90 (93,8)
SL9	HUJJB	Cólon ascendente (46)	3	Het	rs63750704	c.388_389delCA	p.Q130Vfs* ₂	<i>Connector</i>	Patogênica	2630	1229 (46,7)

Legenda: Het/Hom: alteração presente em heterozigose (Het) ou homozigose (Hom).

Tabela 5.4: Predição *in silico* para a variante c.2078G>A presente no éxon 13 do gene *MSH2*.

Éxon	cDNA	Ptn	P	SIFT Prediction		PROVEAN		Polyphen		
				S	Predição	S	Predição	Hum Div	Hum Var	Predição
13	c.2078G>A	p.C693Y	1108	0	Danosa	-8,7	Danosa	0,1	1,0	Prov. Danosa

Legenda: S:score; P:paciente; Ptn: Proteína; Prov.: provavelmente danosa.

5.5.2) Variantes encontradas no gene *MSH6*

Foram encontradas 28 alterações no gene *MSH6*, sendo: 14 classificadas como benignas, 1 provavelmente não patogênica e 8 VUS localizadas em regiões intrônicas. Além disso, foram identificadas cinco VUS nos éxons 1, 2, 4 e 8. Dos pacientes que apresentaram VUS em regiões codificantes, dois não apresentaram perda de expressão de nenhuma das proteínas de MMR e dois tiveram perda de expressão da proteína MLH1. Um único paciente apresentou perda de expressão de *MSH6*. O resultado de IHQ para a proteína *MSH2* no paciente 2278 foi inconclusivo. A instabilidade das regiões de microsatélites foi avaliada para apenas um dos indivíduos (903) com VUS para *MSH6*, entretanto este paciente demonstrou não ter instabilidade (MMS). As Tabelas 5.5 e 5.6 fornecem, respectivamente, informações a respeito das variantes benignas e patogênicas identificadas em *MSH6*. As predições *in silico* realizadas para os pacientes com VUS foram descritas na Tabela 5.7.

Das alterações benignas identificadas em *MSH6*, quatro foram mutações silenciosas e duas foram alterações *missense* em regiões codificantes. As outras dezesseis alterações encontradas foram divididas em alterações silenciosas ou INDELS de significado desconhecido, localizadas nos íntrons e em regiões microsatélites de DNA. As alterações patogênicas e VUS identificadas nos éxons foram divididas entre quatro variantes *missense* nos éxons 1 (c.181G>A - p.A61T), 4 (c.719G>A - p.R240Q e c.2006T>C - p.I669T) e 8 (c.3772C>G - p.Q1258E) e uma variante *frameshift* no éxon 2 (c.402dupT - p.F133fs).

Quatro das cinco VUS exônicas, encontradas em heterozigose, foram *missense* e preditas como benignas (c.181G>A- p.A61T; c.719G>A- p.R240Q; c.2006T>C- p.I669T e c.3772C>G - p.Q1258E) por pelo menos 2 dos 3 programas utilizados. Um único paciente (2278) apresentou uma alteração *frameshift* (c.402dupT - p.F133fs), presente em heterozigose no éxon 2 do gene *MSH6* e descrita pela primeira vez na literatura. A média de cobertura para os alelos alterados das variantes encontradas nas regiões codificantes de *MSH6* foi de 808,5

vezes. Os mapas do NGS e eletroferogramas (Sanger) para cada uma das variantes patogênicas de *MSH6* se encontram disponíveis nos Anexos 19 e 20.

Tabela 5.5: Alterações benignas e de significado incerto (VUS) encontradas nas regiões codificantes e flanqueadoras de *MSH6*.

Região	Het/Hom	ID (dbSNP)	cDNA	Proteína	Domínio	Classificação	Número de pacientes alterados
Éxon 1	Het	rs1042821	c.116G>A	p.G39E	-	Benigna	17
Éxon 1	Het	rs1042820	c.186C>A	p.(=)	-	Benigna	30
Íntron 1	Het	rs55927047	c.260+22C>G	p.(=)	-	Benigna	17
Íntron 4	Het	rs1800931	3173-101G>C	p.(=)	-	Benigna	2
Éxon 2	Het	rs1800932	c.276A>G	p.P92P	PWWP	Benigna	29
Éxon 2	Het	rs3211299	c.431G>T	p.S144I	PWWP	Benigna	1
Éxon 3	Het	rs1800935	c.540T>C	p.(=)	Connector	Benigna	33
Éxon 4	Het	rs56371757	c.2272C>T	p.(=)	Lever	Benigna	1
Íntron 3	Het	rs558111682	c.627+28A>T	p.(=)	-	Benigna	1
Íntron 4	Het/Hom	rs2072447	c.3173-101G>C	p.(=)	-	Benigna	57
Éxon 5	Het	rs2020910	c.3306T>A	p.(=)	-	Benigna	2
Íntron 5	Het/Hom	rs2020911	c.3438+14A>T	p.(=)	-	Benigna	40
Íntron 6	Het	rs189436849	c.3557-40T>A	p.(=)	-	Provavelmente não patogênica	1
Íntron 7	Hom	rs565682904	c.3646+62A>G	NA	-	VUS	1
Íntron 7	Het/Hom	rs3136359	c.3646+91T>C	p.(=)	-	Benigna	57
Íntron 8	Het/Hom	rs3136367	c.3801+54C>G	p.(=)	-	Benigna	57
Íntron 8	Het	rs373409282	c.3802-22G>A	-	-	VUS	1
Íntron 6	Het	rs749755872	c.3557-16insT	-	-	VUS	20
Íntron 6	Het	rs749755872	c.3557-16delT	-	-	VUS	41
Íntron 7	Het	rs2234731	c.3646+29_3646+32delCTAT	-	-	VUS	53
Íntron 9	Het	-	c.4002-27delT	-	-	VUS	56
Íntron 10	Het	rs7965066568	c.4002-27_4002-26delTT	-	-	VUS	24
Íntron 8	Het	rs142812704	c.3802-45_3802-44insT	-	-	VUS	1

Legenda: Het/Hom: Het/Hom: alteração presente em heterozigose (Het) ou homozigose (Hom).

Tabela 5.6: Alterações patogênicas ou de significado incerto encontradas em regiões codificantes do gene *MSH6*.

Caso	Centro	Tumor primário (idade)	Éxon	Het/Hom	ID (dbSNP)	cDNA	Proteína	Domínio	Classificação	DP	Cobertura alelo alterado (%)
903	HCPA	Retossigmóide (48)	1	Het	-	c.181G>A	p.A61T	-	VUS	59	14 (23,7)
2113	HCPA	Retossigmóide (36)	8	Het	rs63750554	c.3772C>G	p.Q1258E	<i>ATPase</i>	VUS	2592	1258 (48,5)
2278	HCPA	Endométrio (39)	2	Het	-	c.402dupT	p.F133fs	PWWP	VUS	841	421 (50,1)
SL9	HUJJB	Cólon ascendente (46)	4	Het	rs542848931	c.719G>A	p.R240Q	-	VUS	65	30 (46,2)
SL3	HUJJB	Reto (38)	4	Het	rs555209664	c.2006T>C	p.I669T	<i>Connector</i>	VUS	3285	1525 (46,4)

Legenda: DP: cobertura total na posição da variante; Het/Hom: heterozigose ou homozigose.

Tabela 5.7: Predições in silico para as variantes de significado incerto encontradas nas regiões codificantes do gene *MSH6*.

Gene	E	cDNA	Proteína	Paciente	SIFT Prediction		PROVEAN		Polyphen		
					S	Predição	S	Predição	HumDiv	HumVar	Predição
MSH6	1	c.181G>A	p.A61T	903	0,0043	Danosa	-0,35	Neutra	0,0	0,001	Benigna
MSH6	4	c.719G>A	p.R240Q	SL9	0,2	Tolerada	-0,1	Neutra	0,2	0,0	Benigna
MSH6	4	c.2006T>C	p.I669T	SL3	0,0	Danosa	-2,4	Neutra	0,1	0,1	Benigna
MSH6	16	c.3772C>G	p.Q1258E	2113	0,8	Tolerada	-0,1	Neutra	0,0	0,0	Benigna

Legenda: E: éxon; S:score.

5.5.3) Variantes encontradas no gene *MLH1*

No total foram encontradas 32 variantes no gene *MLH1*, sendo: 12 alterações benignas, 8 VUS em regiões intrônicas, 3 alterações patogênicas, 1 alteração provavelmente patogênica e 8 VUS em regiões exônicas. As alterações patogênicas e VUS localizadas em regiões codificantes estavam presentes nos éxons 3, 8, 10, 11, 15, 16 e 17. Seis dos indivíduos portadores de mutações patogênicas (50%) demonstraram perda de expressão de pelo menos uma das proteínas de MMR, e 5 dos 7 indivíduos avaliados por MSI, apresentaram 4 ou 5 marcadores alterados para as regiões de microssatélites. As Tabelas 5.8 e 5.9 descrevem as variantes benignas e patogênicas/VUS encontradas para cada um dos pacientes e a Tabela 5.10 demonstra os resultados de predição *in silico* realizados para as VUS.

Das alterações benignas identificadas em *MLH1*, dez foram descritas como silenciosas, duas eram alterações *missense* em regiões codificantes e uma INDEL foi encontrada na região 3' UTR de *MLH1*. Oito VUS eram INDELS (n=5) ou de substituição de bases (n=3) localizadas em regiões não codificantes e repetitivas.

Das alterações patogênicas e VUS identificadas no éxons ou sítios de *splicing*, cinco foram classificadas como variantes *missenses* presentes nos éxons 3 (c.277A>G - p.S93G), 4 (c.350C>T - p.T117R), 8 (c.677G>A - R226Q) e 11 (c.1007G>A - p.G336D), duas estavam presente no sítio de reconhecimento de *splicing* dos éxons 8 (c.677+1G>A) e 10 (c.791-6_794delGTTTAGATC); uma variante *nonsense* foi encontrada no éxon 17 (c.1975C>T - R659X) e cinco eram mutações *frameshift* localizadas nos éxons 3 (c.225delT - p.L75Vfs), 11 (c.888_889delAG - p.E297Afs), 15 (1681_1682insT - Y561Ls) e 16 (c.1863delG - p.M621Ifs) e uma alteração *missense* (c.1852_1853delAAinsGC - p.K618A).

As alterações c.791-6_794delGTTTAGATC, c.1863delG, c.225delT e 1681_1682insT, encontradas em heterozigose, ainda não foram descritas na literatura e não foram avaliadas por nenhum dos três programas de predição utilizados, por estarem presentes em sítio de *splicing* ou serem mutações *frameshift*.

Tabela 5.8: Alterações benignas e de significado incerto (VUS) encontradas nas regiões codificantes e flanqueadoras de *MLH1*.

Região	Het/Hom	ID(dbDSNP)	cDNA	Proteína	Domínio	Classificação	Número de pacientes alterados
Íntron 3	Het	rs139620056	c.307-29C>A	p.(=)	-	Benigna	1
Íntron 5	Het	rs4647246	c.453+25A>G	p.(=)	-	Benigna	1
Íntron 5	Het	rs4647255	c.454-51T>C	p.(=)	-	Benigna	1
Éxon 8	Het/Hom	rs1799977	c.655A>G	p.I219V	<i>Mismatch</i>	Benigna	27
Íntron 3	Het	rs139620056	c.307-29C>A	p.(=)	-	Benigna	1
Íntron 11	Het	rs59684491	c.1039-31A>T	-	-	VUS	1
Íntron 11	Het	rs6771325	c.1039-29A>T	-	-	VUS	18
Íntron 11	Het	rs193922367	c.1039-8T>A	p.(=)	-	Benigna	1
Íntron 13	Het	rs41562513	c.1558+14G>A	p.(=)	-	Benigna	3
Íntron 12	Het	rs7633154	c.1410-54C>T	p.(=)	-	Benigna	3
Íntron 14	Hom/Het	rs9876116	c.1668-19A>G	p.(=)	-	Benigna	28
Éxon 19	Het	rs2020873	c.2152C>T	p.H718Y	<i>C-terminal</i>	Benigna	1
Íntron 11	Het/Hom	rs763206692	c.1039-34_1039-26delTATATATTT	-	-	VUS	4
Íntron 11	Het	rs755228836	c.1732-38G>A	-	-	VUS	1
Éxon 17	Het	rs1800146	c.1959G>T	p.(=)	-	Benigna	2
3`UTR	Het	rs193922366	c.*35_*37delCTT	p.(=)	-	Benigna	1
Íntron 1	Het	rs752136033	c.117-46_117-42delTAGAG	-	-	VUS	1
Íntron 11	Het	.	c.1039-27insA	-	-	VUS	1

Legenda: Het/Hom: alteração presente em heterozigose (Het) ou homozigose (Hom).

Tabela 5.9: Alterações patogênicas ou de significado incerto (VUS) encontradas em regiões codificantes e em sítios de splicing do gene *MLH1*.

Caso	Centro	Tumor primário (idade)	E	Het/Hom	ID(dbDSNP)	cDNA	Proteína	Domínio	Classificação	DP	Cobertura alelo alterado (%)
864	HCPA	Cólon ascendente e íleo (51)	10	Het	-	c.791-6_794delGTTTAGATC	Perda éxon 10	C-terminal	VUS	3714	736 (20,0)
865	HCPA	Cólon ascendente (52)	16	Het	-	c.1863delG	p.M621Ifs	C-terminal	VUS	203	113 (55,7)
867	HCPA	Cólon transverso (48)	8	Het	rs267607778	c.677+1G>A	Perda éxon 8	N-terminal	Provavelmente patogênica	588	261 (44,4)
1085	HCPA	Cólon transverso (42)	3	Het	-	c.225delT	p.L75Vfs	HATPase	VUS	2823	1261 (44,7)
875	HCPA	Cólon ascendente (37)	17	Het	rs63751310	c.1975C>T	p.R659X	C-terminal	Patogênica	4568	2513 (55,0)
872	HCPA	Cólon ascendente:ceco (46)	11	Het	-	c.890A>C + c.888_889delAG	p.E297Afs	N-terminal	VUS	3917	1180 (30,1)
2113	HCPA	Retossigmoide (36)	15	Het	-	c.1681_1682insT	p.Y561Lfs	C-terminal	VUS	2592	1258 (48,5)
891	HCPA	Cólon descendente (40)	4	Het	rs63750781	c.350C>T	p.T117R	HATPase	Patogênica	3277	1511 (46,1)
ADR	INCA	Cólon ascendente (42)	8	Het	rs63751711	c.677G>A	p.R226Q	N-terminal	Patogênica	750	378 (50,4)
9	ACCA M	Cólon ascendente (34)	8	Het	rs63751711	c.677G>A	p.R226Q	N-terminal	Patogênica	2766	1191 (43,1)
FCG	INCA	Cólon transverso (39)	3	Het	rs41295282	c.277A>G	p.S93G	HATPase	VUS	2908	1552 (53,4)
16	ACCA M	Cólon ascendente (53)	11	Het	rs587781750	c.1007G>A	p.G336D	Linker	VUS	3894	2018 (51,8)
2106	HCPA	Retossigmoide (50)	16	Het	-	c.1852_1853delAAinsGC	p.K618A	C-terminal	VUS	1403	658 (46,9)
SL4	HUJJB	Reto (34)	16	Het	-	c.1852_1853delAAinsGC	p.K618A	C-terminal	VUS	1592	698 (43,8)

Legenda: E: éxon; Het/Hom: alteração presente em heterozigose (Het) ou homozigose (Hom).

Tabela 5.10: Predições in silico para as variantes missense de significado incerto encontradas nas regiões codificantes do gene *MLH1*.

E	cDNA	Proteína	Ptn	SIFT Prediction		PROVEAN		Polyphen		
				S	P	S	P	Hum Div	Hum Var	P
3	c.277A>G	p.S93G	FCG	0,0	Danosa	-2,8	Deletéria	0,0	0,1	Benigna
11	c.890A>C	p.E297A	872	0,0	Danosa	-4,7	Deletéria	0,9	0,6	Pos. Danosa
11	c.1007G>A	p.G336D	16	0,0	Danosa	-4,1	Deletéria	0,9	0,7	Pos. Danosa
16	c.1852_1853 delAAinsGC	p.K618A	2106/SL4	0,02	Danosa	-5,0	Deletéria	0,9	0,9	Prov. Danosa

Legenda: E: éxon; P: predição; Ptn: proteína; Pos. Danosa: possivelmente danosa; S:score.

Duas alterações VUS foram encontradas em *cis* no éxon 11 do gene *MLH1* para o mesmo paciente (872): c.890A>C (p.E297A) e c.888_889delAG, conforme demonstrado na Figura 5.7. A substituição c.890A>C (p.E297A), que ocorreu no códon 297, tem significado clínico desconhecido, mas é predita como danosa ou possivelmente danosa pelos três programas de predição utilizados. Por se tratar de mutação *frameshift*, a alteração c.888_889delAG não pôde ser predita pelos programas utilizados. As alterações são encontradas em heterozigose com cobertura média de 1180 vezes para os alelos alternativos e são capazes de alterar a formação do domínio N-terminal de MLH1.

Das alterações patogênicas e VUS encontradas nas regiões codificantes/sítios de *splicing* de *MLH1*, quatro foram descritas pela primeira vez no presente estudo (c.791-6_794delGTTTAGATC, c.1863delG, c.225delT e c.1681_1682insT), cinco já tinham significado clínico conhecido (patogênicas ou provavelmente patogênicas) e duas foram previamente descritas como VUS (c.1007G>A e c.277A>G). Os mapas do NGS e eletroferogramas (Sanger) para cada uma das variantes patogênicas de *MLH1* se encontram nos Anexos 21, 22, 23 e 24.

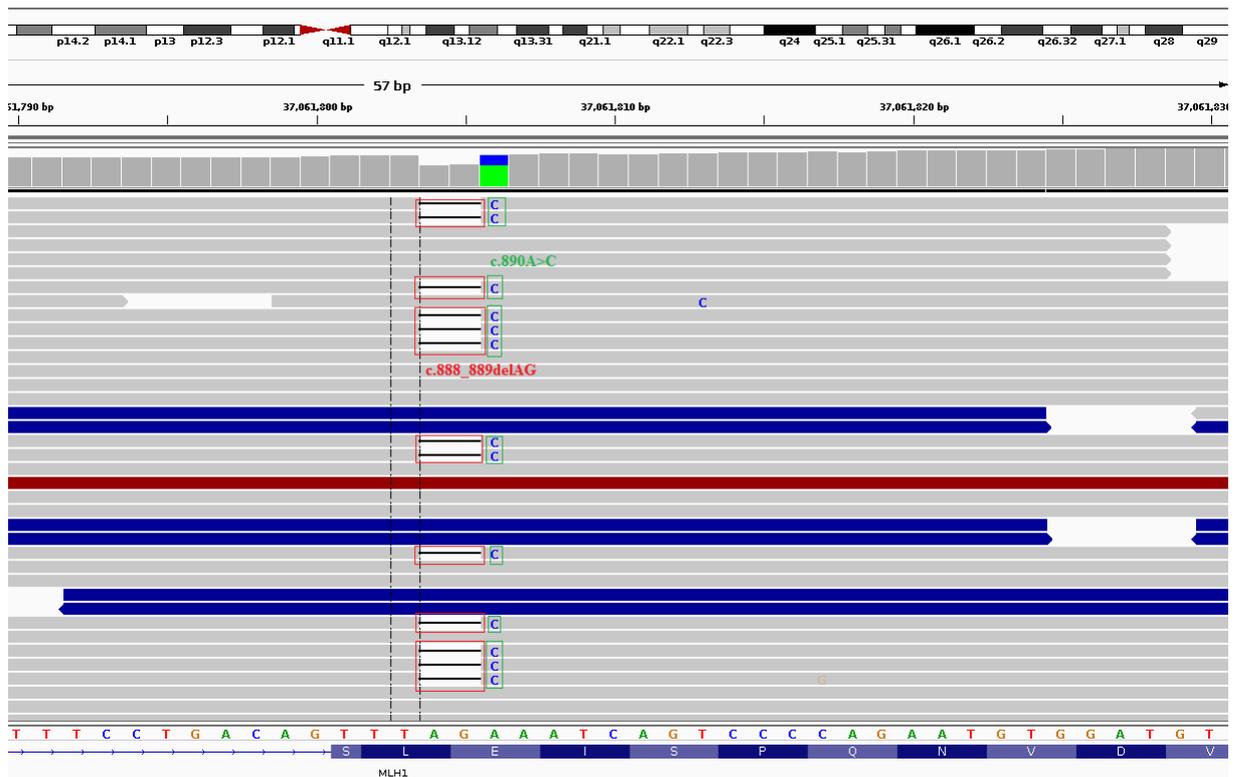


Figura 5.7: Mapa das leituras da região cromossômica L3: 37061803-37061806 do paciente 872, que possui duas alterações em cis no éxon 11 do gene *MLH1* (c.888_889delAG e c.890A>C), figura construída utilizando o programa *Integrative Genome Viewer* (IGV).

5.5.4) Variantes encontradas no gene *PMS2*

O mapeamento de *PMS2* foi realizado contra a sequência de referência do próprio gene, e não com o genoma completo, como foi feito para os genes *MSH2*, *MSH6* e *MLH1*. Esta estratégia foi necessária para evitar o pareamento das sequências de pseudogenes com alta homologia à *PMS2*, como exemplo de *PMS2CL*. A partir do *Variant Calling* realizado para as regiões exônicas e flanqueadoras de *PMS2*, no total foram encontradas 12 variantes presentes nos éxons 2, 7, 11 e 15 dos 41 pacientes rastreados para este gene. Destas alterações, onze foram previamente descritas como benignas e uma apresentou significado clínico desconhecido, como pode ser observado na Tabela 5.11.

Entre as alterações benignas, seis eram *missense* (c.2570C>G - p.G857A, c.1621T>C - p.E541K, c.1531A>G - p.T511A, c.1454G>T - p.T485K, c.1408G>A - p.P470S e c.59C>T - p.R20G) e estavam presentes nos éxons 2, 11 e 15, três eram alterações silenciosas nos éxons 7 (c.780C>G), 15 (c.2466A>G) e íntron 11 (c.2007-4G>A) e duas eram INDELS localizadas em regiões repetitivas do íntron 7 de *PMS2* (c.706-4delA e c.706-5_706-4delAA). A única VUS encontrada (c.-77delA), estava presente em homozigose na região 5'UTR em 30 pacientes, conforme mostrado na Tabela 5.10. A cobertura média dos alelos alterados foi de 142,93 vezes. Devido à localização da alteração (3' UTR), não foi possível prever *in silico* se a alteração compromete a função da proteína.

Tabela 5.11: Variantes benignas e de significado incerto (VUS) encontradas nos pacientes rastreados para o gene *PMS2*.

Região	Het/Hom	ID(dbDSNP)	cDNA	Proteína	Domínio	Classificação	Número de pacientes alterados
Éxon 15	Het	rs1802683	c.2570C>G	p.G857A	-	Benigna	12
Éxon 15	Het	rs10000	c.2466A>G	p.(=)	-	Benigna	2
Íntron 11	Het	rs140788589	c.2007-4G>A	p.(=)	-	Benigna	8
Éxon 11	Het	rs2228006	c.1621T>C	p.E541K	-	Benigna	33
Éxon 11	Het	rs2228007	c.1531A>G	p.T511A	-	Benigna	1
Éxon 11	Het	rs1805323	c.1454G>T	p.T485K	-	Benigna	1
Éxon 11	Het	rs1805321	c.1408G>A	p.P470S	-	Benigna	21
Éxon 7	Het	rs1805319	c.780C>G	p.(=)	-	Benigna	34
Éxon 2	Het	rs10254120	c.59C>T	p.R20G	-	Benigna	4
Íntron 7	Het	rs60794673	c.706-4delA	NA	-	Benigna	25
Íntron 7	Het	rs776641246	c.706-5_706-4delAA	NA	-	Benigna	14
5'UTR	Het/Hom	-	c.-77delA	-	-	VUS	30

Legenda: Het/Hom: alteração presente em heterozigose (Het) ou homozigose (Hom).

5.5.5) Análise do pseudogene *PMS2CL*

Com o objetivo de verificar o grau de similaridade das regiões amplificadas de *PMS2* com *PMS2CL*, utilizamos a estratégia de mapeamento das leituras geradas do NGS contra a sequência de referência do próprio pseudogene (gi|224589819: 6774936-6791232) e do gene funcional (gi|224589819: c6048737-6012870), sem alinhá-las contra o genoma humano completo (GRCh37), e posteriormente, realizar a identificação de variantes para as regiões exônicas de 1 a 6 de *PMS2CL*, parálogas aos éxons 9 e 11-15 do gene funcional. O objetivo desta estratégia foi forçar o alinhamento de possíveis leituras que estavam mapeando contra o gene funcional e verificar se o alinhamento contra o genoma humano completo estava realmente gerando resultados falso-positivos para *PMS2CL*.

O alinhamento contra a sequência de referência do próprio pseudogene, gerou profundidades médias de coberturas altas (~1917X), compatível a média gerada para a região paróloga do gene funcional (~2700X). Esta situação já era esperada, considerando a alta homologia (>98%) entre essas sequências e que esta análise forçou o mapeamento de todas as leituras correspondentes ao gene funcional na região paróloga de *PMS2CL*. As profundidades médias de coberturas para os éxons 1-6 de *PMS2CL* os éxons parálogos do gene funcional são demonstradas na Tabela 5.12.

Tabela 5.12: Comparação entre as profundidades médias de coberturas para os éxons 9, 11-15 de *PMS2* e as regiões correspondentes do pseudogene *PMSCL*.

PMS2				PMS2CL			
Éxons	Cobertura média	Mínimo	Máximo	Éxons	Cobertura média	Mínimo	Máximo
9	6736,21	0,00	52499,78	1	2152,12	9,27	999,25
11	3525,20	0,00	27886,26	2	2396,96	49,66	1613,23
12	2055,77	0,00	28032,74	3	1798,38	6,07	1430,68
13	1447,08	0,00	7265,21	4	1122,80	0,00	787,99
14	1362,59	0,00	7705,94	5	1431,85	4,08	972,81
15	1072,75	0,48	6401,18	6	2597,01	4,18	2033,73
Média global	2699,93				1916,52		

Após o alinhamento com a sequência de referência do próprio pseudogene, não foram identificadas variantes para *PMS2CL*.

5.7) Rastreamento molecular das famílias com SL

No total, foram rastreados 11 indivíduos pertencentes a 5 famílias distintas, que possuíam alterações patogênicas em *MSH2* ou *MLH1*. A mutação c.2152C>T, presente no

éxon 13 de *MSH2*, foi rastreada em três famílias (cinco familiares do paciente 1112, dois familiares do paciente 871 e um familiar do paciente 2106). A deleção c.1738_1741deGAAA no éxon 11 de *MSH2* foi rastreada em um familiar do paciente 2114. Um familiar do paciente 872 teve o éxon 11 de *MLH1* sequenciado para avaliar a presença das alterações em *cis* c.890A>C e c.888_889_delCA, conforme foi demonstrado na Tabela 5.13.

Tabela 5.13: Rastreamento molecular de familiares dos pacientes que apresentaram alterações patogênicas em *MSH2* e *MLH1*.

Centro	Família	Paciente	Parentesco probando	I	Câncer	Gene	E	Alteração rastreada	R
HCPA	Familiar 1112	2284 (III.6)	Tia materna	-	Não	MSH2	13	c.2152C>T (p.Q718*)	-
HCPA		2149 (IV.1)	Primo paterno	-	Não	MSH2			-
HCPA		1552 (IV.10)	Prima materna	30	Não	MSH2			+
HCPA		1553 (IV.2)	Irmã	-	Não	MSH2			-
HCPA		1561 (III.9)	Tio materno	50	CCR	MSH2			-
HCPA	Familiar 871	2235 (V.11)	Sobrinha	23	Não	MSH2	13	c.2152C>T (p.Q718*)	+
HCPA		1880 (III.1)	Tio materno	60	Não	MSH2			-
HCPA	Familiar 2106	2277 (III.7)	Filha	57	Não	MSH2	13	c.2152C>T (p.Q718*)	-
HCPA	Família 2114	2300 (IV.1)	Filha	16	Não	MSH2	11	c.1738_1741deGAAA (p.E580Lfs)	+
HCPA	Familiar 872	1097 (III.2)	Irmão	45	CCR	MLH1	11	c.888_889_delCA (p.E297Afs)	+

Legenda: CCR: câncer colorretal; I: idade do paciente; E: éxons rastreado; R: resultado do rastreamento; - : resultado negativo (familiar não apresenta a alteração), + : resultado positivo (familiar possui a alteração). A identificação do paciente entre parênteses se encontra de acordo com os heredogramas nos Anexos 26 a 30.

A família pertence ao probando 1112 (IV.3) apresentou doze casos de câncer, sendo sete casos de câncer colorretal, um caso de câncer de pulmão, um caso de carcinoma de corpo uterino, um caso de câncer nos rins e um caso em que o tipo de tumor não foi identificado, conforme apresentado no heredograma (Anexo 26). Dos cinco indivíduos desta família (IV.1, IV.2, IV.10, III.9 e III.6) rastreados para a alteração *MSH2* c.2152C>T, um deles demonstrou herdar a alteração (IV.10). Este indivíduo possui 30 anos de idade e ainda não foi diagnosticado com nenhum tipo de câncer.

A segunda família rastreada para esta mesma alteração (c.2152C>T) foi a família pertencente ao probando 871 (IV.6). Esta família apresentou oito casos de cânceres, sendo cinco casos de câncer colorretal, um caso de câncer esôfago, um caso de câncer de útero e um

caso cujo câncer é desconhecido, conforme é demonstrado no heredograma do Anexo 27. Dos dois indivíduos rastreados, um herdou a mutação (V.11). A familiar afetada com a alteração c.2152C>T (V.11) é parente de segundo grau da paciente probanda (filha da irmã), possui 23 anos de idade e não manifestou o fenótipo da doença.

A terceira família rastreada para a alteração c.2152C>T foi a do probando 2106, acometida com nove casos de câncer, sendo dois casos de câncer colorretal, dois casos de câncer de útero, um caso de câncer renal, um caso de câncer de bexiga, um caso de câncer de mama, um caso de câncer na base da língua e um caso de câncer de fígado, conforme demonstrado no heredograma do Anexo 28. A única paciente rastreada, que é filha da probanda (III.7), não herdou a alteração e não desenvolveu nenhum tipo de tumor.

A família 2114 foi rastreada para a alteração *MSH2* c.1738_1741delGAAA no éxon 11 do gene *MLH1*. A família pertencente ao probando 2114 (III.2) foi acometida com dois casos de câncer colorretal (II.2), conforme é demonstrado no Anexo 29. A única paciente submetida ao rastreamento molecular foi a filha da paciente probanda (IV.1), que herdou a deleção, embora não tenha manifestado o fenótipo da doença.

A família pertencente ao probando 872 foi rastreada para as variantes presentes em *cis MLH1* c.890A>C e *MLH1* c.888_889_delCA. Esta família apresentou nove casos de câncer do espectro tumoral da SL, sendo sete casos de câncer colorretal, um caso de câncer de útero e um caso desconhecido. O probando desenvolveu ainda tumores no ovário, endométrio e pâncreas, conforme demonstrado no heredograma do Anexo 30. O único familiar rastreado (III.2) foi portador das alterações, é irmão da probanda (III.1) e desenvolveu o fenótipo da doença (câncer colorretal aos 45 anos de idade).

5.8) Validação das alterações patogênicas e VUS pela metodologia de Sanger

Todas as alterações encontradas nos genes de MMR por NGS foram confirmadas pela metodologia de *Sanger*, exceto as alterações *MSH6* c.181G>A e *PSM2* c.780C>G.

A posição da alteração *MSH6* c.181G>A apresentou profundidade média de cobertura >30X e o alelo alterado estava presente em mais de 20% das sequências (14 vezes). Entretanto, as sequências provenientes do sequenciamento de *Sanger* para a alteração ficaram com uma qualidade baixa nas regiões correspondentes as alterações, impossibilitando sua validação.

A alteração benigna *PSM2* c.780C>G, presente em 34 dos 41 pacientes rastreados para este gene não foi validada, uma vez que os pacientes que a herdaram também apresentaram as deleções intrônicas c.706-4delA e c.706-5_706-4delAA não patogênicas, que flanqueavam o éxon 7 de *PMS2*, dificultando a visualização dos eletroferogramas. A construção de um

iniciador mais próximo da alteração foi impossibilitada, visto que esta região é rica em sequências repetitivas, o que dificultaria o sequenciamento correto do éxon. Os eletroferogramas das mutações patogênicas ou VUS se encontram disponíveis nos Anexos 16 a 24, exceto algumas alterações de *MSH2* e *MLH1*, que não foram disponibilizadas pelo Instituto responsável pela validação destas alterações.

5.9) Análise de custo

Uma análise de custo foi realizada para comparar o preço de sequenciamento dos genes MMR entre a metodologia de Sanger e NGS. Por Sequenciamento de Sanger foi considerada a realização de PCRs convencionais individualizadas por região codificante de cada um dos éxons dos genes e, para NGS, foi considerada a implantação de PCRs Multiplexes ou de longo alcance acoplada ao NGS. A Tabela 5.14 demonstra estes resultados. O sequenciamento de Sanger para o rastreamento dos quatro genes MMR teve um custo de aproximadamente R\$1600,90 por paciente. Este custo foi reduzido para R\$743,20 com a implantação da estratégia de NGS.

Tabela 5.14: Análise do custo efetividade para o rastreamento molecular dos genes MMR utilizando as metodologias de Sanger e NGS, em Reais, por Paciente.

Sequenciamento de Sanger					
Consumíveis	Preço (R\$)	N Reações	Custo Total (65)	Unidade (R\$)	Preço/Paciente
TaqDNA Platinum 500C (Thermo)	436,0	4095	4761,1	4,2	267,3
dNTP 100mM cada (0.1mM)	2684,8	3120	2094,2	0,7	42,3
Primers 100pmol (cada)	25,0	132	3300,0	0,2	26,4
Takara La Taq LR-PCR(Takara) 250 U	3121,1	123	3121,1	12,5	37,5
GE HealthCare Purification kit 250R	1861,0	3858	28718,8	7,0	441,8
BigDye Terminator Kit 100R (suficientes para 800R)	4988,0	8190	51064,7	6,2	785,6
Total					1600,9
Sequenciamento de Nova Geração					
Consumíveis	Preço (R\$)	N Reações	Custo Total (65)	Unidade (R\$)	Preço/Paciente
TaqDNA Platinum 500C (Thermo)	436,0	1057	1228,9	1,1	21,9
dNTP 100mM cada (0.1mM)	2684,8	975	2617,7	2,7	53,7
Primers 100pmol (cada)	25,0	110	2750,0	2,5	49,0
Takara La Taq LR-PCR(Takara) 250 U	3121,1	123	3121,1	19,5	58,5
GE HealthCare Purification kit 250R	1861,0	1057	7868,3	7,4	148,9
Nextera XT (24R)	5500,0	65	14895,8	229,2	229,2
Qubit dsDNA HS Assay Kit (100R)	404,0	65	262,6	4,0	4,0
TruSeq Rapid Duo cBot (2 lanes)	1446,2	96	1446,2	15,1	15,1
TruSeq®RapidSBSKit–HS(2 lanes)	6162,6	96	6162,6	64,2	64,2
TruSeq® Rapid PE Cluster Kit (2 lanes)	4437,1	96	4437,1	46,2	46,2
Kapa Library Quantification kit(1000X10ul)	4200,0	195	3412,5	17,5	52,5
Total					743,2

6) DISCUSSÃO

6.1) Metodologias de PCR multiplex e LR-PCR acopladas ao NGS

A padronização das reações de PCR *multiplex* e de Longo Alcance em nosso estudo permitiu reduzir em 3X o número total de reações realizadas por pacientes para sequenciamento dos genes *MSH2*, *MLH1*, *MSH6* e *PMS2* (de 63 para 21 reações por paciente). Esta redução teve como consequência uma diminuição do tempo de rastreamento molecular dos genes de MMR. Além disso, esta metodologia permitiu economizar a quantidade de material genético utilizado por paciente e teve também impacto no custo do sequenciamento por amostra. Em comparação à utilização do Sequenciamento de Sanger, a metodologia empregada permitiu reduzir em 2,15 vezes o custo de reação por amostra. (R\$743,20 X R\$1600,90).

O NGS oferece ainda uma vantagem em relação ao Sequenciamento de Sanger, que é o aumento da profundidade média de cobertura das amostras. Com o Sequenciamento de Sanger, uma profundidade média de cobertura máxima de 2 vezes (sentido direto e reverso) pode ser obtida, enquanto no NGS as amostras são amplificadas de maneira clonal e são cobertas mais de duas vezes. O aumento desta média, reduz possíveis erros de incorporação de nucleotídeos introduzidos pelas Taq DNA polimerases, diminuindo a identificação de variantes falso-positivas e negativas nas amostras (DE LEENEER *et al.*, 2011).

Pela primeira vez na literatura foi empregada a metodologia de PCR multiplex acoplada ao NGS para o rastreamento molecular dos genes MMR. Esta metodologia já havia sido utilizada previamente por *Leeneer et al.*, para o sequenciamento dos genes *BRCA1/2*. No estudo, os autores compararam a distribuição de profundidade média de cobertura de produtos de PCRs singleplexes com produtos de PCR multiplexes sequenciados na plataforma 454 (Roche) e demonstraram que as PCRs multiplexes parecem fornecer uma distribuição mais uniforme da profundidade média de cobertura com relação as PCRs singleplexes (DE LEENEER *et al.*, 2011).

O grande desvio padrão de cobertura observado para um mesmo *amplicon* de amostras distintas, pode ser consequência da metodologia empregada para quantificação dos produtos de PCR (Nanodrop). Embora quantidades equivalentes de amplicons tenham sido utilizadas para a preparação das bibliotecas, o ideal seria quantificar as amostras utilizando metodologias baseadas em fluorescência, que utilizam sondas que se ligam especificamente a moléculas de DNA dupla-fita e fornecem concentrações de DNA mais próximas de realidade. Apesar de encontrarmos profundidades de cobertura variáveis para as mesmas regiões, a metodologia empregada demonstrou ser eficiente para o rastreamento molecular dos genes

MMR, uma vez que alterações presentes no DNA foram identificadas com uma profundidade média de cobertura alta (média de 1029,41 vezes para os alelos alterados) e foram validadas pela metodologia de Sanger.

As tecnologias de NGS estão se tornando cada vez mais utilizadas e acessíveis para um grande número de laboratórios. Entretanto, ainda existem poucos estudos que implementam esta metodologia para a identificação de alterações nos genes MMR em pacientes com SL. Dos estudos existentes, a maioria utiliza de painéis de câncer baseados em métodos de captura híbrida (HOPPMAN-CHANEY *et al.*, 2010; JANSEN *et al.*, 2016; TALSETH-PALMER *et al.*, 2016) ou por meio de PCR (PRITCHARD; STEPHENS; DONNELLY, 2000; HITCH *et al.*, 2014; SIE *et al.*, 2015; RING *et al.*, 2016; SUSSMAN *et al.*, 2016) para a preparação das bibliotecas. Dois únicos estudos descritos utilizaram produtos de PCRs singleplexes para o rastreamento molecular dos genes MMR (DA SILVA *et al.*, 2010; HANSEN *et al.*, 2014). A Tabela 6.2 faz um resumo dos principais trabalhos que empregam NGS para a identificação de variantes nos genes MMR.

Ao comparar o nosso estudo com outros trabalhos da literatura que empregam NGS, observamos que a utilização de painéis de câncer, métodos de captura e de produtos de PCRs *singleplexes* para o NGS, gerou profundidades médias de coberturas menores do que o observado em nosso estudo (13.538X). Quando comparado com trabalhos que também utilizaram produtos de PCRs (PCRs *singleplexes*) para o sequenciamento das bibliotecas de DNA, as médias de coberturas obtidas foram 192 e 39 vezes maiores do que as médias obtidas por Hansen *et al.*, 2014 (81X) e Silva *et al.*, 2015 (400X), respectivamente.

Além das questões referentes a cobertura, a utilização de painéis de câncer a partir de captura híbrida ou PCR para o rastreamento dos genes MMR, muitas vezes, não é suficientemente apropriada para o rastreamento de sequências que apresentam um grande número de pseudogenes, como é o caso de *PMS2*. Os kits comerciais não utilizam de estratégias que evitem a amplificação dos pseudogenes, e isto gera um grande número de resultados falso positivos (HANSEN *et al.*, 2014). Ao empregar o método de captura para o sequenciamento do gene *PMS2*, por exemplo, não há como garantir que as sondas desenhadas se liguem especificamente a *PMS2*, uma vez que a similaridade entre as regiões parálogas de *PMS2CL* é maior do que 98% (TALSETH-PALMER *et al.*, 2016).

Os painéis de câncer fornecem a possibilidade de que um grande número de genes envolvidos com a doença seja sequenciado em uma única reação. Entretanto, em alguns casos, são utilizados painéis que também contêm informações sobre outros genes, não diretamente relacionados com a doença. Estes painéis podem fornecer o diagnóstico para mutações identificadas em outras patologias e não necessariamente relacionadas com a doença em

estudo, o que pode trazer diversas consequências para o aconselhamento genético do indivíduo. As sociedades *European Society of Human Genetics* e *EuroGentest* recomendam que o rastreamento molecular de indivíduos que apresentem doenças genéticas seja realizado apenas com base em genes conhecidos e que já foram associados a doença, para que problemas no aconselhamento sejam evitados. Em casos que o laboratório fornece o diagnóstico para outros genes é recomendado que as alterações encontradas para as regiões não relacionadas sejam informadas ao paciente (MATTHIJS *et al.*, 2015).

De todos os estudos já descritos na literatura, a maior parte utiliza as plataformas da Illumina (HiSeq 2000, HiSeq2500, MiSeq) para rastreamento molecular dos genes MMR. Estas plataformas vêm se destacando pelo menor custo de sequenciamento por amostra, geração de grande número de dados por corrida (>40Gb) e por lidarem melhor com a presença de regiões homopoliméricas nas sequências de DNA, principalmente quando comparado às plataformas que utilizam como método de detecção a diferença de concentração dos prótons incorporados durante a síntese de DNA (*Ion PGM System*) (CAPORASO *et al.*, 2012; SHIN *et al.*, 2016). *Quail et al.*, demonstram que, embora a plataforma PGM seja capaz de identificar um maior número de SNPs por amostra, o número de resultados falso-positivos é significativamente maior nestas plataformas do que nas plataformas Illumina (*QUAIL et al.*, 2012).

Tabela 6.1: Resumo dos principais trabalhos que utilizaram o NGS para rastreamento molecular dos genes envolvidos com a SL ou síndromes hereditárias relacionadas, ressaltando os tipos de amostras utilizadas, o modo de preparação das bibliotecas, as plataformas utilizadas, o modo de preparação das bibliotecas, a plataforma utilizada para o sequenciamento, os genes rastreados, a porcentagem de leituras mapeadas e a cobertura média encontrada.

Referência	Preparação bibliotecas	Plataforma	Genes rastreados	% leituras mapeadas	Cobertura média	Amostras	Resultados
Rohlin <i>et al.</i> , 2016 Fam. Cancer	SureselectXT Custom 3-5.9 Mb library kit (Agilent Technology, Santa Clara, EUA)- Captura	Illumina HiSeq2000 (Illumina, San Diego, CA, EUA)	<i>APC, MUTYH, BMPRIA, SMAD4, STK11, PTEN, MLH1, MSH2, MSH6, PMS2, EPCAM, CDH1, MLH3, MSH3, PMS1, AXIN2, CTNNB1, CHEK2</i> e <i>MET</i>	-	417X	Amostras de DNA de sangue de pacientes com suspeita de SL ou PAF (n=91)	17 pacientes alterados para os genes <i>MSH2, MSH6, SMAD4, CHEK2, PMS2, MLH3, AXIN2, BMPRIA, APC, MUTYH</i>
Ring <i>et al.</i> , 2016 Modern Pathology, 1-9	RainDance Thunderstorm emulsion PCR system (RainDance Technologies, Billerica, MA, EUA)-Painel PCR	Illumina HiSeq2500 (Illumina, San Diego, CA, EUA)	<i>MLH1, MSH2, MSH6, PMS2, EPCAM, PTEN, STK11, TP53, BRCA1, BRCA2, APC, ATM, BARD1, BMPRIA, BRIP1, CDH1, CDK4, CDKN2A, CHEK2, MUTYH, NBN, PALB2, RAD51C, RAD51D, SMAD4</i>	-	-	Amostras de DNA de pacientes com Câncer endometrial (n=381)/ DNA	13 pacientes (9%) alterados para <i>CHEK2, APC, ATM, BARD1, BRCA1, BRCA2, BRIP1, NBN, PTEN, RAD51c</i> . 22 pacientes (6%) para MMR (3 <i>MLH1</i> , 5 <i>MSH2</i> , 2 <i>EPCAM-MSH2</i> , 6 <i>MSH6</i> , 6 <i>PMS2</i>).
Jasen <i>et al.</i> , 2016 PLoS ONE 11(6): e0157381.	SureSelect Target Enrichment System (Agilent Technologies, Santa Clara, CA, EUA)- Captura	Illumina HiSeq2000	<i>MLH1, MSH2, MSH3, MSH6, PMS2, EPCAM, AXIN1, AXIN2, MSH3, MUTYH, APC, POLE/POLD1-EDM, BRCA1, BRCA2, ATM, CHEK2</i>	95%	101X	Amostras de DNA e RNA de sangue de pacientes de câncer CCR (n=37), endometrial (n=7) e outros (n=1)	8 pacientes (29%) mutações somáticas <i>MLH1, MSH2, MSH3</i> ou <i>MSH6</i> 2,1 2, 1, 1, 1 pacientes alterados para <i>MUTYH, APC, MLH1, AXIN1, AXIN2</i> e <i>MSH3</i> , respectivamente.

Susswein <i>et al.</i> , 2016 Genetics in Medicine, v.18 (8):823-832	SureSelect targeted capture (Agilent, Santa Clara, CA, EUA)- Captura	Illumina MiSeq ou HiSeq	<i>APC, BMPRIA, BRCA1, BRCA2, CDH1, CDKN2A, EPCAM, MLH1, MSH2, MSH6, MUTYH, PMS2, PTEN, SMAD4, STK11, TP53, VHL, ATM, CHEK2, PALB2, AXIN2, BARD1, BRIP1, CDK4, FANCC, NBN, RAD51C, RAD51D, XRCC2</i>	-	-	Extração de DNA de sangue para pacientes com câncer mama, ovário, CCR, estômago, pâncreas, endométrio (n=10030)	901 pacientes (9%) com mutação patogênica (27,2% <i>BRCA1/2</i> ; 15,5% genes SL; 9,1% outros genes de alto risco; 48,2% genes de risco moderado ou desconhecido).
Talseth-Palmer <i>et al.</i> , 2016 Cancer Med, v.5(5): 929-41	NimbleGen human custom array (Roche NimbleGen, Madison, Winsconsin, USA)- Captura	Illumina HiSeq2000	<i>MLH1, MSH2, MSH6, PMS2, MSH3, PMS1, MLH3, EXO1, RFC1, RFC2, RFC3, RFC4, RFC5, PCNA, LIG1, RPA1, RPA2, RPA3, POLD1, POLD2, POLD3 e POLD4</i>	97.2%	110X	Extração de DNA de sangue periférico de pacientes com câncer CCR (n=29) e endometrial (n=12)	3 pacientes alterados (1 <i>MSH2</i> , 1 <i>MLH1</i> , 1 <i>MSH6</i>)
Silva <i>et al.</i> , 2015 PLoS One, v.5;10(10):e0139753	Amplificação individual dos éxons seguida do NGS	Ion Torrent (Applied Biosystems, Foster City, CA, EUA)	<i>MLH1, MSH2, MSH6, PMS1 e PSM2</i>	-	400X	Extração de DNA de sangue periférico de pacientes (n=116) suspeitos de ter SL	45 pacientes alterados (25 mutações pontuais em <i>MSH2</i> , 15 em <i>MLH1</i> , 4 em <i>MSH6</i> e 1 em <i>PMS2</i>)
Conor <i>et al.</i> , 2015 Fam Cancer, v.14(1):69-75	-	Illumina HiSeq2000	<i>MSH6</i>	-	-	-	1 (<i>MSH6</i>)
Hansen <i>et al.</i> , 2014 Mol Genet Genomic Med., v.2(2):186- 200.	Amplificação individual dos éxons e preparação com o kit Amplicon Library Preparation (Roche, Pleasanton, CA, EUA)- PCR	GS Junior Studio (Roche, Pleasanton, CA, EUA)	<i>MLH1, MSH2, MSH6 e PSM2</i>	80%	81X	Extração de DNA de sangue periférico de pacientes com CCR (n=55)	27 pacientes alterados (4 <i>MLH1</i> , 7 <i>MSH2</i> , 9 <i>MSH6</i> , 4 <i>PMS2</i>)

Pritchard <i>et al.</i> , 2012 J Mol Diagn., v.14(4):357-66	Coloseq- Painel PCR	Illumina HiSeq2000	<i>MLH1, MSH2, MSH6, PMS2, EPCAM, APC</i> e <i>MUTYH</i> e 24 outros genes envolvidos menos envolvidos com CCR	62%	320X	Extração de DNA de sangue de pacientes com mutações conhecidas em MMR (n=23), sugestivos de SL (n=32), pacientes sem história familiar(n=19)	28 pacientes alterados
Hoppman-Chaney <i>et al.</i> , 2010 Clin Chem., v.56(8):1297-306	NimbleGen 385K Custom Sequence Capture (Roche, Pleasanton, CA, EUA)-Captura	454 Life Sciences (Roche, Pleasanton, CA, EUA) e Gall (Illumina, San Diego, CA, EUA)	<i>MLH1, MSH2, MSH6, APC</i> e outros 17 genes envolvidos com SL	62,3% (454) e 58% (Gall)	30X (545) e 226X (Gall)	Extração de DNA de sangue de pacientes com câncer CCR(n=8)	7 pacientes (<i>MLH1, MSH2, MSH6, APC</i>)

6.2) Limitações das metodologias de PCRs multiplexes e LR-PCRs acopladas ao NGS

Apesar de apresentar diversas vantagens sobre as metodologias normalmente empregadas para o sequenciamento dos genes MMR, a metodologia empregada neste estudo apresenta algumas limitações.

As regiões flangeadoras de alguns dos éxons dos genes de MMR, são ricas em sequências repetitivas de DNA. Por meio da metodologia empregada, foi possível encontrar deleções ou inserções sequenciais de nucleotídeos nestas regiões repetitivas, a exemplo das alterações de 1 a 7 Adeninas na posição c.942+3 do íntron 5 do gene *MSH2* e deleções de 9 Timinas e Adeninas na posição c.1039-34_1039-26 do íntron 11 de *MLH1* (c.1039-34_1039-26delTATATATTT). Apesar de apresentarem taxas de coberturas elevadas, estas variantes podem ser consideradas alterações falso-positivas, por provavelmente representarem erros de incorporação de nucleotídeos pela enzima *Taq DNA Polimerase* utilizada para as reações de PCR *multiplex*. As enzimas *Taq DNA Polimerases* sintéticas não apresentam uma atividade exonucleásica revisora 3'-5'. Desta maneira, o fenômeno de escorregamento em regiões repetitivas de DNA é bastante comum, o que promove a adição ou remoção de nucleotídeos nestas regiões do DNA (ZAVODNA *et al.*, 2014).

A plataforma HiSeq2500, utilizada para o sequenciamento dos fragmentos de PCR multiplexes/ LR-PCRs, utiliza uma estratégia que incorpora um único nucleotídeo marcado com fluorescência por ciclo e a leitura é realizada de maneira sequencial (CARVALHO; SILVA, 2010). Desta forma, em comparação com as metodologias de Sanger e de NGS que utilizam PCRs de emulsão (454 Roche), o sequenciamento por síntese (HiSeq Illumina) reduz a incorporação errônea de nucleotídeos às regiões microssatélites do DNA durante a leitura das bases. Entretanto, como utilizamos produtos de PCR para o NGS, o efeito da adição ou remoção de nucleotídeos pelo escorregamento da *Taq DNA polimerase* não foi eliminado das sequências, incorporando alterações falso-positivas nas regiões repetitivas de DNA.

Mais de 15% das mutações causais de SL ocorrem como consequência da presença de grandes rearranjos gênicos nos genes MMR (PRITCHARD *et al.*, 2012). Além disso, 22-44% dos pacientes que apresentam alterações no gene *PMS2*, possuem variações no número de cópias de DNA (SENER *et al.*, 2008). Embora as frequências de rearranjos gênicos sejam grandes para os genes MMR, a metodologia empregada em nosso estudo não é capaz de detectar estes tipos de variantes nos pacientes rastreados.

A avaliação da presença de rearranjos em nossas amostras, foi previamente identificada por MLPA para os genes *MSH2*, *MSH6*, *MLH1* e *EPCAM*, sendo que *PMS2* não

foi avaliado. Do número total de pacientes que participaram do projeto, 12,3% (n=8) apresentaram deleções nos genes *MSH2*, *EPCAM* ou *MLH1*. Nestes casos foi também observada a deleção do gene *MUTYH*, por meio da sonda P072 (para *MSH6*), em um dos pacientes que preencheram os critérios para SL (CPB) (Anexo 25). Deleções neste gene estão associados a uma outra síndrome de CCR hereditária, que é a Polipose Hereditária associada a *MUTYH* (MAP).

Indivíduos MAP positivos são muitas vezes classificados como pacientes Lynch-Like, por não apresentarem mutações germinativas ou metilação somática em nenhum dos genes de MMR, mas clinicamente preencherem os critérios clínicos para SL, desenvolverem tumores extracolônicos e apresentarem instabilidade em regiões microssatélites (CASTILLEJO *et al.*, 2014). O gene *MUTYH* codifica para a enzima DNA glicosilase, responsável pelo reparo de DNA por excisão de bases que removem nucleotídeos anormais ou quimicamente modificados, como a 7,8- dihidro-oxi-guanina (8-oxoG). O complexo *MutS* (*MSH2/MSH6*) pode se ligar a *mismatches* 8-oxoG:A formados durante o processo de replicação e estimular a atividade da enzima DNA glicosilase, o que explica a forte associação fenotípica existente entre pacientes MAP e Lynch-like (GU *et al.*, 2002). A deleção do *MUTYH*, encontrada em CPB, indica que este paciente é portador de MAP e não de SL.

No total, 3,1% dos fragmentos que foram sequenciados apresentaram profundidade média de cobertura <30X, especialmente regiões correspondentes aos éxons 1, 2, 7 e 10 de *MSH2* e ao éxon 1 de *MSH6*. Nós acreditamos que o motivo das baixas profundidades médias de coberturas nestas regiões tenha sido causado por um erro na quantificação em cada uma destas amostras, na etapa que precede a preparação de bibliotecas, o que provocou a sub-representação destes fragmentos nas respectivas bibliotecas.

Para o éxon 10 de *MSH2*, a baixa cobertura em 4,6% das amostras sequenciadas para esta região, pode ainda ser explicada pelo grande número de amplicons que participam da mesma reação de multiplex para este éxon (mix 2). PCR *multiplexes* com um maior número de iniciadores, podem resultar em diminuição da eficiência das PCRs individuais para cada um dos fragmentos, reduzindo a qualidade da sequência (DE LEENEER *et al.*, 2011; SCHEIBLE *et al.*, 2014).

6.3) Padronização das análises de bioinformática

Um dos desafios enfrentados na utilização das plataformas de NGS é a análise dos dados gerados, que normalmente é complexa e necessita de entendimento prévio na área de bioinformática. Alguns pacotes comerciais que integram toda a análise dos dados e fornecem interfaces gráficas mais didáticas ao usuário são normalmente empregadas. Entretanto, estes

pacotes apresentam parâmetros pré-delimitados que podem afetar os resultados da análise, tornando-a muitas vezes limitada.

Portanto, a análise manual dos dados com o auxílio de diferentes programas de bioinformática, sem a utilização de pacotes prontos, é a melhor opção. Apesar de trabalhosa, por exigir conhecimento prévio na área, a análise manual permite que o usuário defina os parâmetros de acordo com o seu objetivo e modele sua análise para obtenção de resultados mais precisos (SCHRIJVER *et al.*, 2010). Considerando estas informações, nosso grupo de pesquisa desenvolveu um *script* que foi útil para a identificação de mutações pontuais nas regiões codificantes dos genes *MSH2*, *MLH1*, *MSH6* e *PMS2* e se encontra disponível no Anexo 13.

6.4) Validação da metodologia padronizada para rastreamento dos genes MMR

Para validar as metodologias de PCR multiplexes e LR-PCR acoplada ao NGS e o script produzido para a identificação de variantes, realizamos o rastreamento molecular dos genes *MSH2*, *MSH6*, *MLH1* e *PMS2* para 65 pacientes. Destes pacientes, trinta e dois (49,2%) apresentaram pelo algum tipo de alteração (benigna/patogênica/VUS) em um dos genes de MMR. O gene *MLH1* demonstrou ser o gene mais polimórfico, apresentando 34,0% (32/96) do número total de alterações encontradas, seguido de *MSH6* (29,8%), *MSH2* (25,5%) e *PMS2* (12,8%), respectivamente. Considerando-se apenas as variantes patogênicas ou VUS encontradas em regiões codificantes (n=25), *MLH1* (48%) e *MSH2* (28%) foram os genes mais afetados, seguido de *MSH6* (20%) e *PMS2* (4%).

Nossos achados corroboram com outros estudos recentes que rastrearam os genes de MMR em populações da América do Sul e populações europeias, norte-americanas e asiáticas. Dominguez-Valentin e colaboradores demonstraram que *MLH1* é um gene mais polimórfico do que *MSH2*. O trabalho rastreou amostras de cinco centros distintos da América Latina (Argentina, Brasil, Chile, Colômbia e Uruguai) e mutações patogênicas em *MLH1* foram encontradas em 60% dos pacientes alterados (DOMINGUEZ-VALENTIN *et al.*, 2013). As frequências encontradas em nossos estudos também se aproximam das médias de alterações patogênicas esperadas para populações europeias, norte- americanas e asiáticas : 42% para *MLH1*, 33% para *MSH2*, 18% para *MSH6* e 8% para *PMS2* (PLAZZER *et al.*, 2013).

Em 2015, Carneiro da Silva e colaboradores rastrearam 116 indivíduos brasileiros residentes do estado de São Paulo, que preencheram os critérios clínicos de Amsterdam e Bethesda para SL. Os genes MMR foram sequenciados utilizando a plataforma Ion PGM System e *MSH2* foi identificado como sendo o gene com maior número de mutações

patogênicas na população de estudo, seguido de *MLH1*, *MSH6*, *PMS2* e *PMS1*, respectivamente (CARNEIRO DA SILVA *et al.*, 2015). Estes achados diferem do espectro mutacional encontrado em nossas amostras, no que se refere à frequência de mutações patogênicas encontradas em *MLH1* e *MSH2*, o que pode ser justificado pelo fato de que nosso estudo sequenciou amostras de quatro diferentes regiões do país (RJ/SP/RS/PA) e o estudo de Carneira da Silva e colaboradores utilizou amostras de apenas uma região do Brasil.

6.4.1) Alterações genéticas identificadas no gene *MSH2*

As variantes encontradas no gene *MSH2* atingiram domínios importantes para a execução do reparo de DNA. A maior parte das variantes benignas estavam presentes em domínios que não influenciavam a formação do complexo *MutS* ou em regiões intrônicas repetitivas de DNA (n=11), sendo possível identificar algumas alterações no domínio de reconhecimento do *Mismatch* (n=4) ou no domínio *Lever* da proteína (n=2). Todas as variantes patogênicas e VUS estavam presentes em regiões importantes para MMR, e foram identificadas nos domínios *ATPásico* (n=3), *Lever* (n=2) e *Connector* (n=2) de *MSH2*.

A identificação de variantes no domínio *ATPásico* de *MSH2* é importante, considerando que alterações genéticas neste domínio podem modificar o reconhecimento da base mal pareada e iniciação do reparo de DNA (MUKHERJEE; FEIG, 2009). Das três variantes encontradas no domínio *ATPásico* de *MSH2* (c.2152C>T - p.Q718X, c.2078G>A - C693Y e c.2021G>A - p.G674D), uma tem significado clínico desconhecido (c.2078G>A - C693Y) e ainda não tinha sido publicada no ClinVar. Foi predita como patogênica pelas análises de predição *in silico* e o paciente portador desenvolveu câncer de cólon ascendente em idade precoce (42 anos de idade).

Metade dos pacientes com mutações patogênicas em *MSH2* (n=6) herdaram a mutação c.2152C>T - p.Q718X (874/1112/2106/871/2309/2318), que é capaz de introduzir um códon de parada na posição 718 (Glutamina) da proteína. Esta variante representa 4,1% de todas as alterações patogênicas encontradas nos quatro genes MMR e atingiu 9,2% dos pacientes rastreados (n=65), sendo identificada exclusivamente nas amostras provenientes do Sul do Brasil (HCPA).

Devido ao grande número de pacientes acometidos com esta variante, suspeita-se que ela tenha efeito fundador no Brasil. Embora ainda não existam estudos que comprovem este efeito, acredita-se que a alteração c.2152C>T (p.Q718X) tenha sido trazida ao Brasil por imigrantes portugueses (ISIDRO *et al.*, 2000). A mutação já foi identificada em amostras brasileiras da região Sudeste do Brasil (estado de São Paulo-ACCAM) (DOMINGUEZ-

VALENTIN *et al.*, 2013; CARNEIRO DA SILVA *et al.*, 2015) e em amostras de população Norte-Americana (Estados Unidos) (TERDIMAN *et al.*, 2002).

A alteração missense c.2021G>A (p.G674D) presente no éxon 13 de *MSH2* foi funcionalmente caracterizada como uma alteração patogênica em modelo de *Saccharomyces cerevisiae* (GAMMIE *et al.*, 2007). Já foi identificada em pacientes de origem Germânica, que preencheram os critérios clínicos para SL, e desenvolveram tumores extracolônicos no pâncreas e na mama (BRIEGER *et al.*, 2011). O paciente portador da alteração em nosso estudo desenvolveu adenocarcinoma no cólon sigmóide, entretanto, não desenvolveu tumores extracolônicos e não apresentou nenhum familiar de primeiro grau afetado.

Duas variantes foram identificadas no domínio *Lever* de *MSH2*, responsável por envolver o nucleotídeo que será reparado (LAMERS *et al.*, 2000): c.1738_1741delGAAA (p.E580Lfs) e c.1046C>T (p.P349R). A alteração c.1738_1741delGAAA foi primeiramente identificada por nosso grupo de pesquisa e provavelmente é uma mutação patogênica, por alterar o quadro de leitura do éxon 11 de *MSH2* (p.E580Lfs). A alteração patogênica c.1046C>T, encontrada no éxon 6 de *MSH2*, já foi descrita em trabalhos que rastream populações de origem Italiana (URSO *et al.*, 2008; PASTRELLO *et al.*, 2011) e mais recentemente, foi identificada em pacientes de origem Argentina (DOMINGUEZ-VALENTIN *et al.*, 2016).

Foram identificadas duas variantes responsáveis por alterar o reconhecimento das proteínas do complexo *MutL* (*Connector*) (MENDILLO *et al.*, 2009): c.942+3A>T (p.V265_Gln314del) e c.388_389delCA (p.Q130Vfs*2). A mutação c.942+3A>T é capaz de alterar o sítio de reconhecimento de *splicing*, levando a deleção completa do éxon 5 de *MSH2* (p.V265_Gln314del). Curiosamente, esta variante foi encontrada em homozigose para o paciente alterado. Mutações bialélicas nos genes de MMR são frequentemente encontradas em pacientes portadores da Síndrome Deficiência constitucional de Mismatch Repair (CMMR-D) (em inglês, *Constitutional Mismatch Repair-Deficiency*) (TOLEDANO *et al.*, 2009). Esta síndrome acomete indivíduos jovens e leva ao desenvolvimento de tumores no trato gastrointestinal, cérebro e/ou hematológicos, além de tumores no endométrio e trato urinário (VASEN *et al.*, 2014). Em nosso estudo, o paciente acometido com a mutação c.942+3A>T (p.V265_Gln314del), apresentou perda das proteínas MSH2/MSH6 por IHC, desenvolveu câncer de cólon descendente precoce (aos 21 anos de idade) e teve um familiar de primeiro grau acometido com câncer de endométrio, o que reforça a possibilidade de que este indivíduo tenha desenvolvido a Síndrome CMMR-D.

A alteração c.942+3A>T (p.V265_Gln314del) apresenta efeito fundador na população da ilha *Newfoundland*, ilha canadense originalmente colonizada por Franceses, Britânicos e

Portugueses (FROGGATT *et al.*, 1999). Outros trabalhos também a identificaram em pacientes de origem Argeliana (influência da colonização francesa) (ZIADA-BOUCHAAR *et al.*, 2016) e em pacientes de origem Germânica (MANGOLD *et al.*, 2005). A alteração já foi relatada em pacientes de origem brasileira (n=3) (DOMINGUEZ-VALENTIN *et al.*, 2013; CARNEIRO DA SILVA *et al.*, 2015).

A mutação *frameshift* c.388_389delCA (p.Q130Vfs*2) provoca mudança no quadro de leitura do RNAm no éxon 2 do gene *MSH2* e leva a formação de uma proteína não funcional. A variante foi previamente encontrada em um paciente de origem brasileira e um de origem argentina (DOMINGUEZ-VALENTIN *et al.*, 2013).

As c.2152C>T (p.Q718X) e c.942+3A>T (p.V265_Gln314del) já foram relatadas como *hot spots* para a população brasileira, e suas frequências foram justificadas pela alta influência da colonização europeia no país (DOMINGUEZ *et al.*, 2011). Além destas mutações, outras variantes encontradas podem ter sido herdadas pela influência da colonização europeia no Brasil, como as alterações c.2021G>A e c.1046C>T (p.P349R), previamente identificadas em populações de origem germânica e italiana, respectivamente.

O rastreamento molecular do gene *MSH2* permitiu identificar um maior número de alterações no éxon 13 deste gene, sendo que 8/12 pacientes (66,6%) estavam mutados para esta região. Estudos com amostras de pacientes brasileiros demonstram também que este é o éxon mais frequentemente mutado em *MSH2*. Valentin e colaboradores identificaram 8 famílias mutadas (29,62%) para o éxon 13, em um total de 27 probandos rastreados (DOMINGUEZ-VALENTIN *et al.*, 2013), enquanto Carneiro da Silva e colaboradores identificaram para 7 probandos mutados (36,48%) para um total de 19 rastreados (CARNEIRO DA SILVA *et al.*, 2015). O éxon 13 pode ser considerado um *hot spot* de *MSH2* para a população brasileira, considerando os nossos achados e os achados de estudos já publicados.

6.4.2) Alterações genéticas identificadas no gene *MSH6*

A maior parte das variantes benignas encontradas no gene *MSH6* (n=19) estavam presentes em domínios não envolvidos com a formação do complexo *MutS* (n=13) ou em regiões intrônicas repetitivas de DNA (n=6), sendo possível identificar algumas alterações no domínio PWWP (n=2), no domínio *Lever* da proteína (n=1) ou no domínio *Connector* (n=2) de *MSH6*. Das VUS presentes nas regiões codificantes de *MSH6*, três foram identificadas em domínios importantes da proteína: PWWP (c.402dupT - p.F133fs), *ATPásico* (c.3772C>G - p.Q1258E) e *Connector* (c.2006T>C - p.I669T).

Das cinco alterações VUS encontradas no gene *MSH6*, quatro foram classificadas como neutras por pelo menos dois dos programas de predição *in silico* utilizados. A única alteração provavelmente patogênica identificada foi a inserção c.402dupT (p.F133fs), encontrada no domínio PWWP (motivo Pro-Trp-Trp-Pro) da proteína MSH6, responsável pela ligação da H4K20me (WANG *et al.*, 2009). Esta mutação altera o quadro de leitura do RNAm, gerando um stop códon prematuro.

Mulheres com mutações patogênicas nos genes *MLH1* e *MSH2* têm um risco de 40 a 50% de desenvolverem câncer endometrial, entretanto, quando as mutações ocorrem no gene *MSH6*, este risco aumenta para 64 a 70%. O paciente portador da alteração c.402dupT (2278) desenvolveu câncer endometrial em idade precoce, o que corrobora com o que é descrito na literatura (RUBIO *et al.*, 2016). Os outros dois pacientes que desenvolveram câncer endometrial em nosso estudo (843 e 854) não apresentaram Variantes de Nucleotídeo Único (SNVs) em nenhum dos genes de MMR, mas mostraram perda de expressão da proteína MSH6 (843) e MSH6/MLH1 (854) por IHC.

O portador da alteração c.3772C>G (p.Q1258E) (2113) herdou uma segunda alteração no gene *MLH1*, que pode ser considerada patogênica e altera a tradução de *MLH1* (c.1681_1682insT).

Das VUS exônicas encontradas em *MSH6*, duas (c.719G>A - p.R240Q e c.2006T>C - p.I669T) foram identificadas no éxon 4 deste gene, mas foram preditas como não patogênicas pelos algoritmos de análise utilizados. Carneiro da Silva e colaboradores também identificaram o éxon 4 como sendo a região com maior número de variantes em *MSH6*, em pacientes da região Sudeste do Brasil (CARNEIRO DA SILVA *et al.*, 2015).

6.4.3) Alterações genéticas identificadas no gene *MLH1*

As variantes identificadas no gene *MLH1* estavam envolvidas com a formação de diferentes domínios da proteína codificada por esse gene. A maior parte das alterações benignas (n=17) estavam localizadas em regiões não relacionadas a formação do complexo *MutL*, sendo que quatro foram identificadas nos sítios de ligação ao *mismatch* (n=1) ou no domínio C-terminal (n=3) desta proteína. Todas as alterações patogênicas e VUS exônicas em sítios de *splicing*, encontradas para o gene *MLH1*, estavam presentes nos domínios envolvidos com a formação do complexo *MutL*, sendo 2 alterações identificadas no domínio N-terminal, 5 no domínio C-terminal, 3 no domínio HATPásico e 1 no domínio de ligação (*Linker*).

O domínio C-terminal de MLH1, que apresenta atividade endonucleásica para quebra da base mal pareada (GUENEAU *et al.*, 2013), foi acometido com cinco variantes patogênicas ou VUS, sendo uma alteração em sítio de *splicing*, duas *frameshift*, uma *missense*

e uma *nonsense*. A mutação patogênica c.1975C>T (p.R659X) foi a única alteração *nonsense* encontrada para o gene *MLH1* e que atingiu o domínio C-terminal da proteína. Esta alteração foi previamente identificada em dois trabalhos realizados com amostras de pacientes finlandeses (NYSTROM-LAHTI *et al.*, 1996; GYLLING *et al.*, 2008) e foi descrita no LOVD em amostras brasileiras (São Paulo-ACCAM) (http://chromium.lovd.nl/LOVD2/colon_cancer/variants.php?select_db=MLH1&action=view&view=1011920%2C0006069%2C0).

Embora ainda não tenha sido descrita por nenhum outro estudo, a alteração *frameshift* c.1863delG (p.M621Ifs), pode ser classificada como provavelmente patogênica, por provocar uma mudança no quadro de leitura do RNAm de *MLH1* na região correspondente ao domínio C-terminal. O paciente portador desenvolveu câncer de cólon ascendente em idade precoce (aos 52 anos de idade).

Três variantes estavam presentes na região correspondente ao domínio HATPase (*Histidine kinase-like domain*) da proteína *MLH1* (c.350C>T - p.T117R), c.225delT - p.L75Vfs e c.277A>G - p.S93G). Este domínio é conservado em diversas proteínas que possuem sítios de ligação com ATP, e facilita a dimerização do domínio ATPásico assim que um nucleotídeo mal pareado é reconhecido pelo complexo *MutL* (DUTTA; INOUE, 2000).

A alteração patogênica c.350C>T (p.T117R), identificada no éxon 4 de *MLH1*, foi primeiramente descrita em um grupo de pacientes da França (DIEUMEGARD *et al.*, 2000). Em 2002, Brieger *et al.*, verificaram a capacidade desta alteração em reduzir a estabilidade da proteína *MLH1* e diminuir os níveis de apoptose de células mutadas (BRIEGER *et al.*, 2002). A perda funcional de *MLH1* na presença de p.T117R foi comprovada por meio de ensaios *in vitro*, que verificaram a capacidade da mutação em alterar os sítios de ligação ao ATP ao complexo *MutL* (TAKAHASHI *et al.*, 2007).

A deleção c.225delT (p.L75Vfs), presente em heterozigose no éxon 3 de *MLH1*, foi primeiramente identificada neste estudo, e uma vez que altera o quadro de leitura deste éxon pode ser considerada patogênica. O paciente que herdou a alteração desenvolveu câncer de cólon transverso aos 42 anos de idade.

A alteração c.277A>G (p.S93G), também encontrada no éxon 3 de *MLH1* é uma VUS, provoca a troca de aminoácidos na estrutura de sua proteína e foi predita como danosa (SIFT) ou deletéria (PROVEAN) por dois dos três bancos de dados utilizados para predição. Essa variante foi encontrada em um paciente que desenvolveu câncer no cólon transverso aos 39 anos de idade.

A alteração 1681_1682insT (p.Y561Lfs) ainda não foi descrita, mas é provavelmente patogênica por causar a mudança do quadro de leitura do RNAm do éxon 15 de *MLH1*. O

paciente portador também possui uma alteração predita como benigna (c.3772C>G) em *MSH6* e foi diagnosticado com câncer no reto e no cólon sigmóide aos 36 anos de idade.

Duas dessas alterações foram identificadas no códon 226 da proteína MLH1: c.677G>A - p.R226Q (presente em dois indivíduos) e c.677+1G>A (presente em um indivíduo). A alteração c.677G>A leva a troca de uma Arginina por uma Glutamina e, a variante c.677+1G>A, modifica o sítio de reconhecimento de *splicing* do éxon 8 de *MLH1*. Já foram previamente detectadas em pacientes de origem holandesa (HENDRIKS *et al.*, 2003) e mais recentemente, c.677G>A também foi identificada em amostras brasileiras provenientes da região Sudeste do país, bem como em um paciente de origem argentina (DOMINGUEZ-VALENTIN *et al.*, 2013; CARNEIRO DA SILVA *et al.*, 2015).

Uma segunda mutação foi encontrada em sítios de reconhecimento de *splicing* de *MLH1* (c.791-6_794delGTTTAGATC). Esta alteração foi pela primeira vez descrita na literatura e demonstra ser patogênica por alterar o sítio de reconhecimento de *splicing* do éxon 10 de *MLH1*, embora a sua patogenicidade não pode ser predita por ensaios *in silico*. O portador da alteração desenvolveu câncer do cólon ascendente e íleo em idade precoce (51).

Foram encontradas duas alterações em *cis* no éxon 11 do gene *MLH1* (c.890A>C e c.888_889delAG - p.E297Afs) para o mesmo paciente (872). A alteração c.890A>C foi predita como danosa pelos algoritmos de predição utilizados. A deleção c.888_889delAG é capaz de provocar uma mudança no quadro de leitura do RNAm para o éxon 11 do gene *MLH1*, podendo ser considerada uma alteração patogênica. O portador foi diagnosticado com CCR aos 47 anos de idade, possuía instabilidade para os cinco marcadores utilizados na MSI e teve perda da expressão da proteína PMS2 por IHC. Pacientes que apresentam alterações *missense* e *frameshift* em *MLH1* podem demonstrar uma expressão normal da proteína codificada por este gene (MLH1) por IHC e uma perda solitária de PMS2, mesmo quando mutações em *PMS2* não são identificadas (ROSTY *et al.*, 2016). A perda solitária de PMS2 e a presença de alterações no gene *MLH1* foi também observada para outros dois pacientes do nosso estudo: deleção dos éxons 17 a 19 de *MLH1* (observada por MLPA) e introdução de um códon de parada no éxon 17 (p.R659*) de *MLH1*.

Dois pacientes (2106 e SL4) foram portadores da variante c.1852_1853delAAinsGC, que leva a substituição de uma Lisina por uma Alanina no códon 618 da proteína de MLH1 (p.K618A). Esta variante foi depositada no banco de dados LOVD como sendo uma alteração benigna. Dominguez-Valentin *et al.*, classificaram a variante como sendo patogênica (DOMINGUEZ-VALENTIN *et al.*, 2013) e nossos ensaios de predições *in silico* também indicaram que a variante é danosa. Entretanto, Kondo e colaboradores já haviam demonstrado que a formação de heterodímeros entre as proteínas MLH1 e PMS2 (MutL) não era interferida

pela troca de aminoácidos p.K618A (KONDO *et al.*, 2003). Apesar de duas fontes bibliográficas (KONDO *et al.*, 2003; DOMINGUEZ-VALENTIN *et al.*, 2013) indicarem o efeito deletério da alteração, os estudos funcionais demonstraram que esta alteração é provavelmente benigna.

A alteração c.1007G>A, presente em heterozigose no éxon 11 do gene *MLH1*, apresenta significado funcional ainda não conhecido, e leva a troca de uma Glicina por um Aspartato no códon 336 (p.G336D) da proteína de *MLH1*. Os três programas de predição utilizados indicam que esta alteração é possivelmente danosa, podendo causar perda de função da proteína. Esta variante foi detectada em um paciente diagnosticado com câncer de cólon ascendente aos 53 anos de idade.

Um maior número de alterações patogênicas foi encontrado nos éxons 3 (n=2), 8 (n=2), 11 (n=2) e 16 (n=2) de *MLH1*, representando 66,6% das mutações patogênicas ou VUS exônicas para este gene. Os éxons 4, 10, 15 e 17 apresentaram uma alteração em cada. A proporção dos éxons mutados em nosso estudo para *MLH1* foi diferente do encontrado para outras amostras brasileiras. Dominguez-Valentin *et al.*, identificaram os éxons 6 (11,7%), 13 (11,7%) e 18 (11,7%) como as regiões com um maior número de mutações patogênicas em *MLH1* (DOMINGUEZ-VALENTIN *et al.*, 2013), enquanto Carneiro da Silva e colaboradores identificaram os éxon 6 (13,3%) e 7 (13,3%) como sendo as regiões mais afetadas (CARNEIRO DA SILVA *et al.*, 2015).

6.4.4) Alterações genéticas identificadas no gene *PMS2* e análise do pseudogene *PMS2CL*

Nenhuma variante patogênica foi identificada para o gene *PMS2*, sendo que a única VUS encontrada neste gene foi a alteração c.-77delA. As predições *in silico* não puderam ser realizadas para esta variante, devido a sua localização (5'UTR). Entretanto, devido ao grande número de portadores (n=30), provavelmente ela não é patogênica. Além disso, a maior parte dos pacientes afetados, herdaram também mutações patogênicas/VUS nos genes de MMR (5 em *MSH2*, 10 em *MLH1* e 1 em *MUTYH*).

A metodologia utilizada neste trabalho foi adequada para a amplificação do gene *PMS2*, uma vez que ao empregar PCRs de longo alcance, à amplificação de sequências parálogas aos éxons 1-5, 7-9, 11-15 de *PMS2* foi dificultada (BORRAS *et al.*, 2013).

O alinhamento das leituras geradas pelo NGS com a sequência de referência do pseudogene e a identificação das variantes para os éxons de *PMS2CL* parálogos a *PMS2*, foi importante para eliminar possíveis resultados falso-positivos de variantes encontradas no pseudogene. Apesar da estratégia de LR-PCR ter sido eficiente para a detecção de variantes no gene *PMS2*, as sequências de *PMS2* e *PMS2CL* recombina-se frequentemente por

crossing over desigual, dificultando a identificação correta das variantes e geração de leituras para estas duas sequências (HAYWARD *et al.*, 2007; VAN DER KLIFT *et al.*, 2016). Devido à alta similaridade presente entre o pseudogene e o gene funcional, não foi possível afirmar com 100% de certeza que o pseudogene não estava amplificando juntamente com o gene funcional. As altas taxas de coberturas encontradas para o pseudogene podem ser referentes a presença do mesmo no meio, entretanto, podem se confundir também com a presença do gene funcional.

O pseudogene *PMS2CL* apresenta similaridade com os éxons 9, 11-15 do gene funcional e sua formação é caracterizada por uma perda de 2,7Kb que ocorre entre dois elementos *Alu* localizados nas extremidades do éxon 10 do gene funcional. Considerando esta diferença entre o gene funcional e o pseudogene, Hayward e colaboradores desenharam dois pares de iniciadores capazes de amplificar os éxons 10 a 15 e os éxons 9 a 15 por LR-PCR, com o objetivo de diferenciar as sequências do *PMS2* e *PMS2CL*, considerando a perda do éxon 10 pelo pseudogene. A estratégia demonstrou eficaz para a identificação de variantes nos éxon 9, 13 e 14 do gene funcional, entretanto, não se mostrou eficiente para a diferenciação do éxon 15 entre as duas sequências, considerando que *PMS2* é idêntico a *PMS2CL* em aproximadamente 900 pb acima do éxon 15. A estratégia também não se mostrou eficiente na identificação de variantes no éxon 12, considerando a similaridade de 100% entre *PMS2* e essa região paráloga no pseudogene (HAYWARD *et al.*, 2007).

Uma segunda forma de diferenciar a aplicação da metodologia de PCR de longo alcance a partir de RNA poderia ser utilizada para a identificação de possíveis PSVs, com a avaliação da frequência de transferência de sequências que ocorrem entre as regiões duplicadas de *PMS2* e *PMS2CL* (KLIFT *et al.*, 2010). Entretanto, em nosso estudo, não foi possível realizar esta etapa, uma vez que não tivemos acesso as amostras de RNA dos pacientes rastreados.

6.5) Novas variantes depositadas nos bancos de dados

De todas as variantes encontradas para os genes MMR (n=96), doze ainda não tinham sido descritas em nenhum dos bancos de dados utilizados como fonte de consulta para este trabalho (LOVD e ClinVar/dbSNP). A maior parte destas variantes (n=6) foram preditas como possivelmente danosas pelos programas de predição *in silico* utilizados (PolyPhen, SIFT, PROVEAN), conforme demonstrado na Tabela 6.1 e foram submetidas aos bancos de dados LOVD e ClinVar.

Tabela 6.2: Alterações encontradas neste estudo que ainda não descritas pelos bancos de dados LOVD e ClinVar.

Gene	Posição (GRCh37)	Alteração	Ptn	E	Predição <i>in silico</i>	LOVD	ClinVar
MSH2	47703578	c.2078G>A	p.C693Y	13	Provavelmente danosa	SIM	NÃO
MSH2	47698176	c.1738_1741del GAAA	p.E580Lfs	11	Provavelmente danosa	SIM	NÃO
MSH6	48010553	c.181G>A	p.A61T	1	Provavelmente benigna	NÃO	NÃO
MSH6	48018201	c.402dupT	p.F133fs	2	Provavelmente danosa	NÃO	NÃO
MSH6	48033890	c.4002-27delT	-	L_9	-	NÃO	NÃO
PMS2	6048727	c.-77delA	NA	3' UTR	-	NÃO	NÃO

7) CONCLUSÕES

Pela primeira vez foi descrito na literatura um ensaio que utiliza de PCRs *multiplex* acopladas ao NGS para a obtenção de variantes em regiões codificantes e flanqueadoras dos genes de MMR. Esta metodologia permitiu reduzir o tempo e os custos para rastreamento molecular dos genes envolvidos com a SL, quando comparado ao Sequenciamento de Sanger, que é a metodologia padrão ouro utilizada atualmente. Com a redução do tempo e custo para a identificação de mutações nos genes de MMR, a metodologia padronizada neste projeto pode contribuir para uma intervenção mais acelerada para os pacientes que preenchem os critérios clínicos para a SL.

A metodologia PCR multiplex/ LR-PCR acoplada ao NGS permitiu a identificação de dos sítios variáveis com alta profundidade média de cobertura. No total, foram identificadas 96 alterações, sendo que *MLH1* foi o gene com o maior de mutações patogênicas e VUS em éxons ou sítios de *splicing* (n=12), seguido de *MSH2* (n=7), *MSH6* (n=5) e *PMS2* (n=1), respectivamente. Do número total de pacientes rastreados (n=65) por NGS, 49,2% (n=32) apresentaram algum tipo de variante pontual em um dos genes de MMR.

Foram encontradas 16 VUS em regiões codificantes ou em sítios de *splicing*. Parte destas variantes (n=6) foram preditas como deletérias, por alterarem o quadro de leitura ou o reconhecimento do sítio de *splicing* das regiões exônicas dos genes MMR. Entretanto, estudos funcionais são necessários para comprovar a patogenicidade destas alterações.

A estratégia LR-PCR utilizada para rastreamento do gene *PMS2* foi adequada para evitar a amplificação de pseudogenes parálogos aos éxons 1-5, 7-9, 11-15 deste gene. Entretanto, devido aos eventos de *crossing-over* desigual que ocorrem entre as sequências de *PMS2* e *PMS2CL*, a identificação do gene funcional foi dificultada.

8) REFERÊNCIAS BIBLIOGRÁFICAS

ACHARYA, S.; WILSON, T.; GRADIA, S.; KANE, M. F.; GUERRETTE, S.; MARSISCHKY, G.T.; KOLODNER, R.; FISHEL, R. hMSH2 forms specific mismatch-binding complexes with hMSH3 and hMSH6. **Proceedings of the National Academy of Sciences of the United States of America**, v. 93, n. 24, p. 13629–34, 1996.

ADZHUBEI, I.; JORDAN, D. M.; SUNYAEV, S. R. Predicting Functional Effect of Human Missense Mutations Using PolyPhen-2. **Current Protocols in Human Genetics**, v. 76, p. 7.20.1-7.20.41, 2013.

AHADOVA, A.; VON KNEBEL DOEBERITZ, M.; BLÄKER, H.; KLOOR, M. CTNNB1-mutant colorectal carcinomas with immediate invasive growth: a model of interval cancers in Lynch syndrome. **Familial Cancer**, p. 1–8, 2016.

AMARO, A.; CHIARA, S.; PFEFFER, U. Molecular evolution of colorectal cancer: from multistep carcinogenesis to the big bang. **Cancer and Metastasis Reviews**, v. 35, p. 63–74, 2016.

BACHER, J. W.; FLANAGAN, L. A.; SMALLEY, R. L.; NASSIF, N. A.; BURGART, L. J.; HALBERG, R. B.; MEGID, W. M. A.; THIBODEAU, S. N. Development of a fluorescent multiplex assay for detection of MSI-High tumors. **Disease markers**, v. 20, n. 4–5, p. 237–50, 2004.

BAN, C.; JUNOP, M.; YANG, W. Transformation of MutL by ATP binding and hydrolysis: a switch in DNA mismatch repair. **Cell**, v. 97, p. 85–97, 1999.

BARTKOVA, J. R. N.; LIONTOS, M.; KARAKAIDOS, P.; KLETSAS, D.; ISSAEVA, N.; VASSILIOU, L. F.; KOLETTAS, E.; NIFOROU, K.; ZOUMPOURLIS, V. C.; TAKAOKA, M.; NAKAGAWA, H.; TORT, F.; FUGGER, K.; JOHANSSON, F.; SEHESTED, M.; ANDERSEN, C. L.; DYRSKJOT, L.; ØRNTTOFT, T.; LUKAS, J.; KITTAS, C.; HELLEDAY, T.; HALAZONETIS, T.D.; BARTEK, J.; GORGOULIS, V.G. Oncogene-induced senescence is part of the tumorigenesis barrier imposed by DNA damage checkpoints. **Nature**, v. 444, n. 7119, p. 633–637, 2006.

BAUER, C. M.; RAY, A.M.; HALSTEAD-NUSSLOCH, B.A.; DEKKER, R.G.; RAYMOND, V.M.; GRUBER, S.B.; COONEY, K.A. Hereditary Prostate Cancer as a Feature of Lynch Syndrome. **Familial Cancer**, v. 10, n. 1, p. 37–42, 2011.

BECK, N. E.; TOMLINSON, I. P. M.; HOMFRAY, T.; FRAYLING, I.; HODGSON, S. V.; HAROCOPOS, C.; BODMER, W. F. Use of SSCP analysis to identify germline mutations in HNPCC families fulfilling the Amsterdam criteria. **Human Genetics**, v. 99, n. 2, p. 219–224, 1997.

BERGER, A. H.; KNUDSON, A. G.; PANDOLFI, P. P. A Continuum Model for Tumour Suppression. **Nature**, v. 476, n. 2, p. 163–169, 2015.

BIANCHI, F.; GALIZIA, E.; CATALANI, R.; BELVEDERESI, L.; FERRETTI, C.; CORRADINI, F.; CELLERINO, R. CAT25 is a mononucleotide marker to identify HNPCC patients. **The Journal of molecular diagnostics : JMD**, v. 11, n. 3, p. 248–252, 2009.

BOISSIÈRE-MICHOT, F.; FRUGIER, H.; HO-PUN-CHEUNG, A.; LOPEZ-CRAPEZ, E.; DUFFOUR, J.; BIBEAU, F. Immunohistochemical staining for p16 and BRAFV600E is useful to distinguish between sporadic and hereditary (Lynch syndrome-related) microsatellite instable colorectal carcinomas. **Virchows Archiv**, v. 1, n. 5, p. 1–10, 25 maio 2016.

BOLAND, C. R.; THIBODEAU, S.N.; HAMILTON, S.R.; SIDRANSKY, D.; ESHLEMAN, J. R.; BURT, R.W.; MELTZER, A.J.; RODRIGUEZ-BIGAS, M.A.; FODDE, R.; RANZANI, G.N.; SRIVASTAVA, S. A. National Cancer Institute Workshop on Microsatellite Instability for Cancer Detection and Familial Predisposition: Development of International Criteria for the Determination of Microsatellite Instability. **Cancer Research**, v. 59, n. 1, p. 249–256, 1999.

BOLAND, C. R.; GOEL, A. Microsatellite Instability in Colorectal Cancer. **Gastroenterology**, v. 138, n. 6, p. 2073–2087, 2011.

BORRAS, E.; PINEDA, M.; CADINANOS, J.; DEL VALLE, J.; BRIEGER, A.; HINRICHSEN, I.; CABANILLAS, R.; NAVARRO, M.; BRUNET, J.; SANJUAN, X.; MUSULEN, E.; VAN DER KLIFT, H.; LAZARO, C.; PLOTZ, G.; BLANCO, I.; CAPELLA, G. Refining the role of pms2 in Lynch syndrome: germline mutational analysis improved by comprehensive assessment of variants. **Journal of Medical Genetics**, v. 50, n. 8, p. 552–563, 2013.

BRIEGER, A.; ENGELS, K.; SCHAEFER, D.; PLOTZ, G.; ZEUZEM, S.; RAEDLE, J.; TROJAN, J. Malignant fibrous histiocytoma is a rare Lynch syndrome-associated tumor in two German families. **Familial Cancer**, v. 10, n. 3, p. 591–595, 2011.

BRIEGER, A.; TROJAN, J.; RAEDLE, J.; PLOTZ, G.; ZEUZEM, S. Transient mismatch repair gene transfection for functional analysis of genetic hMLH1 and hMSH2 variants. **Gut**, v. 51, p. 677–684, 2002.

BRONNER, C. E.; BAKER, S. M.; MORRISON, P. T.; WARREN, G.; SMITH, L. G.; LESCOE, M. K.; KANE, M.; EARABINO, C.; LIPFORD, J.; LINDBLOM, A. Mutation in the DNA mismatch repair gene homologue hMLH1 is associated with hereditary non-polyposis colon cancer. **Nature**, v. 368, n. 6468, p. 258–261, 1994.

BURN, J.; GERDES, A.; MACRAE, F.; MECKLIN, J.; MOESLEIN, G.; OLSCHWANG, S.; ECCLES, D.; EVANS, D. G.; MAHER, E. R.; BERTARIO, L.; BISGAARD, M.; DUNLOP, M. G.; HO, J. W. HODGSON, S. V.; LINDBLOM, A.; LUBINSKI, J.; MORRISON, P. J.; MURDAY, V.; RAMESAR, R.; SIDE, L.; SCOTT, R. J.; THOMAS, H. J.; VASEN, H. F.; BARKER, G.; CRAWFORD, G.; ELLIOTT, F.; MOVAHEDI, M.; PYLVANAINEN, K.; WIJNEN, J. T.; FODDE, R.; LYNCH, H. T.; MATHERS, J. C.; BISHOP, D. T. Long-term effect of aspirin on cancer risk in carriers of hereditary colorectal cancer: an analysis from the CAPP2 randomised controlled trial. **The Lancet**, v. 378, n. 9809, p. 2081–2087, 2011.

CAMPOS, F. G.; FIGUEIREDO, M. N.; REAL MARTINEZ, C. A. Colorectal cancer risk in hamartomatous polyposis syndromes. **World Journal of Gastrointestinal Surgery**, v. 7, n. 3, p. 25, 2015.

CAPORASO, J. G.; LAUBER, C. L.; WALTERS, W.; BERG-LYONS, D.; HUNTLEY, J.; FIERER, N.; OWENS, S. M.; BETLEY, J.; FRASER, L.; BAUER, M.; GORMLEY, N.; GILBERT, J.; SMITH, G.; KNIGHT, R. Ultra-high-throughput microbial community analysis

on the Illumina HiSeq and MiSeq platforms. **The ISME Journal**, v. 6, n. 8, p. 1621–1624, 2012.

CARNEIRO DA SILVA, F.; FERREIRA, J. R. O.; TORREZAN, G. T.; FIGUEIREDO, M. C. P.; SANTOS, E. M. M.; NAKAGAWA, W. T.; BRIANESE, R. C.; PETROLINI DE OLIVEIRA, L.; BEGNANI, M. D.; AGUIAR-JUNIOR, S.; ROSSI, B. M.; FERREIRA, F. O.; CARRARO, D. M. Clinical and Molecular Characterization of Brazilian Patients Suspected to Have Lynch Syndrome. **PloS one**, v. 10, n. 10, p. e0139753, 2015.

CARVALHO, M. C. D. C. G. DE; SILVA, D. C. G. DA. Sequenciamento de DNA de nova geração e suas aplicações na genômica de plantas. **Ciência Rural**, v. 40, n. 3, p. 735–744, 2010.

CASTILLEJO, A.; VARGAS, G.; CASTILLEJO, M. I.; NAVARRO, M.; BARBERÁ, V. M.; GONZÁLEZ, S.; HERNÁNDEZ-ILLÁN, E.; BRUNET, J.; RAMÓN Y CAJAL, T.; BALMAÑA, J.; OLTRA, S.; IGLESIAS, S.; VELASCO, A.; SOLANES, A.; CAMPOS, O.; SÁNCHEZ HERAS, A. B.; GALLEGO, J.; CARRASCO, E.; GONZÁLEZ JUAN, D.; SEGURA, A.; CHIRIVELLA, I.; JUAN, M. J.; TENA, I.; LÁZARO, C.; BLANCO, I. PINEDA, M.; CAPELLÁ, G.; SOTO, J. L. Prevalence of germline MUTYH mutations among Lynch-like syndrome patients. **European Journal of Cancer**, v. 50, n. 13, p. 2241–2250, 2014.

CHADWICK, R. B.; PYATT, R. E.; NIEMANN, T. H.; RICHARDS, S. K.; JOHNSON, C. K.; STEVENS, M. W.; MEEK, J. E.; HAMPEL, H.; PRIOR, T. W. DE LA CHAPELLE, A. Hereditary and somatic DNA mismatch repair gene mutations in sporadic endometrial carcinoma. **Journal of medical genetics**, v. 38, n. 7, p. 461–466, 2001.

CHEN, S. WANG, W.; LEE, S.; NAFA, K.; LEE, J.; ROMANS, K.; WATSON, P.; GRUBER, S. B.; EUHUS, D.; KINZLER, K. W.; JASS, J.; LOND, D.; LINDOR, N. M.; CASEY, G.; ELLIS, N. Prediction of Germline Mutations and Cancer Risk in the Lynch Syndrome. **JAMA : the journal of the American Medical Association**, v. 296, n. 12, p. 1479–1487, 2006.

CHOI, Y. A fast computation of pairwise sequence alignment scores between a protein and a set of single-locus variants of another protein. **Proceedings of the ACM Conference on Bioinformatics, Computational Biology and Biomedicine - BCB '12**, p. 414–417, 2012.

DA SILVA, F. C.; DE OLIVEIRA, L. P.; SANTOS, E. M.; NAKAGAWA, W. T.; AGUIAR JUNIOR, S.; VALENTIN, M. D.; ROSSI, B. M.; DE OLIVEIRA FERREIRA, F. Frequency of extracolonic tumors in Brazilian families with Lynch syndrome: Analysis of a hereditary colorectal cancer institutional registry. **Familial Cancer**, v. 9, n. 4, p. 563–570, 2010.

DASHTI, S. G.; CHAU, R.; OUAKRIM, D.A.; BUCHANAN, D.D.; CLENDENNING, M.; YOUNG, J. P.; WINSHIP, I. M.; ARNOLD, J.; AHNEN, D. J.; HAILE, R. W.; CASEY, G.; GALLINGER, S.; THIBODEAU, S. N.; LINDOR, N. M.; LE MARCHAND, L.; NEWCOMB, P.A.; POTTER, J.D; BARON, J.A.; HOPPER, J. L.; JENKINS, M. A.; WIN, A.K. Female Hormonal Factors and the Risk of Endometrial Cancer in Lynch Syndrome. **Jama**, v. 314, n. 1, p. 61, 2015.

DE LEENEER, K.; HELLEMANS, J.; DE SCHRIJVER, J.; BAETENS, M.; POPPE, B.; VAN CRIEKINGE, W.; DE PAEPE, A.; COUCKE, P.; CLAES, K. Massive parallel

amplicon sequencing of the breast cancer genes BRCA1 and BRCA2: Opportunities, challenges, and limitations. **Human Mutation**, v. 32, n. 3, p. 335–344, 2011.

DE VOS, M.; HAYWARD, B. E.; PICTON, S.; SHERIDAN, E.; BONTHRON, DA.T. Novel PMS2 pseudogenes can conceal recessive mutations causing a distinctive childhood cancer syndrome. **American journal of human genetics**, v. 74, n. 5, p. 954–964, 2004.

DENG, G.; BELL, I.; CRAWLEY, S.; GUM, J.; TERDIMAN, J. P.; ALLEN, B.; TRUTA, B.; SLEISENGER, M. H.; KIM, Y. BRAF Mutation Is Frequently Present in Sporadic Colorectal Cancer with Methylated hMLH1, but Not in Hereditary Nonpolyposis Colorectal Cancer. **Clinical Cancer Research**, v. 10, n. 415, p. 191–195, 2004.

DIEUMEGARD, B.; GRANDJOUAN, S.; SABOURIN, J. C.; LE BIHAN, M. L.; LEFRÈRE, I.; BELLEFQIH PIGNON, J. P.; ROUGIER, P.; LASSER, P.; BÉNARD, J.; COUTURIER, D.; BRESSAC-DE PAILLERETS, B. Extensive molecular screening for hereditary non-polyposis colorectal cancer. **British journal of cancer**, v. 82, n. 4, p. 871–80, 2000.

DOMINGUEZ, M.; DA SILVA, F. C.; SANTOS, E. M. M.; LISBOA, B. G.; DE OLIVEIRA FERREIRA, F.; GOMY, I.; NAKAGAWA, W. T.; JUNIOR, S. A.; REDAL, M.; VACCARO, C.; VALLE, A. D.; SARROCA, C.; CARRARO, D. M.; ROSSI, B. M. Characterization of germline mutations of MLH1 and MSH2 in unrelated south American suspected Lynch syndrome individuals. **Familial Cancer**, v. 10, n.4, p. 641–647, 2011.

DOMINGUEZ-VALENTIN, M.; NILBERT, M.; WERNHOFF, P.; LÓPEZ-KÖSTNER, F.; VACCARO, C.; SARROCA, C.; PALMERO, E. I.; GIRALDO, A.; ASHTON-PROLLA, P.; ALVAREZ, K.; FERRO, A.; NEFFA, F.; CARIS, J.; CARRARO, D. M. ROSSI, B. M. Mutation spectrum in South American Lynch syndrome families. **Hereditary cancer in clinical practice**, v. 11, n. 1, p. 18, 2013.

DOMINGUEZ-VALENTIN, M.; WERNHOFF, P.; CAJAL, A. R.; KALFAYAN, P. G.; PIÑERO, T. A.; GONZALEZ, M. L.; FERRO, A.; SAMMARTINO, I.; CAUSADA CALO, N. S.; VACCARO, C. A. MLH1 Ile219Val Polymorphism in Argentinean Families with Suspected Lynch Syndrome. **Frontiers in Oncology**, v. 6, n. August, p. 1–5, 2016.

DUTTA, R.; INOUYE, M. GHKL, an emergent ATPase/kinase superfamily. **Trends in Biochemical Sciences**, v. 25, n. 1, p. 24–28, 2000.

EDELMANN, W.; UMAR, A.; YANG, K.; HEYER, J.; KUCHERLAPATI, M.; LIA, M.; KNEITZ, B.; AVDIEVICH, E.; FAN, K.; WONG, E.; CROUSE, G.; KUNKEL, T.; LIPKIN, M.; KOLODNER, R. D.; KUCHERLAPATI, R. The DNA mismatch repair genes Msh3 and Msh6 cooperate in intestinal tumor suppression. **Cancer Research**, v. 60, n. 4, p. 803–807, 2000.

FANG, W. H.; MODRICH, P. Human strand-specific mismatch repair occurs by a bidirectional mechanism similar to that of the bacterial reaction. **Journal of Biological Chemistry**, v. 268, n. 16, p. 11838–11844, 1993.

FEARON, E. F.; VOGELSTEIN, B. A Genetic model for Colorectal Tumorigenesis. **Cell**, v. 61, n. June, 1, p. 759–767, 1989.

FEITELSON, M. A.; ARZUMANYAN, A.; KULATHINAL, R. J.; BLAIN, S. W.; HOLCOMBE, R. F.; MAHAJNA, J.; MARINO, M.; MARTINEZ-CHANTAR, M. L.;

NAWROTH, R.; SANCHEZ-GARCIA, I.; SHARMA, D.; SAXENA, N. K.; SINGH, N.; VLACHOSTERGIOS, P. J.; GUO, S.; HONOKI, K.; FUJII, H.; GEORGAKILAS, A. G.; AMEDEI, A.; NICCOLAI, E.; AMIN, A.; ASHRAF, S. S.; BOOSANI, C. S.; GUHA, G.; CIRIOLO, M. R.; AQUILANO, K.; CHEN, S.; MOHAMMED, S. I.; AZMI, A. S.; BHAKTA, D.; HALICKA, D.; NOWSHEEN, S. Sustained proliferation in cancer: Mechanisms and novel therapeutic targets. **Seminars in Cancer Biology**, v. 02.006, p. 1–30, 2015.

FERLAY, J.; SOERJOMATARAM, I.; DIKSHIT, RA.; ESER, S.; MATHERS, C.; REBELO, M.; PARKIN, D. M.; FORMAN, D.; BRAY, F. Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. **International Journal of Cancer**, v. 386, n. 136, p. 359–386, 2015.

FISHEL, R.; LESCOE, M. K.; RAO, M. R. S.; COPELAND, N. G.; JENKINS, N. A.; GARBER, J.; KANE, M.; KOLODNER, R. The human mutator gene homolog MSH2 and its association with hereditary nonpolyposis colon cancer. **Cell**, v. 75, n. 5, p. 1027–1038, 1993.

FROGGATT, N. J.; GREEN, J.; BRASSETT, C.; EVANS, D. G. R.; BISHOP, D.T.; KOLODNER, R.; MAHER, E. R. A common MSH2 mutation in English and North American HNPCC families: origin, phenotypic expression, and sex specific differences in colorectal cancer. **J Med Genet**, v. 36, p.97–102, 1999.

GAMMIE, A. E.; ERDENIZ, N.; BEAVER, J.; DEVLIN, B.; NANJI, A.; ROSE, M. D. Functional characterization of pathogenic human MSH2 missense mutations in *Saccharomyces cerevisiae*. **Genetics**, v. 177, n. 2, p. 707–721, 2007.

GILSING, A. M. J.; SCHOUTEN, L. J.; GOLDBOHM, R. A.; DAGNELIE, P. C.; VAN DEN BRANDT, P.; WEIJENBERG, M. P. Vegetarianism, low meat consumption and the risk of colorectal cancer in a population based cohort study. **Scientific Reports**, v. 5, n. April, p. 13484, 2015.

GOODENBERGER, M. L.; THOMAS, B. C.; RIEGERT-JOHNSON, D.; BOLAND, C. R.; PLON, S. E.; CLENDENNING, M.; WIN, A. K.; SENTER, L.; LIPKIN, S. M.; STADLER, Z. K.; MACRAE, F.; LYNCH, H. T.; WEITZEL, J. N.; DE LA CHAPELLE, A.; SYNGAL, S.; LYNCH, P.; PARRY, S.; JENKINS, M.; GALLINGER, S.; HOLTER, S.; ARONSON, M.; NEWCOMB, P.; BURNETT, T.; LE MARCHAND, L.; PICHURIN, P.; HAMPEL, H.; TERDIMAN, J. P.; LU, K. H.; THIBODEAU, S.; LINDOR, N. M. PMS2 monoallelic mutation carriers: the known unknown. **Genetics in Medicine**, v. 18, n. 1, p. 23–29, 2015.

GORGOLIS VG, VASSILIOU LV, KARAKAIDOS P, ZACHARATOS P, KOTSINAS A, LILOGLOU T, VENERE M, DITULLIO RA JR, KASTRINAKIS NG, LEVY B, KLETSAS D, YONETA A, HERLYN M, KITTAS C, H. T. Activation of the DNA damage checkpoint and genomic instability in human precancerous lesions. **Nature**, v. 434, p. 907–913, 2005.

GRILLEY, M.; HOLMES, J.; YASHAR, B.; MODRICH, P. Mechanisms of DNA-mismatch correction. **Mutation Research**, v. 236, n. March, p. 253–267, 1990.

GROOTHUIZEN, F. S.; SIXMA, T. K. The conserved molecular machinery in DNA mismatch repair enzyme structures. **DNA Repair**, v. 38, p. 14–23, 2016.

GU, Y.; PARKER, A.; WILSON, T. M.; BAI, H.; CHANG, D. Y.; LU, A. Human MutY homolog, a DNA glycosylase involved in base excision repair, physically and functionally

interacts with mismatch repair proteins human MutS homolog 2/human MutS homolog 6. **Journal of Biological Chemistry**, v. 277, n. 13, p. 11135–11142, 2002.

GUENEAU, E.; DHERIN, C.; LEGRAND, P.; TELLIER-LEBEGUE, C.; GILQUIN, B.; BONNESOEUR, P.; LONDINO, F.; QUEMENER, C.; LE DU, M.; MÁRQUEZ, J.; MOUTIEZ, M.; GONDROY, M.; BOITEUX, S.; CHARBONNIER, J. Structure of the MutL α C-terminal domain reveals how Mlh1 contributes to Pms1 endonuclease site. **Nature structural & molecular biology**, v. 20, n. 4, p. 461–8, 2013.

HALAZONETIS, T. D.; GORGOULIS, V. G.; BARTEK, J. An oncogene-induced DNA damage model for cancer development. **Science**, v. 319, n. 5868, p. 1352–1355, 2008.

HAMPEL, H. Genetic counseling and cascade genetic testing in Lynch syndrome. **Familial Cancer**, p. 1–5, 2016.

HANAHAH, DOUGLAS AND WEINBERG, R. A. The Hallmarks of cancer. **Cell**, v. 100, n. 3, p. 57–70, 2010.

HANAHAH, D.; WEINBERG, R. A. Hallmarks of cancer: the next generation. **Cell**, v. 144, n. 5, p. 646–674, 2011.

HANSEN, M. F.; NECKMANN, U.; LAVIK, L. A. S.; VOLD, T.; GILDE, B.; TOFT, R. K.; SJURSEN, W. A massive parallel sequencing workflow for diagnostic genetic testing of mismatch repair genes. **Molecular genetics & genomic medicine**, v. 2, n. 2, p. 186–200, 2014.

HAYWARD, B. E.; DE VOS, M.; VALLELEY, E. M. A.; CHARLTON, R. S.; TAYLOR, G. R.; SHERIDAN, E.; BONTHRON, D. T. Extensive Gene Conversion at the PMS2 DNA Mismatch Repair Locus. **Human Mutation**, v. 28, n. 5, p. 424–430, 2007.

HAMPEL, H.; FRANKEL, W. L.; MARTIN, E.; ARNOLD, M.; KHANDUJA, K.; KUEBLER, P.; NAKAGAWA, H.; SOTAMAA, K.; PRIOR, T.W.; WESTMAN, J.; PANESCU, J. Screening for the Lynch Syndrome (Hereditary Nonpolyposis Colorectal Cancer). **The New England Journal of medicine**, v. 352, n. 18, p. 1815–1860, 2005.

HENDRIKS, Y.; FRANKEN, P.; DIERSSSEN, J. W.; DE LEEUW, W.; WIJNEN, J.; DREEF, E.; TOPS, C.; BREUNING, M.; BRÖCKER-VRIENDS, A.; VASEN, H.; FODDE, R.; MORREAU, H. Conventional and tissue microarray immunohistochemical expression analysis of mismatch repair in hereditary colorectal tumors. **The American journal of pathology**, v. 162, n. 2, p. 469–77, 2003.

HERCEG, Z. Epigenetics and cancer: Towards an evaluation of the impact of environmental and dietary factors. **Mutagenesis**, v. 22, n. 2, p. 91–103, 2007.

HILL, M. J.; MORSON, B. C.; BUSSEY, H. J. R. Etiology of Adenoma-Carcinoma Sequence in Large Bowel. **The Lancet**, v. 311, n. 8058, p. 245–247, 1978.

HINGORANI, M. M. **Mismatch binding, ADP-ATP exchange and intramolecular signaling during mismatch repair** DNA Repair, 2016.

HITCH, K.; JOSEPH, G.; GULTINAN, J.; KIANMAHD, J.; YOUNGBLOM, J.; BLANCO, A. Lynch Syndrome Patients' Views of and Preferences for Return of Results Following Whole Exome Sequencing. **Journal of Genetic Counseling**, v. 23, n. 4, p. 539–551, 2014.

HOPPMAN-CHANEY, N.; PETERSON, L. M.; KLEE, E. W.; MIDDHA, S.; COURTEAU, L. K.; FERBER, M. J. Evaluation of oligonucleotide sequence capture arrays and comparison of next-generation sequencing platforms for use in molecular diagnostics. **Clinical Chemistry**, v. 56, n. 8, p. 1297–1306, 2010.

ILLUMINA. **Nextera XT DNA Sample Preparation Guide**. ed, San Diego, 2012, 48 p.

ISIDRO, G.; VEIGA, I.; MATOS, P.; ALMEIDA, S.; BIZARRO, S.; MARSHALL, B.; BAPTISTA, M.; LEITE, J.; REGATEIRO, F.; SOARES, J.; CASTEDO, S.; BOAVIDA, M. G. Four novel MSH2 / MLH1 gene mutations in portuguese HNPCC families. **Human Mutation**, v. 15, n. June, p. 116, 2000.

IYER, R. R.; PLUCIENNIK, A.; BURDETT, V.; MODRICH, P. L. DNA mismatch repair: functions and mechanisms. **Chemical reviews**, v. 106, n. 2, p. 302–23, 2006.

JANSEN, A. M. L.; GEILENKIRCHEN, M. A.; WEZEL, T. V.; JAGMOHAN, S. C. Whole Gene Capture Analysis of 15 CRC Susceptibility Genes in Suspected Lynch Syndrome Patients. **Plos One**, v. 11, n. 6, p. 1–15, 2016.

JASPERSON, K.; BURT, R. W. The Genetics of Colorectal Cancer. **Surgical Oncology Clinics of North America**, v. 24, n. 4, p. 683–703, 2015.

JASPERSON, K. W.; TUOHY, T. M.; NEKLASON, D. W.; BURT, R. W. Hereditary and Familial Colon Cancer. **Gastroenterology**, v. 138, n. 6, p. 2044–2058, 2010.

JIRICNY, J. The multifaceted mismatch-repair system. **Nature Reviews Molecular Cell Biology**, v. 7, n. 5, p. 335–346, 2006.

JULIÉ, C.; TRÉSALLET, C.; BROUQUET, A.; VALLOT, C.; ZIMMERMANN, U.; MITRY, E.; RADVANYI, F.; ROULEAU, E.; LIDEREAU, R.; COULET, F.; OLSCHWANG, S.; FREBOURG, T.; ROUGIER, P.; NORDLINGER, B.; LAURENT-PUIG, P.; BOILEAU, C.; FRANC, B.; MUTI, C.; HOFMANN-RADVANYI, H. Identification in daily practice of patients with lynch syndrome (hereditary nonpolyposis colorectal cancer): Revised Bethesda guidelines-based approach versus molecular screening. **American Journal of Gastroenterology**, v. 103, n. 11, p. 2825–2835, 2008.

JUNG, S. M.; YOON, Y. S.; LIM, S.; YU, C. S.; KIM, J. C. Clinicopathological features of familial adenomatous polyposis in Korean patients. **World Journal of Gastroenterology**, v. 22, n. 17, p. 4380, 2016.

KANG, S. Y.; PARK, C. K.; CHANG, D. K.; KIM, J. W.; SON, H. J.; CHO, Y. B.; YUN, S. H.; KIM, H. C.; KWON, M.; KIM, K. M. Lynch-like syndrome: Characterization and comparison with EPCAM deletion carriers. **International Journal of Cancer**, v. 136, n. 7, p. 1568–1578, 2015.

KASTRINOS, F.; MUKHERJEE, B.; TAYOB, N.; WANG, F.; SPARR, J.; RAYMOND, V. M.; BANDIPALLIAM, P.; STOFFEL, E. M.; GRUBER, S. B.; SYNGAL, S. Risk of pancreatic cancer in families with Lynch syndrome. **Jama**, v. 302, n. 16, p. 1790–5, 2009.

KLIFT, H. M. V. D.; TOPS, A. C. M. J.; BIK, E. C.; BOOGAARD, M. W.; BORGSTEIN, A.; HANSSON, K. B. M.; AUSEMS, M. G. E. M.; GARCIA, E. G.; GREEN, A.; HES, F. J.; IZATT, L.; HEST, L.; P. VAN ALONSO, A.M.; VRIENDS, A. H. J. T.; WAGNER, A.; ZELST-STAMS, W. A. G. V.; VASEN, H. F. A.; MORREAU, H.; DEVILEE, P.; WIJNEN, J. T. Quantification of Sequence Exchange Events between PMS2 and PMS2CL Provides a Basis for Improved Mutation Scanning of Lynch Syndrome Patients *Human Mutation*, v. 31, n. 5, p. 578–587, 2010.

KLOOR, M.; STAFFA, L.; AHADOVA, A.; VON KNEBEL DOEBERITZ, M. Clinical significance of microsatellite instability in colorectal cancer. *Langenbeck's Archives of Surgery*, v. 399, n. 1, p. 23–31, 2014.

KNUDSON, A. G. Hereditary Cancer , Oncogenes , and Antioncogenes. *Cancer Research*, v. 45, n. April, p. 1437–1443, 1985.

KNUDSON JR, A. G. Mutation and Cancer: Statistical Study of Retinoblastoma. *Proceedings of the National Academy of Sciences of the United States of America*, v. 68, n. 4, p. 820–823, 1971.

KONDO, E.; SUZUKI, H.; HORII, A.; FUKUSHIGE, S. A yeast two-hybrid assay provides a simple way to evaluate the vast majority of hMLH1 germ-line mutations. *Cancer Research*, v. 63, n. 12, p. 3302–3308, 2003.

KRAVOCHUCK, S. E.; CHURCH, J. M. Hereditary non-polyposis colorectal cancer/Lynch syndrome in three dimensions. *ANZ Journal of Surgery*, p. 1445–2197, mar. 2016.

KWOK, C.-T.; VOGELAAR, I. P.; VAN ZELST-STAMS, W.; MENSENKAMP, A. R.; LIGTENBERG, M. J.; RAPKINS, R.W.; WARD, R. L.; CHUN, N.; FORD, J. M.; LADABAUM, U.; MCKINNON, W. C.; GREENBLATT, M. S.; HITCHINS, M. P. The MLH1 c.-27C>A and c.85G>T variants are linked to dominantly inherited MLH1 epimutation and are borne on a European ancestral haplotype. *European journal of human genetics : EJHG*, v. 22, n. 5, p. 617–24, 2014.

LAMERS, M. H.; PERRAKIS, A.; ENZLIN, J. H.; WINTERWERP, H. H.; DE WIND, N.; SIXMA, T. K. The crystal structure of DNA mismatch repair protein MutS binding to a G x T mismatch. *Nature*, v. 407, n. 6805, p. 711–7, 2000.

LATCHFORD, A. R.; PHILLIPS, R. K. S. Gastrointestinal polyps and cancer in Peutz-Jeghers syndrome: Clinical aspects. *Familial Cancer*, v. 10, n. 3, p. 455–461, 2011.

LI, J.; DAI, H.; FENG, Y.; TANG, J.; CHEN, S.; TIAN, X.; GORMAN, E.; SCHMITT, E. S.; HANSEN, T.; WANG, J.; PLON, S. E.; ZHANG, V. W.; WONG, L. C. A. Comprehensive Strategy for Accurate Mutation Detection of the Highly Homologous PMS2. *The Journal of Molecular Diagnostics*, v. 17, n. 5, p. 545–553, 2015.

LI, S. K. H.; MARTIN, A. Mismatch Repair and Colon Cancer : Mechanisms and Therapies Explored. *Trends in Molecular Medicine*, v. 22, n. 4, p. 274–289, 2016.

LIGTENBERG, M. J. L.; KUIPER, R. P.; CHAN, T. L.; GOOSSENS, M.; HEBEDA, K. M.; VOORENDT, M.; LEE, T. Y. H.; BODMER, D.; HOENSELAAR, E.; HENDRIKS-CORNELISSEN, S. J. B.; TSUI, W. Y.; KONG, C. K.; BRUNNER, H. G.; VAN KESSEL, A.

G.; YUEN, S. T.; VAN KRIEKEN, J. H. J. M.; LEUNG, S. Y.; HOOGERBRUGGE, N. Heritable somatic methylation and inactivation of MSH2 in families with Lynch syndrome due to deletion of the 3' exons of TACSTD1. **Nature genetics**, v. 41, n. 1, p. 112–117, 2009.

LUBBE, S. J.; DI BERNARDO, M. C.; CHANDLER, I. P.; HOULSTON, R. S. Clinical implications of the colorectal cancer risk associated with MUTYH mutation. **Journal of clinical oncology**, v. 27, n. 24, p. 3975–80, 2009.

LYNCH, H. T.; OMARA, M. D.; SHAW, M. W.; MICH, A. A.; MAGNUDSON, C. W.; LARSEN, A. L.; OMAHA, A. J.; KRUSH, M.S. Hereditary Factors in Cancer: Study of Two Large Midwestern Kindreds. **Archives of Internal Medicine**, v. 117, n. February, p. 206–2012, 1966.

LYNCH, H. T.; KRUSH, A. J. Cancer family “G” Revisited: 1895-1970. **Cancer**, v. 27, n. 6, p. 1505–1511, 1971.

LYNCH, P. M.; LYNCH, H. T.; HARRIS, R. E. Hereditary proximal colonic cancer. **Diseases of the Colon & Rectum**, v. 20, n. 8, p. 661–668, 1977.

MANGOLD, E.; PAGENSTECHER, C.; FRIEDL, W.; MATHIAK, M.; BUETTNER, R.; ENGEL, C.; LOEFFLER, M.; HOLINSKI-FEDER, E.; MÜLLER-KOCH, Y.; KELLER, G.; SCHACKERT, H. K.; KRÜGER, S.; GOECKE, T.; MOESLEIN, G.; KLOOR, M.; GEBERT, J.; KUNSTMANN, E.; SCHULMANN, K.; RÜSCHOFF, J.; PROPPING, P. Spectrum and frequencies of mutations in MSH2 and MLH1 identified in 1,721 German families suspected of hereditary nonpolyposis colorectal cancer. **International Journal of Cancer**, v. 116, n. 5, p. 692–702, 2005.

MARGULIES, M.; EGHOLM, M.; ALTMAN, W. E.; ATTIYA, S.; BADER, J. S.; BEMBEN, L.; BERKA, J.; BRAVERMAN, M. S.; CHEN, Y.; CHEN, Z.; DEWELL, S. B.; DU, L.; FIERRO, J. M.; GOMES, X. V.; GODWIN, B. C.; HE, W.; HELGESEN, S.; HO, C. H.; HO, C. H.; IRZYK, G. P.; JANDO, S. C.; ALENQUER, M. L. I.; JARVIE, T. P.; JIRAGE, K. B.; KIM, J.; KNIGHT, J. R.; LANZA, J. R.; LEAMON, J. H.; LEFKOWITZ, S. M.; LEI, M.; LI, J.; LOHMAN, K. L.; LU, H.; MAKHIJANI, V. B.; MCDADE, K. E.; MCKENNA, M. P.; MYERS, E. W.; NICKERSON, E.; NOBILE, J. R.; PLANT, R.; PUC, B. P.; RONAN, M. T.; ROTH, G. T.; SARKIS, G. J.; SIMONS, J. F.; SIMPSON, J. W.; SRINIVASAN, M.; TARTARO, K. R.; TOMASZ, A.; VOGT, K.; VOLKMER, G.; WANG, S. H.; WANG, Y.; WEINER, M. P.; YU, P.; BEGLEY, R. F.; ROTHBERG, J. M. Genome sequencing in microfabricated high-density picolitre reactors. **Nature**, v. 437, n. 7057, p. 376–80, 2005.

MATTHIJS, G.; SOUCHE, E.; ALDERS, M.; CORVELEYN, A.; ECK, S.; FEENSTRA, I.; RACE, V.; SISTERMANS, E.; STURM, M.; WEISS, M.; YNTEMA, H.; BAKKER, E.; SCHEFFER, H.; BAUER, P. Guidelines for diagnostic next-generation sequencing. **European journal of human genetics**, v. 24, n. 2016, p. 2–5, 2015.

MENDILLO, M. L. HARGREAVES, V. V.; JAMISON, J. W.; MO, A. O.; LI, S.; PUTNAM, C. D.; WOODS, V. L.; KOLODNER, R. D. A conserved MutS homolog connector domain interface interacts with MutL homologs. **Proceedings of the National Academy of Sciences of the United States of America**, v. 106, n. 52, p. 22223–22228, 2009.

MENSENKAMP, A. R.; VOGELAAR, I. P.; VAN ZELST-STAMS, W. A. G.; GOOSSENS, M.; OUCHENE, H.; HENDRIKS-CORNELISSEN, S. J. B.; KWINT, M. P.;

HOOGERBRUGGE, N.; NAGTEGAAL, I. D.; LIGTENBERG, M. J. L. Somatic mutations in MLH1 and MSH2 are a frequent cause of mismatch-repair deficiency in lynch syndrome-like tumors. **Gastroenterology**, v. 146, n. 3, p. 643–646.e8, 2014.

MINISTÉRIO DA SAÚDE. **ABC DO CÂNCER: Abordagens Básicas para o Controle do Câncer**. Rio de Janeiro, RJ, 2012, 129p.

MINISTÉRIO DA SAÚDE. **Estimativa 2016: incidência de câncer no Brasil / Instituto Nacional de Câncer José Alencar Gomes da Silva**. Rio de Janeiro, RJ, 2015, 126p.

MIYAKI, M.; KONISHI, M.; TANAKA, K.; KIKUCHI-YANOSHITA, R.; MURAOKA, M.; YASUNO, M.; IGARI, T.; KOIKE, M.; CHIBA, M.; MORI, T. Germline mutation of MSH6 as the cause of hereditary nonpolyposis colorectal cancer. **Nature Genetics**, v. 17, n. 3, p. 271–272, 1997.

MUKHERJEE, S.; FEIG, M. Conformational change in MSH2-MSH6 upon binding DNA coupled to ATPase activity. **Biophysical Journal**, v. 96, n. 11, p. L63–L65, 2009.

NEGRINI, S.; GORGOULIS, V. G.; HALAZONETIS, T. D. Genomic instability--an evolving hallmark of cancer. **Nature reviews. Molecular cell biology**, v. 11, n. 3, p. 220–228, 2010.

NG, P. C.; HENIKOFF, S. SIFT : predicting amino acid changes that affect protein function. v. 31, n. 13, p. 3812–3814, 2003.

NICOLAIDES, N. C.; PAPADOPOULOS, N.; LIU, B.; WEI, Y. F.; CARTER, K. C.; RUBEN, S. M.; ROSEN, C. A.; HASELTINE, W. A.; FLEISCHMANN, R. D.; FRASER, C. M. Mutations of two PMS homologues in hereditary nonpolyposis colon cancer. **Nature**, v.371, p. 75-80, 1994.

NOLL, A.; PAREKH, P.; KARLITZ, J. Diagnosis of Lynch syndrome before colorectal resection : does it matter ? **Techniques in Coloproctology**, v. 20, n. 4, p. 203–205, 2016.

PAJARES, J. A.; PEREA, J. Multiple primary colorectal cancer: Individual or familial predisposition? **World journal of gastrointestinal oncology**, v. 7, n. 12, p. 434–44, 2015.

PASTRELLO, C.; PIN, E.; MARRONI, F.; BEDIN, C.; FORNASARIG, M.; TIBILETTI, M. G.; OLIANI, C.; PONZ DE LEON, M.; URSO, E. D.; DELLA PUPPA, L.; AGOSTINI, M.; VIEL, A. Integrated analysis of unclassified variants in mismatch repair genes. **Genet Med**, v. 13, n. 2, p. 115–124, 2011.

PLAZZER, J. P.; SIJMONS, R. H.; WOODS, M. O.; PELTOMAKI, P.; THOMPSON, B.; DEN DUNNEN, J. T.; MACRAE, F. The InSiGHT database: Utilizing 100 years of insights into Lynch Syndrome. **Familial Cancer**, v. 12, n. 2, p. 175–180, 2013.

PLENTZ, R. R. WIEMANN, S. U.; FLEMMING, P.; MEIER, P. N.; KUBICKA, S.; KREIPE, H.; MANNS, M. P.; RUDOLPH, K. L. Telomere shortening of epithelial cells characterises the adenoma-carcinoma transition of human colorectal cancer. **Gut**, v. 52, n. 9, p. 1304–1307, 2003.

POYNTER, J. N.; SIEGMUND, K. D.; WEISENBERGER, D. J.; LONG, T. I.; THIBODEAU, S. N.; LINDOR, N.; YOUNG, J.; JENKINS, M. A.; L, J.; BARON, J. A.;

BUCHANAN, D.; CASEY, G.; LEVINE, A. J.; LE, L.; GALLINGER, S.; BAPAT, B.; POTTER, J. D.; NEWCOMB, P. A. Molecular characterization of MSI-H colorectal cancer by MLHI promoter methylation, immunohistochemistry and mismatch repair germline mutation screening. **Biomarkers**, v. 17, n. 11, p. 3208–3215, 2009.

PRITCHARD, C. C.; SMITH, C.; SALIPANTE, S. J.; LEE, M. K.; THORNTON, A. M.; NORD, A. S.; GULDEN, C.; KUPFER, S. S.; SWISHER, E. M.; BENNETT, R. L.; NOVETSKY, A. P.; JARVIK, G. P.; OLOPADE, O. I.; GOODFELLOW, P. J.; KING, M.; TAIT, J. F.; WALSH, T. ColoSeq Provides Comprehensive Lynch and Polyposis Syndrome Mutational Analysis Using Massively Parallel Sequencing. **The Journal of Molecular Diagnostics**, v. 14, n. 4, p. 357–366, 2012.

PRITCHARD, J. K.; STEPHENS, M.; DONNELLY, P. Inference of population structure using multilocus genotype data. **Genetics**, v. 155, n. 2, p. 945–959, 2000.

QUAIL, M.; SMITH, M. E.; COUPLAND, P.; OTTO, THOMAS. D.; HARRIS, S. R.; CONNOR, T; RICHARD BERTONI, A.; SWERDLOW, H. P.; GU, Y. A tale of three next generation sequencing platforms: comparison of Ion torrent, pacific biosciences and illumina MiSeq sequencers. **BMC Genomics**, v. 13, n. 1, p. 1, 2012.

RAYMOND, V. M; EVERETT, J. N.; FURTADO, L. V.; GUSTAFSON, S. L.; JUNGBLUTH, C. R.; GRUBER, S. B.; HAMMER, G. D.; STOFFEL, E. M.; GREENSON, J. K.; GIORDANO, T. J.; ELSE, T. Adrenocortical carcinoma is a lynch syndrome-associated cancer. **Journal of clinical oncology : official journal of the American Society of Clinical Oncology**, v. 31, n. 24, p. 3012–3018, 2013.

RICCIARDIELLO, L.; AHNEN, D. J.; LYNCH, P. M. Chemoprevention of hereditary colon cancers: time for new strategies. **Nature Reviews Gastroenterology & Hepatology**, v. 13, p. Ahead of Print, 2016.

RING, K. L.; BRUEGL, A. S.; ALLEN, B. A.; ELKIN, E. P.; SINGH, N.; HARTMAN, A.; DANIELS, M. S.; BROADDUS, R. R. Germline multi-gene hereditary cancer panel testing in an unselected endometrial cancer cohort. **Nature Publishing Group**, p. 1–9, 2016.

RODRIGUEZ-BIGAS, M. A.; BOLAND, C. R.; HAMILTON, S. R.; HENSON, D. E.; JASS, J. R.; KHAN, P M.; LYNCH, H.; PERUCHO, M.; SMYRK, T.; SOBIN, L.; SRIVASTAVA, S. A National Cancer Institute Workshop on Hereditary Nonpolyposis Colorectal Cancer Syndrome: Meeting Highlights and Bethesda Guidelines. **Journal of the National Cancer Institute**, v. 89, n. 23, p. 1758–1762, 1997.

ROSSI, B. M. Síndrome de Lynch. **Manual Operacional- Rede Nacional de Câncer Familiar**: In: Ministério da Saúde, INCA. Rio de Janeiro: ESDEVA , 2009, p. 41-49

ROSTY, C.; CLENDENNING, M.; WALSH, M. D.; ERIKSEN, S. V.; SOUTHEY, M. C.; WINSHIP, I. M.; MACRAE, F. A.; BOUSSIOUTAS, A.; POPLAWSKI, N. K.; PARRY, S.; ARNOLD, J.; YOUNG, J. P.; CASEY, G.; HAILE, R. W.; GALLINGER, S.; LE MARCHAND, L.; NEWCOMB, P. A.; POTTER, J. D.; DERYCKE, M.; LINDOR, N. M.; THIBODEAU, S. N.; BARON, J. A.; WIN, A. K.; HOPPER, J. L.; JENKINS, M. A.; BUCHANAN, D. D. Germline mutations in PMS2 and MLH1 in individuals with solitary loss of PMS2 expression in colorectal carcinomas from the Colon Cancer Family Registry Cohort. **BMJ Open**, v. 6, n. 2, p. e010293, 2016.

RUBIO, I.; BÁÑEZ-FEIJOO, E.; ANDRÉS, L.; AGUIRRE, E.; BALMAÑA, J.; LLORT, G.; GONZÁLEZ-SANTIAGO, S.; MAORTUA, H.; TEJADA, M. I. ; MARTINEZ-BOUZAS, C. Analysis of Lynch Syndrome Mismatch Repair Genes in Women with Endometrial Cancer. **Oncology (Switzerland)**, v. 91, p. 171–176, 2016.

SCHEIBLE, M.; LOREILLE, O.; JUST, R.; IRWIN, J. Short tandem repeat typing on the 454 platform: Strategies and considerations for targeted sequencing of common forensic markers. **Forensic Science International: Genetics**, v. 12, p. 107–119, 2014.

SCHLUSSEL, A. T.; JR, RONALD G.; SETO-DONLON, S.; EGGERDING, F.; DONLON, T.; BERENBERG, J.; LYNCH, H. T. The evolution of colorectal cancer genetics — Part 2 : clinical implications and applications. **Journal of gastrointestinal oncology**, v. 5, n. 5, p. 336–344, 2014.

SCHRIJVER, J. M.; LEENEER, K.; LEFEVER, S.; SABBE, N.; PATTYN, F.; NIEUWERBURGH, F. V.; COUCKE, P.; DEFORCE, D.; VANDESOMPELE, J.; BEKAERT, S.; HELLEMANS, J.; CRIEKINGE, W. V. Analysing 454 amplicon resequencing experiments using the modular and database oriented Variant Identification Pipeline. **BMC Bioinformatics**, v. 11, n. 269, p. 1–12, 2010.

SENER, L.; CLENDENNING, M.; SOTAMAA, K.; HAMPEL, H.; GREEN, J.; LINDBLOM, A.; THIBODEAU, S. N.; NORALANE, M.; YOUNG, J.; WINSHIP, I.; WHITE, D. M.; JOHN, L.; BAGLIETTO, L.; JENKINS, M. A.; CHAPELLE, A. The clinical phenotype of Lynch syndrome due to germline PMS2 mutations. **Gastroenterology**, v. 135, n. 2, p. 419–428, 2008.

SEPPÄLÄ, T. T.; BÖHM, J. P.; FRIMAN, M.; LAHTINEN, L.; VÄYRYNEN, V. M. J.; LIPO, T. K. E.; RISTIMÄKI, A. P.; KAIRALUOMA, M. V. J.; KELLOKUMPU, I. H.; KUOPIO, T. H. I.; MECKLIN, J-P. Combination of microsatellite instability and BRAF mutation status for subtyping colorectal cancer. **British journal of cancer**, v. 112, n. 12, p. 1966–1975, 2015.

SHIN, S.; HWANG, I. S.; LEE, S. T.; CHOI, J. R. Evaluation of an amplicon-based next-generation sequencing panel for detection of BRCA1 and BRCA2 genetic variants. **Breast Cancer Research and Treatment**, v. 158, n. 3, p. 433–440, 2016.

SIE, A. S.; PRINS, J.B.; VAN ZELST-STAMS, W.A.G.; VELTMAN, J.A.; FEENSTRA, I.; HOOGERBRUGGE, N. Patient experiences with gene panels based on exome sequencing in clinical diagnostics: high acceptance and low distress. **Clinical Genetics**, v. 87, n. 4, p. 319–326, 2015.

SIRAJ, A. K.; PRABHAKARAN, S.; BAVI, P.; BU, R.; BEG, S.; HAZMI, M. AL.; AL-RASHEED, M.; AL-ASSIRI, M.; SAIRAFI, R.; AL-DAYEL, F.; AL-SANEA, N.; UDDIN, S.; AL-KURAYA, K. S. Prevalence of Lynch syndrome in a Middle Eastern population with colorectal cancer. **Cancer**, v. 121, n. 11, p. 1762–1771, 2015.

SMITH, J.; MUN THO, L.; XU, N.; A. GILLESPIE, D. The ATM-Chk2 and ATR-Chk1 pathways in DNA damage signaling and cancer. **Advances in Cancer Research**, v. 108, n. C. P. 73- 112, 2010.

SONG, M.; GARRETT, W. S.; CHAN, A. T. Nutrients, Foods, and Colorectal Cancer Prevention Mingyang. **Gastroenterology**, v. 148, n. 6, p. 1244–1260, 2015.

SUN, Z.; YU, X.; WANG, H.; ZHANG, S.; ZHAO, Z.; XU, R. Clinical significance of mismatch repair gene expression in sporadic colorectal cancer. **Experimental and Therapeutic Medicine**, v. 8, p. 1416–1422, 2014.

SURAWEERA, N.; DUVAL, A.; REPERANT, M.; VAURY, C.; FURLAN, D.; LEROY, K.; SERUCA, R.; IACOPETTA, B.; HAMELIN, R. Evaluation of tumor microsatellite instability using five quasimonomorphic mononucleotide repeats and pentaplex PCR. **Gastroenterology**, v. 123, n. 6, p. 1804–1811, 2002.

SUSSMAN, D. A.; BERERA, S.; KORU-SENGUL, T.; MIAO, F.; CARRASQUILLO, O.; NADJI, M.; ZHANG, Y.; HOSEIN, P. J.; MCCAULEY, J. L.; ABREU, M. T. Colorectal Tumors from Different Racial and Ethnic Minorities Have Similar Rates of Mismatch Repair Deficiency. **Clinical Gastroenterology and Hepatology**, 2016.

TAKAHASHI, M.; SHIMODAIRA, H.; ANDREUTTI-ZAUGG, C.; IGGO, R.; KOLODNER, R. D.; ISHIOKA, C. Functional analysis of human MLH1 variants using yeast and in vitro mismatch repair assays. **Cancer Research**, v. 67, n. 10, p. 4595–4604, 2007.

TALSETH-PALMER, B. A.; BAUER, D. C.; SJURSEN, W.; EVANS, T. J.; MCPHILLIPS, M.; PROIETTO, A.; OTTON, G.; SPIGELMAN, A. D.; SCOTT, R. J. Targeted next-generation sequencing of 22 mismatch repair genes identifies Lynch syndrome families. **Cancer Medicine**, v. 5, n. 5, p. 929–941, 2016.

TARIQ, K.; GHIAS, K. Colorectal cancer carcinogenesis : a review of Mechanisms. **Cancer Biology Medicine**, v. 13, p. 120–135, 2016.

TEN BOSCH, J. R.; GRODY, W. W. Keeping Up With the Next Generation. **The Journal of Molecular Diagnostics**, v. 10, n. 6, p. 484–492, 2008.

TEN BROEKE, S. W.; BROHET, R. M.; TOPS, C. M.; VAN DER KLIFT, H. M.; VELTHUIZEN, M. E.; BERNSTEIN, I.; CAPELLA MUNAR, G.; GOMEZ GARCIA, E.; HOGERBRUGGE, N.; LETTEBOER, T. G. W.; MENKO, F. H.; LINDBLOM, A.; MENSENKAMP, A. R.; MOLLER, P.; VAN OS, T.; RAHNER, N.; REDEKER, B. J. W.; SIJMONS, R. H.; SPRUIJT, L.; SUERINK, M.; VOS, Y. J.; WAGNER, A.; HES, F. J.; VASEN, H. F.; NIELSEN, M.; WIJNEN, J. T. Lynch syndrome caused by germline PMS2 mutations: Delineating the cancer risk. **Journal of Clinical Oncology**, v. 33, n. 4, p. 319–325, 2015.

TERDIMAN, J. P.; LEVIN, T. R.; ALLEN, B.; GUM, J. R.; FISHBACH, A.; CONRAD, P. G.; MILLER, G.; WEINBERG, V.; BACHMAN, R.; BERGOFFEN, J.; STEMBRIDGE, A.; TORIBARA, N. W.; SLEISENGER, M. H.; KIM, Y. S. Hereditary nonpolyposis colorectal cancer in young colorectal cancer patients: High-risk clinic versus population-based registry. **Gastroenterology**, v. 122, n. 4, p. 940–947, 2002.

THIBODEAU, S. N.; BREN, G.; SCHAID, D. Microsatellite instability in cancer of the proximal colon. **Science (New York, N.Y.)**, v. 260, n. May, p. 816–819, 1993.

TOLEDANO, H.; GOLDBERG, Y.; KEDAR-BARNES, I.; BARIS, H.; PORAT, R. M.; SHOCHAT, C.; BERCOVICH, D.; PIKARSKY, E.; LERER, I.; YANIV, I.; ABELIOVICH, D.; PERETZ, T. Homozygosity of MSH2 c.1906G>C germline mutation is

associated with childhood colon cancer, astrocytoma and signs of Neurofibromatosis type I. **Familial Cancer**, v. 8, n. 3, p. 187–194, 2009.

TOMSIC, J.; SENTER, L.; LIYANARACHCHI, S.; CLENDENNING, M.; VAUGHN, C. P.; JENKINS, M. A.; HOPPER, J. L.; YOUNG, J.; AMOWITZ, W.; DE LA CHAPELLE, A. Recurrent and founder mutations in the PMS2 gene. **Clinical Genetics**, v. 83, n. 3, p. 238–243, 2013.

TOSS, A.; TOMASELLO, C.; RAZZABONI, E.; CONTU, G.; GRANDI, G.; CAGNACCI, A.; SCHILDER, R. J.; CORTESI, L. Hereditary Ovarian Cancer: Not Only BRCA 1 and 2 Genes. **BioMed research international**, v. 2015, n. April, p. 1–11, 2015.

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.; LINDBLUM, A.; LYNCH, H. T.; PELTOMAKI, P.; RAMSEY, S. D.; RODRIGUEZ-BIGAS, M. A.; VASEN, H. F. A.; HAWK, E. T.; BARRETT, J. C.; FREEDMAN, A. N.; SRIVASTAVA, S. Revised Bethesda Guidelines for hereditary nonpolyposis colorectal cancer (Lynch syndrome) and microsatellite instability. **Journal of the National Cancer Institute**, v. 96, n. 4, p. 261–8, 2004.

URSO, E.; PUCCIARELLI, S.; AGOSTINI, M.; MARETTO, I.; MESCOI, C.; BERTORELLE, R.; VIEL, A.; RUGGE, M.; NITTI, D. Proximal colon cancer in patients aged 51–60 years of age should be tested for microsatellites instability. A comment on the revised bethesda guidelines. **International Journal of Colorectal Disease**, v. 23, n. 8, p. 801–806, 2008.

VAN DEN BRANDT, P. A.; GOLDBOHM, R. A. Nutrition in the prevention of gastrointestinal cancer. **Best practice & research. Clinical gastroenterology**, v. 20, n. 3, p. 589–603, 2006.

VAN DER KLIFT, H. M.; MENSENKAMP, A. R.; DROST, M.; BIK, E. C.; VOS, Y. J.; GILLE, H. J. J. P.; REDEKER, B. E. J. W.; TIERSMA, Y.; ZONNEVELD, J. B. M.; GARCÍA, E. G.; LETTEBOER, T. G. W.; OLDERODE-BERENDS, MARAN J. W.; VAN HEST, L. P.; VAN OS, T. A.; VERHOEF, S.; WAGNER, A.; VAN ASPEREN, C. J.; TEN BROEKE, S. W.; HES, F. J.; DE WIND, N.; NIELSEN, M.; DEVILEE, P.; LIGTENBERG, M. J. L.; WIJNEN, J. T.; TOPS, C. M. J. Comprehensive Mutation Analysis of PMS2 in a Large Cohort of Proband Suspected of Lynch Syndrome or Constitutional Mismatch Repair Deficiency Syndrome. **Human mutation**, v. 37, n. 11, p. 1162–1179, 2016.

VASEN, H.; WATSON, P.; MECKLIN, J.; LYNCH, H. New clinical criteria for hereditary nonpolyposis colorectal cancer (HNPCC, Lynch syndrome) proposed by the International Collaborative Group on HNPCC. **Gastroenterology**, v. 116, p. 1453–1456, 1999.

VASEN, H. F.; BLANCO, I.; AKTAN-COLLAN, K.; GOPIE, J. P.; ALONSO, A.; ARETZ, S.; BERNSTEIN, I.; BERTARIO, L.; BURN, J.; CAPELLA, G.; COLAS, C.; ENGEL, C.; FRAYLING, I. M.; GENUARDI, M.; HEINIMANN, K.; HES, F. J.; HODGSON, S. V.; KARAGIANNIS, J. A.; LALLOO, F.; LINDBLUM, A.; MECKLIN, J. P.; MOLLER, P.; MYRHOJ, T.; NAGENGAST, F. M.; PARC, Y.; PONZ DE LEON, M.; RENKONEN-SINISALO, L.; SAMPSON, J. R.; STORMORKEN, A.; SIJMONS, R. H.; TEJPAR, S.; THOMAS, H. J.; RAHNER, N.; WIJNEN, J. T.; JARVINEN, H. J.; MOSLEIN, G. Group Revised guidelines for the clinical management of Lynch syndrome (HNPCC): recommendations by a group of European experts. **Gut**, v. 62, n. 6, p. 812–823, 2013.

VASEN, H. F. A.; MECKLIN, J.; KHAN, P. M.; LYNCH, H. T. The International Collaborative Group on Hereditary Non-Polyposis Colorectal Cancer (ICG-HNPCC). **Diseases of the colon and rectum**, v. 34, n. 5, p. 424–425, 1991.

VASEN, H. F. A.; GHORBANOGLI, Z.; BOURDEAUT, F.; CABARET, O.; CARON, O.; DUVAL, A.; ENTZ-WERLE, N.; GOLDBERG, Y.; ILENCIKOVA, D.; KRATZ, C. P.; LAVOINE, N.; LOEFFEN, J.; MENKO, F. H.; MULERIS, M.; SEBILLE, G.; COLAS, C.; BURKHARDT, B.; BRUGIERES, L.; WIMMER, K. Guidelines for surveillance of individuals with constitutional mismatch repair-deficiency proposed by the European Consortium “Care for CMMR-D” (C4CMMR-D). **Journal of Medical Genetics**, v. 51, n. 5, p. 283–293, 2014.

VATANDOOST, N.; GHANBARI, J. Early detection of colorectal cancer: from conventional methods to novel biomarkers. 2015.

VAUGHN, C. P.; ROBLES, J.; SWENSEN, J. J.; MILLER, C. E.; LYON, E.; MAO, R.; BAYRAK-TOYDEMIR, P.; SAMOWITZ, W. S. Clinical analysis of PMS2: Mutation detection and avoidance of pseudogenes. **Human Mutation**, v. 31, n. 5, p. 588–593, 2010.

VAUGHN, C. P.; HART, K. J.; SAMOWITZ, W. S.; SWENSEN, J. J. Avoidance of Pseudogene Interference in the Detection of 3 0 Deletions in PMS2. **Human mutation**, 2011.

VOGT, J.; WERNSTEDT, A.; RIPPERGER, T.; PABST, B.; ZSCHOCKE, J.; KRATZ, C.; WIMMER, K. PMS2 inactivation by a complex rearrangement involving an HERV retroelement and the inverted 100-kb duplison on 7p22.1. **European Journal of Human Genetics**, v. 24, n. 11, p. 1598–1604, 2016.

VOS, M.; HAYWARD, B. E.; PICTON, S.; SHERIDAN, E.; BONTHRON, D. T. Novel PMS2 Pseudogenes Can Conceal Recessive Mutations Causing a Distinctive Childhood Cancer Syndrome. **American journal of human genetics**, v. 74, p. 954–964, 2004.

WALTHER, A.; JOHNSTONE, E.; SWANTON, C.; MIDGLEY, R.; TOMLINSON, I.; KERR, D. Genetic prognostic and predictive markers in colorectal cancer. **Nature Reviews Cancer**, v. 9, n. 7, p. 489–99, 2009.

WANG, M.; ALDUBAYAN, S.; CONNOR, A. A.; WONG, B.; MCNAMARA, K.; KHAN, T.; SEMOTIUK, K.; KHALOUEI, S.; HOLTER, S.; ARONSON, M.; COHEN, Z.; GALLINGER, S.; CHARAMES, G.; POLLETT, A.; LERNER-ELLIS, J. Genetic testing for Lynch syndrome in the province of Ontario. **Cancer**, v. 10, p. 1–8, 2016.

WANG, Y.; REDDY, B.; THOMPSON, J.; WANG, H.; NOMA, K.; JOHN, R. Regulation of Set9-mediated H4K20 methylation by a PWWP domain protein. **Mol Cell**, v. 33, n. 4, p. 428–437, 2009.

WANG, Z.; CUMMINS, J. M.; SHEN, D.; CAHILL, D. P.; JALLEPALLI, P. V.; WANG, T. L.; PARSONS, D. W.; TRAVERSO, G.; AWAD, M.; SILLIMAN, N.; PTAK, J.; SZABO, S.; WILLSON, J. K. V.; MARKOWITZ, S. D.; GOLDBERG, M. L.; KARESS, R.; KINZLER, K. W.; VOGELSTEIN, B.; VELCULESCU, V. E.; LENGAUER, C. Three Classes of Genes Mutated in Colorectal Cancers with Chromosomal Instability. **Cancer Research**, v. 64, n. 9, p. 2998–3001, 2004.

WARTHIN, A. S. Heredity with reference to carcinoma as shown by the study of the cases examined in the pathological laboratory of the University of Michigan, 1895-1913. **Archives of Internal Medicine**, v. 12, p. 546–555, 1913.

WATSON, P.; VASEN, H. F. A.; MECKLIN, J.; BERNSTEIN, I.; JÄRVINEN, H. J.; MYRHØJ, T.; SUNDE, L.; WIJNEN, J. T.; LYNCH, H. T. The Risk of Extra-colonic, Extra-endometrial Cancer in the Lynch Syndrome. **International Journal**, v. 123, n. 2, p. 444–449, 2009.

WERNSTEDT, A.; VALTORTA, E.; ARMELAO, F.; TOGNI, R.; GIRLANDO, S.; BAUDIS, M.; HEINIMANN, K.; VALTORTA, E.; ARMELAO, F.; TOGNI, R.; GIRLANDO, S.; BAUDIS, M.; HEINIMANN, K.; MESSIAEN, L.; STAEHLI, N.; ZSCHOCKE, J.; MARRA, G.; WIMMER, K.; MESSIAEN, L.; STAEHLI, N.; ZSCHOCKE, J.; MARRA, G.; WIMMER, K. Improved multiplex ligation-dependent probe amplification analysis identifies a deleterious PMS2 allele generated by recombination with crossover between PMS2 and PMS2CL. **Genes, Chromosomes and Cancer**, v. 51, n. 9, p. 819–831, set. 2012.

WU, X.; PATTERSON, S.; HAWK, E. Chemoprevention - History and general principles. **Best Practice and Research: Clinical Gastroenterology**, v. 25, n. 4–5, p. 445–459, 2011.

YAMAGISHI, H.; KURODA, H.; IMAI, Y.; HIRAISHI, H. Molecular pathogenesis of sporadic colorectal cancers. **Chinese Journal of Cancer**, v. 35, n. 1, p. 1–8, 2016.

YIU, A. J.; YIU, C. Y. Biomarkers in Colorectal Cancer. **ANTICANCER RESEARCH**, v. 36, p. 1093–1102, 2016.

ZAHARY, M. N.; KAUR, G.; HASSAN, M. R. A.; SINGH, H.; NAIK, V. R.; ANKATHIL, R. Germline mutation analysis of MLH1 and MSH2 in malaysian lynch syndrome patients. **World Journal of Gastroenterology**, v. 18, n. 8, p. 814–820, 2012.

ZAVODNA, M.; BAGSHAW, A.; BRAUNING, R.; GEMMELL, N. J. The accuracy, feasibility and challenges of sequencing short tandem repeats using next-generation sequencing platforms. **Plos One**, v. 9, n. 12, p. 1–14, 2014.

ZIADA-BOUCHAAR, H.; SIFI, K.; FILALI, T.; HAMMADA, T.; SATTA, D.; ABADI, N. First description of mutational analysis of MLH1, MSH2 and MSH6 in Algerian families with suspected Lynch syndrome. **Familial Cancer**, v. Epub, 2016.

ZUMSTEIN, V.; VINZENS, F.; ZETTL, A.; HEINIMANN, K.; BOLLI, M. Systematic immunohistochemical screening for Lynch syndrome in colorectal cancer: a single centre experience of 486 patients. **Swiss Medical Weekly**, v. 146, n. May, p. 1–6, 2016.

9) ANEXOS

ANEXO 1: Termo de Consentimento Livre e Esclarecido (TCLE) para participação do estudo.

**TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO
PROJETO: DIAGNÓSTICO CLÍNICO E LABORATORIAL DA SÍNDROME DE
LYNCH: UM PROJETO DE CUSTO-EFETIVIDADE E DE VIABILIDADE DE
IMPLANTAÇÃO NO SUS**

Nome do Voluntário: _____

Você está sendo convidado(a) a participar de um projeto para avaliar o custo de execução de testes genéticos que permitem detectar alterações no DNA associadas ao desenvolvimento de câncer colo-retal. Para realização desse projeto estruturou-se um consórcio de pesquisadores e Hospitais Brasileiros para avaliar a eficácia, o custo e o impacto do diagnóstico de câncer de cólon (intestino) em indivíduos Brasileiros. Serão convidados a participar pacientes com câncer colorretal atendidos em qualquer um de cinco hospitais participantes (Hospital de Clínicas de Porto Alegre, INCA no Rio de Janeiro, Hospital AC Camargo em São Paulo, Hospital Universitário João de Barros Barreto em Belém do Pará, e Hospital do Câncer de Barretos). O projeto pretende avaliar as características pessoais e familiares da pessoa e sua relação com as alterações observadas em diferentes testes laboratoriais, que poderão indicar se o câncer de intestino em uma determinada pessoa tem origem hereditária ou genética.

Serão estudadas algumas características genéticas dos tumores de intestino (a partir do projeto do bloco de parafina que contém o tecido tumoral retirado em cirurgia) que incluem os assim chamados exames de rastreamento do câncer genético de intestino. Além disso, serão realizados testes para verificar se há alterações genéticas (mutações) que podem causar o câncer de intestino (rearranjos genômicos, mutações em genes de reparo e mutações em outros genes que causam o câncer de intestino) a partir do DNA retirado das células do sangue de cada paciente. Em resumo, os resultados deste projeto poderão trazer mais informações sobre a influência de fatores genéticos no tratamento de tumores de intestino.

Para que você possa decidir se quer participar ou não deste projeto, precisa conhecer seus benefícios, riscos e implicações.

OBJETIVO DO PROJETO

Avaliar o impacto e a viabilidade do diagnóstico de síndrome de Lynch, a forma mais comum de câncer colorretal hereditário, no Sistema Único de Saúde mediante investigação clínica e laboratorial de indivíduos com a doença e provenientes de quatro regiões Brasileiras.

PROCEDIMENTOS DO PROJETO

Se você concordar em participar desse projeto serão coletados cerca de 10 mL de seu sangue para extração e análise de DNA (seu material genético), você responderá a alguns questionários. Além disso, você autorizará o estudo de uma amostra de seu tumor de intestino, que será retirada de blocos de parafina contendo amostras do tumor retiradas durante a cirurgia. Se porventura os resultados deste projeto mostrarem um risco maior para câncer hereditário de intestino, você e seus familiares serão convidados para uma nova consulta de aconselhamento genético em que os resultados da pesquisa lhe serão fornecidos e lhe serão transmitidas informações sobre recomendações de cuidados médicos e de prevenção.

MÉTODOS ALTERNATIVOS

Sua participação nesse projeto é voluntária. Caso você não deseje participar deste projeto de pesquisa, basta que você não assine este Termo de Consentimento, e nenhuma amostra de sangue ou de seu tumor será coletada. Não existem métodos alternativos para sua participação nesse projeto de pesquisa, se você não concordar com os procedimentos que serão realizados para a execução desse projeto você é totalmente livre para não participar do mesmo.

RISCOS

O seu tratamento será exatamente o mesmo caso você participe ou não deste projeto. A coleta de sangue para o projeto será uma punção venosa adicional que não é necessária para o seu tratamento. Estas punções venosas podem resultar em dor no local da punção ou em manchas roxas transitórias chamadas de equimoses.

BENEFÍCIOS

Os resultados desse projeto de pesquisa não trarão benefícios para o seu tratamento. Mas os resultados obtidos poderão informar se você tem ou não maior risco de desenvolver outros tumores e se esse risco pode ser compartilhado por outros familiares seus.

Os resultados desse projeto também serão importantes para que possamos entender melhor como se comportam os tumores de intestino hereditários. Essas informações poderão nos ajudar acompanhar e tratar melhor as pessoas em risco para este tipo de câncer.

CARÁTER CONFIDENCIAL DOS REGISTROS

Além da equipe de saúde que cuidará de você, seus registros médicos poderão ser consultados pelo Comitê de Ética em Pesquisa do Instituto Nacional de Câncer (CEP-INCA) e equipe de

pesquisadores envolvidos. Seu nome não será revelado ainda que informações de seu registro médico sejam utilizadas para propósitos educativos ou de publicação, que ocorrerão independentemente dos resultados obtidos.

TRATAMENTO MÉDICO EM CASO DE DANOS

Todo e qualquer dano decorrente do desenvolvimento deste projeto de pesquisa, e que necessite de atendimento médico, ficará a cargo da instituição. Seu tratamento e acompanhamento médico independem de sua participação neste projeto.

CUSTOS

Não haverá qualquer custo ou forma de pagamento para o paciente pela sua participação no projeto.

BASES DA PARTICIPAÇÃO

É importante que você saiba que a sua participação neste projeto é totalmente voluntária e que você pode recusar-se a participar ou interromper sua participação a qualquer momento sem penalidades ou perda de benefícios aos quais você tem direito. Caso você decida interromper sua participação no projeto, a equipe assistente deve ser comunicada e a coleta de amostras para os exames relativos ao projeto será imediatamente interrompida, e os resultados obtidos não serão utilizados na execução desse projeto.

O médico responsável por sua internação pode interromper sua participação no projeto a qualquer momento, mesmo sem a sua autorização. Nesse caso o motivo dessa interrupção será discutido com você.

GARANTIA DE ESCLARECIMENTOS

Nós estimulamos a você ou seus familiares a fazerem perguntas a qualquer momento do projeto. Neste caso, por favor, ligue para o Dr. Miguel Ângelo Martins Moreira no telefone 21-3207 6586, ou para o Doutor Fernando Regla Vargas no telefone 21 3207 6584. Se você tiver perguntas com relação a seus direitos como participante do projeto clínico, também pode contar com um terceiro contato imparcial, o Comitê de Ética em Pesquisa do INCA, Rua André Cavalcanti, N° 37, telefones 21 – 3207-6551 ou 21-3207-6565. ou também pelo e-mail: cep@inca.gov.br .

ANEXO 2: Carta comprovando a aprovação do estudo pelo CEP.



A(o): Dr. Miguel Ângelo Martins Moreira
Pesquisador(a) Principal



Rio de Janeiro, 18 de agosto de 2011.

Registro CEP nº 82/11 (Este nº. deve ser citado nas correspondências referentes a este estudo)

CAAE – 0254.1.001.007-11

Título do Estudo: Diagnóstico Clínico e Laboratorial da Síndrome de Lynch: Um Estudo de Custo-Efetividade e de Viabilidade de Implantação no SUS

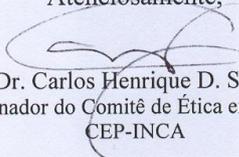
Prezado(a) Pesquisador(a),

Informo a composição do Comitê de Ética em Pesquisa do Instituto Nacional de Câncer que analisou e aprovou o estudo acima especificado.

Ana Lúcia Amaral Eisenberg
Ana Valéria Paranhos Miceli
Aníbal Guimarães
Carlos Henrique D Silva
Célia Helena Fernandes Costa
Cláudio Gustavo Stefanoff
Danielle Nogueira Ramos
Héilton Spindola Antunes
Marcelle Coelho de Rosário
Márcia de Cássia Cassimiro
Maria da Glória dos Santos Nunes
Michele Quarti
Murilo Vilaça
Paulina Henrique Moutinho
Rocio Hassan
Rosane Vianna Jorge
Urquilei dos Santos Pinheiro
Wilza Barbosa Felipe

Médica Patologista – DIPAT
Psicóloga - HC I
Doutor em Saúde Pública
Médico HC IV – Coordenador Substituto
Enfermeira o Serviço de Oncologia – HC I
Coordenador do Laboratório do BNT - CGTC
Doutora em Epidemiologia e Bioestatística - CONPREV
Dentista do Serviço de Pesquisa Clínica - CGTC
Advogada Extra INCA
Mestre em Saúde Coletiva Extra INCA
Enfermeira da Divisão de Enfermagem – HC I
Enfermeira - Extra INCA
Mestre em Educação - Extra INCA
Representante de Usuários – CES/RJ
Bióloga – CEMO
Doutora em Química Biológica da Divisão de Farmacologia - CGTC
Representante de Usuários – CES/RJ
Enfermeira da Comissão de Infecção Hospitalar - HC III

Atenciosamente,


Dr. Carlos Henrique D. Silva
Coordenador do Comitê de Ética em Pesquisa
CEP-INCA

ANEXO 3: Resultados de IHC e MSI para pacientes que preencheram os critérios clínicos e moleculares para SL.

Características clínicas			Características moleculares	
CASO	CENTRO	CRITERIO	IHQ	MSI
864	HCPA	Amsterdam	N	MSI-H (5)
865	HCPA	Amsterdam	N	MSS
867	HCPA	Amsterdam	M1,M6, P2	MSI-H (5)
1083	HCPA	Amsterdam	999	999
1085	HCPA	Amsterdam	N	MSI-H (5)
875	HCPA	Amsterdam	P2	Incl
874	HCPA	Amsterdam	N	MSI-H (4)
885	HCPA	Bethesda	N	MSS
896	HCPA	Bethesda	P2	MSI-H (5)
1108	HCPA	Bethesda	M6	MSI-H (4)
898	HCPA	Bethesda	999	999
903	HCPA	Bethesda	M1	MSS
1094	HCPA	Bethesda	M2, M6	MSS
893	HCPA	Bethesda	N	MSS
1107	HCPA	Amsterdam	N	MSI-H (5)
1109	HCPA	Amsterdam	N	999
1985	HCPA	Amsterdam	N	999
872	HCPA	Amsterdam	P2	MSI-H (5)
1112	HCPA	Amsterdam	M1, M2, M6	MSI-H (5)
2284	HCPA	Amsterdam	999	999
1342	HCPA	Amsterdam	999	999
2113	HCPA	Amsterdam	M1	999
2114	HCPA	Bethesda	M2	999
2106	HCPA	Amsterdam	M2, M6	999
871	HCPA	Amsterdam	M2, M6	Incl
882	HCPA	Bethesda	M1	MSS
889	HCPA	Bethesda	M1,M6	MSS
891	HCPA	Bethesda	M1,M2,M6,P2	MSI-H (5)
900	HCPA	Bethesda	M1,M6	MSI-H (5)
855	HCPA	Bethesda	M2,M6	999
854	HCPA	Bethesda	M6	999
843	HCPA	Bethesda	M1, M6	999
2309	HCPA	Bethesda	M2, M6	999
2313	HCPA	Bethesda	M2, M6	999
2318	HCPA	Amsterdã	999	999
2278	HCPA	Amsterdã	N (incl M2)	999
ADR	INCA	Amsterdam	N	999
FCG	INCA	Bethesda	N	999
CPB	INCA	Bethesda	N	999

MOG	INCA	Amsterdam	P2	999
UCC	INCA	Bethesda	N	MSS
AP	INCA	Amsterdam	N	999
S	INCA	Bethesda	Incl	999
RBO	INCA	Bethesda	Incl	999
SDL	INCA	Bethesda	N	999
1	ACCAM	Bethesda	N	998
3	ACCAM	Bethesda	N	999
4	ACCAM	Bethesda	M1,P2	999
5	ACCAM	Bethesda	N	999
6	ACCAM	Bethesda	N	999
7	ACCAM	Amsterdam	M1,P2	999
8	ACCAM	Bethesda	M2	999
9	ACCAM	Amsterdam	N	999
10	ACCAM	Bethesda	N	999
14	ACCAM	Amsterdam	M6	999
16	ACCAM	Amsterdam	N	999
MT4A	ACCAM	Amsterdam	M6, P2	999
13	ACCAM	Bethesda	M1	999
MRV	ACCAM	Amsterdam	MS, M6	999
SL9	HUJJB	Amsterdam	M6	999
SL7	HUJJB	Bethesda	M6	999
SL4	HUJJB	Bethesda	N (incl P2)	999
SL5	HUJJB	Bethesda	N	999
SL1	HUJJB	Bethesda	N	999
SL6	HUJJB	Bethesda	N	999
SL3	HUJJB	Bethesda	N	999

Legenda:

999- Teste não foi realizado para paciente.

Incl: Teste inconclusivo para paciente.

N: paciente não apresenta nenhuma alteração, é normal para determinado teste

M2, M1, M6, P2: Proteínas MSH2, MLH1, MSH6 e PMS2, respectivamente.

ANEXO 4: Iniciadores utilizados para a amplificação das regiões codificantes do gene *MSH2*.

Éxon	Iniciador direto (5'-->3')	Iniciador reverso (5'-->3')	Amplicon (pb)
1	CTTCAACCAGGAGGTGAGGAGGT	GAAAGGAGCCGCGCCACAAG	301
2	ATGTAATATCTCAAATCTGTAATGT	ATAAGTAAATTA AAAAGGAAGATAA	227
3	TGTTCAAGAGTTTGTTAAATTTTT	TGGAATCTCCTCTATCACTAGACT	359
4	TCTTATTCCTTTTCTCATAGTAG	TATTGTAATTCACATTTATAATCC	224
5	CAGTGGTATAGAAATCTTCG	ACCATTCAACATTTTTTAACCC	252
6	ATTGAGCTTGCCATTCTTTCTA	TGCAGGTTACATAAAACTAACGAA	220
7	AATATTTTACATTAATCAAGTTA	ATATATTGTATGAGTTGAAGGAA	284
8	AATGAGATCTTTTTATTTGTTTGTT	ACTGCTTAAATTA AAAAGTATATTG	182
9	GGATTTTGTTCAGTTTGTCTGTTTG	TTCCAACCTCCAATGACCCATTC	179
10	ATACTTTTTCTTTTCTTCTTGATTA	GGTTAAAAATATAATAACGACTTG	210
11	TAAATAAACTGTTATTTTCGATTTG	AGCCAGGTGACATTCAGAACATTAT	169
12	TTTCTGTTTTTATTTTTATACACG	AAACGTTACCCCCACAAAG	309
13	AAACTTGCTTCTGATATAATTTGT	CATTTCTATCTTCAAGGGACTAGGA	273
14	TTCAGGTAATTATGTGCTTCA	TACTGAATTTAGAGTACTCCAATAGT	302
15	TCTCATGCTGCTCCCTCACG	TAAGTTAACTATGAAAACAACTG	247
16	ATTTTAATTACTAATGGGACATT	GCTTATCAATATTACCTTCATTC	246

ANEXO 5: Iniciadores utilizados para a amplificação das regiões codificantes do gene *MSH6*.

Éxon	Iniciador direto (5'-3')	Iniciador reverso (5'-3')	Amplicon (pb)
1	TCCGTCCGACAGAACGGTTG	CCCCAAATGCTCCAGACTCG	477
2	GCCAGAAGACTTGGAATTGTTTATTTG	ACACAAACACACACACATGGCAGTAG	327
3	TCTTTCTAGGTTCAAATCAAAGGA	CCCCATCACCCCTAACATAAA	277
4a	GTCTTACATTATGGTTTTCC	ATGAATACCAGCCCCAGTTC	790
4b	CTGTACCACATGGATGCTCT	AACATCACCCCAATGCCATCAC	662
4c	CTGTTCTCTTCAGGAAGGTC	AGCCATTGCTTTAGGAGCCG	492
4d	GGGTTGATACTTGCCATACTCC	CAGCTGGCAAACAGCACTAC	949
5	GGGGAGATCGTTGGACTGTAATTG	TTGCTTCCTATTAAGTCACTGGCTG	521
6	GTTTATGAAACTGTTACTACC	GCAAATATCTTTTATCACAT	237
7	GCCAATAATTGCATAGTCTCTTAATG	GCCACAATGGTGAGTGCGTG	368
8	CCTTTTTTGTTTAAATCCT	ATATGTCCTAGGCGCACAGC	210
9	GATTATTCTCAAATGTTGCTGTGCG	TTCCTAGCCAGGCAAACCTCCC	481
10	ATTTTAAGGGAAGTTTGCC	GTTTATTAGATCATAATGTT	165

ANEXO 6: Iniciadores utilizados para a amplificação das regiões codificantes do gene *MLH1*.

Éxon	Iniciador direto (5'-3')	Iniciador reverso (5'-3')	Amplicon (pb)
1	GAGGTGATTGGCTGAAGGC	GCGGCCCGTTAAGTCGTAG	235
2	TTGTTATCATTGCTTGGCTCA	GGTCCTGACTCTTCCATGAAG	232
3	TGGGAATTCAAAGAGATTTGG	TGACAGACAATGTCATCACAGG	225
4	CCCTTTGGTGAGGTGACAGT	GACAGGATTACTCTGAGACC	224
5	TTCCCCTTGGGATTAGTATC	TATACAAACAAAGCTTCAAC	183
6	CGCTTTTGCCAGGACATC	GACAAATCTCAGAGACCCAC	209
7	GCTCTGACATCTAGTGTGTG	CCTTATCTCCACCAGCAAAC	225
8	ATGATGTTTCAGTCTCAGCC	GTGATGGAATGATAAACCAAG	274
9	GTAATGTTTGAGTTTTGAG	GTGGATTTCCCATGTGGTTC	284
10	GGAAAGTGGCGACAGGTAAG	CCTTGTGAGTCTTGGTTGAG	300
11	CACCACATACACCATATGTG	GGCACATTTCAAAGGCCCCAG	282
12	CTTATTCTGAGTCTCTCCAC	AAAGCCAAAGTTAGAAGGCAG	515
13	GCTCCTCCAAAATGCAACC	AACCTTGGCAGTTGAGGCC	288
14	TGCTTTGGTCAATGAAGTGG	TTGTGCCTGTGCTCCCTGG	286
15	TTCTCCATTTTGTCCCAAC	TGGAGAGCTACTATTTTCAGAAACG	245
16	TGGATTCTTGGGAATTCAGG	ATTACAGCCATGAGCCACC	319
17	CCTGGGAAAGCACTGGAG	TCCAGCACACATGCATGTAC	228
18	GGTATTGAATTTCTTTGGACCAG	GATGGGCAAGTTTCATCTCC	216
19	GGACACCAGTGTATGTTGGG	ATCAGAAGGCAAGTATAAGT	365

ANEXO 7: Iniciadores utilizados para a amplificação das regiões codificantes do gene *PMS2*.

Fragmento	Éxon	Iniciador direto (5'- 3')	Iniciador reverso (5'- 3')	Amplicon (pb)
LRPCR1	1 a 5	AGCCTGTGGTATGCACAGGTGGAAG	CTTCACCTGTGCATACCACAGGCT	10264
LRPCR2	6 a 9	GGTCCAGGTCTTACATGCATACTGT	CTGACTGACATTTAGCTTGTTGACA	9440
LRPCR3	11 a 15	GCGTTGATATCAATGTTACTCCAGA	CCTTCCATCTCCAAAACCAGCAAGA	18436
E6	6	TGCTTCCCTTGATTTGTGCGATGAT	CTACTGGAAGGGACAATGGAAACC	364
E10	10	AATTAGCCAGTGTGGTGGCACTTG	AGCTTTAGAAGCTGTTTGTACAC	416
E1_N	1	ACGTCGAAAGCAGCCAATGGAGTT	CAGGTAGAAAGGAAATGCATTCAGT	464
E2_N	2	ACAGTGTTGAGTCATTTCCACAGT	TTCTTAGCATAACACCTGCCTGGCA	455
E7_N	7	ATTGTA CTCCAGCCTGGGCAATAG	ATTGTAGTTCTCTTGCCAGCAATC	310
E11_N	11	TCACATAAGCACGTCCTCTCACCAT	GCAACAGAGCAAGACTCTGTCTCAA	1022
E15_N	15	GCTGAGATCTAGAACCTAGGCTTCT	ACACACGAGCGCATGCAAACATAGA	470

ANEXO 8: Concentrações finais dos iniciadores (pmol) de MgCl₂ (μM) utilizadas nas reações de multiplexes ou LR-PCRs.

Gene	Mixes	Éxons	Concentração final iniciadores (pmol)	Concentração Final MgCl₂ (μM)
MSH2	1.1	9, 12, 15	0.4	1,5
	1.2	4, 7	0.4	2,5
	1.3	6, 13	0.4	1,5
	1.4	5, 8	0.4	1,5
	2	3, 10, 11, 14, 16	0.4	1,5
	3	1, 2	0.4	1,5
MSH6	8	3, 6, 7	0.4	1,5
	9	2, 4, 5	0.4	2,5
	10	8, 9, 10	0.4	1,5
	11	1	0.4	2
MLH1	4	5, 9, 19	2	1,5
	5	2, 4, 7, 10, 15, 18	2	1,5
	6	1, 6, 11	2	1,5
	7	14, 17	2	1,5
	7.2	13, 16	0.4	1,5
	7.6	3, 8, 12	0.4	2,5
PMS2	LR-1	1 ao 5	0,8	2
	E6	6	0,8	1,5
	LR-2	7 ao 9	0,4	2
	E10	10	0,8	1,5
	LR-3	11 ao 15	0,4	2

ANEXO 9: Condições de ciclagem para o gene *MSH2* utilizando de PCRs

MSH2											
Condições de ciclagem Mixes 1 (1.1/1.2/1.3/1.4)				Condições de ciclagem Mix 2				Condições de ciclagem Mix 3			
Etapa	Temperatura (°C)	Tempo	N°Ciclos	Etapa	Temperatura (°C)	Tempo	N°Ciclos	Etapa	Temperatura (°C)	Tempo	N°Ciclos
Desnaturação	94°C	5min	1X	Desnaturação	94°C	5min	1X	Desnaturação	94°C	5min	1X
Desnaturação	94°C	30seg		Desnaturação	94°C	30seg	10X	Desnaturação	94°C	15seg	
Anelamento	47°C	30seg	7X	Anelamento	47°C	30seg		Anelamento	58,5°C	15seg	35X
Extensão	72°C	30seg		Extensão	72°C	30seg		Extensão	72°C	15seg	
Desnaturação	94°C	30seg		Desnaturação	94°C	30seg	10X	Extensão final	72°C	10 min	1X
Anelamento	48°C	30seg	6X	Anelamento	48°C	30seg		Hold	8°C	Inf.	
Extensão	72°C	30seg		Extensão	72°C	30seg					
Desnaturação	94°C	30seg		Desnaturação	94°C	30seg	10X				
Anelamento	49°C	30seg	6X	Anelamento	49°C	30seg					
Extensão	72°C	30seg		Extensão	72°C	30seg					
Desnaturação	94°C	30seg		Desnaturação	94°C	30seg	10X				
Anelamento	50°C	30seg	6X	Anelamento	50°C	30seg					
Extensão	72°C	30seg		Extensão	72°C	30seg					
Desnaturação	94°C	30seg		Extensão final	72°C	10 min	1X				
Anelamento	51°C	30seg	6X	Hold	8°C	Inf.					
Extensão	72°C	30seg									
Desnaturação	94°C	30seg									
Anelamento	52°C	30seg	6X								
Extensão	72°C	30seg									
Extensão final	72°C	10 min	1X								
Hold	8°C										

ANEXO 10: Condições de ciclagem para o gene *MLH1* utilizando de PCRs multiplexes.

MLH1											
Condições de ciclagem Mix 4				Condições de ciclagem Mix 6				Condições de ciclagem Mix 7.2			
Etapa	Temperatura (°C)	Tempo	N°Ciclos	Etapa	Temperatura (°C)	Tempo	N°Ciclos	Etapa	Temperatura (°C)	Tempo	N°Ciclos
Desnaturação	94°C	5min	1X	Desnaturação	94°C	5min	1X	Desnaturação	94°C	5min	1X
Desnaturação	94°C	30seg		Desnaturação	94°C	30seg		Desnaturação	94°C	30seg	
Anelamento	50°C	30seg	35X	Anelamento	56°C	15seg	35X	Anelamento	57°C	30seg	35X
Extensão	72°C	30seg		Extensão	72°C	15seg		Extensão	72°C	30seg	
Extensão	72°C	15 min	1X	Extensão	72°C	15 min	1X	Extensão	72°C	15 min	1X
Hold	8°C	Inf.		Hold	8°C	Inf.					
Condições de ciclagem Mix 5				Condições de ciclagem Mix 7				Condições de ciclagem Mix 7.6			
Etapa	Temperatura (°C)	Tempo	N°Ciclos	Etapa	Temperatura (°C)	Tempo	N°Ciclos	Etapa	Temperatura (°C)	Tempo	N°Ciclos
Desnaturação	94°C	5min	1X	Desnaturação	94°C	5min	1X	Desnaturação	94°C	5min	1X
Desnaturação	94°C	30seg		Desnaturação	94°C	30seg		Desnaturação	94°C	30seg	35X
Anelamento	53°C	30seg	35X	Anelamento	59°C	30seg	35X	Anelamento	59°C	30seg	
Extensão	72°C	30seg		Extensão	72°C	30seg		Extensão	72°C	30seg	
Extensão	72°C	15 min	1X	Extensão	72°C	15 min	1X	Extensão	72°C	15 min	1X
Hold	8°C	Inf.		Hold	8°C	Inf.					

ANEXO 11: Condições de ciclagem para o gene *MSH6* utilizando de PCRs multiplexes.

MSH6							
Condições de ciclagem Mix 8				Condições de ciclagem Éxon 8F/ 10R			
Etapa	Temperatura (°C)	Tempo	N°Ciclos	Etapa	Temperatura (°C)	Tempo	N°Ciclos
Desnaturação	94°C	5min	1X	Desnaturação	94°C	5min	1X
Desnaturação	94°C	30seg		Desnaturação	94°C	30seg	
Anelamento	52°C*	30seg	35X	Anelamento	54°C	30seg	35X
Extensão	72°C	30seg		Extensão	72°C	1min	
Extensão	72°C	15 min	1X	Extensão	72°C	15 min	1X
Hold	8°C	Inf.		Hold	8°C	Inf.	1X
Condições de ciclagem Mix 9				Condições de ciclagem Éxon 1			
Etapa	Temperatura (°C)	Tempo	N°Ciclos	Etapa	Temperatura (°C)	Tempo	N°Ciclos
Desnaturação	94°C	5min	1X	Desnaturação	94°C	5min	1X
Desnaturação	94°C	30seg		Desnaturação	94°C	30seg	35X
Anelamento	62°C	30seg	35X	Anelamento	62°C	30seg	
Extensão	72°C	1min30seg		Extensão	72°C	30seg	
Extensão	72°C	15 min	1X	Extensão	72°C	15 min	1X
Hold	8°C	Inf.	1X	Hold	8°C	Inf.	1X

ANEXO 12: Condições de ciclagem para o gene *PMS2* utilizando de PCRs multiplexes.

PMS2											
Ciclagem LR_PCR1				Ciclagem LR_PCR2				Ciclagem LR_PCR3			
Etapa	Temperatura (°C)	Tempo	N°Ciclos	Etapa	Temperatura (°C)	Tempo	N°Ciclos	Etapa	Temperatura (°C)	Tempo	N°Ciclos
Desnaturação	94°C	1min	1X	Desnaturação	94°C	1min	1X	Desnaturação	94°C	1min	1X
Desnaturação	94°C	30seg		Desnaturação	94°C	30seg		Desnaturação	94°C	30seg	
Anelamento	65°C	30seg	35X	Anelamento	60°C	30seg	35X	Anelamento	65°C	30seg	35X
Extensão	72°C	15 min		Extensão	72°C	15 min		Extensão	72°C	15 min	
Extensão	72°C	10 min	1X	Extensão	72°C	10 min	1X	Extensão	72°C	10 min	1X
Hold	8°C	Inf.	1X	Hold	8°C	Inf.	1X	Hold	8°C	Inf.	1X
Ciclagem Éxon 6				Ciclagem Éxon 10							
Etapa	Temperatura (°C)	Tempo	N°Ciclos	Etapa	Temperatura (°C)	Tempo	N°Ciclos				
Desnaturação	94°C	5min	1X	Desnaturação	94°C	5min	1X				
Desnaturação	94°C	30seg		Desnaturação	94°C	30seg					
Anelamento	62°C	30seg	35X	Anelamento	62°C	30seg	35X				
Extensão	72°C	30seg		Extensão	72°C	30seg					
Extensão	72°C	15 min	1X	Extensão	72°C	15 min	1X				
Hold	8°C	Inf.	1X	Hold	8°C	Inf.					

ANEXO 13: Script para identificação de variantes nos genes MMR.

#Quality Filter

```
perl prinseq -fastq -fastq2 -out_format 3 -no_qual_header -min_len 30 -min_qual_mean 20 -
ns_max_p 60 -trim_tail_right 5 -trim_tail_left 5 -trim_qual_step 1 -trim_ns_right 5 -
trim_ns_left 5 -trim_qual_right 20 -trim_qual_left 20 -trim_qual_type mean -trim_qual_rule lt
-trim_qual_window 5 -out_bad null
```

Mapping GRCh37 reference

```
bwa aln -t Ref_GRCh37.fasta R1.fastq > R1.sai
bwa aln -t Ref_GRCh37.fasta.fasta R2.fastq > R2.sai
bwa sampe Ref_GRCh37.fasta.fasta R1.sai R2.sai R1.fastq R2.fastq > sample.sam
```

Grouping and sorting files.sam

```
java -jar AddOrReplaceReadGroups.jar INPUT= sample.sam OUTPUT=
sample_grouped.bam RGLB=FCG RGPL=illumina RGPU=run RGSM=sample
samtools sort sample_grouped.bam sample_sortedGrouped.bam
```

Mark Duplicates

```
java -jar MarkDuplicates.jar I= sample_sortedGrouped.bam O= sample_dedup.bam
METRICS_FILE= sample_metrics.txt
```

Indel realignment

```
java -Xmx4g -jar GATK -T RealignerTargetCreator -R Ref_GRCh37.fasta -I
sample_dedup.bam -o sample_forIndelRealigner.intervals
java -Xmx4g -jar GATK -T IndelRealigner -R Ref_GRCh37.fasta -I sample_dedup.bam -
targetIntervals sample_forIndelRealigner.intervals -o sample_realignedBam.bam
```

#Base Recalibration

```
java -Xmx4g -jar GATK -T BaseRecalibrator -I sample_realignedBam.bam -R
Ref_GRCh37.fasta -knownSites dbsnp_138.b37.vcf -knownSites 1000G_omni2.5.b37.vcf -
knownSites hapmap_3.3.b37.vcf -o sample_recal_data.table
java -Xmx4g -jar GATK -R Ref_GRCh37.fasta -T PrintReads -o sample_recalibrated.bam -I
sample_realignedBam.bam --read_filter MappingQualityZero
```

Variant Calling

```
java -jar GATK -T HaplotypeCaller -R Ref_GRCh37.fasta -I sample_recalibrated.bam --  
dbsnp dbsnp_138.b37.vcf -stand_call_conf 30 -stand_emit_conf 30 -L
```

#Hard Filter

```
java -jar GATK -T SelectVariants -R Ref_GRCh37.fasta.fasta -V sample_raw.snps.indels.vcf  
-L -selectType SNP -o sample_raw_snps.vcf
```

```
java -jar GATK -T VariantFiltration -R Ref_GRCh37.fasta.fasta -V sample_raw_snps.vcf --  
filterExpression "QD < 2.0 || FS > 60.0 || MQ < 40.0 || MQRankSum < -12.5 ||  
ReadPosRankSum < -8.0" -o sample_filtered_snps.vcf
```

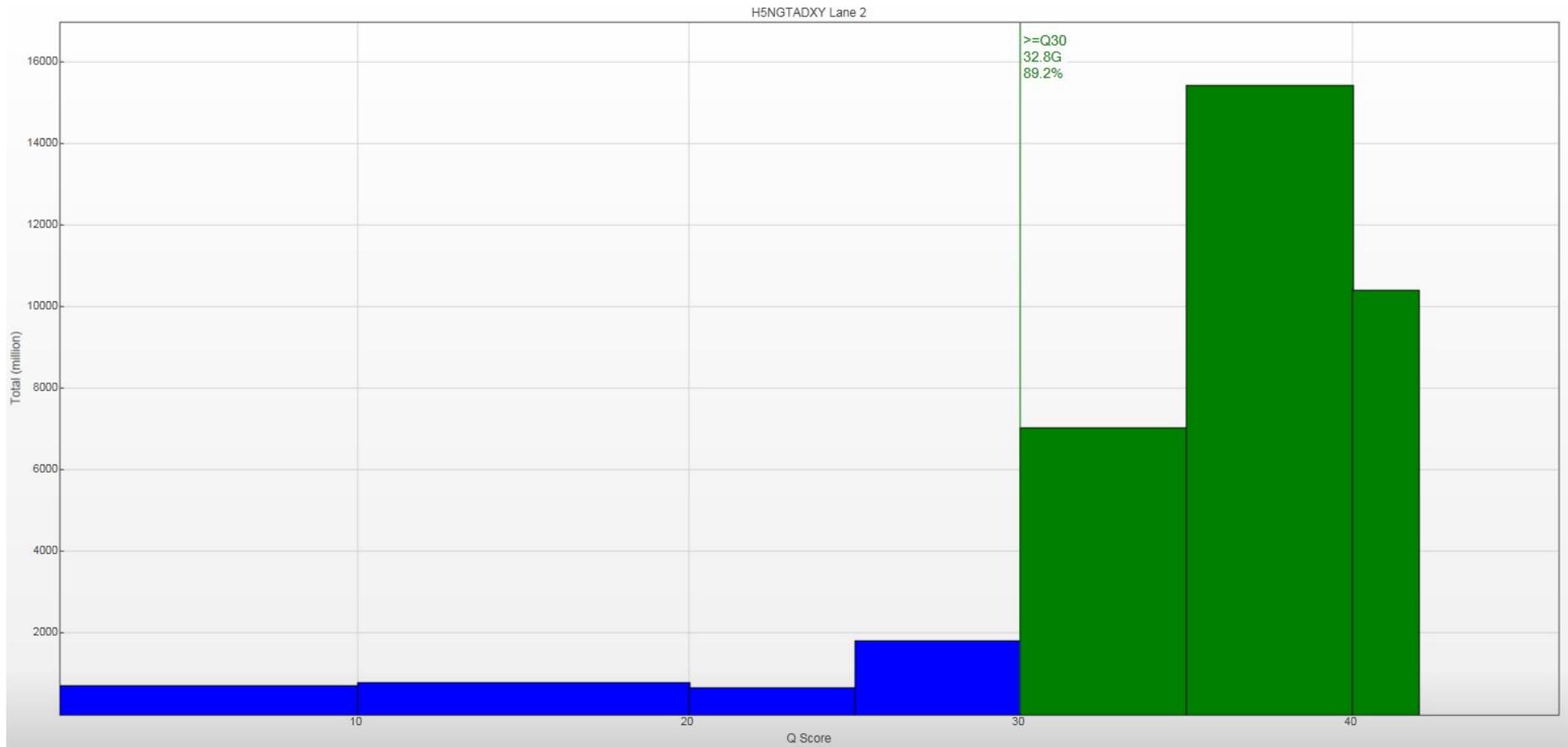
```
java -jar GATK -T SelectVariants -R Ref_GRCh37.fasta.fasta -V sample_raw.snps.indels.vcf  
-L -selectType INDEL -o sample_raw_indels.vcf
```

```
java -jar GATK -T VariantFiltration -R Ref_GRCh37.fasta.fasta -V sample_raw_indels.vcf --  
filterExpression "QD < 2.0 || FS > 200.0 || ReadPosRankSum < -20.0" -o  
sample_filtered_indels.vcf
```

#Depth Coverage

```
coverageBed -d -abam sample_recalibrated.bam -b exons.bed > sample_coberturaBed.txt
```

ANEXO 14: Gráfico que demonstra a porcentagem de leituras que tiveram qualidade das bases acima de 30.



ANEXO 15: Profundidade média de cobertura por gene par cada um dos pacientes rastreados para os genes *MSH2*, *MSH6*, *MLH1* e *PMS2*.

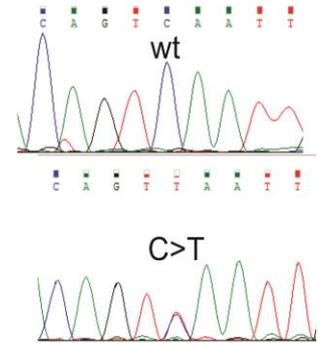
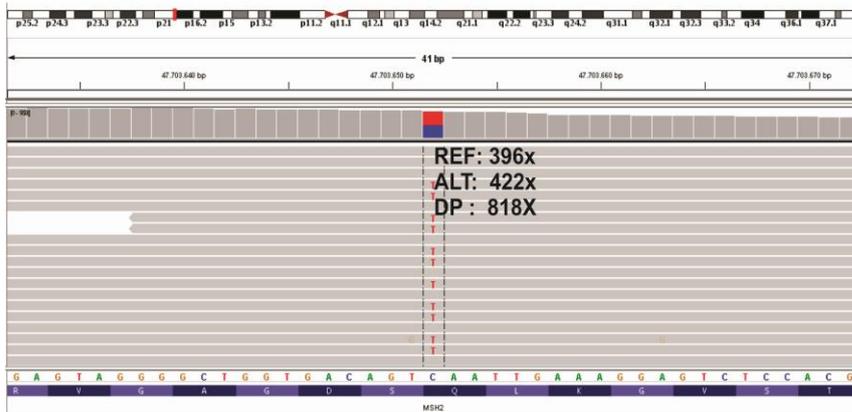
Gene	MSH2	MSH6	MLH1	PMS2
874	668,46	1846,29	951,11	1573,29
885	3002,19	16844,52	4492,67	1449,88
889	1568,35	16711,77	2754,07	844,12
891	3402,02	6406,00	3822,11	894,14
893	4254,27	29100,99	6927,72	
898	3235,69	3927,43	4184,88	1491,98
900	2306,50	17164,50	6121,56	4305,76
903	4038,62	32874,16	5872,82	
1085	3228,09	10688,61	3490,13	1506,88
1094	1231,74	5288,63	1002,07	1658,81
1107	1869,00	23171,71	3363,20	
1108	4553,84	20131,54	6224,52	
1109	5794,41	30293,40	7803,54	
1112	6404,31	6244,86	7621,96	832,23
1342	2656,47	11426,97	1961,98	2675,72
1985	6568,15	32685,35	10808,82	
2106	7384,40	34301,87	10399,63	
2113	7466,94	28719,14	9256,43	
2114	2366,73	7219,67	3107,93	0,15
2309	2219,66	24565,84	4285,53	1139,27
2313	1406,78	14183,50	2209,88	920,72
2318	3758,46	27290,72	6849,19	
2278	3688,65	27634,61	1062,33	
2284	15603,03	14787,27	967,26	1463,51
S	1737,36	12067,18	2964,65	8649,48
CPB	2093,72	20329,72	3665,52	5796,49
UCC	3054,45	9422,88	3760,51	5842,35
RBO	11967,15	50496,36	14922,35	14293,77
FCG	2007,08	12113,63	4369,10	5876,88
MOG	1789,36	23162,02	2976,99	10590,67
ADR	3718,96	5545,24	3869,08	10997,35
MRV	4621,91	38672,04	7868,79	3151,21
AP	393,30	3513,05	1078,61	10466,75
MT-4	2724,05	17334,78	3011,47	1833,48
P1	7230,31	45842,61	4897,38	1600,18
P3	23637,25	98664,74	33865,29	
P4	32117,62	59078,99	48867,17	
P5	12820,37	54858,39	36097,18	
P6	15605,87	88280,69	26523,75	
P7	22224,55	59371,25	26617,24	
P8	19178,94	73102,04	16917,37	
P9	7000,07	34751,03	7854,77	10027,88

P10	15184,36	29511,87	10951,50	
P13	18356,70	96392,89	21496,97	
P14	10278,94	64507,36	35928,82	
P16	17021,62	85798,99	28538,60	
SL001	8229,10	85670,55	16192,23	
SL003	6313,83	107870,32	10616,68	
SL004	5087,70	91488,38	6365,35	
SL005	3991,61	30809,92	5246,46	
SL006	12252,54	84203,74	24495,43	
SL007	10266,51	66752,10	11711,90	
SL009	33023,64	88588,75	25329,08	
843	425,21	4205,41	6806,99	4487,15
854	4599,33	22379,95	49482,43	1622,32
855	11382,48	46966,22	19210,59	12199,83
864	10570,86	34290,70	10469,97	3729,51
865	5149,97	23803,01	4206,37	6649,22
867	1445,33	3346,03	1260,30	2000,44
871	669,15	8333,51	1274,31	61,92
872	16792,78	82893,76	41983,20	3167,73
875	19469,44	94901,08	28602,55	26891,59
882	4344,32	61558,02	8913,60	2749,03
896	2526,14	15490,93	5607,38	5989,68
1083	29214,14	16286,61	33104,17	4,42

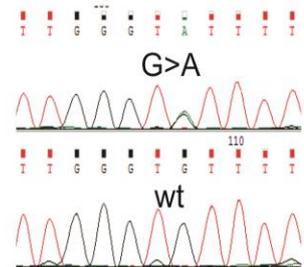
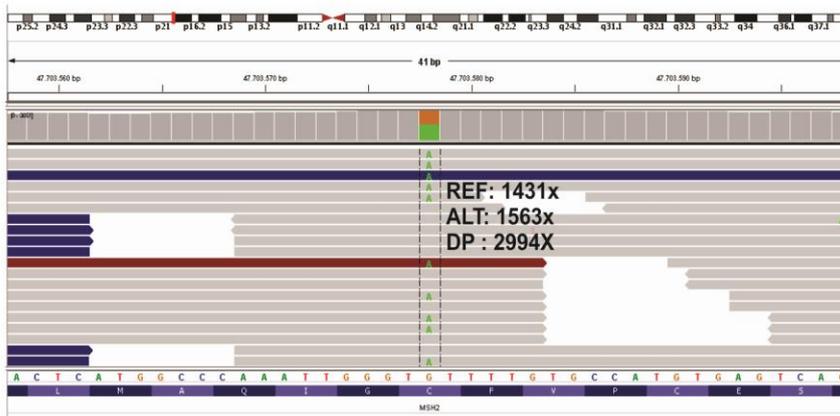
Obs: Os pacientes não sequenciados para *PMS2* apresentam a posição a este gene em branco.

ANEXO 16: Mapas e eletroferogramas para as variantes patogênicas /VUS c.2152C>T, c.2078C>A e c.1046G>C encontradas no gene *MSH2*.

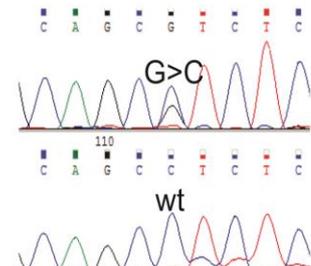
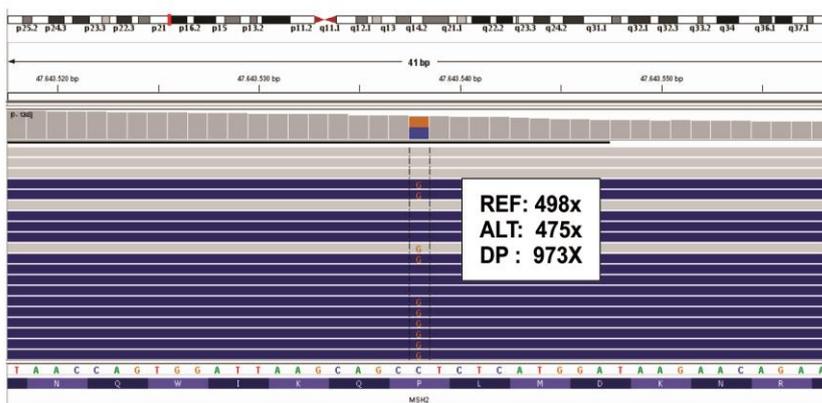
c.2152C>T



c.2078G>A

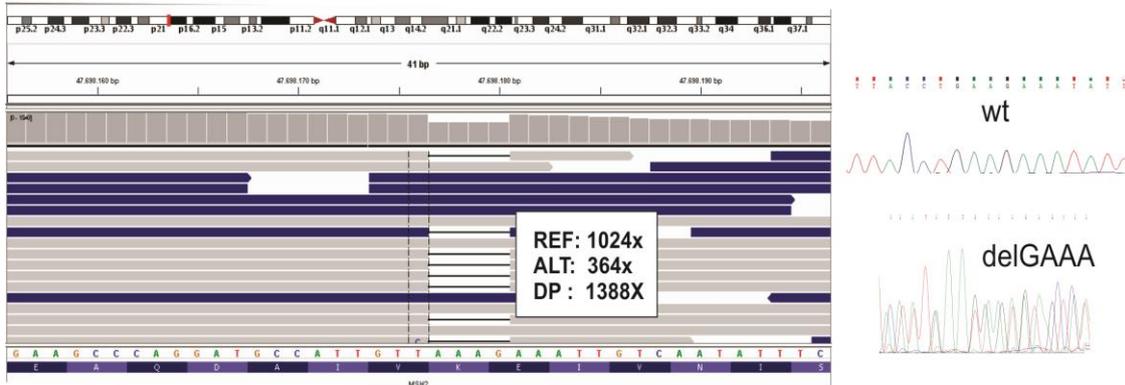


c.1046G>C

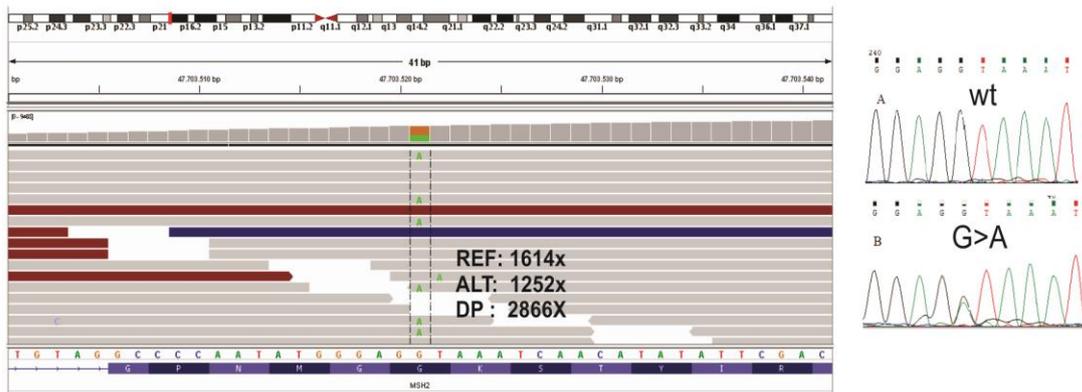


ANEXO 17: Mapas e eletroferogramas para as variantes patogênicas c.1738delGAAA, c.2021G>A e c.338_389del1CA encontradas no gene *MSH2*.

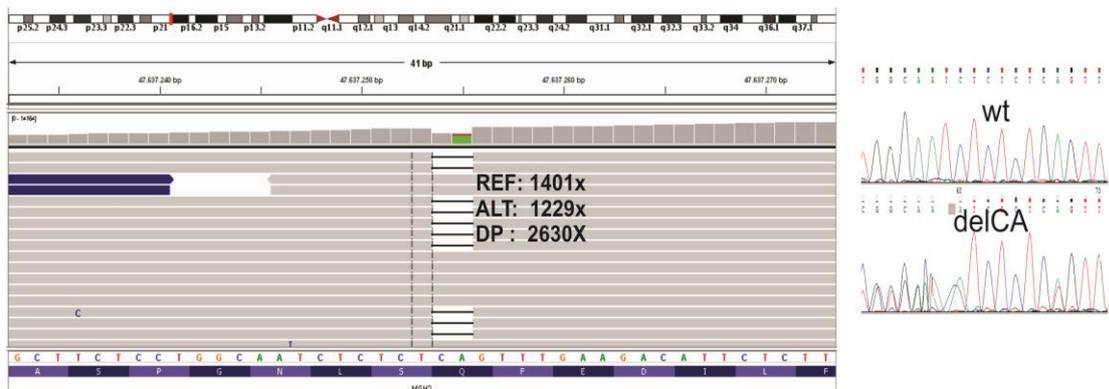
c.1738delGAAA



c.2021G>A



c.338_389del1CA

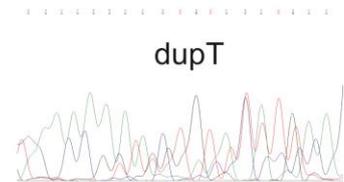
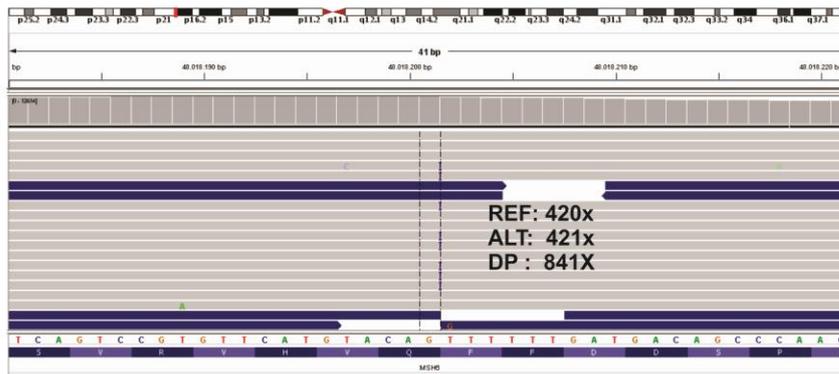


ANEXO 19: Mapas e eletroferogramas para cada uma das variantes patogênicas/VUS encontradas para no gene *MSH6*.

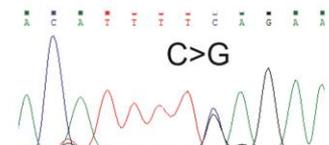
c.181G>A



c.402dupT

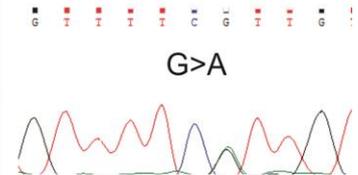
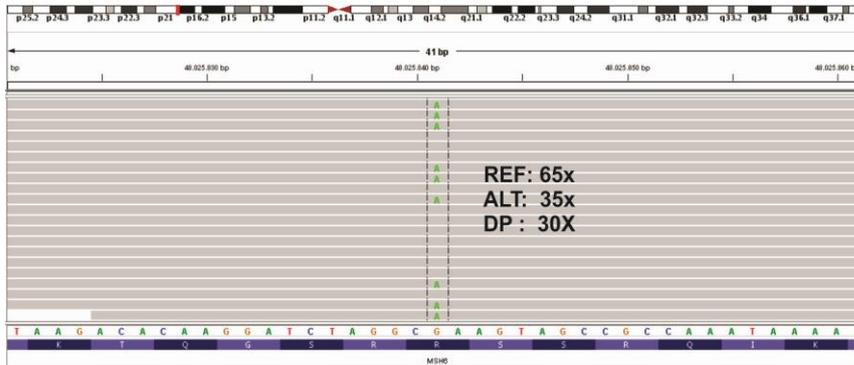


c.3772C>G

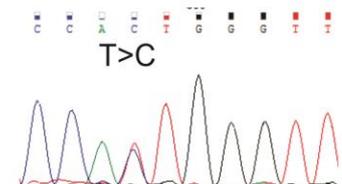


ANEXO 20: Mapas e eletroferogramas para cada uma das variantes patogênicas/VUS encontradas no gene *MSH6*.

c.719G>A

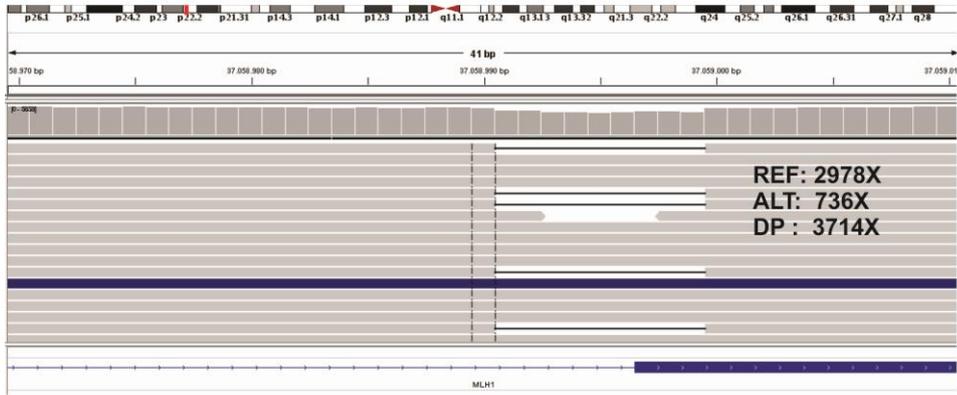


c.2006T>C

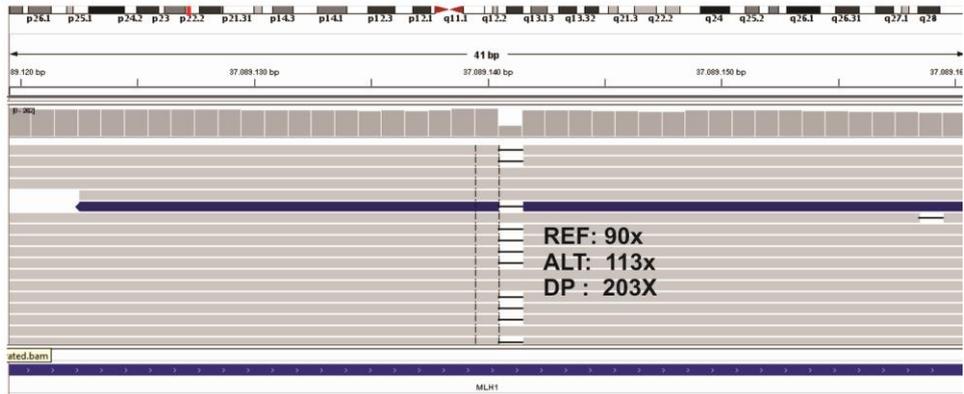


ANEXO 21: Mapas e eletroferogramas para cada uma das variantes patogênicas/VUS c.791-6_796delGTTTAGATC, c.1863delG e c.677+1G>A encontradas no gene *MLH1*.

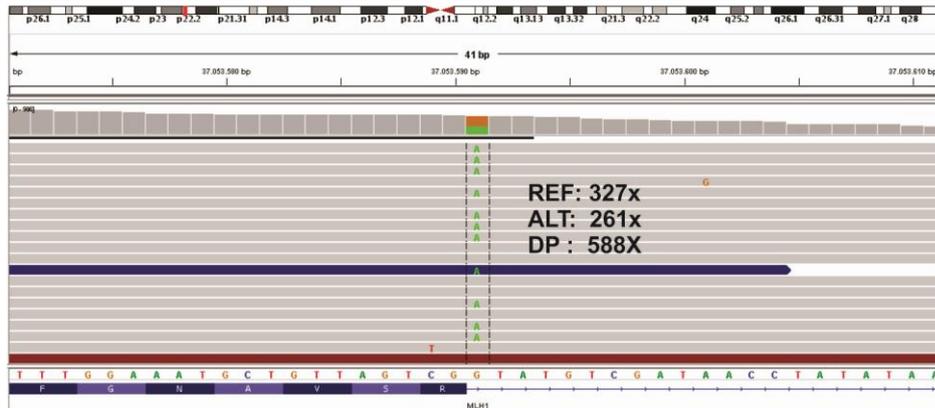
c.791-6_794delGTTTAGATC



c.1863delG

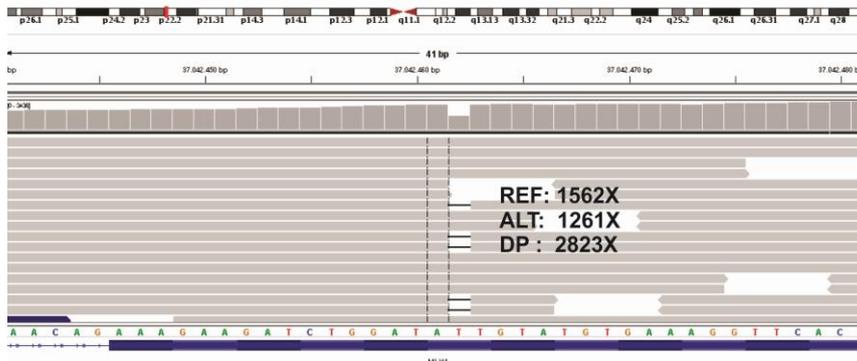


c.677+1G>A

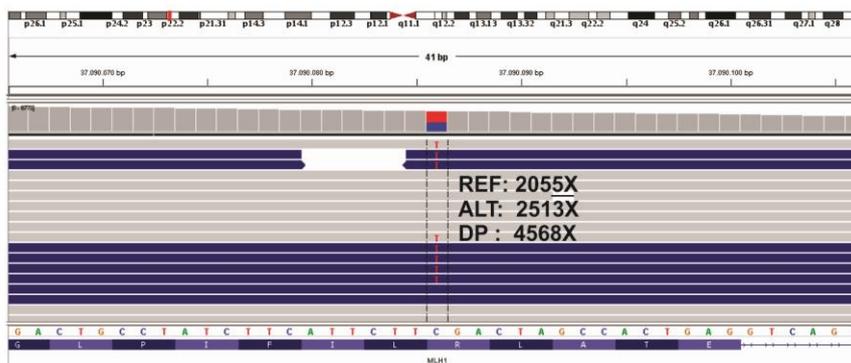


ANEXO 22: Mapas e eletroferogramas para cada uma das variantes patogênicas/VUS c.224delT, c.1975C>T, c.890C>A e c.888_889delCA encontradas no gene *MLH1*.

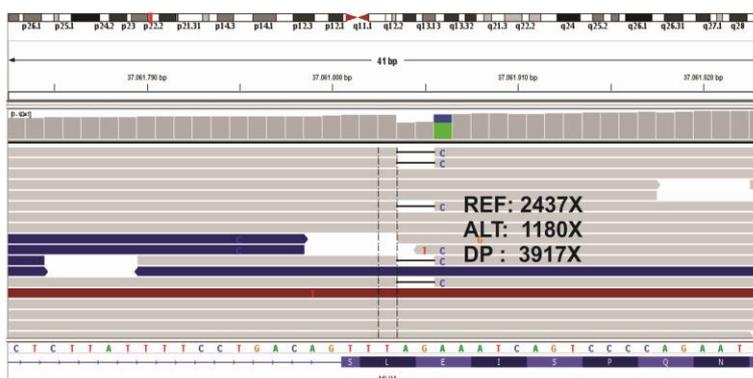
c.224delT



c.1975C>T

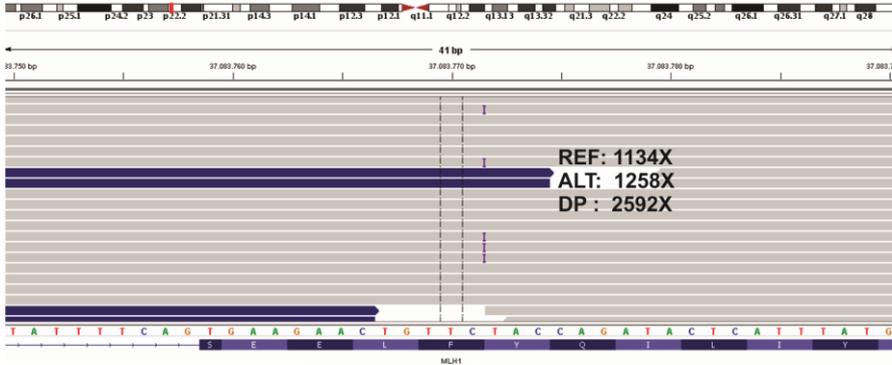


c.890C>A + c.888_889delCA

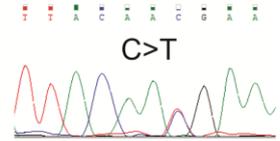
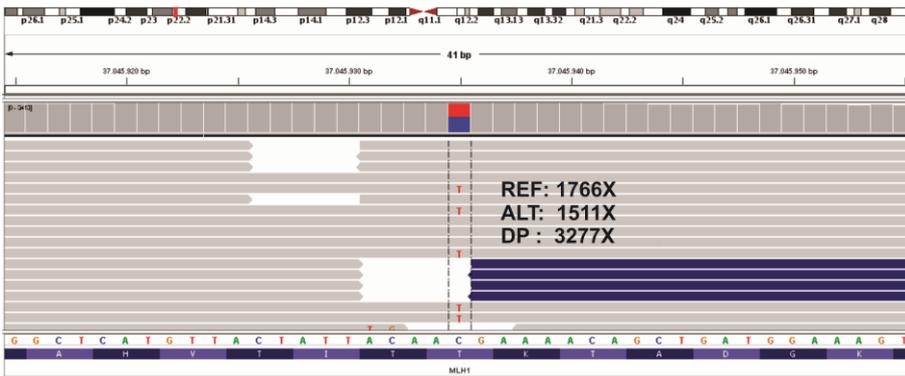


ANEXO 23: Mapas e eletroferogramas para cada uma das variantes patogênicas/VUS c.1681_1682insT, c.350C>T e c.677G>A encontradas no gene *MLH1*.

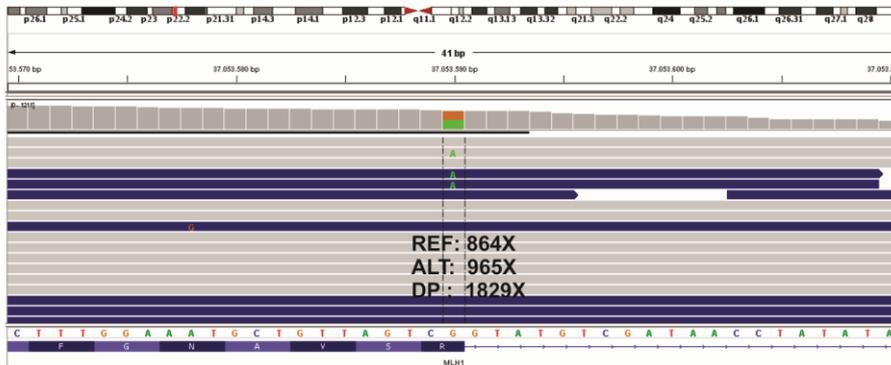
c.1681_1682insT



c.350C>T

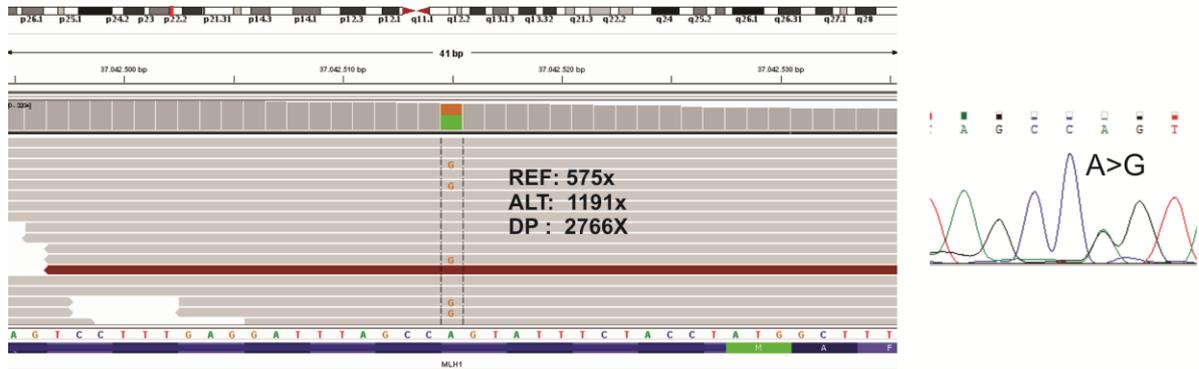


c.677G>A

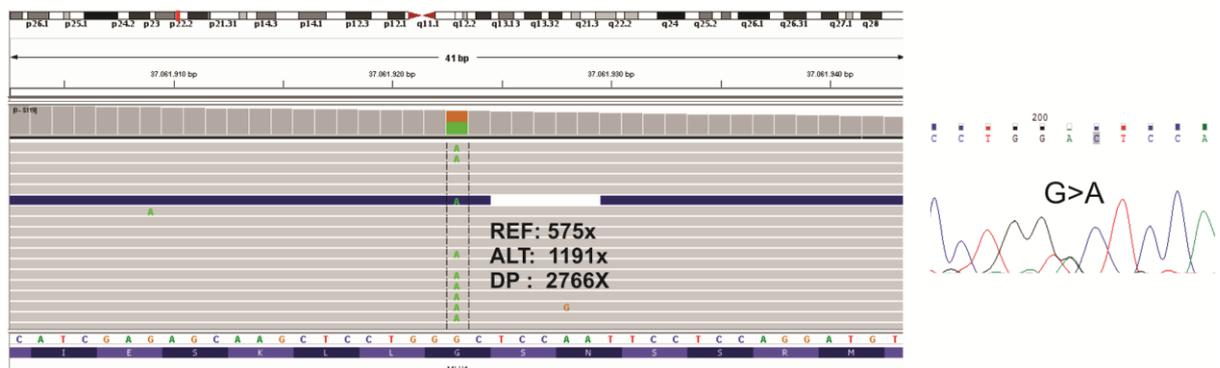


ANEXO 24: Mapas e eletroferogramas para cada uma das variantes patogênicas/VUS c.277G>A e c.1007G>A encontradas no gene *MLH1*.

c.277G>A

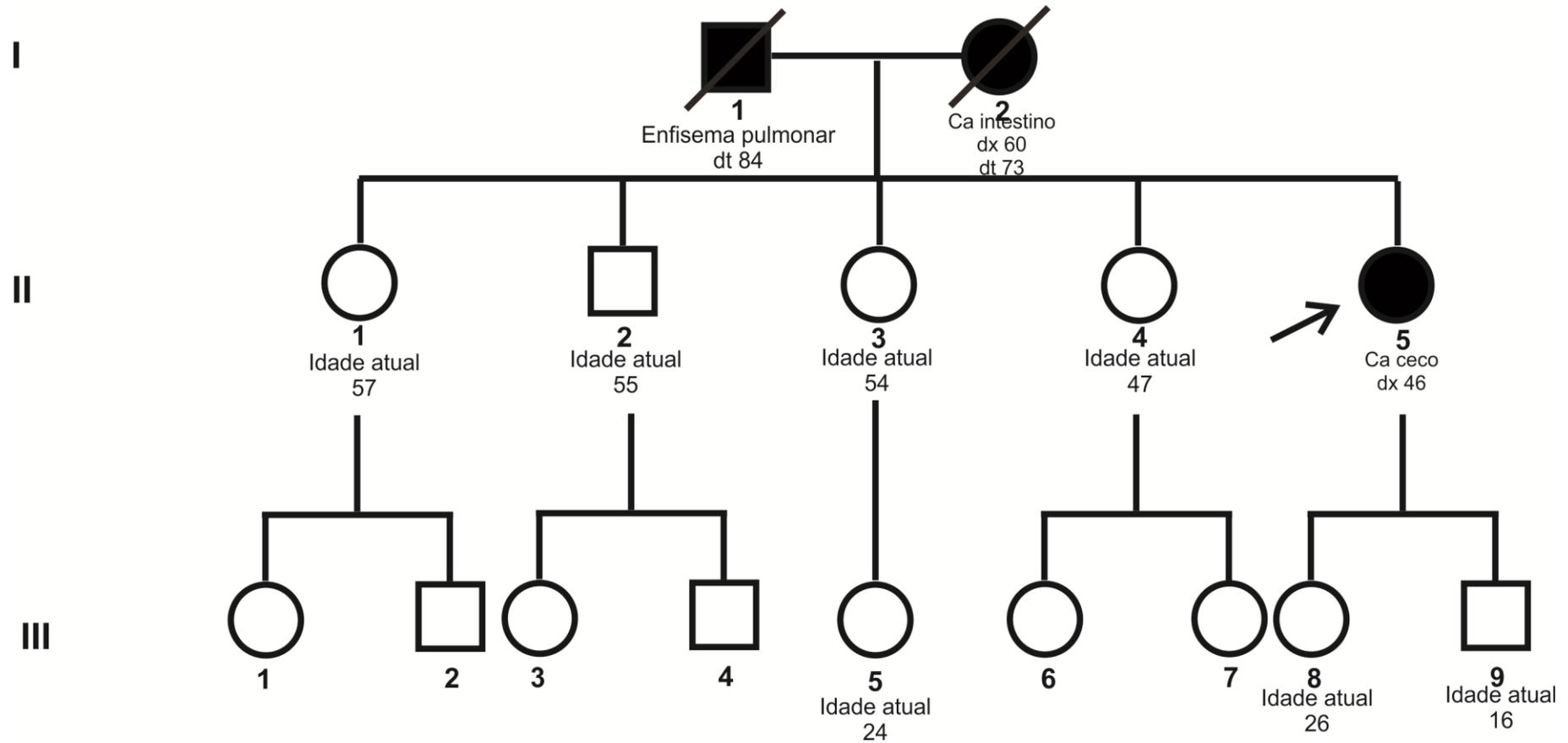


c.1007G>A

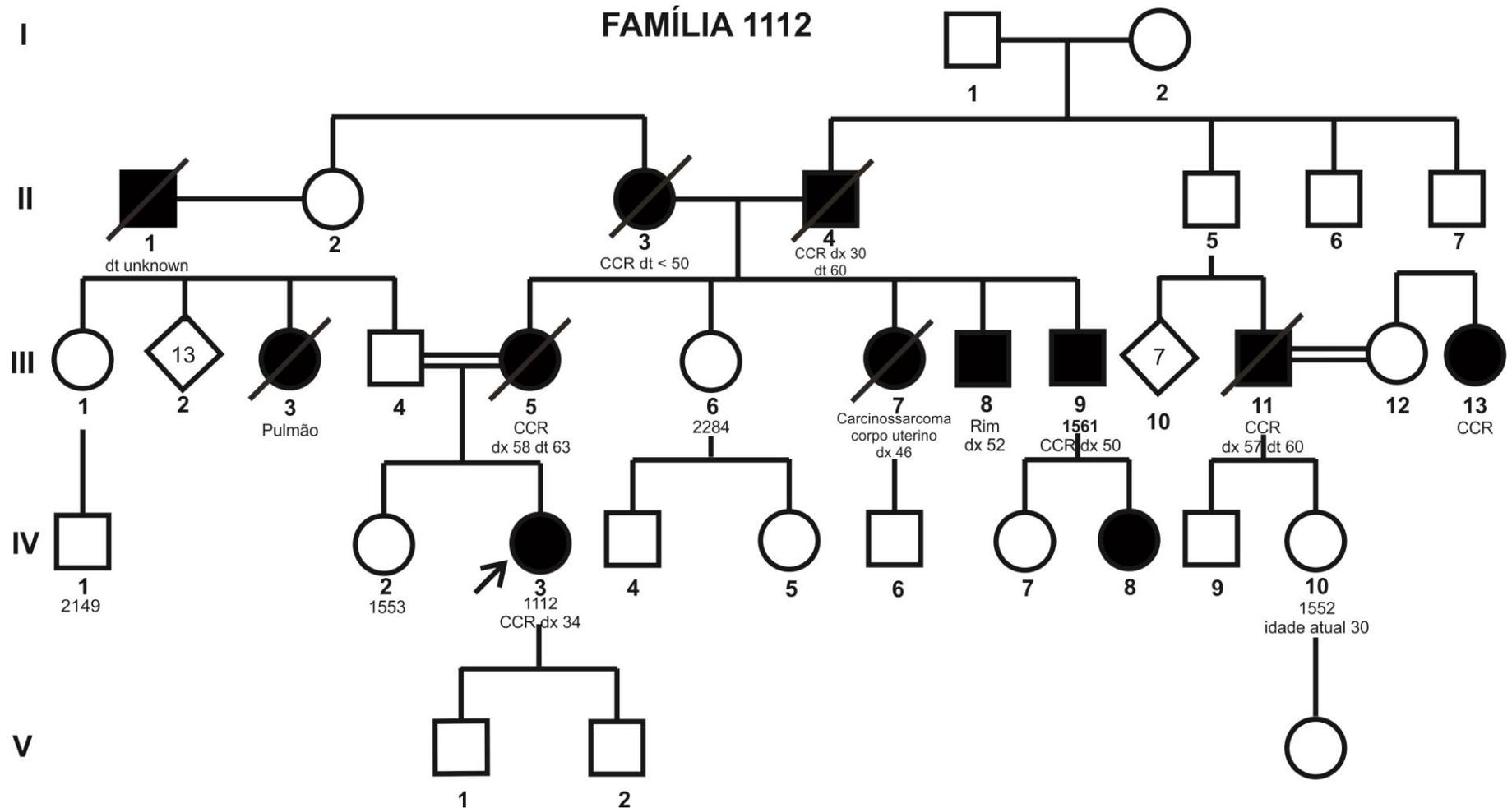


ANEXO 25: Heredograma da família pertencente ao probando CPB, deletado para o gene *MUTYH*.

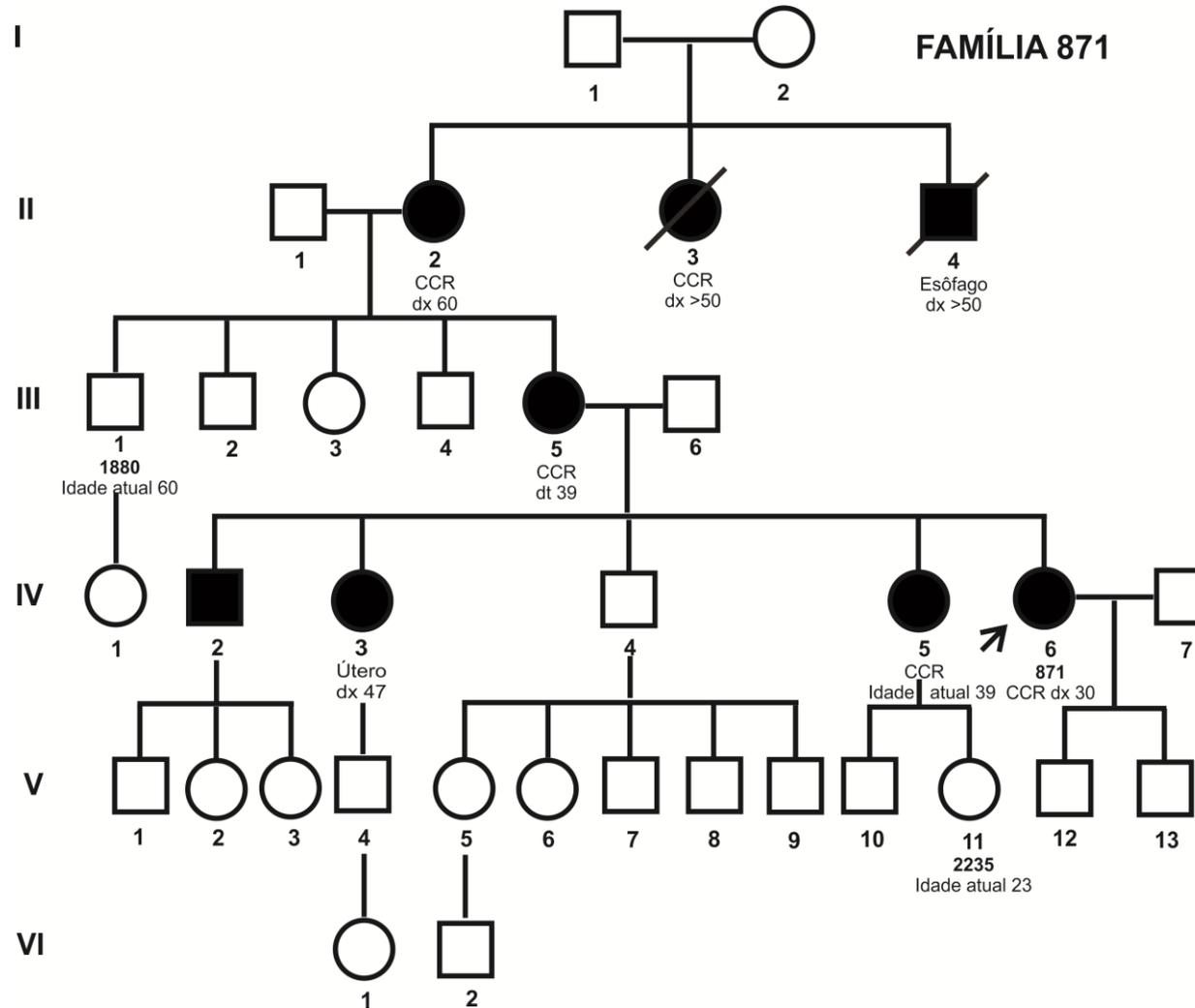
FAMÍLIA CPB



ANEXO 26: Heredograma da família pertencente 1112, portadora da alteração p. Q718*.



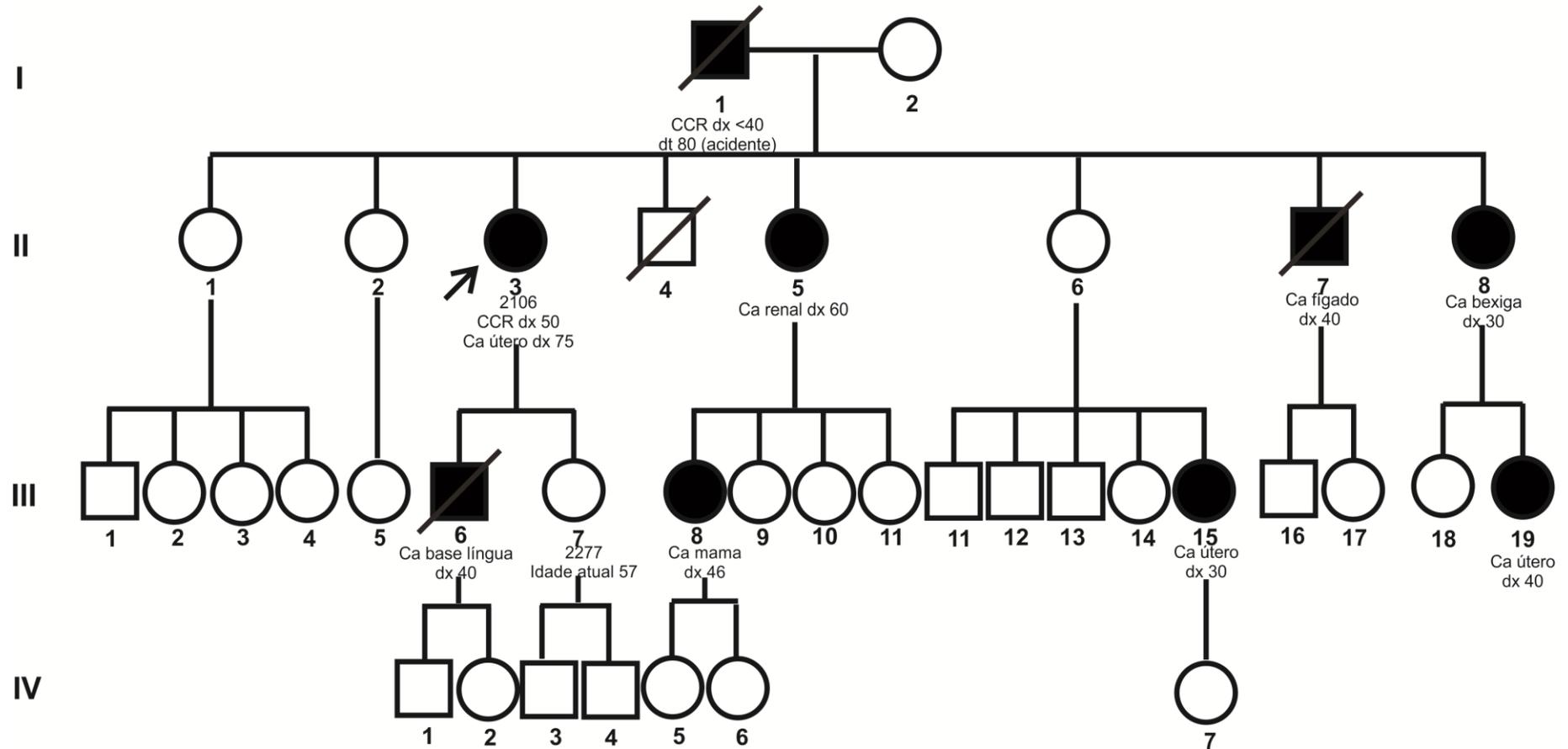
ANEXO 27: Heredograma da família 871, portadora da alteração p. Q718*.



Os pacientes 1180 e 2235 foram também rastreados para a alteração. As idades de diagnóstico e óbito são representadas pelas siglas dx e dt, respectivamente.

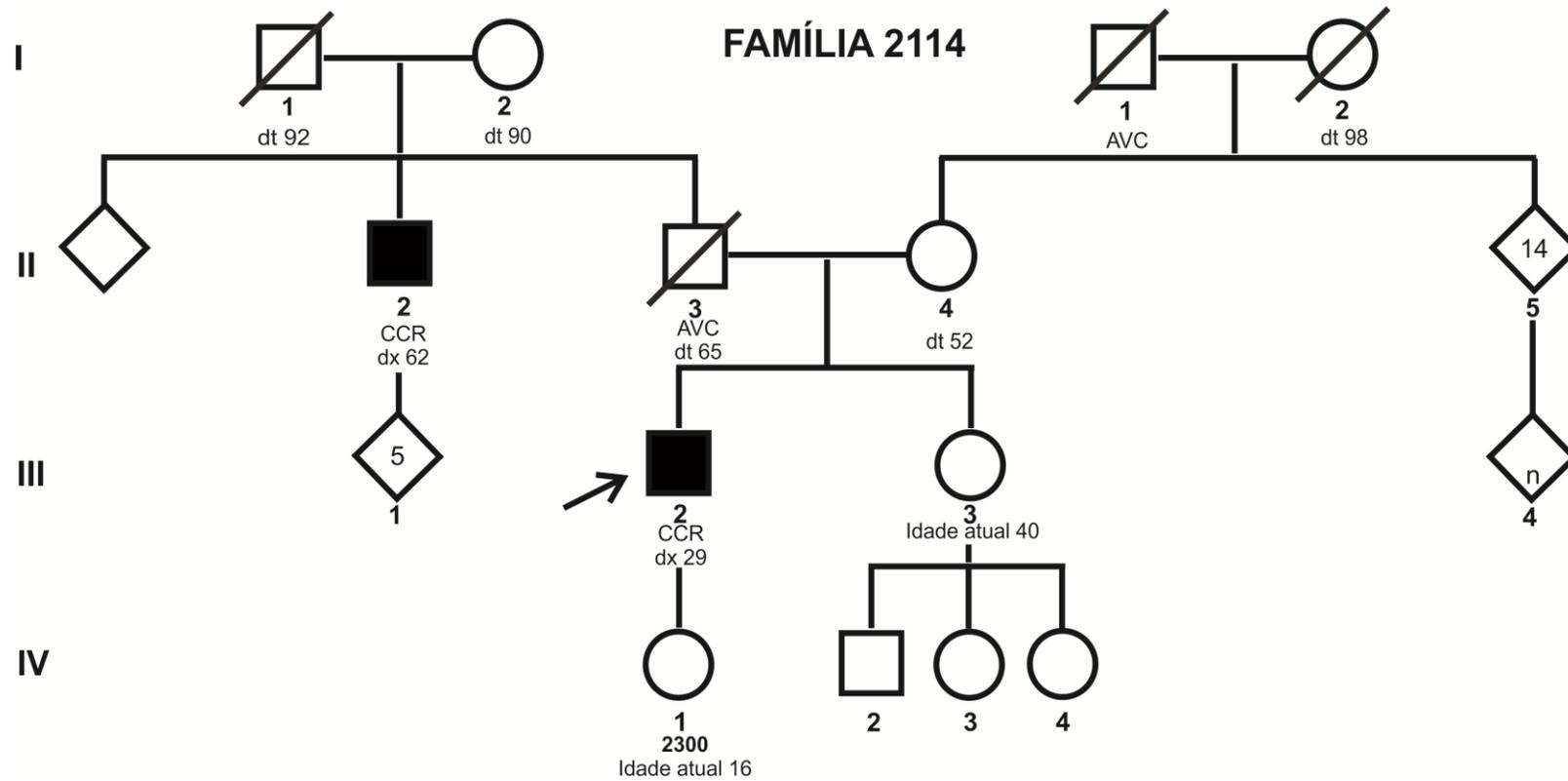
ANEXO 28: Heredograma da família 2106 portadora da alteração c.2152C>T.

FAMÍLIA 2106



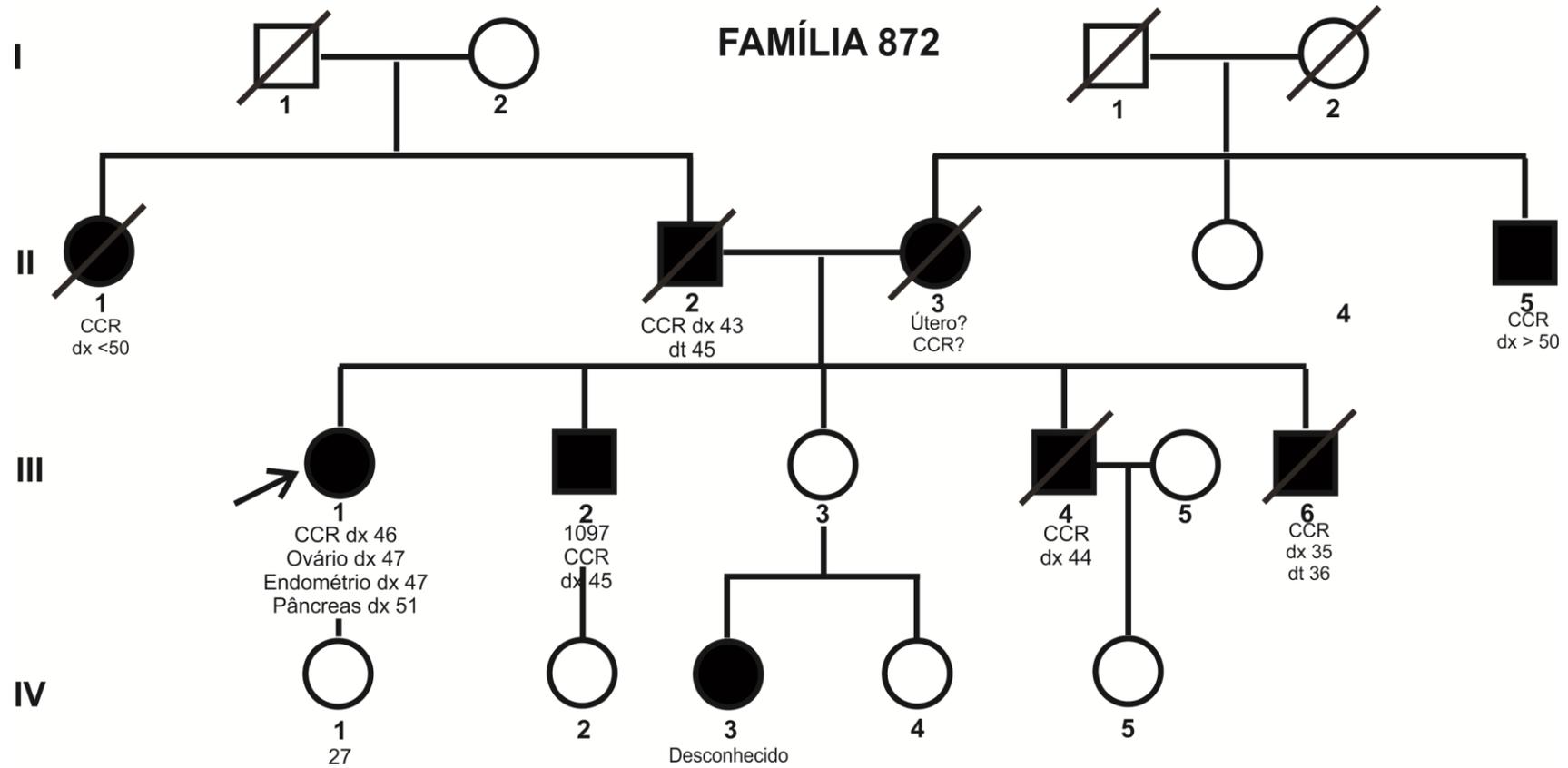
O paciente 2277 foi rastreado para a alteração. As idades de diagnóstico e óbito são representadas pelas siglas dx e dt, respectivamente.

ANEXO 29: Heredograma da família 2114 portadora da alteração c.1738_1741deGAAA.



O paciente 2300 foi rastreado para a alteração. As idades de diagnóstico e óbito são representadas pelas siglas dx e dt, respectivamente.

ANEXO 30: Heredograma da família 872, portadora das alterações em *cis* c.890A>C (p.E297A) e c.888_889_delCA.



O paciente 1097 foi rastreado para a alteração. As idades de diagnóstico e óbito são representadas pelas siglas dx e dt, respectivamente.