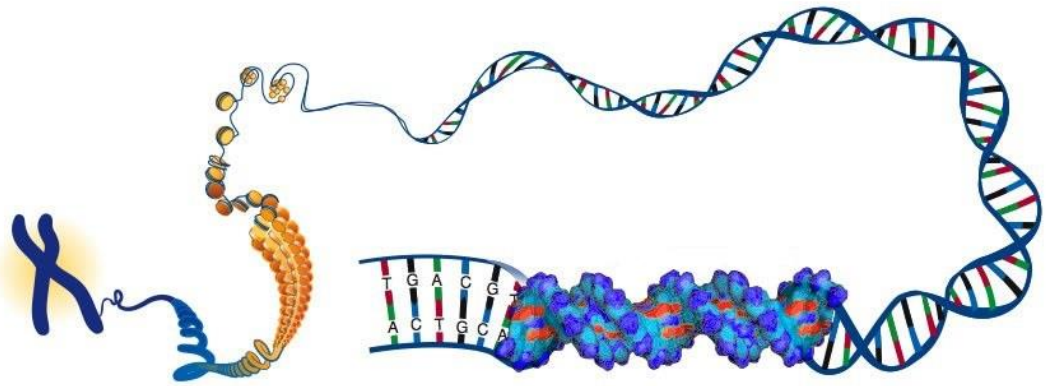
Diana Touceira Pinto

THE ROLE OF *MLH1* CONSTITUTIONAL METHYLATION IN LYNCH SYNDROME

Dissertação de candidatura ao grau de ***Mestre em Oncologia*** – especialização em Oncologia Molecular – submetida ao Instituto de Ciências Biomédicas de Abel Salazar da Universidade do Porto.

Orientador: Manuel António Rodrigues Teixeira, MD, PhD
Director do Serviço de Genética e do Centro de Investigação
Instituto Português de Oncologia do Porto
Professor Catedrático Convidado do Departamento de
Patologia e Imunologia Molecular
Instituto de Ciências Biomédicas de Abel Salazar da Universidade do Porto

Coorientador: Carla Alexandra Cavaco Pinto, MSc
Serviço de Genética e do Centro de Investigação
Instituto Português de Oncologia do Porto



“The purpose of life is to live it, to taste experience to the utmost, to reach out eagerly and without fear for newer and richer experience.”

Eleanor Roosevelt

AGRADECIMENTOS

Em primeiro lugar, gostaria de agradecer ao meu orientador, ao Professor Manuel Teixeira por me mostrar como uma investigação deve ser conduzida, pela oportunidade proporcionada em poder ter seguido esta área e pela ajuda na superação dos vários obstáculos. Todo o conhecimento e ideias partilhadas contribuíram largamente para o meu desenvolvimento, não só como pessoa, mas principalmente como investigadora. Obrigada por toda a sua simpatia, disponibilidade, atenção e boa disposição.

À minha coorientadora, Carla, por ter partilhado comigo toda a sua experiência e conhecimento e por me ter ajudado a superar todas as adversidades que foram surgindo. Tenho a agradecer, acima de tudo, a constante preocupação demonstrada e acompanhamento dado, tanto a nível laboral como pessoal. Nos momentos menos conseguidos, o teu apoio foi essencial em cada conselho dado, em cada lição que me incutiste. Se me perguntassem se tenho algum modelo a seguir na minha vida profissional, sem dúvida que serias tu. Quando for grande quero ser como tu!! És das pessoas mais profissionais, organizadas, inteligentes que já conheci. Obrigada por me teres proporcinado o melhor ano e meio da minha vida!!

Cabe-me a mim agradecer a ambos (orientador e coorientadora), expressando o meu sincero agradecimento, considero-me uma privilegiada por me terem proporcionado a oportunidade de trabalhar e aprender convosco.

À Prof.ª Berta, atual diretora do Mestrado, e a todos os docentes por terem partilhado todo o seu conhecimento sobre esta vasta área que é a Oncologia.

À Joana, acima de tudo pelo enorme companheirismo e amizade. Foram 24 sobre 24 horas juntas durante este último ano, fica acima de tudo, uma grande cumplicidade. Foste um pilar essencial nesta difícil fase da minha vida pessoal. E foste, sem sombra de dúvidas, o meu grande apoio a nível laboral também. Foi, de facto, um privilégio ter partilhado esta fase contigo. Obrigada pela forma absolutamente fantástica que sempre tiveste para me motivar, para me fazer ver as coisas como elas realmente são. Obrigada por suportares o meu mau humor, o meu feitio complicado e por te disponibilizares sempre para me ajudar. A forma empenhada como te prontificaste a discutir os resultados obtidos e deste sugestões pertinentes, foram sem dúvida alguma, cruciais na etapa final e decisiva desta tese. Considero que esta tese também é tua, sem dúvida sem o teu contributo não teria chegado à “meta”. Sempre, desde o primeiro minuto, que

funcionamos como uma equipa e trabalhamos como tal. Aliás, nós costumavamos dizer, em jeito de brincadeira: “Isto de fazer duas teses ao mesmo tempo não é nada fácil”. De facto, não foi fácil, mas assim aprendemos juntas, mas também aprendemos uma com a outra, a cada crítica, cada chamada de atenção, desavença, que fizeram que crescessêmos como pessoas e como profissionais.

Ao Rui, por toda a partilha de ideias e ajuda que sempre me deu ao longo deste trabalho, porque sempre se prontificou a ajudar quando mais precisei. És um exemplo, tanto a nível profissional, como pessoal, és a pessoa mais empreendedora e criativa que conheci. A tua humildade é notável e isso, é sem dúvida, de enaltecer e apreciar. Obrigada por toda a ajuda, por cada palavra, pela tua interminável simpatia e boa disposição. Jamais esquecerei as nossas horas de almoço, as nossas piadas que fizeram com que dias impossíveis se tornassem em dias toleráveis. Foste simplesmente impecável, obrigada!!

À Manela, por estar sempre lá quando eu precisei, por toda a ajuda, pela sua constante boa disposição. Talvez tenha sido abusivo da minha parte procurá-la tantas vezes para esclarecer dúvidas, mas teria sido ingénua em desperdiçar a oportunidade de trabalhar e aprender com alguém tão inteligente e competente para partilhar o conhecimento como tu. És grande Manela!

À Catarina Santos, por ser das pessoas com o maior coração que já conheci. Pela disponibilidade e preocupação demonstrada ao longo deste trabalho. Foi um gosto ter partilhado este momento contigo. Obrigada por enriqueceres e animares as minhas horas de almoço.

À Isabel, por me animar constantemente, sem estes momentos tenho a certeza que teria sido tudo muito mais complicado, mais maçudo. Obrigada por tornar os momentos mais complicados e tristes em momentos mais descontraídos, de transformar problemas em gargalhadas! Sem dúvida que as nossas conversas *nonsense* me vão fazer muita falta.

À Maria e à Marta, pela partilha de experiências e conhecimento, mas acima de tudo pela força, encorajamento e amizade. Obrigada pelos momentos hilariantes que me proporcionaram, pelas conversas “de batanetes”, que sem dúvida, tornaram as coisas com uma carga pesada em coisas muito mais simples. Obrigada por todos os conselhos e pela ajuda sempre que precisei, por toda a partilha de conhecimento e experiências. Nada teria sido igual sem vocês por perto. Gosto muito de vocês!!

À Paula e à Sofia, por serem grandes conselheiras e por partilharem os vossos intermináveis conhecimentos, que foram, sem dúvidas importantes em vários momentos. Obrigada pelo carinho, amizade e preocupação. Obrigada por me animarem e incentivarem, sempre!

A todo o Serviço de Genética do IPO-Porto em geral, por me terem recebido da melhor forma possível e por todo o apoio que me deram ao longo desta etapa. Como eu costumo dizer: -“Há pessoas que entram na nossa vida por uma razão, e ficam por muitas”. Espero que assim seja.

Aos meus amigos de longa data, em especial à Fontes, Sara (minha açoreana), Raquel, Nazaré, Tiago, Liliana, Juliana, Ju, Marina, Carla, Rita e Zé, pela preocupação, pelo encorajamento e, acima de tudo por acreditarem em mim, sempre!

Às “poderosas” e verdadeiras amigas que conheci no decorrer deste Mestrado, Ana e Anabela, obrigada pelos divertidos momentos que me proporcionaram, obrigada pela ajuda, por cada palavra de incentivo, por cada conselho, por nunca me deixarem desistir.

À minha madrinha Eleonor, pelos magníficos conselhos, por todas as palavras, pela constante preocupação, pela ajuda, por estar sempre ao meu lado quando mais precisei. És um exemplo: como pessoa, como profissional, como amiga. És uma grande mulher!

Aos meus avós maternos, sem eles não teria sido possível a realização e concretização deste grande sonho. Obrigada por todo o encorajamento, por nunca desistirem de mim. Obrigada por colmatarem todas as lacunas que foram surgindo na minha vida e por me ensinarem que desistir não pode fazer parte dos meus planos, nunca!

Aos restantes amigos e familiares que sempre me incentivaram, me deram força e acreditaram que eu seria capaz.

Ao Chico, por tudo! Por aturar o meu péssimo feitio, pela ajuda constante, por todo o incentivo. Por simplesmente fazer parte da minha vida e me provar todos os dias que eu sou capaz, por maiores que sejam as dificuldades e adversidades. Tens o dom de tornar tudo muito mais fácil. Obrigada por me acalmares nas situações mais difíceis e festejares comigo as minhas vitórias como se fossem as tuas.

Por último, à minha Deusa, à minha mãe. Em primeiro, por seres a melhor mãe do Mundo e por me teres incutido os melhores valores. És mãe, pai, melhor amiga, conselheira, companheira, és tudo!! Em segundo, por seres uma super mulher, porque sempre lutaste por mim e por ti, és uma protetora. Obrigada por nunca teres deixado que me faltasse nada. Tenho tanto, mas tanto orgulho em ti! És e serás uma mãe coragem, um exemplo a seguir. Sem ti nada disto teria sido concretizável, obrigada por seres a minha incondicional admiradora e por achares que sou sempre capaz!

Por tudo isto, este trabalho é dedicado a ti.

TABLE OF CONTENTS

AGRADECIMENTOS 5

RELEVANT ABBREVIATIONS 17

ABSTRACT 23

RESUMO 27

I. INTRODUCTION 31

1. Overview of colorectal cancer 31

2. Colorectal cancer epidemiology 31

3. Etiology and risk factors 33

4. Anatomy and functionality of the large intestine 35

5. Diagnosis and staging of colorectal cancer 36

5.1. Screening and diagnosis 36

5.2. Histopathology and staging 37

6. Prognosis and treatment 39

7. Colorectal carcinogenesis 40

7.1. Chromosomal instability (CIN) pathway 41

7.2. Microsatellite instability (MSI) pathway 43

7.3. CpG island methylator phenotype (CIMP) pathway 44

8. Hereditary colorectal cancer 46

8.1. Lynch syndrome 47

8.2. The MMR system 49

9. The role of epigenetics in Lynch syndrome 50

9.1. Constitutional epimutation of *MLH1* 51

9.2. Importance of *MLH1* epimutations in clinical practice 52

II. AIMS OF THE STUDY 55

III. MATERIALS AND METHODS 59

1. Patients and samples collection 59
2. DNA and RNA extraction from peripheral blood samples 62
3. DNA extraction from formalin-fixed paraffin-embedded tissue 62
4. DNA extraction from buccal mucosa swabs 62
5. *MLH1* methylation analysis....63
6. *MLH1* promoter sequencing 64
7. *MLH1* transcript quantification 65
8. Screening of the *MLH1* c.655A>G SNP....66
9. *MLH1* allelic expression analysis....67
10. Study of microsatellite instability 67
IV. RESULTS 71
1. Identification of Lynch syndrome cases harboring constitutional *MLH1* promoter methylation....71
2. Confirmation of correct MS-MLPA probe hybridization 75
3. Clinico-pathological features of patients with constitutional methylation of the *MLH1* promoter....75
4. Quantification of the *MLH1* transcript....78
5. *MLH1* allelic expression....79
6. Microsatellite instability analysis 82
V. DISCUSSION 85
1. *MLH1* constitutional methylation....85
2. *MLH1* expression pattern 87
3. Molecular characterization of the tumors 88
VI. CONCLUSION 93
VII. FUTURE PERSPECTIVES 97
VIII. REFERENCES....101

FIGURE INDEX

Figure 1 - Estimated CRC incidence and mortality worldwide in 2012 for both sexes, all ages…….....32

Figure 2 - Estimated age-standardised incidence and mortality rates of CRC for A) women and B) men in Europe, all ages…….....33

Figure 3 - Estimated age-standardised incidence and mortality rates of CRC for both sexes, in Portugal, all ages.....33

Figure 4 - Anatomy of the large intestine. Anterior view of large intestine showing major regions: Cecum, ascending colon, transverse colon, descending colon, sigmoid and rectum.....35

Figure 5 - Genetic model for colorectal tumorigenesis. The accumulation of mutations in oncogene and tumor suppressor genes, which is thought to contribute to tumor progression and the propensity to metastasize.....41

Figure 6 - Different genetic pathways in colorectal cancer pathogenesis. Three distinct parallel pathways are implicated in CRC pathogenesis: (A) Chromosomal instability (CIN) pathway driven by inactivating mutations in tumor suppressor genes and activating mutations in proto-oncogenes, (B) Microsatellite instability (MSI) pathway characterized by inactivation of the mismatch repair (MMR) genes. The failure of MMR genes resulting in mutations in specific target genes involved in proliferation and cellular differentiation, and (C) Serrated pathway (or CIMP) driven by promoter hypermethylation of several genes and presenting *BRAF* mutation.....43

Figure 7 - Circle graph depicting the marked genotypic and phenotypic heterogeneity in hereditary colorectal cancer syndromes….....46

Figure 8 - Steps in MMR system. The replicative DNA polymerase misincorporates a nucleotide during DNA replication…….....50

Figure 9 - Detailed methylation patterns of *MLH1* promoter, assessed by MS-MLPA. The five regions of the CpG islands targeted by the selected probes are shown.....72

Figure 10 - Electropherogram plot representing MS-MLPA in a PBL with constitutional methylation of the *MLH1* promoter. The C- and D-Deng regions, the most important regions associated with the transcription of *MLH1* gene…….....73

Figure 11 - Family pedigrees of four patients positive for constitutional methylation of the *MLH1* promoter. Family history of the four patients with constitutional methylation (A) #3; B) #10; C) #27; D) #38 with age (in years) and health status.....78

Figure 12 - Electrophoretic pattern of RT-PCR multiplex products. This was obtained for the four positive patients for constitutional methylation and five normal controls, where M is the molecular weight marker (100bp) and *B2M* is beta-2-microglobulin.....79

Figure 13 - Electropherogram of fragment analysis of the RT-PCR multiplex products. The relative peak areas of a positive case for constitutional methylation (A) and control (B) are shown……..…….....80

Figure 14 - Electropherogram of sequencing analysis demonstrating the allelic expression of the SNP rs1799977 in exon 8 of *MLH1* in cDNA samples. Patient #10 (A), Control (B), (both reverse).....80

Figure 15 - SNuPE analysis at *MLH1* rs1799977 (c.655A>G) in gDNA (upper panels) and in cDNA (bottom panels) of an heterozygous patient, case #10 with constitutional methylation (left panels) and of a control (right panels). Partial transcriptional silencing of the A allele in the cDNA of the patient was observed.....81

Figure 16 - SNuPE analysis at *MLH1* rs1799977 (c.655A>G) in gDNA (upper panels) and in cDNA (bottom panels) of an heterozygous patient, case #27 with constitutional methylation (left panels) and of a control (right panels). Partial transcriptional silencing of the A allele in the cDNA of the patient was observed……..…82

Figure 17 - Fragment analysis showing microsatellite sequences of mononucleotide BAT26 in cases #3 (A), #10 (B), #27 (C), and #38 (D)…….....83

TABLE INDEX

Table 1 - Overview of the AJCC TNM cancer staging system for CRC.....38

Table 2 - Staging groups for CRC, according to AJCC guidelines……..….39

Table 3 - Revised minimum criteria for clinical definition of Lynch syndrome (Amsterdam and Bethesda criteria)…….....48

Table 4 - Clinico-pathological data of 38 index cases fulfilling the clinical criteria for Lynch syndrome…….....60,61

Table 5 - Primer sequences for *MLH1* promoter sequencing and primer sequences for rs1799977 in cDNA and gDNA *MLH1* gene.....66

Table 6 - *MLH1* methylation levels (%) using MS-MLPA in samples from different germline origins in the four probands.....74

Table 7 - Clinical and molecular features of patients with constitutional methylation of the *MLH1* promoter.....75

RELEVANT ABBREVIATIONS

RELEVANT ABBREVIATIONS

5-FU - 5-fluorouracil
AICR - American Institute for Cancer Research
AJCC - American Joint Committee on Cancer
APC - adenomatous polyposis coli
ASR - age-standardized rate
BAX - BCL2-associated X protein
BRAF - B-Raf proto-oncogene, serine/threonine kinase
CA19-9 - carbohydrate antigen 19-9
CACNA1G - calcium channel, voltage-dependent, T type, alpha 1G subunit
CAPOX - **CAP** - Capecitabine; **OX**- Oxaliplatin
cDNA - complementary deoxyribonucleic acid
CEA - carcinoembryonic antigen
CIMP - CpG island methylator phenotype
CIMP- H - CpG island methylator phenotype high
CIMP- L - CpG island methylator phenotype low
CIN - chromosomal instability
CRC - colorectal cancer
CSPY - colonoscopy
CT - computed tomography
CTC - computed tomographic colonography
CTNNB1 - cadherin-associated protein beta 1
DCBE - double-contrast barium enema
DCC - DCC netrin 1 receptor
ddNTP - dideoxyribonucleotide triphosphate
DNA - deoxyribonucleic acid
dNTP - deoxynucleoside triphosphate
EGFR - epidermal growth factor receptor
ERK - extracellular signal-regulated kinase
EXO1 - exonuclease 1
FAP - familial adenomatous polyposis
FDA - Food and Drug Administration
FDG-PET - fluorodeoxyglucose-positron emission tomography

FIT - fecal immunochemical test

FOBT - fecal occult blood test

FOLFIRI - **FOL-** Leucovorin Calcium (Folinic Acid); **F**- Fluorouracil; **IRI**- Irinotecan Hydrochloride

FOLFOX - **FOL**- Leucovorin Calcium (Folinic Acid); **F**- Fluorouracil; **OX**- Oxaliplatin

FSIG - flexible sigmoidoscopy

gDNA - genomic deoxyribonucleic acid

GDP - guanosine 5´ diphosphate

GISTs - gastrointestinal stromal tumors

GTP - Guanosine 5' triphosphate

IARC - International Agency for Research on Cancer

IGF2 - insulin-like growth factor 2

IHC - immunohistochemistry

JPS - juvenile polyposis syndrome

KRAS - kristen rat sarcoma viral oncogene homolog

LOH - loss of heterozygosity

LS- Lynch syndrome

MAP - *MUTYH*-associated polyposis

MAPK - mitogen-activated protein kinase

mCRC - metastatic colorectal cancer

MGMT - O (6)-methylguanine-DNA methyltransferase

MLH1 - mutL homolog 1

MMR - mismatch repair

MRI - magnetic resonance imaging

mRNA - messenger ribonucleic acid

MSH2 - mutS homolog 2

MSH3 - mutS homolog 3

MSH6 - mutS homolog 6

MSI - microsatellite instability

MSI-H- microsatellite instability high

MSI-L - microsatellite instability low

MSS - microsatellite stable

MYC - v-myc avian myelocytomatosis viral oncogene homolog

NCCN - Nacional Comprehensive Cancer Network

NEUROG1 - neurogenin 1

NSAIDs - non-steroidal anti-inflammatory drugs

PBL - peripheral blood lymphocytes

PCR - polymerase chain reaction

PET - positron emission tomography

PJS - Peutz-Jeghers syndrome

PMS1 - PMS1 protein homolog 1

PMS2 - mismatch repair endonuclease PMS2

PTEN - phosphatase and tensin homolog

RNA - ribonucleic acid

RPA - replication protein A

RUNX3 - runt-related transcription factor 3

sDNA - stool DNA test

SMAD2 - SMAD family member 2

SMAD4 - SMAD family member 4

SNP - single nucleotide polymorphism

SNuPE - single-nucleotide primer extension

SOCS1 - suppressor of cytokine signaling 1

TGF-α - transforming growth factor α

TGF-β - transforming growth factor β

TGFBR2 - transforming growth factor, beta receptor II

TNM - tumor, node, metastasis

TP53 - tumor protein p53

VEGF - vascular endothelial growth factor

VEGF-A - vascular endothelial growth factor A

VEGFR - vascular endothelial growth factor receptor

WCRF - World Cancer Research Fund

WNT - Wingless-Type MMTV Integration Site Family

ABSTRACT

ABSTRACT

Colorectal cancer (CRC) is the fourth most frequent type of cancer in both sexes worldwide. Most of the CRC are sporadic, but approximately 10% of the cases are hereditary. Lynch syndrome (LS) is the most common hereditary syndrome that predisposes patients to CRC, corresponding to 2-5% of all CRCs. It is described as an autosomal dominant disease caused by germline mutations in the DNA repair genes - Mismatch Repair (MMR) genes. These include *MLH1*, *MSH2*, *MSH6* and *PMS2*, and about 90% of the mutations described in this syndrome occur in *MLH1* or *MSH2*. There are cases described in the literature with clinical criteria for LS without germline mutation in MMR genes, which have "constitutional epimutations" or simply "'epimutations" in the *MLH1* or in the *MSH2* genes. This phenomenon consists in transcriptional silencing of the promoter of these genes by epigenetic mechanisms rather than by genetic mutations that directly affect the sequence of the gene.

In a series of 38 patients from families who meet clinical criteria for LS, with loss of MLH1 expression in the tumor and with no germline mutations in the *MLH1*, we screened for constitutional methylation of the *MLH1* gene promoter using methylation-specific multiplex ligation-dependent probe amplification (MS-MLPA). We found four (4/38; 10.5%) patients with constitutional methylation in the *MLH1* gene promoter in mosaicism. RNA studies demonstrated a decreased *MLH1* expression in the cases with constitutional methylation when compared with controls, and in two cases (heterozygous for a coding polymorphism) we could demonstrate that this reduction in expression was monoallelic. These results indicate that the constitutional *MLH1* promoter methylation directly correlates with their reduction of expression. All tumors of the patients harboring constitutional methylation were microsatellite instability-high (MSI-H) and did not present the p.Val600Glu *BRAF* mutation. In addition, we were able to study three relatives of one of the probands (parents and sister). Constitutional methylation was not detected in any of the three family members, suggesting that the methylation arose *de novo* in this proband. Additional studies will be necessary for a correct classification of the origin of these epimutations. This classification will allow the identification of the risk for relatives to inherit the methylation and develop tumors associated with LS, thus allowing adapting screening and/or the prophylactic measures according to risk.

RESUMO

RESUMO

O cancro colorretal (CCR) é a quarta neoplasia mais frequente em ambos os sexos em todo o mundo. A grande maioria dos CCRs são esporádicos, no entanto cerca de 10% dos casos são hereditários. A síndrome de Lynch é a forma mais comum de CCR hereditário, correspondendo a 2-5% de todas as neoplasias colorretais. É descrita como uma doença autossómica dominante, causada por mutações germinativas nos genes de reparação do DNA - genes *Mismatch Repair* (MMR). Estes incluem os genes *MLH1*, *MSH2*, *MSH6* e *PMS2*, e cerca de 90% das mutações descritas nesta síndrome ocorrem no *MLH1* ou no *MSH2*. Existem casos descritos na literatura com critérios clínicos para síndrome de Lynch, sem mutação germinativa identificada nos genes MMR, que apresentam ''epimutações constitucionais'', ou simplesmente ''epimutações'' nos genes *MLH1* ou *MSH2*. Estas consistem no silenciamento transcripcional do promotor destes genes por mecanismos epigenéticos, e não por mutações genéticas que afetam diretamente a sequência do próprio gene.

Numa série de 38 pacientes pertencentes a famílias que cumprem critérios clínicos de síndrome de Lynch, com perda de expressão de MLH1 no tumor e sem mutações germinativas no *MLH1*, foi pesquisada metilação constitucional do promotor do gene *MLH1*, através da técnica de *methylation-specific multiplex ligation-dependent probe amplification* (MS-MLPA). Foram encontrados quatro (4/38; 10,5%) pacientes com metilação constitucional no gene *MLH1* em mosaicismo. Estudos de RNA demonstraram redução da expressão do gene *MLH1* nos casos com metilação constitucional quando comparados com controlos, e em dois casos (heterozigóticos para um polimorfismo codificante) foi possível demonstrar que essa redução de expressão era monoalélica. Estes resultados indicam que a metilação constitucional do promotor do *MLH1* está correlacionada com a redução de expressão desse mesmo gene. Observamos que todos os tumores dos pacientes positivos para a metilação constitucional apresentavam alta instabilidade de microssatélites e não apresentavam a mutação p.Val600Glu do gene *BRAF*. Adicionalmente, foi possível estudar três familiares de um dos casos índice (pais e irmã). Não foi detetada metilação constitucional em nenhum dos três familiares, o que sugere que a metilação neste paciente surgiu *de novo*. Estudos adicionais serão necessários para uma correta classificação da origem destas epimutações. Esta classificação permitirá identificar o risco dos familiares herdarem a metilação e de desenvolverem tumores associados com a síndrome de Lynch, permitindo, assim, adoptar medidas de rastreio e/ou profilaxia adequadas ao risco.

I. INTRODUCTION

I. INTRODUCTION

1. Overview of colorectal cancer

Colorectal cancer (CRC) is a complex disease with a heterogeneous etiology, caused by a combination of numerous factors with environmental, genetic and epigenetic origins (Gryfe, 2009). It is characterized by an accumulation of genetic and epigenetic alterations leading to an invasive state, and usually occurs in one of three patterns: inherited, familial or sporadic (Berg and Søreide, 2011; Markowitz and Bertagnolli, 2009; Worthley *et al.*, 2007). Inherited forms are responsible for about 5-10% of all CRCs and are related to recognized hereditary conditions (Gryfe, 2009).

2. Colorectal cancer epidemiology

Currently, cancer is the leading cause of death in economically developed countries and the second leading cause of death in developing countries. CRC is increasing in economically developing countries as a result of population aging and growth, adoption of cancer-associated lifestyle choices, such as smoking, physical inactivity, and type of diet (Jemal *et al*., 2011).

Considering both sexes, CRC was the fourth most common cancer in the world, accounting for 9.7% (1,360,602 cases) of all new cancer cases diagnosed in 2012, with an estimated age-standardized rate (ASR) incidence of 17.2/100,000 (Ferlay *et al*., 2013). CRC represents the fourth leading cause of cancer mortality with 8.5% (693,933 deaths) estimated worldwide, with an estimated ASR mortality of 8.4/100,000 (Ferlay *et al*., 2013). More specifically, CRC is the third most incident cancer in men, with an estimated ASR incidence of 20.6/100,000 (746,000 cases, 10.0% of the total) and the second in women, with an estimated ASR incidence of 14.3/100,000 (614,000 cases, 9.2% of the total) worldwide (figure 1) (Ferlay *et al*., 2013). CRC represents the fourth most common cause of death from cancer in men and the third in women, with an estimated ASR mortality of 10.0/100,000 and 6.9/100,000, respectively (Ferlay *et al*., 2013). Almost 55% of the cases occur in more developed regions. There is wide geographical variation in incidence across the world and the geographical patterns are very similar in men and women with incidence rates varying ten-fold in both sexes worldwide (figure 1) (Ferlay *et al*., 2013).

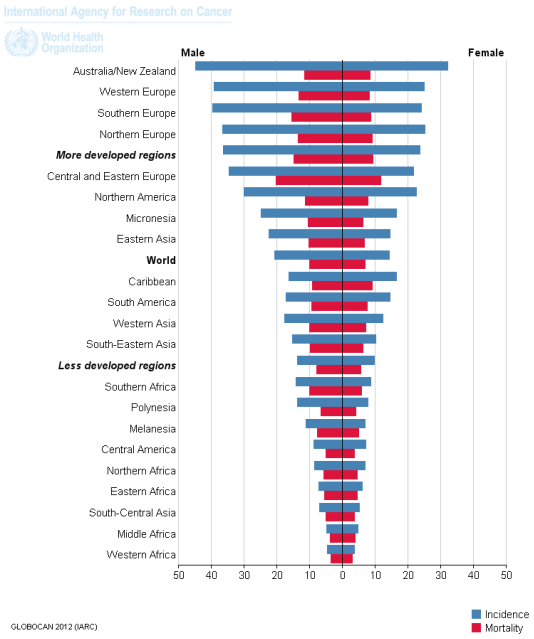


Figure 1 - Estimated CRC incidence and mortality worldwide in 2012 for both sexes, all ages (Ferlay *et al.,* 2013).

In 2012, CRC was the third most common malignancy in Europe in both sexes (excluding non-melanoma skin cancers), with 13.1% (447,136 newly diagnosed cases), with an estimated ASR incidence of 29.5/100,000 of all new cancer cases (Ferlay *et al*., 2013). In Europe, CRC is slightly more incident in men than in women. In women it is the second most incident cancer, with an estimated ASR incidence of 23.6/100,000, and the third most common in men with an estimated ASR incidence of 37.3/100,000 (Ferlay *et al*., 2013). The main difference is that, in Europe, CRC is considered the second cause of cancer-related death, with an estimated ASR mortality of 12.5/100,000 in both sexes (figure 2) (Ferlay *et al*., 2013; Siegel *et al.*, 2013).

In Portugal, differences in incidence and mortality rates of this disease have been changing over the past few years. These rates have decrease due to screening, improvement in diagnosis and treatment, and, above all, by adopting healthier lifestyles (Jemal *et al*., 2011; Siegel *et al*., 2014). In 2012, it was the third highest incident malignancy, corresponding to 14.5% (7,129 new cases) with an estimated ASR incidence of 31.7/100,000 diagnosed in both sexes, after breast cancer in women and prostate in men. It was also the leading cause of death by cancer, which accounts for 15.7% (3,797 deaths) of all cancer deaths, with an estimated ASR mortality of 13.6/100,000 in both sexes (figure 3) (Ferlay *et al*., 2013).

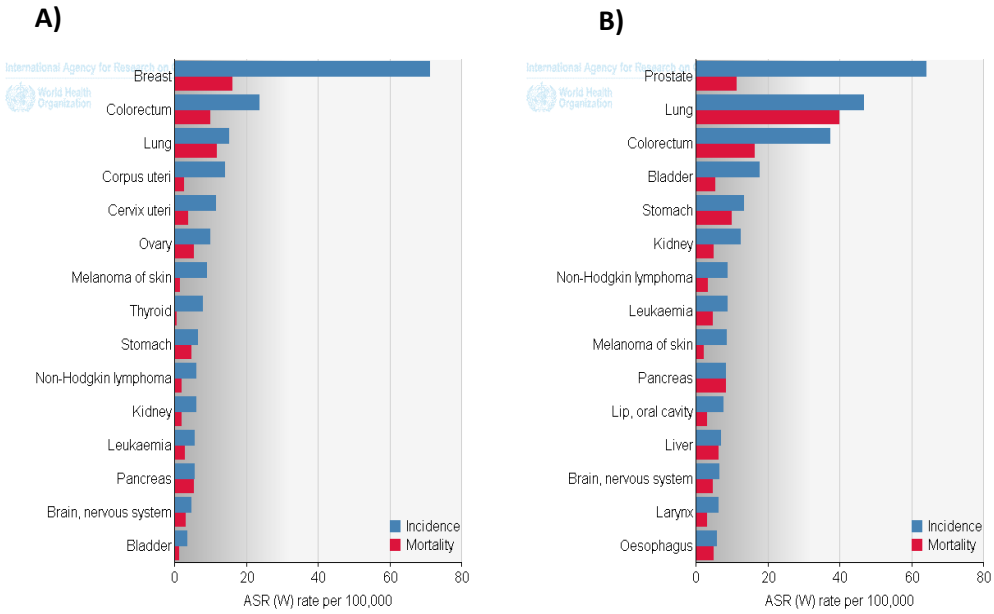


Figure 2 - Estimated age-standardized incidence and mortality rates of CRC for A) women and B) men in Europe, all ages. ASR, age-standardized rates (Ferlay *et al.,* 2013).

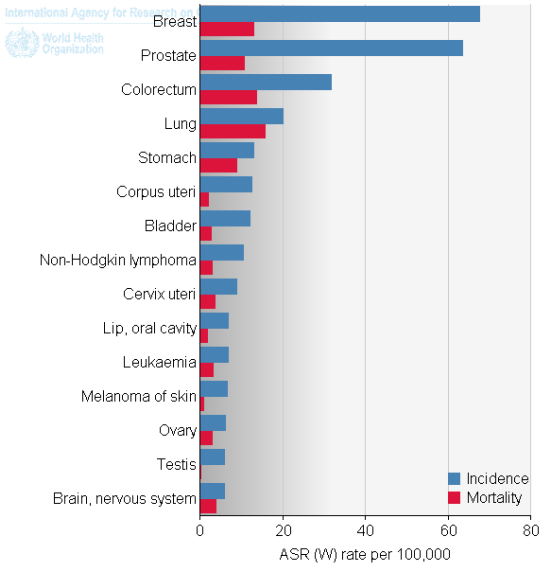


Figure 3 - Estimated age-standardized incidence and mortality rates of CRC in both sexes, in Portugal, all ages. ASR, age-standardized rates (Ferlay *et al.,* 2013).

3. Etiology and risk factors

The etiologic risk factors can be divided into environmental and genetic factors. Theoretically, environmental factors can be prevented and/or changed by the individual whereas the genetic factors are not under such control. Multiple dietary and lifestyle

factors are known to be related to an increased risk of developing CRC (Theodoratou *et al.*, 2014).

Several epidemiological studies have suggested that diet, physical inactivity, obesity, tobacco use, alcohol consumption, diabetes and the presence of inflammatory bowel diseases play an important role in the etiology of the disease (Huxley *et al.*, 2009). In fact, it was suggested that nearly 66-77% of CRC might be preventable by an appropriate combination of diet, physical activity and a healthy lifestyle (Giovannucci, 2002; Platz *et al.*, 2000). High intake of animal fat may increase the risk of CRC, although previous studies have possibly overemphasized the risk. Other dietary components of etiological importance include vegetables, fibre and vitamin C (all of which probably lower the risk), whereas high alcohol consumption appears to increase it. The World Cancer Research Fund (WCRF), in combination with the American Institute for Cancer Research (AICR), in 2007, released the second expert report that summarizes the current scientific evidence on diet and nutrition, where it is estimated that these are responsible for about 30-50% of the worldwide incidence of CRC. Diet and nutrient factors are widely believed to act as pro- and anti-tumor risk modifiers across the entire multistep process of CRC tumorigenesis, which includes tumor initiation, promotion, and progression (Vargas and Thompson, 2012). Furthermore, a number of reports have suggested that regular use of aspirin and/or non-steroidal anti-inflammatory drugs (NSAIDs) may be protective (Bosetti *et al*., 2012).

Age and gender are also risk factors, since increasing age is associated with advanced neoplasia, and CRC incidence rates are lower in women than in men, with a particularly striking discrepancy between pre-menopausal women and age-matched men (Purim *et al*., 2013; Stegeman *et al*., 2013). The likelihood of diagnosis of CRC increases progressively (90%) in people aged 50 or older (Haggar and Boushey, 2009).

Another important risk factor is family history (Stegeman *et al*., 2013). Up to 20% of people who develop CRC have other family members who have been affected by this disease (Giardiello *et al*., 2014). Individuals with history of CRC or adenomatous polyps in one or more first-degree relatives are at an increased risk. This risk is higher in people with a stronger family history, such as history of CRC or adenomatous polyps in any first-degree relative younger than age 60; or history of CRC or adenomatous polyps in two or more first-degree relatives at any age. The reasons for the increased risk are not clear, but it is likely due to inherited mutations, shared environmental factors, or some combination of both (Dunlop *et al*., 2013). Approximately 5-10% of CRCs are a consequence of recognized hereditary conditions (Jasperson *et al*., 2010). Lynch syndrome (LS) and familial adenomatous polyposis (FAP) are the most common hereditary syndromes that predisposes patients to CRC (Hughes and Huang, 2011).

4. Anatomy and functionality of the large intestine

The colon and rectum are portions of the digestive system, also called gastrointestinal system. Cancer develops less commonly in the small intestine than in the colon or in the rectum. The small intestine joins the large intestine in the lower right abdomen. The leading function of colon is to absorb water and mineral nutrients from the food matter (Tortora and Derrickson, 2012). The large intestine presents 1.5m long and 6.5cm in diameter. The first portion, the cecum, forms a blind-ended pouch from which extends the appendix. The colon is divided into ascending, transverse, descending, and sigmoid portions. The ascending colon lies retroperitoneally, lacks a mesentery and it is continuous with the transverse colon at the hepatic (right) flexure of colon near the right inferior margin of the liver. The transverse colon has its own mesentery, the transverse mesocolon and it becomes continuous with the descending colon at the splenic (left) flexure. The terminal portion of the descending colon is S-shaped, forming the sigmoid colon, which empties into a relatively straight segment of the large intestine, the rectum, which ends at the anus (figure 4) (Widmaier, 2011). The ascending and transverse sections are collectively referred to as the proximal colon, while the descending and sigmoid colon are referred to as the distal colon (Tortora and Derrickson, 2012). Relatively to the anatomy of the rectum is usually divided into three portions: the lower rectum, the midrectum, and the upper rectum. The determination of the boundary between rectum and sigmoid colon is important in defining adjuvant therapy (DeVita *et al*., 2011).

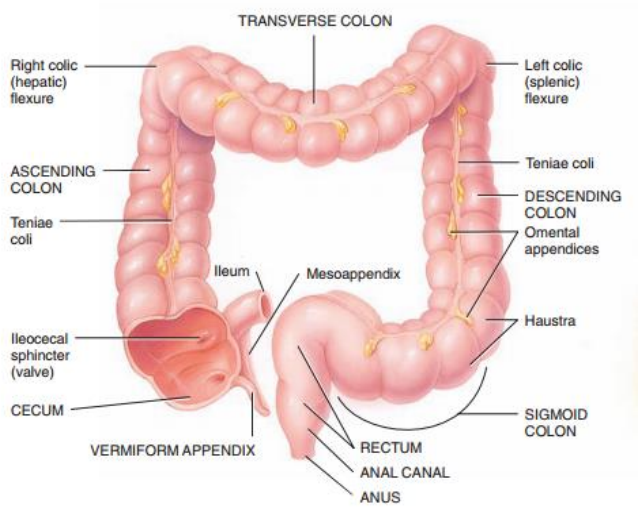


Figure 4 - Anatomy of the large intestine. Anterior view of large intestine showing major regions: Cecum, ascending colon, transverse colon, descending colon, sigmoid and rectum (adapted from Tortora and Derrickson, 2012).

5. Diagnosis and staging of colorectal cancer

5.1. Screening and diagnosis

Many symptoms of CRC have been described, with the main ones being rectal bleeding, diarrhea, constipation, loss of weight, cramping, abdominal pain, anemia, decreased appetite and weakness, and, in particular, obstructive symptoms (DeVita *et al*., 2011; Labianca *et al*., 2013). Colorectal cancer diagnosis can be difficult because most of these symptoms are non-specific (Cappell, 2008).

There is no test available for use in primary care that has sufficient discrimination to provide the basis for referral decisions, although primary care investigation sometimes includes fecal occult blood testing and estimation of hemoglobin (Astin *et al*., 2011). For example, Carcinoembryonic Antigen (CEA) is a tumor marker often elevated in CRC, but it is insufficiently specific to be a reliable indicator of the disease (He and Efron, 2011). Use of recommended CRC screening tests can both detect earlier or prevent CRC by promoting the removal of precancerous polyps (Bujanda *et al*., 2012; Cunningham *et al*., 2010; Winawer *et al*., 2006). Recommended strategies for CRC screening fall into two broad categories: 1) stool DNA tests (sDNA): e.g. fecal occult blood test (FOBT), fecal immunochemical test (FIT) and fecal DNA testing; and 2) structural examinations: which comprises flexible sigmoidoscopy (FSIG), colonoscopy (CSPY), double-contrast barium enema (DCBE) and computed tomographic colonography (CTC) (Bujanda *et al*., 2012; Levin *et al*., 2008). Regarding the structural examinations, endoscopy methods can detect abnormalities and remove them in one procedure. The two main endoscopy procedures are FSIG and CSPY. With FSIG only approximately one-half of the colorectum can be examined, whereas CSPY generally visualizes the entire colorectum (Bujanda *et al*., 2012; Hassan *et al*., 2005). According to the literature, there are specific guidelines described for CRC screening for individuals at average risk. Colonoscopy is also the most appropriate and sensitive test for CRC diagnosis. To evaluate the presence and extent of metastases, particularly in the liver, a physical examination combined with chest x-Ray, liver function tests, CEA level, computed tomography (CT) combined with Fluorodeoxyglucose-positron emission tomography (FDG-PET) or magnetic resonance imaging (MRI) are, currently, the best options (Heitman *et al*., 2010; Townsend, 2007; van Cutsem *et al*., 2010).

5.2. Histopathology and staging

The type of colorectal tumor describes the cells from which the tumor arises. Typically, they may be divided into epithelial and non-epithelial. The first group encompasses the adenomas, carcinomas and carcinoid tumors, whilst the second group includes the malignant lymphomas and gastrointestinal stromal tumors (GISTs) (Hamilton and Aaltonen, 2000).

Colorectal cancers can be classified as well differentiated, moderately differentiated, or poorly differentiated based on the degree of preservation of normal glandular architecture and cytologic features (Cappell, 2008). Approximately 20% of CRCs are poorly differentiated and they confer a poor prognosis (Hassan *et al*., 2005). Most colorectal carcinomas are gland-forming with variations in size and conformation of the glandular structures, and may be divided into several types according to those (table 1) (Hamilton and Aaltonen, 2000; Quirke *et al*., 2011).

Adenocarcinoma is the most common type of colorectal carcinoma (nearly 85%) and 10-15% are commonly diagnosed as mucinous adenocarcinoma (Lanza *et al*., 2011). Mucinous cancers are defined histologically by the presence of abundant extracellular mucin, with more than 50% of the tumor mass being mucinous (Marzouk and Schofield, 2011). Other tumor types (signet-ring cell, small cell, squamous cell, medullary carcinoma and undifferentiated carcinomas) present a much lower incidence (Lanza *et al*., 2011).

Approximately 20-25% of CRC patients initially presents metastases (Lanza *et al*., 2011). CRC metastasizes firstly to the liver because of the venous drainage of the colon via the portal system (Poston *et al*., 2008). Other sites, including the lungs, peritoneum, pelvis, and adrenals, typically become involved only after hepatic or lymphatic metastases occur. Rectal cancers, which are below the peritoneal reflection, lack a serosa and, therefore, penetrate firstly into adjacent pelvic structures (Hamilton and Aaltonen, 2000; Hugen *et al*., 2014).

Pathologic tumor staging assessment in CRC remains a crucial step for prognostic evaluation and treatment decision. There are several staging systems. The Astler-Coller, the Duke's, and the Tumor, Node and Metastases (TNM) system, being the latter the most commonly used, and the recommended by the American Joint Committee on Cancer (AJCC) (Centelles, 2012; Shia *et al*., 2012). Specifically, this staging describes the mural depth of the primary tumor (T), the presence of local lymph node metastases (N), and the presence of distant metastases (M) (table 1).

Table 1 - Overview of the AJCC TNM cancer staging system for colorectal carcinomas (adapted from Shia *et al*., 2012).

	Primary tumor (T)
TX	Primary tumor cannot be assessed
T0	No evidence of primary tumor
Tis	Carcinoma *in situ*: intraepithelial or invasion of lamina propria
T1	Tumor invades submucosa
T2	Tumor invades muscularis propria
T3	Tumor invades muscularis propria into pericolorectal tissues
T4a	Tumor penetrates to the surface of the visceral peritoneum
T4b	Tumor directly invades or is adherent to other organs or structures
	Regional lymph nodes (N)
NX	Regional lymph nodes cannot be assessed
N0	No regional lymph node metastasis
N1	Metastasis in 1-3 regional lymph nodes
N1a	Metastasis in 1 regional lymph node
N1b	Metastasis in 2-3 regional lymph nodes
N1c	Tumor deposit(s) in the subserosa, mesentery, or non-peritonealized pericolic or perirectal tissues without regional nodal metastasis
N2	Metastasis in 4 or more regional lymph nodes
N2a	Metastasis in 4-6 or more regional lymph nodes
N2b	Metastasis in 7 or more regional lymph nodes
	Distant metastasis (M)
M0	No distant metastasis (no pathologic M0; use clinical M to complete stage group)
M1	Distant metastasis
M1a	Metastasis confined to 1 organ or site (eg liver, lung, ovary, non-regional node)
M1b	Metastases in more than 1 organ/site or the peritoneum

TNM- Tumor, Node and Metastases.

The three letters T, N and M are combined with numbers in order to indicate increasing severity and progression of the disease (Edge and Compton, 2010). After combining the information of each letter, the stage can be assessed in a process called stage grouping, in which the stage is expressed in Roman numerals from I to IV (table 2) (Centelles, 2012).

Table 2 - Staging groups for CRC, according to AJCC guidelines (adapted from Centelles, 2012).

	T	N	M
Stage 0	Tis	N0	M0
Stage I	T1-T2	N0	M0
Stage IIA	T3	N0	M0
Stage IIB	T4a	N0	M0
Stage IIC	T4b	N0	M0
Stage IIIA	T1-T2 T1	N1 N2a	M0 M0
Stage IIIB	T3-T4 T2-T3 T1-T2	N1 N2a N2b	M0 M0 M0
Stage IIIC	T4a T3-T4 T4b	N2a N2b N1-N2	M0 M0 M0
Stage IV	Any T Any T	Any N Any N	M1a M1b

AJCC, American Joint Committee on Cancer; **CRC**, Colorectal Cancer.

6. Prognosis and treatment

The prognosis of CRC is clearly related to the degree of tumor invasion through the bowel wall and the presence or absence of nodal involvement. Important parameters include grading, lymphatic, venous or perineural invasion, lymphoid inflammatory response, and involvement of resection margins. Many other possibly prognostic factors, such as proliferation index and aneuploidy, are under evaluation for their single or combined value in high risk conditions. Bowel obstruction and perforation are clinical indicators of a poor prognosis. Elevated pretreatment serum levels of CEA and/or carbohydrate antigen 19-9 (CA19-9) have a negative prognostic significance (Labianca *et al*., 2010). The most widely studied marker CEA may be useful in the preoperative staging and postoperative follow-up of patients with large bowel cancer, but has a low predictive value for diagnosis in asymptomatic patients due to its relatively low sensitivity and specificity (Schmoll *et al*., 2012).

Relatively to treatment, surgery is the most common treatment for locoregional colon and rectal carcinoma (Siegel *et al*., 2013). According to the Nacional Cancer Comprehensive Network (NCCN), in colon cancer adjuvant chemotherapy is administered to reduce the risk of recurrence, in particular 5-fluorouracil (5-FU) regimens in combination with leucoverin and oxaliplatin (FOLFOX) are the most common choice for stage III and high-risk stage II patients. In rectal cancer, neoadjuvant combined-modality therapy, including chemotherapy alone or in combination with radiation therapy, is often given to

patients with late-stage disease before or after surgery to reduce local, and distant recurrence. Postoperative oxaliplatin-containing regimens such as FOLFOX are typically used in the setting of rectal cancer postoperatively (NCCN guidelines, 2015).

Regarding the metastatic CRC (mCRC) systemic treatment, the chemotherapy options currently used are the FOLFOX regimen, the FOLFIRI regimen (leucovorin, 5-FU and irinotecan) and the CAPOX regimen (capecitabine and oxaliplatin), which have shown a comparable activity and efficacy as first-line treatments for mCRC, but have a different toxicity profile (Nallapareddy, 2011; Wolpin and Mayer, 2008).

Moreover, there are three targeted monoclonal antibody therapies approved by the United States Food and Drug Administration (FDA) to treat patients with mCRC that can be used in combination or as single agents. This monoclonal antibodies are agents that act in two central cellular molecules: the vascular endothelial growth factor A (VEGF-A), targeted by bevacizumab, or the epidermal growth factor receptor (*EGFR*), targeted by cetuximab or panitumumab (Siegel *et al*., 2013; Nallapareddy, 2011). Their benefit was rapidly seen in combination with standard chemotherapy (Douillard *et al*., 2010; Saltz *et al*., 2008; Van Cutsem *et al.,* 2009).

7. Colorectal carcinogenesis

The model of colorectal carcinogenesis by Fearon and Vogelstein in 1990 was the first multistep genetic model described based on the observation of specific genetic changes in benign and malignant lesions (Fearon and Vogelstein, 1990). Some of the key features of this model is a result of four premises: (a) mutational activation of oncogenes along with mutational inactivation of key tumor suppressor genes plays a critical role in the development of CRC; (b) mutations in at least four to five genes are required for tumor formation to occur; (c) the total accumulation of genetic mutations as opposed to their specific order with respect to one another is the more critical event; and (d) mutant tumor suppressor genes have been shown to exert a biologic effect even when present in the heterozygote state (Chu, 2008). Following this, the genetic mutations that are necessary for the initiation and tumor progression in CRC occur among a variety of genes, such as adenomatous polyposis coli (*APC*), kristen rat sarcoma viral oncogene homolog (*KRAS*) and tumor protein p53 (*TP53*) (Ilyas *et al*., 1999). Only a fraction of adenomas progress to cancer, and progression probably occurs over years to decades. For instance, adenomas roughly 1cm in size may have an approximately 10 to 15% chance of progressing to carcinoma over a 5-20 year period (figure 5) (Berg and Søreide, 2011; Fearon, 2011; Fearon and Vogelstein, 1990; Kanthan *et al*., 2012; Markowitz and Bertagnolli, 2009; Søreide *et al*., 2009; Worthley *et al*., 2007).

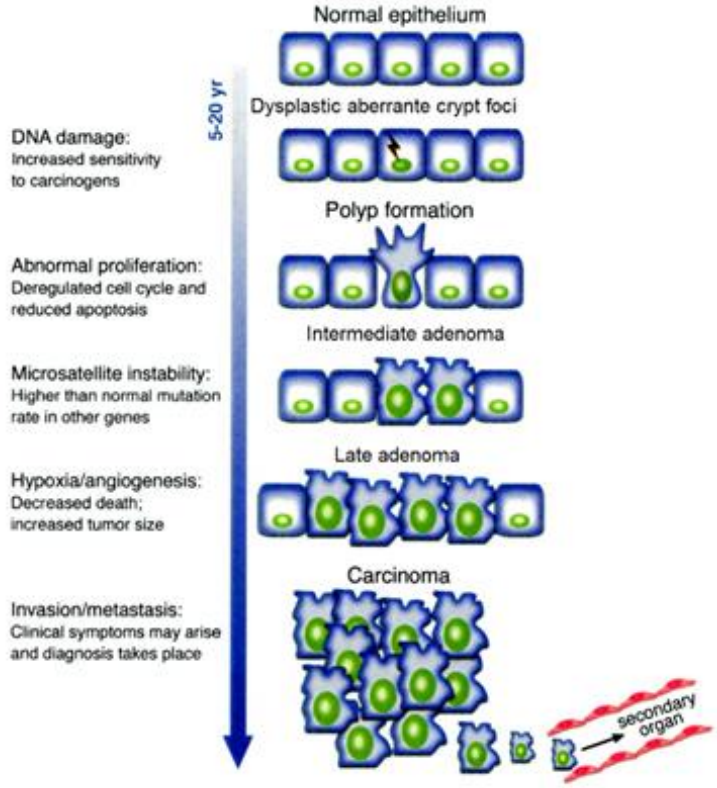


Figure 5 - Genetic model for colorectal tumorigenesis. The accumulation of mutations in oncogene and tumor suppressor genes, which is thought to contribute to tumor progression and the propensity to metastasize (adapted from Brown and DuBois, 2005).

Since the appearance of the model proposed by Fearon and Vogelstein in 1990 at least three distinct, but not mutually exclusive pathways have been described: chromosomal instability (CIN), microsatellite instability (MSI) and CpG island methylator phenotype (CIMP) (Al-Sohaily *et al*., 2012; Kanthan *et al*., 2012; Worthley *et al*., 2007).

7.1. Chromosomal instability (CIN) pathway

Chromosomal instability (CIN) is the most common type of genomic instability observed, accounts for 70 to 85% of sporadic CRC (Markowitz and Bertagnolli, 2009; Perea *et al*., 2011). These tumors commonly are aneuploid, present chromosomal amplifications, high frequency of loss of heterozygosity (LOH), mutations in tumor suppressor genes and proto-oncogenes, such as *APC, TP53,* and *KRAS,* and frequent allelic loss at 18q. Causes leading to CIN are still unknown, but it was suggested that these may be related to defects in chromosome segregation, telomere dysfunction or defects in the DNA repair mechanisms. Some of the main karyotypic alterations present in the CIN pathway are gains on chromosomes 7, 8q and 13q, deletions on chromosomes 1,

4, 5, 8q, 18q and 17p, which contains the *TP53* gene, and focal gains or losses in regions containing important cancer genes, such as vascular endothelial growth factor (*VEGF*), v-myc avian myelocytomatosis viral oncogene homolog (*MYC*) and phosphatase and tensin homolog (*PTEN*) (Markowitz and Bertagnolli, 2009; Sheffer *et al*., 2009; Worthley *et al*., 2007).

APC mutations are an early event in CRC carcinogenesis and mutation frequency in sporadic adenomas varies from 30-70%, whereas in colorectal carcinomas it varies from 60-80% (Worthley *et al*., 2007). This gene is described as a "gatekeeper" of cellular proliferation in CRC and is associated with both sporadic CIN CRCs and FAP (Al-Sohaily *et al.*, 2012; Colussi *et al.*, 2013; Ilyas *et al.*, 1999; Worthley *et al.*, 2007). This gene is involved in the Wingless-Type MMTV Integration Site Family (Wnt) signaling pathway and it is an important component of a degradation complex responsible for regulating β-catenin levels. This is a multifunctional protein with important roles in Wnt signaling pathway: intercellular adhesion, cytoskeleton stabilization, cell cycle regulation, and apoptosis (Armaghany *et al*., 2012; Coppedè *et al*., 2014). Other genetic alterations involved in β-catenin regulation include gain-of-function mutations in its coding gene, the cadherin-associated protein beta 1 (*CTNNB1*), present in up to 50% of tumors lacking an *APC* mutation (Colussi *et al.*, 2013; Kanthan *et al.*, 2012; Pino and Chung, 2010).

KRAS proto-oncogene encodes a signal transduction protein, which in its active state forms a complex with a guanosine triphosphate (GTP) group. This complex is inactivated by hydrolysis of GTP to guanosine diphosphate (GDP). The *KRAS* encodes a 21 kDa protein involved in the G-protein signal transduction pathway, modulating cellular proliferation and differentiation. The frequency of mutations in the *KRAS* proto-oncogene in sporadic CRC is 30 to 50% (Calistri *et al*., 2005; Deschoolmeester *et al*., 2010). The most common mutations found in CRC are in exon 2 and to a lesser extent in exon 3 (Calistri *et al*., 2005). If *KRAS* is mutated, the resulting complex is less sensitive to hydrolysis, remaining in a constitutively active state, leading to cell proliferation by a variety of signaling pathways, including the mitogen-activated protein kinases (MAPKs) pathway (Takayama *et al*., 2006).

Another important event in CRC carcinogenesis is the occurrence of 18q deletion in 50-70% of CRC cases (Armaghany *et al*., 2012). In this chromosomal region several important genes are located, such as SMAD family member 2 (*SMAD2*), SMAD family member 4 (*SMAD4*) and DCC netrin 1 receptor (*DCC*). The *SMAD2* and the *SMAD4* are transcription factors involved in the transforming growth factor β (TGF-β) signaling pathway, with importance in cell proliferation, differentiation, migration and apoptosis. The *DCC* gene encodes a transmembrane receptor of the immunoglobulin superfamily that promotes apoptosis when its ligand, Netrin-1, is absent (Al-Sohaily *et al*., 2012).

Additionally, *TP53* gene is frequently mutated in CRC, and mutations in this gene are associated with the transition between late adenoma to carcinoma. The frequency of mutation in this gene in CRC is 35-75% (Naccarati *et al*., 2012). *TP53* encodes a transcription factor with tumor suppressor activity, which activates a number of genes involved in cell cycle arrest, senescence, autophagy and apoptosis. When genetic damage occurs, DNA repair genes are activated (Armaghany *et al*., 2012; Iacopetta, 2003). However, when the damage cannot be repaired, *TP53* induces several pro-apoptotic genes leading to cell death. Due to that characteristic, *TP53* is often called "guardian of the genome" (Arends, 2013; Mills, 2005) (figure 6).

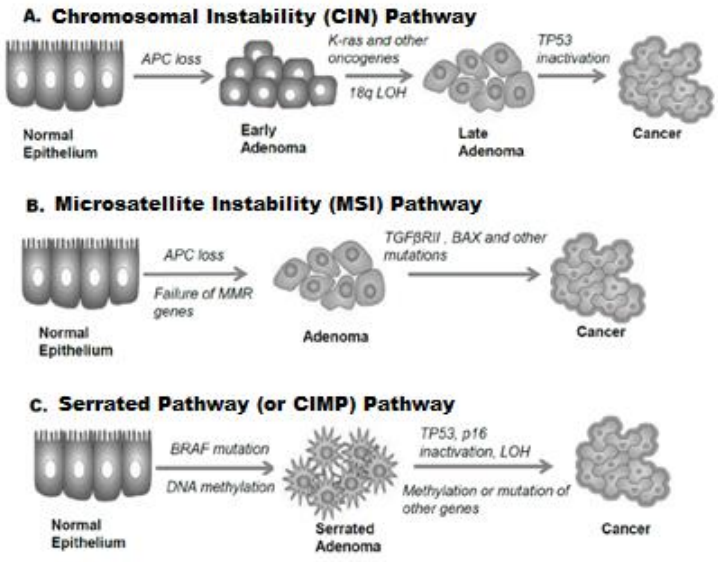


Figure 6 - Different genetic pathways in colorectal cancer pathogenesis. Three distinct parallel pathways are implicated in CRC pathogenesis: (A) Chromosomal instability (CIN) pathway driven by inactivating mutations in tumor suppressor genes and activating mutations in proto-oncogenes, (B) Microsatellite instability (MSI) pathway characterized by inactivation of the mismatch repair (MMR) genes. The failure of MMR genes resulting in mutations in specific target genes involved in proliferation and cellular differentiation, and (C) Serrated pathway (or CIMP) driven by promoter hypermethylation of several genes and presenting *BRAF* mutation (adapted from Mundade *et al*., 2014).

7.2. Microsatellite instability (MSI) pathway

The microsatellite instability (MSI) or "mutator" pathway is the other main mechanism for genomic instability in CRC (Worthley *et al*., 2007). Spread throughout the genome, there are short repeat nucleotide sequences that are prone to errors during

replication, due to its repetitive nature. These sequences, named microsatellites, are present in both protein-coding and non-coding regions of the DNA. Those errors are recognized by the DNA Mismatch Repair (MMR) system, which is then responsible for repairing all base-base mismatches and ensuring a correct DNA synthesis during replication (Al-Sohaily *et al*., 2012; Kanthan *et al*., 2012). The MMR system is composed of multiple interacting proteins, such as mutS homolog 2 (MSH2), mutS homolog 3 (MSH3), mutS homolog 6 (MSH6), mutL homolog 1 (MLH1), PMS1 homolog 1 (PMS1) and PMS1 homolog 2 (PMS2). Mutations in the genes encoding these proteins lead to the inactivation of the MMR system, which fails to repair these errors, causing their accumulation in repetitive sequences resulting in a widespread microsatellite instability (MSI) (Al-Sohaily *et al*., 2012; Ilyas *et al*.,1999; Kanthan *et al*.,2012; Markowitz and Bertagnolli, 2009). In order to assess the MSI tumor status, the Bethesda panel (BAT25, BAT26, D5S346, D2S123 and D17S250) was created to classify tumors, these can be classified as MSI-high (MSI-H, ≥30%) or MSI-low (MSI-L, <30%), and microsatellite stable (MSS, 0%) (Boland *et al*., 1998).

Besides being the hallmark of LS, the MSI pathway is also involved in the genesis of approximately 15% of sporadic CRC cases and is mostly caused by epigenetic silencing of the *MLH1* gene promoter (Al-Sohaily *et al*., 2012; Bogaert and Prenen, 2014; Colussi *et al*., 2013; Kanthan *et al*., 2012). Tumors that develop through this pathway are more likely to arise in the proximal colon, are poorly differentiated and exhibit lymphocytic infiltrations and confer a better prognosis. The MSI is characterized by accumulation of mutations in microsatellite sequences that may occur in coding regions of several genes, such as transforming growth factor, beta receptor II (*TGFBR2*) and BCL2-associated X protein (*BAX*). *TGFBR2* mutations are found in more than 80% of those cases, and inactivating mutations in this gene are involved in the adenoma transition to high-grade dysplasia or invasive carcinoma (Colussi *et al*., 2013; Kanthan *et al*., 2012) (figure 6).

7.3. CpG island methylator phenotype (CIMP) pathway

CpG island methylator phenotype (CIMP) pathway (also named serrated pathway) occurs in approximately 20 to 30% of CRC and it was reported that clinical features of CIMP CRCs are similar to those associated with MSI (Mundade *et al*., 2014). This pathway consists of the aberrant hypermethylation of the CpG dinucleotide sequences. This specific sites, the CpG islands, are regions containing high levels of cytosine-guanine pairs with phosphate bonds present in the genome (Perea *et al*., 2011). These regions are localized in the promoter regions of genes involved in several functions, such as cell cycle regulation, angiogenesis, DNA repair, invasion, adhesion or even apoptosis (Colussi *et*

al., 2013; Wong *et al*., 2007). In normal conditions, these CpG sites are unmethylated, and when methylation occurs it may inhibit gene expression and result in gene inactivation. Specific promoter methylation occurs, physiologically, to silence particular genes, while decontrolled methylation may occur pathologically as an important step in carcinogenesis. In fact, the epigenetic silencing of a gene by CpG island methylation is considered biologically equivalent to acquiring an inactivating mutation and it may occur as the first, second, or both hits in silencing tumor suppressor genes (Worthley *et al*., 2007). In CRC, the influence of epigenetics is seen both by global hypomethylation of the genome and by hypermethylation of the promoter region of specific genes. This provides an alternative mechanism for loss of function of tumor suppressor genes, such as *p16*, *APC, MLH1* and O(6)-methylguanine-DNA methyltransferase (*MGMT*) (Kanthan *et al*., 2012).

B-Raf proto-oncogene, serine/threonine kinase (*BRAF*), a member of the RAF kinase family, is a serine/threonine-specific protein kinase that plays a key role in regulating the mitogen-activated protein kinase/elk-related tyrosine kinase (MAPK/ERKs) signaling pathway, which is important to cell division, differentiation and secretion. The point mutation p.Val600Glu in *BRAF* (also described as V600E), leads to the constitutive activation of this kinase and to its insensitivity to negative feedback mechanisms, conducting to enhanced of the MAPK/ERK signaling. This overactive signaling cascade reaches cellular DNA within the nucleus and triggers downstream effectors to induce uncontrolled cell proliferation, evasion of immune response, angiogenesis, as well as resistance to apoptosis. This *BRAF* mutation is present in most CRCs with CIMP phenotype and those who do not have it may have a mutation in *KRAS* (Berg and Søreide, 2011; Markowitz and Bertagnolli, 2009; Worthley *et al*., 2007).

It is possible to classify these tumors as CIMP-high (CIMP-H) or CIMP-low (CIMP-L), based on the number of methylated markers, being the most common used the calcium channel, voltage-dependent, T type, alpha 1G subunit (*CACNA1G*), the insulin-like growth factor 2 (*IGF2*), the neurogenin 1 (*NEUROG1*), the runt-related transcription factor 3 (*RUNX3*), and the suppressor of cytokine signaling 1 (*SOCS1*) (Al-Sohaily *et al*., 2012; Colussi *et al*., 2013; Weisenberger *et al*., 2006). Most CIMP-H CRCs contain mutations in the *BRAF* gene, and usually are located in the proximal colon and confer poor prognosis (Bogaert and Prenen, 2014; Worthley *et al*., 2007). On the other hand, CIMP-L tumors are usually associated with *KRAS* mutation and *MGMT* methylation (Bogaert and Prenen, 2014;Colussi *et al*., 2013; Worthley *et al*., 2007) (figure 6).

8. Hereditary colorectal cancer

Colorectal cancer usually occurs in one of three patterns: inherited, familial and sporadic (Lynch and Shaw, 2013). The vast majority of CRCs are “sporadic” accounting for 60% of all CRC and comprising patients with no notable family history and, by definition, with no identifiable germline gene mutation. Sporadic cancers are caused by a series of genetic abnormalities in tumor suppressor genes and oncogenes that give cells an evolutionary advantage over their neighbors and arise at a median age of 70-75 years (figure 7) (Kheirelseid *et al*., 2013). Familial CRC occurs in an estimated 30% of the cases and refers to patients who have at least one blood relative with CRC or an adenoma, but with no specific germline mutation or clear pattern of inheritance (figure 7) (Kheirelseid *et al*., 2013).

Hereditary CRC syndromes are responsible for about 10% of all CRC cancers, which result from germline inheritance of mutations in highly penetrant cancer susceptibility genes (figure 7) (Jasperson *et al.*, 2010; Rustgi, 2007). The most common hereditary CRC syndromes identified so far are Lynch syndrome (LS), familial adenomatous polyposis (FAP), MUTYH associated polyposis (MAP), and the hamartomatous polyposis syndromes Peutz-Jeghers syndrome (PJS), juvenile polyposis syndrome (JPS) and Cowden syndrome (Al-Sohaily *et al*., 2012; Lynch and Shaw, 2013; Zbuk and Eng, 2007).

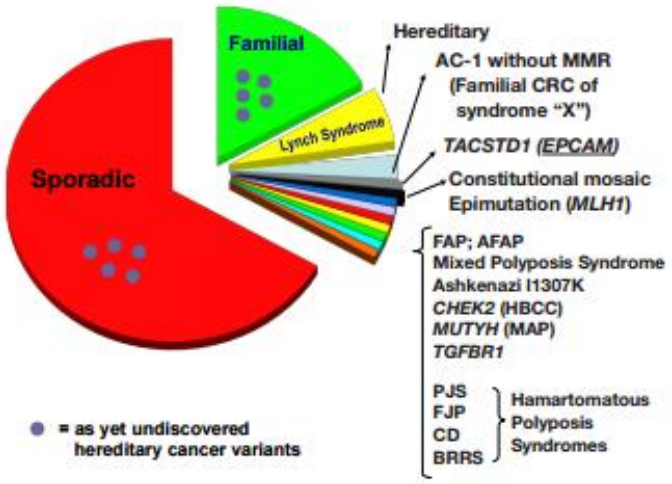


Figure 7 - Circle graph depicting the marked genotypic and phenotypic heterogeneity in hereditary colorectal cancer syndromes. AC-1, Amsterdam Criteria 1; **MMR**, mismatch repair; **FAP**, familial adenomatous polyposis; **AFAP**, attenuated familial adenomatous polyposis; **HBCC**, hereditary breast and colorectal cancer; **PJS,** Peutz-Jeghers syndrome; **FJP**, familial juvenile polyposis; **CD**, Cowden’s disease; **BRRS**, Bannayan-Ruvalcaba-Riley syndrome (adapted from Lynch and Shaw, 2013).

8.1. Lynch syndrome

Lynch syndrome is an autosomal dominant disease caused by heterozygous loss-of-function germline mutations in DNA MMR genes, most frequently in *MLH1* or *MSH2,* which account for approximately 90% of the all mutations. Mutations in *MSH6* and *PMS2* have collectively been associated with a small percentage of LS cases (Hitchins *et al*., 2005; Hitchins *et al*., 2007). However, the tumors in LS only occur after somatic biallelic gene inactivation resulting in total loss of DNA MMR activity (Giardiello *et al*., 2014). This syndrome is recognized as the most common hereditary CRC, accounting for nearly 2-5% of all colorectal malignancies (Lynch *et al*., 2015). The lifetime risk of CRC in patients with the recognized LS-related mutations may be as high as 70 to 80% (Gryfe, 2009; Pande *et al*., 2012).

Patients with LS develop CRC at a younger age than the general population (~45 *vs.* 65 years), develop predominantly proximal colon cancers, and are at a higher risk for synchronous CRCs and extra-colonic tumors (including endometrial, ovarian, gastric, small bowel, pancreatic, hepatobiliary, skin, brain, and urethral tumors) (Pande *et al*., 2012). A genotype-phenotype correlation has been observed in which *MLH1* mutation carriers are at higher risk of young onset CRC cancer, *MSH2* at higher risk of extracolonic cancers, *MSH6* at increased risk of endometrial cancer, and *PMS2* carriers show a lower absolute lifetime risk of CRC and endometrial cancer (15-20%) compared with other mutation carriers (Balmaña *et al*., 2013).

Table 3 - Revised minimum criteria for clinical definition of Lynch syndrome.

Amsterdam Criteria II (adapted from Vasen *et al*., 2007)
At least three relatives must have a cancer associated with LS*, all of the following criteria should be present:
One must be a first degree relative of the other two;
At least two successive generations must be affected;
At least one relative with cancer associated with LS should be diagnosed before age 50 years;
FAP should be excluded in the CRC cases(s) (in any);
Tumors should be verified by histopathological examination.
Bethesda Criteria (Revised) (adapted from Umar *et al*., 2004)
CRC diagnosed in a patient aged <50 years;
Presence of synchronous, metachronous or other Lynch syndrome-related tumors, regardless of age;
CRC with MSI-H phenotype diagnosed in a patient aged <60 years, with specific pathological features;
Patient with CRC and a first-degree relative with a Lynch syndrome-related tumor[#], with one of the cancers diagnosed at age <50 years;
Patient with CRC with two or more first degree or second-degree relatives with a Lynch syndrome-related tumor regardless of age.

* Lynch syndrome-associated cancer includes those of endometrium, small bowel, ureter or renal pelvis.
[#] Lynch syndrome-related tumors include colorectal, endometrial, stomach, ovarian, pancreas, ureter, renal pelvis, biliary tract and brain tumors, sebaceous gland adenomas and carcinoma of the small bowel.

The clinical diagnosis of LS is based on the Amsterdam or on the Bethesda criteria (table 3) (Guidelines NCCN, 2015). The Amsterdam criteria are very specific, but too stringent (low sensitivity) and there are several studies that have shown that 40% of LS families with an identified gene mutation did not meet Amsterdam criteria (Kastrinos and Syngal, 2011). Bethesda guidelines were developed to improve the sensitivity, although they have a lower specificity and require a prescreening by tumor molecular testing either looking for MSI or abnormal immunohistochemistry (IHC) to detect altered expression of MMR proteins (Umar *et al*., 2004).

8.2. The MMR system

Lynch syndrome is characterized by germline mutations in the MMR genes. The human MMR system recognize errors in the DNA nucleotide sequence by matching the complementary chromosome strands and repairs base-base mismatches that occur during DNA replication in proliferating cells. There are several human MMR proteins, including human homologs of MutS, MutL, exonuclease 1 (EXO1), single-strand DNA-binding protein replication protein A (RPA), proliferating cellular nuclear antigen (PCNA), DNA polymerase δ (POLD)/DNA polymerase ε (POLE) and DNA ligase I (figure 8) (Jun *et al*., 2006). Human MutS and MutL homologues are heterodimers and MSH2 heterodimerizes with MSH6 or MSH3 to form MutSα or MutSβ, respectively, both of which are ATPases that play a critical role in mismatch recognition and initiation of repair. The MutSα preferentially recognizes base-base mismatches and insertion/deletion (ID) mispairs of one or two nucleotides, while MutSβ preferentially recognizes larger ID mispairs. The MLH1 heterodimerizes with PMS2, PMS1, or MLH3 to form MutLα, MutLβ, or MutLγ, respectively. The MutLα is required for MMR activity and the MutLγ plays a role in meiosis, but no specific biological role has been identified for MutLβ. The MutLα possesses an ATPase activity and defects in this activity inactivate the MMR system in human cells. In a reconstituted human MMR system, MutLα regulates termination of mismatch-provoked excision. PCNA interacts with MSH2 and MLH1 and is thought to play important roles in the initiation and DNA resynthesis steps of MMR. PCNA also interacts with MSH6 and MSH3 via a conserved PCNA interaction motif termed the PIP box. RFC that is bound at the 5′ terminus of the discontinuity prevents degradation in the 5′→3′ direction (away from the mismatch). Once the mismatch is removed and the EXO1 activity is inhibited by bounding to RPA and MutLα, the gap is filled by Pol δ. DNA ligase I seals the remaining nick to complete the repair process (Rasmussen *et al*., 2012) (figure 8).

Cells with biallelic mismatch repair gene mutations cannot repair spontaneous DNA errors and progressively accumulate mutations with succeeding DNA replications throughout the genome, resulting in genetic instability (Jiricny, 2006; Martín-López and Fishel, 2013). Development of mismatch repair-deficient cancers exhibiting microsatellite instability (MSI) occurs with the loss-of-function of the remaining normal allele of the affected gene within somatic tissues (Crépin *et al*., 2012).

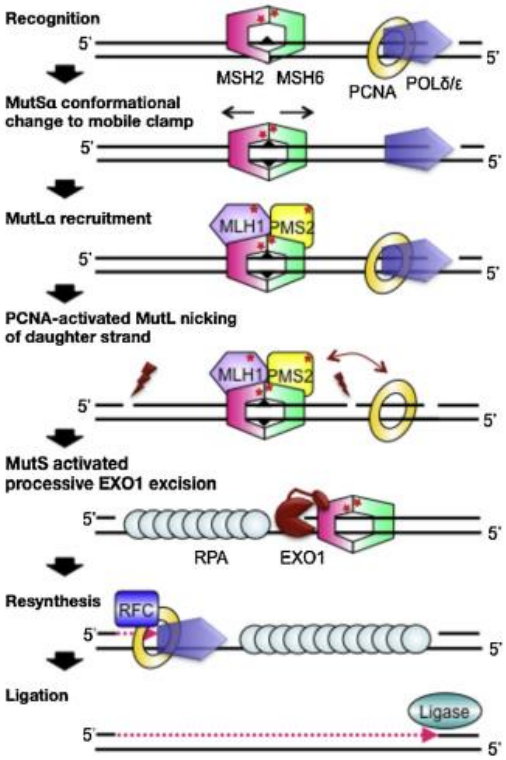


Figure 8 - Steps in MMR system. The replicative DNA polymerase misincorporates a nucleotide during DNA replication. The MSH2-MSH6 heterodimer recognizes and binds to the mismatch, followed by mismatch validation by ADP → ATP exchange. This complex recruits the MLH1-PMS2 heterodimer that, in turn, exchanges ADP for ATP. The latent endonuclease activity of PMS2 introduces a nick in the daughter strand, 5' of the misincorporation. The nick serves as entry point for the exonuclease 1 (EXO1) that degrades a patch of the daughter strand that includes the misincorporation. The remaining single-stranded DNA gap is covered by the single-strand DNA binding protein replication protein A (RPA) and filled by the replicative DNA polymerase (adapted from Rasmussen *et al*., 2012).

9. The role of epigenetics in Lynch syndrome

In the past decade, another distinct mechanism affecting the two key MMR genes *MLH1* and *MSH2*, was unraveled in a subset of patients meeting the clinical criteria for LS without a germline MMR mutation, termed as ‘‘constitutional epimutation’’, or just ‘‘epimutation” (Castillejo *et al*., 2015; Hitchins and Ward, 2009). In this mechanism, the expression of either of these two genes is interrupted within normal tissues via promoter constitutional methylation instead of a direct genetic alteration (Hitchins, 2013; Lynch *et al*., 2015). These gene expression changes are brought about by the addition of various biochemical modifications to the DNA backbone, which include methylation of cytosine

bases within cytosine-guanine dinucleotides (CpG), modifications to the histone core of the nucleosomes, as well as the positioning of nucleosomes along the DNA sequence. Methylation of clusters of CpG sites spanning the gene promoter (CpG islands) will be the sole chromatin modification referred to since this is the most easily and frequently studied of all epigenetic marks, and is a clear characteristic of a gene that has been epigenetically silenced (Dawson and Kouzarides, 2012; Herceg and Vaissière, 2011).

Constitutional epimutations of the *MSH2* gene are secondary to germline deletions in the *EPCAM* gene in *cis*, being transmitted in an autosomal dominant inheritance (Ligtenberg *et al*., 2009). However, constitutional epimutations of the *MLH1* gene are more variable, and the pattern of transmission of these distinct forms of *MLH1* epimutation presumably reflects their mechanistic basis (Castillejo *et al*., 2015; Hitchins, 2013).

9.1. Constitutional epimutation of *MLH1*

The *MLH1* gene is located on the short arm of chromosome 3 at position p21.3. The *MLH1* gene is 72,558 bases in length and consists of 19 coding exons; the translated protein contains 756 amino acids. The protein MLH1 dimerizes mainly with the protein product of the PMS2 gene to coordinate the binding of other proteins involved in MMR system, as referred previously (Hegde *et al*., 2014).

The first case to be identified with constitutional epimutation of *MLH1* was described by Gazzoli and coworkers in 2002. They recognized the presence of dense methylation of a single allele of the *MLH1* promoter in the peripheral blood lymphocytes (PBL) of a patient who had developed a CRC demonstrating MSI and MLH1 loss at age of 25 years (Gazzoli *et al*., 2002; Vasen *et al*., 2013). Since *MLH1* inactivation follows Knudson's two-hit-model, the observed loss of heterozygosity of the unmethylated allele in the tumor of this patient provided the first signal that these epigenetic aberrations in normal tissues could predispose to LS (Gazzoli *et al*., 2002; Hitchins, 2013).

Although these epigenetic silencing mechanisms are quite distinct from germline sequence mutations, they also confer an elevated risk of developing mismatch repair deficient tumors at a young age of onset (Lynch *et al*., 2015). They therefore represent an alternative etiological mechanism to genetic mutation for this cancer predisposition syndrome (Hitchins, 2010). The *MLH1* epimutation may be dichotomized into two categories: (1) those that tend to arise spontaneously and are reversible between generations, though occasionally transmitted to the next generation in a non-Mendelian pattern (primary *MLH1* epimutation); and (2) Mendelian epimutations that follow a classic autosomal dominant inheritance pattern due to an underlying *cis*-genetic cause (secondary/genetically facilitated *MLH1* epimutation).

Since a single parental allele is affected, it has been proposed that *MLH1* epimutations may originate in the germline. Certainly the demonstration that monoallelic promoter methylation is present in tissues derived from all three embryonic germ cell lineages strongly suggests that *MLH1* epimutations are already present in the early embryo, prior to the differentiation of the germ cell layers (Goel *et al*., 2011).

According to the study of Hitchins and coworkers, the mature spermatozoa, from four *MLH1* epimutation carriers with apparent soma-wide epimutation, showed absence of this alteration (Hitchins, 2013). Interestingly, in the majority of sporadic cases studied, in whom the epimutation has arisen *de novo* in the proband, *MLH1* epimutation tended to occur on the maternal *MLH1* allele. Therefore, it remains plausible that some *MLH1* epimutations occur in the oocyte. The precise timing and cellular origin of this defect remains to be determined and may even differ from one case to another. On the other hand, a certain degree of allelic epigenetic mosaicism has been observed in most carriers, since the affected allele remains unmethylated or partially methylated in a proportion of somatic cells, whilst other copies are fully methylated (Seisenberger *et al*., 2012; Kwok *et al*., 2014; Pineda *et al*., 2012).

9.2. Importance of *MLH1* epimutations in clinical practice

Constitutional *MLH1* epimutation are infrequent in comparison to the incidence of germline MMR mutations and with sporadic CRC demonstrating MSI. Nevertheless, this is an etiological mechanism that confers a high risk of development LS-type cancers in carriers (Lynch *et al*., 2015). Given that molecular diagnosis is important to identify these carriers and to provide (and to their family members) an appropriate clinical surveillance (Hitchins, 2013), it is important not only to identify an epimutation as the cause for cancer susceptibility, but also to define the type of epimutation (that is its mechanistic basis), since this will dictate its risk of intergenerational inheritance (Goel *et al*., 2011). It is therefore opportune that screening for this etiological mechanism is implemented on a molecular diagnostic routine basis. It is important to establish a consensus selection, and an efficient diagnostic assay. Testing for *MLH1* epimutation on a triaged basis by incorporating this into existing algorithms for patient inclusion and molecular detection of constitutive MMR defects might be improved (Banno *et al*., 2012; Hitchins, 2013).

II. AIMS OF THE STUDY

II. AIMS OF THE STUDY

There are several families that fulfill the clinical criteria for Lynch syndrome and do not present any MMR germline mutation. Given that, it is important to characterize these families.

Essentially, the specific aims of this work were:

- To investigate the prevalence of *MLH1* constitutional methylation in a series of CRC patients with no expression of the MLH1 protein in the tumor and with no *MLH1* germline mutation;
- Deduce the pattern of constitutional methylation inheritance in Lynch syndrome families;
- To establish the correlation between *MLH1* constitutional epimutation and family history and tumor phenotype.

III. MATERIALS AND METHODS

III. MATERIALS AND METHODS

1. Patients and samples collection

A consecutive series of PBL samples from 38 patients (index cases) 17 males (44.7%) and 21 females (55.3%) fulfilling the clinical criteria for LS with colorectal carcinomas where included. All patients presented loss of MLH1 protein expression and had no *MLH1* pathogenic germline mutation identified, this is shown in table 4. These patients were diagnosed and surgically treated at the Portuguese Institute of Oncology-Porto assessed through Genetic Cancer Counselling, and referred to the Genetics Department between 1997 and 2014. Three of these patients (3/38; 7.9%) presented the p.Val600Glu *BRAF* mutation in their tumors. The majority of these patients (37/38; 97.4%) met the Bethesda criteria and one patient (1/38; 2.6%) met the Amsterdam criteria. Clinico-pathological information was obtained from medical records (table 4).

Whenever possible, family members of the index patients were also studied, and, when available, swab buccal samples and paraffin embedded tissue samples (with different germ layers origins) were also analyzed in patients harboring constitutional epimutation.

Table 4 - Clinico-pathological data of 38 index cases fulfilling the clinical criteria for Lynch syndrome.

Patient	Gender	Tumor localization (diagnosis age)	IHC MMR	Clinical Criteria	*BRAF*
#1	F	Ascending colon (40)	MLH1/PMS2 absence	BC	NA
#2	F	Stomach (75) Cecum (75) Breast (78)	MLH1 absence*	BC	NA
#3	M	Descending colon (38 and 48)	MLH1 absence*	BC	WT
#4	M	Ascending colon (25)	MLH1/PMS2 absence	BC	WT
#5	M	Sigmoid colon (51)	MLH1/PMS2 absence	BC	WT
#6	M	Rectum (53)	MLH1/PMS2 with decreased immunoreactivity	BC	WT
#7	M	Ascending colon (43)	MLH1/PMS2 absence	BC	WT
#8	F	Rectum (16)	MLH1 absence (normal PMS2)	BC	WT
#9	F	Rectum (43)	MLH1 absence*	BC	V600E
#10	F	Ascending colon (26) Stomach (60)	MLH1/PMS2 absence	BC	WT
#11	M	Ascending colon (65)	MLH1/PMS2 absence	BC	WT
#12	F	Ascending colon (62)	MLH1/PMS2 absence	BC	WT
#13	M	Sigmoid colon (44)	MLH1/PMS2 absence	BC	WT
#14	F	Ascending colon (69)	MLH1/PMS2 absence	BC	WT
#15	F	Uterus and ovary (38)	MLH1/PMS2 absence	BC	WT
#16	F	Breast (60) Ascending colon (66)	MLH1/PMS2 absence	BC	WT
#17	M	Ascending colon (25)	MLH1/PMS2 absence	BC	WT
#18	M	Sigmoid colon (43)	MLH1/PMS2 absence	BC	NA
#19	M	Sigmoid colon (47)	MLH1/PMS2 absence	BC	WT
#20	M	Ascending colon (23)	MLH1/PMS2 absence	BC	WT
#21	F	Endometrium (57) Ascending Colon (74)	MLH1/PMS2 absence	BC	V600E

		Lung (74)			
#22	F	Sigmoid colon (47)	MLH1/PMS2 absence	BC	WT
#23	F	Ascending colon (59)	MLH1/PMS2 absence	BC	WT
#24	M	Rectum (45) Ascending colon (61)	MLH1/PMS2 absence	BC	WT
#25	F	Ascending colon (62) Endometrium (63)	MLH1/PMS2 absence	BC	WT
#26	F	Ascending colon (41)	MLH1/PMS2 absence	BC	WT
#27	M	Rectum (33)	MLH1/PMS2 absence	BC	WT
#28	F	Stomach (78)	MLH1/PMS2 absence	BC	WT
#29	F	Breast (30)	MLH1/PMS2 absence[#]	AC	WT
#30	M	Sigmoid colon (61)	MLH1/PMS2 absence	BC	WT
#31	F	Descending colon (65)	MLH1/PMS2 absence	BC	V600E
#32	F	Ascending colon (54)	MLH1/PMS2 absence	BC	WT
#33	F	Ascending colon (42)	MLH1/PMS2 absence	BC	NA
#34	M	Ascending colon (60)	MLH1/PMS2 absence	BC	WT
#35	M	Ascending colon (44)	MLH1/PMS2 absence	BC	NA
#36	F	Endometrium (50)	MLH1/PMS2 absence	BC	WT
#37	F	Sigmoid colon (56)	MLH1/PMS2 with decreased immunoreactivity	BC	WT
#38	M	Ascending colon (48) Transverse colon (48) Descending colon (48)	MLH1/PMS2 absence	BC	WT

AC, Amsterdam criteria; **BC**, Bethesda criteria; **F**, female; **M**, male **NA,** not analyzed.

*The analysis was not performed for PMS2.

[#]IHC was performed on tumor of a relative.

2. DNA and RNA extraction from peripheral blood samples

Peripheral blood samples were collected in sterile tubes containing EDTA anticoagulant. In order to achieve the lysis of erythrocytes, a hypotonic solution was added (AKE: NH_4Cl [Merck, Darmstadt, Germany] 155 mM; $KHCO_3$ [Merck] 10 mM; EDTA [Sigma-Aldrich, Steinheim, Germany] 0.1 mM; pH=7.4) at 3-5 mL of blood in a ratio of 9 to 10 times this volume, followed by incubation at 4°C for 30 minutes. The samples were then centrifuged at 1500g [Sigma centrifuge 4K15]. The supernatant was removed and the procedure was repeated as described above until the pellet was free of hemoglobin. The cell pellet was resuspended in 1 mL of PBS [Merck], transferred to a microcentrifuge tube and centrifuged at 3000g. The supernatant was removed and the cell pellet was stored at 4°C for until DNA and/or RNA extraction.

DNA was obtained using the Magna Pure LC 2.0 [Roche Applied Science, Indianapolis, Indiana], and total RNA extraction was performed using TRIzol® Reagent [Invitrogen Carlsbad, CA, USA] according to the manufacturer's instructions and standard protocol (Chomczynski, 1993). DNA and RNA quality and concentration was evaluated using a NanoDrop ND-1000® [NanoDrop Technologies, Wilmington, DE, USA].

3. DNA extraction from formalin-fixed paraffin-embedded tissue

Tumor areas containing at least 50% of tumor cells were delimited, by a pathologist, in the hematoxylin and eosin (H&E) stained slides of each sample. The corresponding unstained slides were immersed in xylene [Sigma] and twice in ethanol 100% [Merck] for 5 minutes, each. Tumor and normal areas, which were previously delimited by comparison with the correspondent H&E stained slides, were macrodissected and transferred to a centrifuge tube. DNA was isolated using the QIAamp® DNA FFPE Tissue Kit [Qiagen, Hilden, Germany], following manufacturer's instructions. Finally, DNA was quantified by spectrophotometry with NanoDrop ND-1000®.

4. DNA extraction from buccal mucosa swabs

The buccal mucosa swabs were collected and preserved in dry medium. PBS [Merck] (1 mL) was added and, after incubation of 3 hours at 4ºC, the samples was centrifuged for 10 minutes at 3000g. After centrifugation, 4 mL of SE [Merck], 400 µL SDS [Gibco Invitrogen, Carlsbad, CA, USA] and 50 µL of proteinase K [Gibco, Invitrogen] were added, and the

mixture was incubated overnight at 55 ºC. After incubation, 1 mL of NaCl 6M [Merck] was added followed by 10 minutes of incubation at 55 ºC. One equivalent volume of chloroform was added, the mixture was homogenized during 30 minutes, and centrifuged for 10 minutes at 3000g, to separate the different phases. The upper phase was collected to a new tube and an equal volume of isopropanol [Merck] was added. DNA was isolated after 10 minutes of centrifugation at 3000g. The supernatant was discarded and 3 mL of ethanol 70% [Merck] was added to wash the DNA during incubation overnight at 4ºC. After this, the DNA was centrifuged for 10 minutes at 3000g, the supernatant was discarded and the DNA was dry out at 50ºC for about 3 hours. DNA was eluted with 200-300µL of nuclease-free water [Quiagen] and incubated for an additional 10 minutes at 50ºC. Finally, the DNA was quantified by spectrophotometry with NanoDrop ND-1000®.

5. *MLH1* methylation analysis

Analysis of *MLH1* promoter methylation was performed on PBL DNA of the index patients. Methylation testing was performed by methylation-specific multiplex ligation-dependent probe amplification (MS-MLPA) using the SALSA MS-MLPA ME011-B1 kit [MRC Holland, Amsterdam, Netherlands], according to the manufactures instructions. This method allows the detection of methylation patterns in a number of tumor suppressor genes in addition to the semi-quantification of the copy number of sequences analyzed. Some key considerations need to be taken into account when designing or adopting an assay for screening of patients for the presence of *MLH1* promoter constitutional epimutation. The molecular method employed needs to be sufficiently sensitive to detect methylation levels down to 5% or below, in order to detect cases with a mosaic form of epimutation (Hitchins and Ward, 2009; Ward *et al*., 2012). Furthermore, the assay should target specific CpG sites within the *MLH1* promoter, namely, the C- or D-Deng regions, since these have been correlated most closely with the loss of transcription and are less susceptible to age-related or other non-specific methylation (Deng *et al*., 2002).

In MS-MLPA, the ligation of MLPA probes is combined with digestion of the genomic DNA with the methylation sensitive endonuclease *HhaI* and comparing the undigested MLPA essay (Nygren *et al.*, 2005). The ME011-B1 MMR probemix has been developed to detect aberrant CpG islands methylation of MMR genes and includes 6 probes for *MLH1* that contain a digestion site specific for the methylation-sensitive *HhaI* enzyme. The six probe pairs in *MLH1* promoter (with the respective *HhaI* sites located at -659,-518, -382, -246, -13 and +206 relative to the start codon;

GenBank accession number NM_000249.3) covers independent regions: regions A to D of the promoter and intron 1 (Deng *et al.*, 2002). The most important methylation region for *MLH1* expression, the C- Deng region, is from -248 to -178nt before the transcription site. The transcription start site that Deng used for reference lies 21nt before the start codon. The second most important region, the D- Deng region, is from -9 to +15nt (Deng *et al*., 1999).

The samples were analyzed on an ABI PRISM™ 310 Genetic Analyzer [Applied Biosystems, Foster City, CA, USA]. For this propose, about 0.5 µL of the MS-MLPA reaction product was added to 15 µL of Hi-Di™ Formamide [Applied Biosystems], and to 0.4 µL of Gene Scan ROX (Carboxy-X-Rhodamine) size standard [Applied Biosystems]. Fluorescently labeled products from MS-MLPA reactions were analyzed using the Gene Mapper® software, version 3.7 [Applied Biosystems]. Electropherogram plots of each sample were analyzed manually.

Whenever possible, swab buccal samples and paraffin embedded tissues were analyzed for *MLH1* promoter methylation using the same approach.

Three independent MS-MPLA reactions were performed in all samples and the average percentage of methylation in the more important regions of the *MLH1* were calculated using Microsoft Office Excel 2003™ and according to the manufacture's protocol (Joensuu *et al*., 2008; MRC-Holand, 2012).

6. *MLH1* promoter sequencing

Screening for mutations within the *MLH1* promoter (that may affect the binding of MLPA probes) was performed by Sanger sequencing in the 38 peripheral blood samples and using two sets of primers, according to Pineda *et al* (2012) (table 5). For this purpose, DNA was amplified in a solution containing 1x *Taq reaction buffer* [Thermo Fisher Scientific, Rockford, USA] (75mM Tris-HCl, 20 mM $(NH_4)2SO_4$), 1.25 mM of $MgCl_2$ [Thermo Fisher Scientific], 0.5 mM dNTP mix [Thermo Fisher Scientific], 0.33 mM of each primer (reverse and forward) [Frilabo, Portugal], 1 U of *Taq DNA polymerase* [Thermo Fisher Scientific] and bidestilled sterile water [B. Braun, Foster City, CA, USA] in a final reaction volume of 30 µL. PCR reaction was performed in a termocycler [Perkin-Elmer, Gene Amp PCR Systern 9700, Waltham, Massachusetts, U.S] according to the following conditions: an initial denaturation step at 95°C for 10 min, followed by 35 cycles of 95°C for 45 seconds, annealing step at 58°C for 45 seconds and a 45 seconds extension step at 72°C. A final extension step was done at 72°C for 10 min. Amplified products were then analyzed by electrophoresis in a 2% (w/v) agarose gel [Gibco BRL] stained with green safe 0.05 µL/mL [Sigma].

Subsequently, the PCR products were purified using the ExoSAP-IT method for the removal of primers and dNTPs in excess. Samples were purified adding 2 µL of ExoSAP solution (Exonuclease I [Thermo Fisher Scientific] (20 U/µL) and Fast Thermosensitive Alkaline Phosphatase [Thermo Fisher Scientific] (1 U/µL), in a proportion of 1:2) to 5 µL of PCR product, followed by incubation at 37ºC for 50 minutes, and enzyme inactivation at 80ºC for 15 minutes.

The purified PCR products were sequenced using BigDye® Terminator v1.1 or v3.1 Sequencing Kits [Applied Biosystems], and according to manufactors instructions. In order to remove excess of dNTPs, labeled ddNTPs and non-incorporated primer, the sequencing products were purified with Illustra Sephadex® G-50 fine [GE Healthcare, Life Sciences, Cleveland, USA], according to standard procedure. After purification, 12 µL of Hi-Di™ Formamide [Applied Biosystems, Foster City, CA, USA] were added to the products to help stabilize the single stranded DNA (ssDNA). The products were then analyzed in either an ABI PRISM™ 310 Genetic Analyzer or a 3500 Genetic Analyzer [Applied Biosystems] by capillary electrophoresis. The electropherograms of each sample were analyzed with the Sequencing Analysis Software v5.4 [Applied Biosystems]. All of them were read at least twice, reviewed manually and with the Mutation Surveyor® DNA Variant Analysis Software v4.0.8 [Softgenetics, State College, PA, USA].

7. *MLH1* transcript quantification

The *MLH1* transcripts were quantified by semiquantitative multiplex RT-PCR. For this propose, we amplified simultaneously a control transcript, the β2-microglobulin (*B2M*), and part of the *MLH1* transcript, using flurescence-labeled primers designed in the Primer-BLAST (Basic Local Alignment Search Tool) program (table 5) and according to the QIAGEN OneStep RT-PCR Kit [Qiagen]. Semiquantitative RT-PCR products were analyzed by fragment analysis on an ABI PRISM™ 310 Genetic Analyzer using the same conditions described above for the methylation analyses. *MLH1* transcript levels were calculated by comparing the relative peak areas of the patients to the relative peak areas of the controls.

Four independent reactions of semiquantitative multiplex RT-PCR were performed in all cases with constitutional methylation and controls. It was then calculated the average percentage of the decrease of *MLH1* transcript levels.

Table 5 - Primer sequences for *MLH1* promoter sequencing and primer sequences for rs1799977 in cDNA and gDNA *MLH1* gene.

Gene	Analysis	Primer name	Forward primer (5'-3')	Reverse primer (5'-3')	Ta (º C)
MLH1					
	Promoter sequencing	*MLH1* promoter	AACCCTTTCACCATGCT CTG	TGAAGAGAGAGCTGCTC GTG	59/58
		MLH1 promoter	TACATGCTCGGGCAGTA CCT	CCTCGTGCTCACGTTC TTC	59
		g*MLH1* ex7_8	TGGGTAAAATATTAATAG GCTGTATGG (6-FAM)#	AAGGTTCCAAAATAATG TGATGG	58
	RT amplification	RT_*MLH1* ex6_9	GGAAGTTGTTGGCAGG TATTCA	CCATTCATTTTGAAGGC TAGGGT	54
	ASE (SNuPE)	*MLH1* rs1799977	GTGGACAATATTCGC TCC	ACTAACAGCATTTCCAA AGA	54
B2M		B2M	CCTCCATGATGCTGCTTA CATGTC (6-FAM)#	ATGTCTCGCTCCGTGGCC TTAGCT	58

6-FAM, 6- Carboxyfluorescein.

8. Screening of the *MLH1* c.655A>G SNP

Screening of the *MLH1* c.655A>G SNP (rs1799977), within exon 8, was performed by Sanger sequencing in DNA and RNA samples from the index patients harboring *MLH1* constitutional methylation.

Briefly, the genomic DNA (gDNA) was amplified for *MLH1* exon 8, using a PCR reaction containing 1x *Taq reaction buffer* [Thermo Fisher Scientific] (75 mM Tris-HCl, 25 mM $(NH_4)2SO_4$), 2.5 mM of $MgCl_2$ [Thermo Fisher Scientific], 0.8 mM dNTP mix [Thermo Fisher Scientific], 2 mM of each primer (reverse and forward) [frilabo], 0.75 U of *Taq Gold DNA polymerase* [Thermo Fisher Scientific] bidestilled sterile water, in a final reaction volume of 25 µL. PCR reaction was performed in a termocycler [Perkin-Elmer, Gene Amp PCR Systern 9700] according to the following conditions: an initial denaturation step at 95°C for 10 min, followed by 35 cycles of 95°C for 45 seconds, annealing step at 58°C for 45 seconds and a 45 seconds extension step at 72°C. A final extension step was done at 72°C for 10 min. The RNA was amplified using the same primers (not labeled) as for semiquantitative *MLH1* RT-PCR amplification and according to the QIAGEN OneStep RT-PCR Kit.

All PCR and RT-PCR products were then analyzed by electrophoresis in a 2% (w/v) agarose gel stained with green safe 0.05 µL/mL, purified and sequenced as described above.

9. *MLH1* allelic expression analysis

For allelic expression analyses of the rs1799977 in heterozygous patients, the relative levels of the A/G alleles were determined in genomic DNA and RNA by single-nucleotide primer extension (SNuPE). Briefly, genomic DNA and total RNA were amplified using the same approach as for rs1799977 sequencing. After amplification, and standard ExoSAP-IT purification, the SNuPE reaction and capillary electrophoresis was performed following the SNaPshot Kit [Applied Biosystems] manufacturer´s protocol.

Four independent reactions of allelic expression of the relative levels of the A/G alleles were determined in gDNA and cDNA of heterozygous patients (cases and controls) and the average percentage of the decrease in the allele relative areas were calculated.

10. Study of microsatellite instability

Microsatellite instability was performed using the Bethesda panel (BAT25, BAT26, D2S123, D5S346 and D17S250), according to the 1997 National Cancer Institute Guidelines (Boland *et al*., 1998) using fluorescence labeled primers (Dietmaier *et al*., 1997). The amplification reaction occurred in a total volume of 20 µL of a solution containing: 1x *Taq reaction buffer* [Thermo Fisher Scientific] (75mM Tris-HCl, 20 mM $(NH_4)2SO_4$), 2.5 mM $MgCl_2$ [Thermo Fisher Scientific], 0.5 mM dNTP's [Thermo Fisher Scientific] (250 µM dTTP, 250 µM dATP, 250 µM dGTP, 250 µM dCTP), 0.15 mM of both primers (forward and reverse) [Frilabo], 0.3 µL Taq DNA polymerase [Thermo Fisher Scientific] (0.6 U) and 30 to 50 ng of genomic DNA, making up to volume with bidestilled water [B. Braun]. In a thermocycler PCR reaction was performed according to the following conditions: an initial denaturation step at 95°C for 10 min, followed by 35 cycles of 95°C for 1 minute, annealing step at 55°C for 1 minute and a 1 minute extension step at 72°C. A final extension step was done at 72°C for 10 min.

Fragments were analyzed for length variations on an ABI PRISM™ 310 Genetic Analyzer DNA sequencer (as previously described) and allele sizes were determined using the Gene Mapper® software. The results were independently scored by two observers and an additional round of analyses confirmed the results.

IV. RESULTS

IV. RESULTS

1. Identification of Lynch syndrome cases harboring constitutional *MLH1* promoter methylation

The methylation status of the *MLH1* promoter was analyzed by methylation-specific multiplex ligation-dependent probe amplification (MS-MLPA) in PBL samples of 38 patients with clinical criteria for LS, with no germline *MLH1* mutations and with no MLH1 immunoexpression on their tumors. In all cases, methylation in the *MLH1* promoter was studied in the five regions analyzed, including C- and D-Deng promoter regions (figure 9), the most correlated with transcriptional regulation (Deng *et al*., 2002).

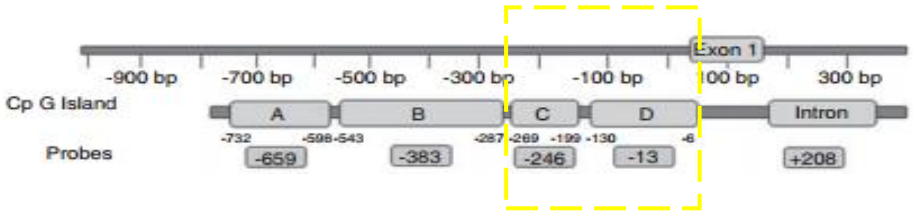


Figure 9 - Detailed methylation patterns of *MLH1* promoter, assessed by MS-MLPA. The five regions of the CpG islands targeted by the selected probes are shown (adapted from Gausachs *et al*., 2012). Yellow rectangle highlight the most important regions.

Constitutional methylation of the *MLH1* promoter was detected in four (4/38; 10.5%) patients (table 6). The mean age at diagnosis of these patients was 36 (range was 26-48 years). The clinical and molecular features of positive patients for constitutional methylation of the *MLH1* promoter are presented in table 7. The methylation level detected in PBL at the C-Deng region (the region directly involved in *MLH1* transcriptional activity) was on average of 28%, indicating that this alteration is present in mosaic. Figure 10 shows an electropherogram plot of a positive case for constitutional methylation in the *MLH1* promoter studied by MS-MLPA.

We also studied *MLH1* promoter methylation in samples representative of all embryonic layers, namely: tumor and normal colon mucosa (endoderm), buccal

mucosa (ectoderm) and muscle (mesoderm). *MLH1* methylation levels reveled by MS-MLPA analysis of these tissues does not differ significantly from those observed in PBL (mesoderm) (Table 6), demonstrating that this epigenetic alteration affects similarly the tissues from different embryonic origins.

We were able to study three relatives (parents and sister) of one of the probands with constitutional methylation of the *MLH1* promoter (patient #27), and none of them present constitutional methylation of the *MLH1* promoter.

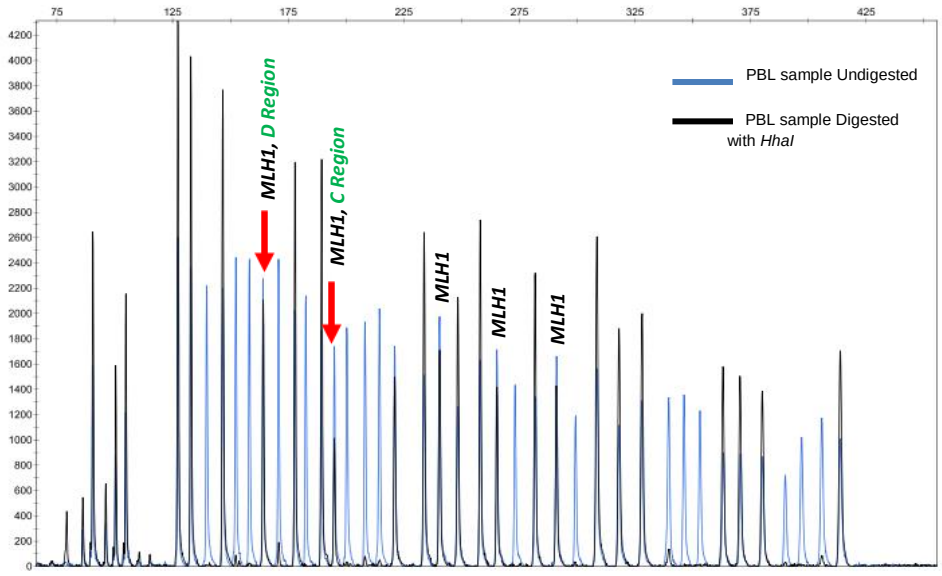


Figure 10 - Electropherogram plot representing MS-MLPA in a PBL with constitutional methylation of the *MLH1* promoter. The C- and D-Deng regions, the most important regions associated with the transcription of *MLH1* gene, are highlighted with the red arrows.

Table 6 - *MLH1* methylation levels (%) using MS-MLPA in samples from different germline origins in the four probands.

Patient	% *MLH1* Methylation									
	PBL		Tumor		Normal Colon Mucosa		Buccal Mucosa		Muscle	
	C Region (-246 nt)	D Region (-13 nt)	C Region (-246 nt)	D Region (-13 nt)	C Region (-246 nt)	D Region (-13 nt)	C Region (-246 nt)	D Region (-13 nt)	C Region (-246 nt)	D Region (-13 nt)
#3	14.1	43.3	20.4	62.9	8.35	27.7	12.4	44.5	11.1	32.3
#10	18.3	49.9	11	32.2	9.1	30.5	NA*	NA*	NA*	NA*
#27	32.6	51.4	13	38.6	14.4	35.4	13.8	40.5	12.2	39.5
#38	46.3	52.3	16.4	40.2	11.5	36	11.6	36.4	NA	NA

NA, not available; **PBL**, peripheral blood lymphocytes.

*Patient died during the study, so it was not possible to study buccal mucosa.

Table 7 - Clinical and molecular features of patients with constitutional methylation of the *MLH1* promoter.

	Age (years)	Sex	Proband	Samples Studied by MS-MLPA	Promoter variants	*BRAF*	Familial history	*MLH1* methylation in relatives
#3	**52**	**M**	Metachronous CRC at age <50 years	PBL, CRC, normal colon mucosa, buccal mucosa and muscle	**ht c.-93G>A**	**WT**	None	NS
#10[a]	**60**	**F**	CRC at age 26 years	PBL, CRC and normal colon mucosa	**None**	**WT**	First degree (maternal) relative with stomach cancer	NS
#27	**36**	**M**	CRC at age <50 years	PBL, CRC, normal colon mucosa, buccal mucosa and muscle	**None**	**WT**	Maternal great-grandmother with breast cancer, a paternal grandfather with prostate cancer, a paternal grandmother with colon cancer and father with bladder cancer	Father*, mother* and sister*
#38	**49**	**M**	Synchronous CRC at age <50 years	PBL, CRC, normal colon mucosa, buccal mucosa	**ht c.269C>G**	**WT**	Father with liver and kidney cancers. Mother with uterine and endometrium cancers	NS

CRC, colorectal cancer; **F**, female; **ht**, heterozygous; **IHC**, immunohistochemistry; **NS,** not studied; **M**, male; **PBL**, peripheral blood lymphocytes; **WT**, wildtype.

*Negative for *MLH1* promoter methylation.

[a] This patient died during the study.

2. Confirmation of correct MS-MLPA probe hybridization

Sanger sequencing of the whole *MLH1* promoter (from c.-1469 to intron 1) was performed in PBL samples from all cases, probands (n=38) and relatives (n=3) in order to find out if there were any variants which might inhibit the binding of the MS-MLPA probes. The results showed that none of the cases had DNA variants affecting the binding sites of the MS-MLPA probes or the *HhaI* restriction sites. Nonetheless, analysing the entire promoter region of the *MLH1* gene, we found that one patient, who did not present constitutional methylation of *MLH1,* had the alteration c.-261G>A (rs587782685) described as a variant of unknown significance (VUS) in the ClinicalVar NCBI database (www.ncbi.nih.gov/clinvar). This change is outside to the probe hybridization site (-246 nt). Another heterozygous alteration, the c.-269C>G (rs35032294) (Zavodna *et al*., 2006), was found in patient #38 (table 8), being also outside to the probe hybridization site. Therefore, although these changes are located near the end of the probe hybridization site, this did not influence the hybridization efficiency because no copy number changes were found by MS-MLPA.

Additionally, a common SNP, the c.-93G>A (rs180073439) relative to the translation start site of *MLH1*, was found in heterozygosity in sixteen cases (16/38; 42%), including patient #3 with constitutional methylation, and in homozygosity in one case (1/30; 2.6%). This common polymorphism is also outside the probe hybridization sites. No other variants were found in the *MLH1* promoter region.

3. Clinico-pathological features of patients with constitutional methylation of the *MLH1* promoter

Case #3 was a male who had two tumors of the descending colon, one at 38 and another at 45 years of age (metachronous tumors). The patient had no family history of cancer as it is shown in the family pedigree (figure 11 A).

Case #10 was a female who was diagnosed with a moderately differentiated adenocarcinoma in the ascending colon at the age of 26 years. At 60 years of age, the patient was diagnosed with a stomach carcinoma. This patient presented scant family history of cancer, namely a maternal uncle affected with gastric cancer at age 70 years. The patient deceased at age 60 years during the course of this study (figure 11 B).

Case #27 was a male who was diagnosed with an invasive adenocarcinoma in the rectum at the age of 33 years. This patient has family history of cancer, namely, a

maternal great-grandmother with breast cancer, a paternal grandfather with prostate cancer, a paternal grandmother with colon cancer, and the father with bladder cancer (figure 11 C). The father and the healthy mother and sister also participated in this study (figure 11 C).

Case #38 was a male who was diagnosed with three different colon synchronous tumors in the ascending, transverse and descending colon at age of 48 years. The patient´s father was affected by kidney and liver cancer at the age of 60 years and the mother was affected by an endometrium cancer at the age 65 years (figure 11 D).

None of the tumors of the four patients with constitutional methylation of the *MLH1* promoter presented the p.Val600Glu *BRAF* mutation.

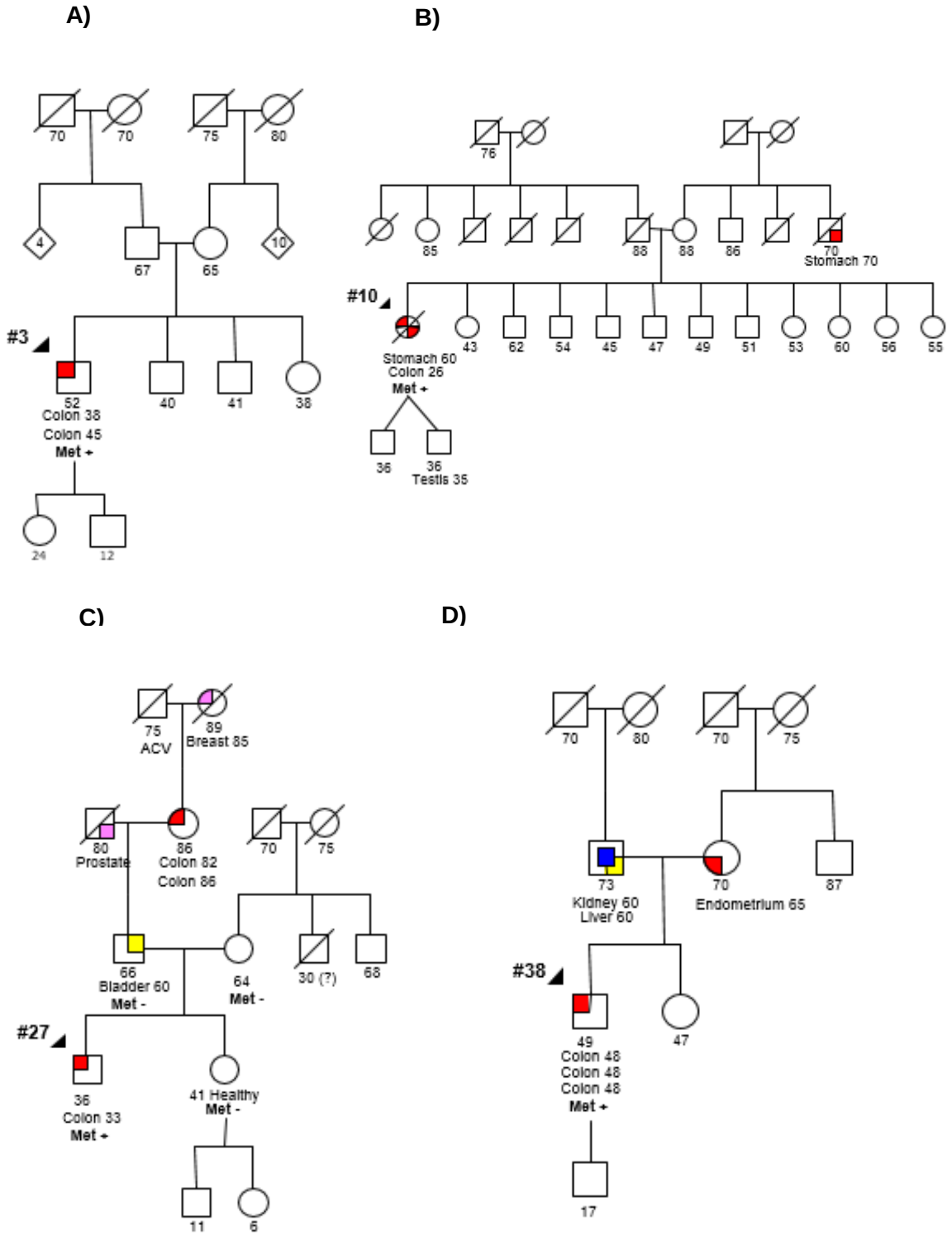


Figure 11 - Family pedigrees of four patients positive for constitutional methylation of the *MLH1* promoter. Family history of the four patients with constitutional methylation (A) #3; B) #10; C) #27; D) #38) with age (in years) and health status. Black arrows indicate the probands. Abbreviations: **met-**, with no methylation of the *MLH1* promoter; **met+**, with methylation of the *MLH1* promoter. Circles, females; squares, males; semi filled, cancer affected.

4. Quantification of the *MLH1* transcript

In order to assess global *MLH1* transcript levels, we measured its relative expression levels by semiquantitative multiplex RT-PCR in PBL samples of patients with constitutional methylation (#3, #10, #27, #38) and controls. We amplified the *MLH1* transcript and an internal control, the beta-2-microglobulin (*B2M*).

As shown by the electrophoretic profile of the multiplex RT-PCR products (figure 12), in the cases positive for constitutional methylation we observed a decrease of *MLH1* RNA expression compared to controls.

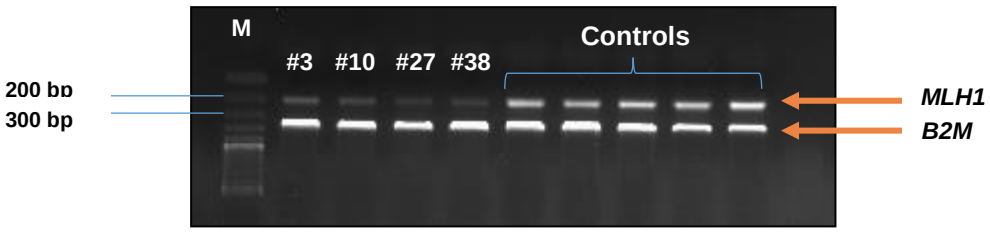


Figure 12 - Electrophoretic pattern of multiplex RT-PCR products. This was obtained for the four positive patients for constitutional methylation and five normal controls, where M is the molecular weight marker (100bp) and *B2M* is beta-2-microglobulin.

To confirm and measure the differences in the transcript levels observed in the electrophoresis of the multiplex RT-PCR products from patients with constitutional methylation and controls, we performed fragment analysis. The results obtained are shown in figure 13.

The *MLH1* transcript levels were calculated by comparing the relative peak areas of the *MLH1* transcript to the relative peak areas of the beta-2-microglobulin (*B2M*).

Controls presented an average of 99% of *MLH1* transcript expression relatively to *B2M*. Considering our four cases with constitutional methylation, patient #3 presented a decrease of 38%; patient #10 a decrease of 37%; patient #27 a decrease of 46% and patient #38 a decrease of 38% of *MLH1* transcript levels when compared to *B2M (*figure 13).

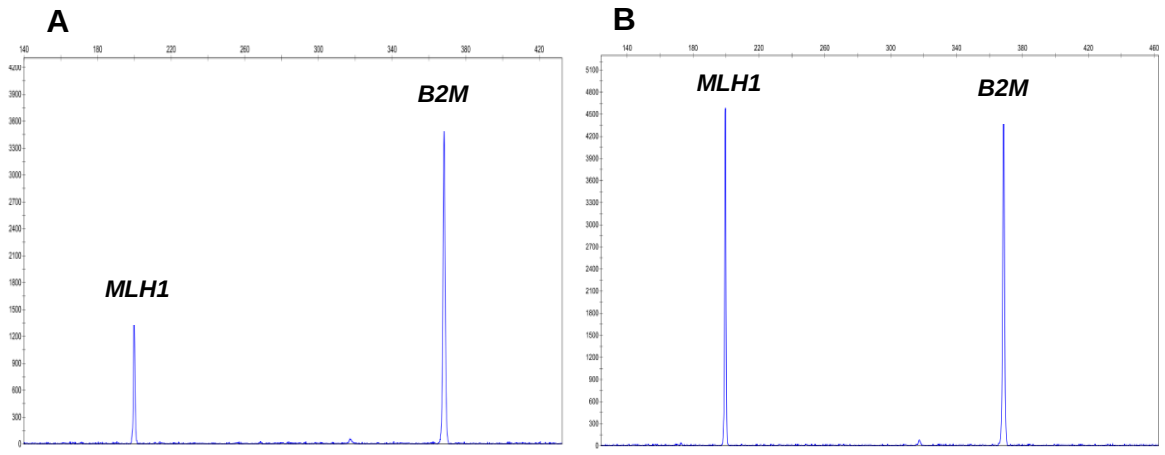


Figure 13 - Electropherogram of fragment analysis of the multiplex RT-PCR products. The relative peak areas of a positive case for constitutional methylation (A) and control (B) are shown. In the case with constitutional methylation, a decreased *MLH1* expression relatively to the internal control was verified.

5. *MLH1* allelic expression

Data available from the *MLH1* germline mutation screening showed that two of the probands positive for constitutional methylation of the *MLH1* promoter (patient #10 and #27) were heterozygous for the coding polymorphism c.655A>G, p.Ile219Val (rs1799977) in exon 8. In order to evaluate if the *MLH1* promoter methylation were monoallelic, the cDNA of this patients was sequenced in PBL samples. In both cases, both alleles were present, but one of them appeared to be less expressed (figure 14 A). This difference was more evident when compared with controls (figure 14 B).

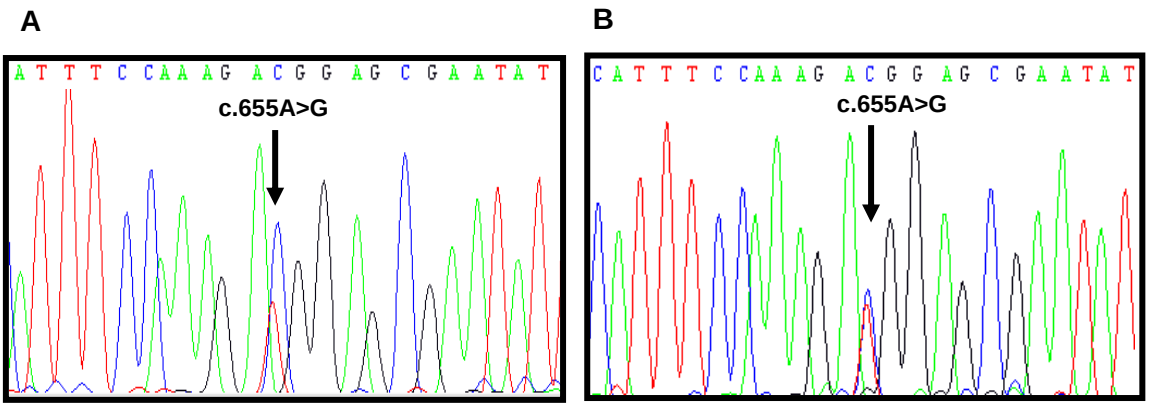


Figure 14 - Electropherogram of sequencing analysis demonstrating the allelic expression of the SNP rs1799977 in exon 8 of *MLH1* in cDNA samples. Patient #10 (A), Control (B) (both reverse).

In order to confirm these results, we performed on gDNA and cDNA SNuPe analysis specific for the rs1799977 (c.655A>G), in the two constitutional methylation-positive patients heterozygous for this polymorphism (cases #10 and #27) and in controls (also heterozygous).

In the gDNA, no significant differences were obtained in the A/G allele relative areas (case #10: 86%; case #27: 86%; and controls: 88%, figures 15 and 16). In the cDNA, we observed that both alleles were expressed in the two probands, but in both cases the A allele showed a signal reduction not observed in the controls (figure 15). These results suggest a decrease in the expression of the A allele. We observed a 45% decrease of the A allele relative to the G allele in patient #10 (figure 15) and a 27% decrease in patient #27 (figure 16). In controls, the difference between the two alleles was 2%. Moreover, the normalized ratios between cDNA and gDNA revealed a 53% and a 32% decrease of A allele in case #10 and in case #27, respectively. In controls, the normalized ratio obtained was of <0.1%.

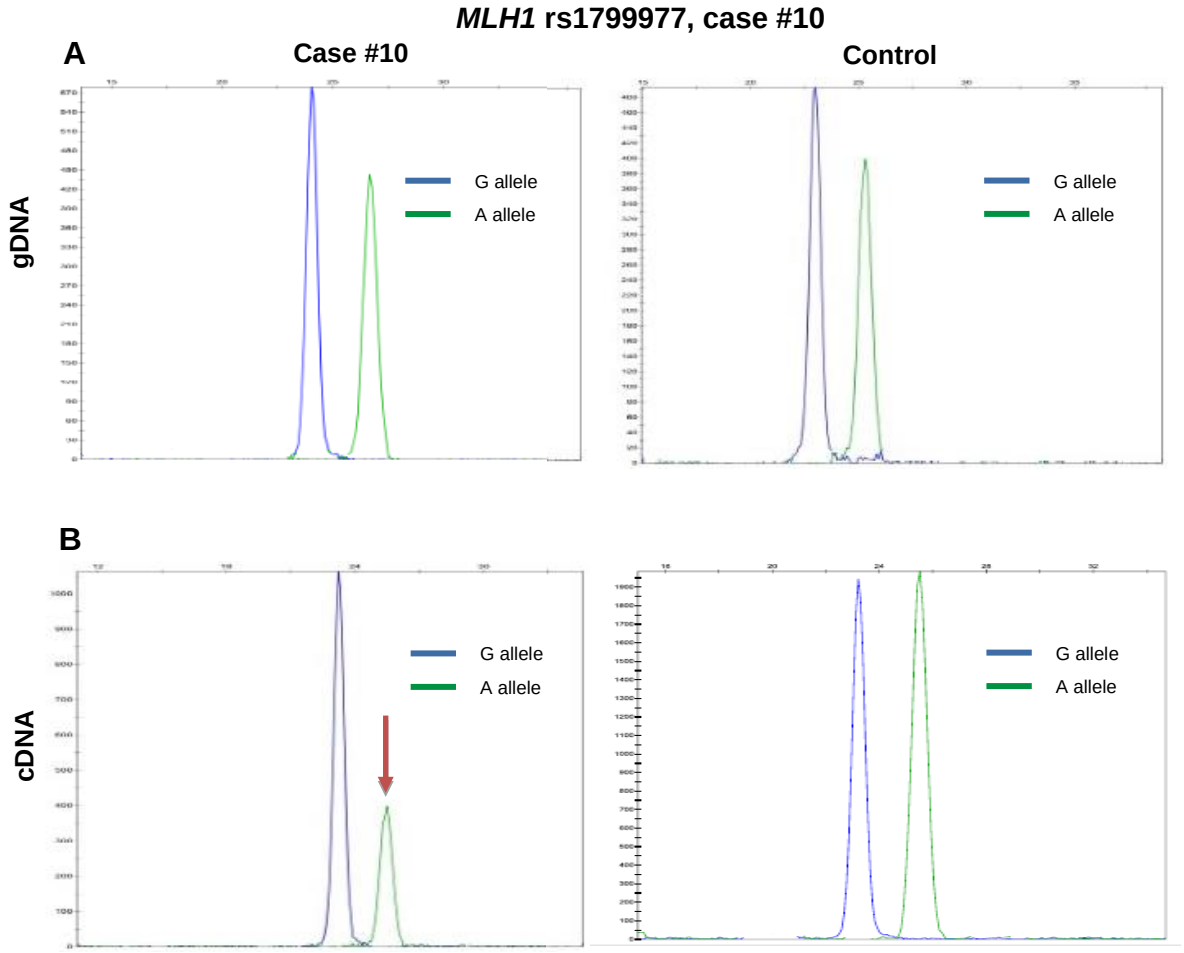


Figure 15 - **SNuPE analysis at *MLH1* rs1799977 (c.655A>G) in gDNA (upper panels) and in cDNA (bottom panels) of an heterozygous patient, case #10 with constitutional methylation (left panels) and of a control (right panels).** Partial transcriptional silencing of the A allele in the cDNA of the patient was observed.

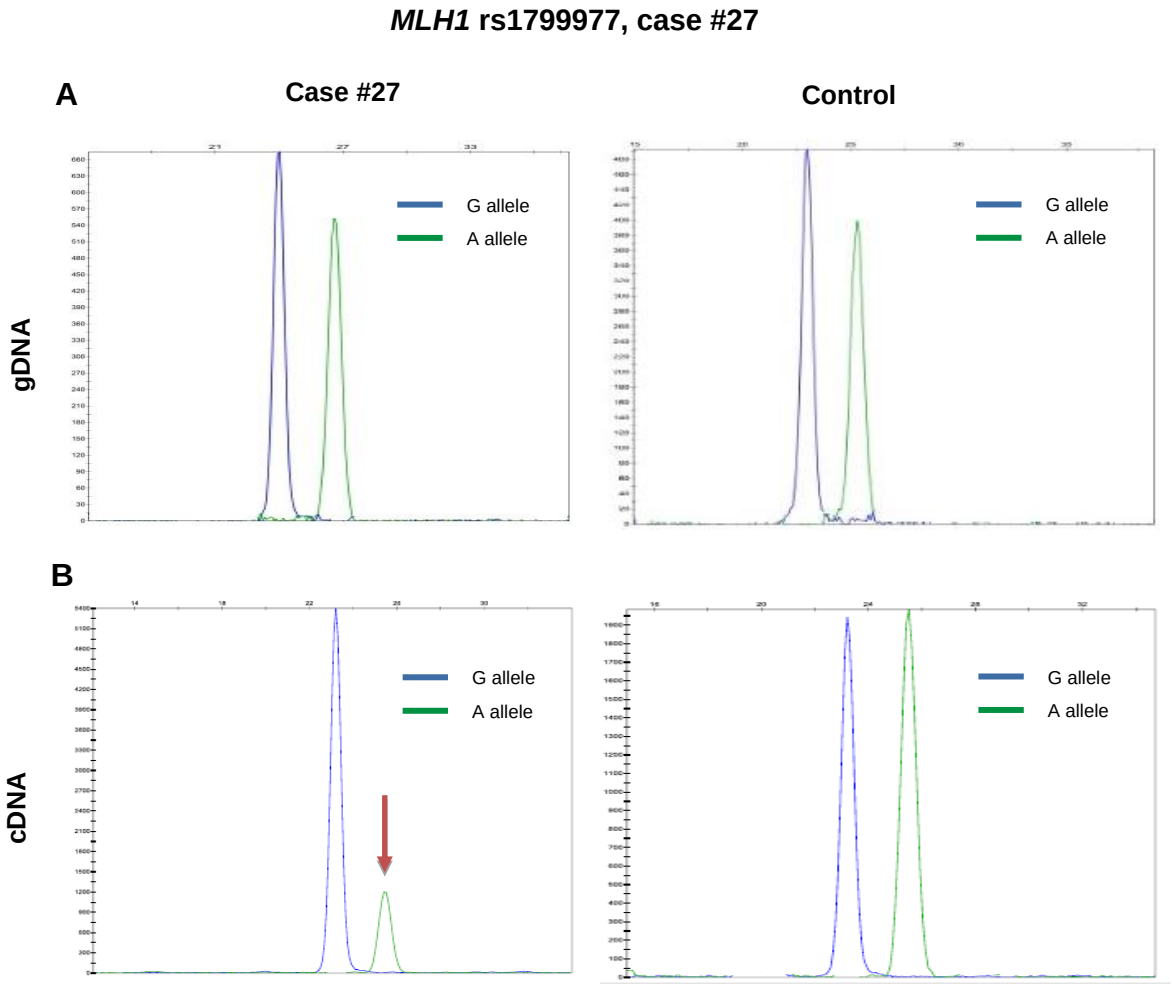


Figure 16 - **SNuPE analysis at *MLH1* rs1799977 (c.655A>G) in gDNA (upper panels) and in cDNA (bottom panels) of an heterozygous patient, case #27 with constitutional methylation (left panels) and of a control (right panels).** Partial transcriptional silencing of the A allele in the cDNA of the patient was observed.

6. Microsatellite instability analysis

Microsatellite instability analysis was performed by fragment analysis (using the Bethesda panel) in all four patients positive for constitutional methylation. This analysis was performed in tumors and normal mucosa samples. All cases with constitutional methylation of the *MLH1* promoter have shown MSI-H in their tumors. In more detail, cases #3, #10, and #27 had instability in 50% of the markers and case #38 showed instability in 100% of the markers (figure 16).

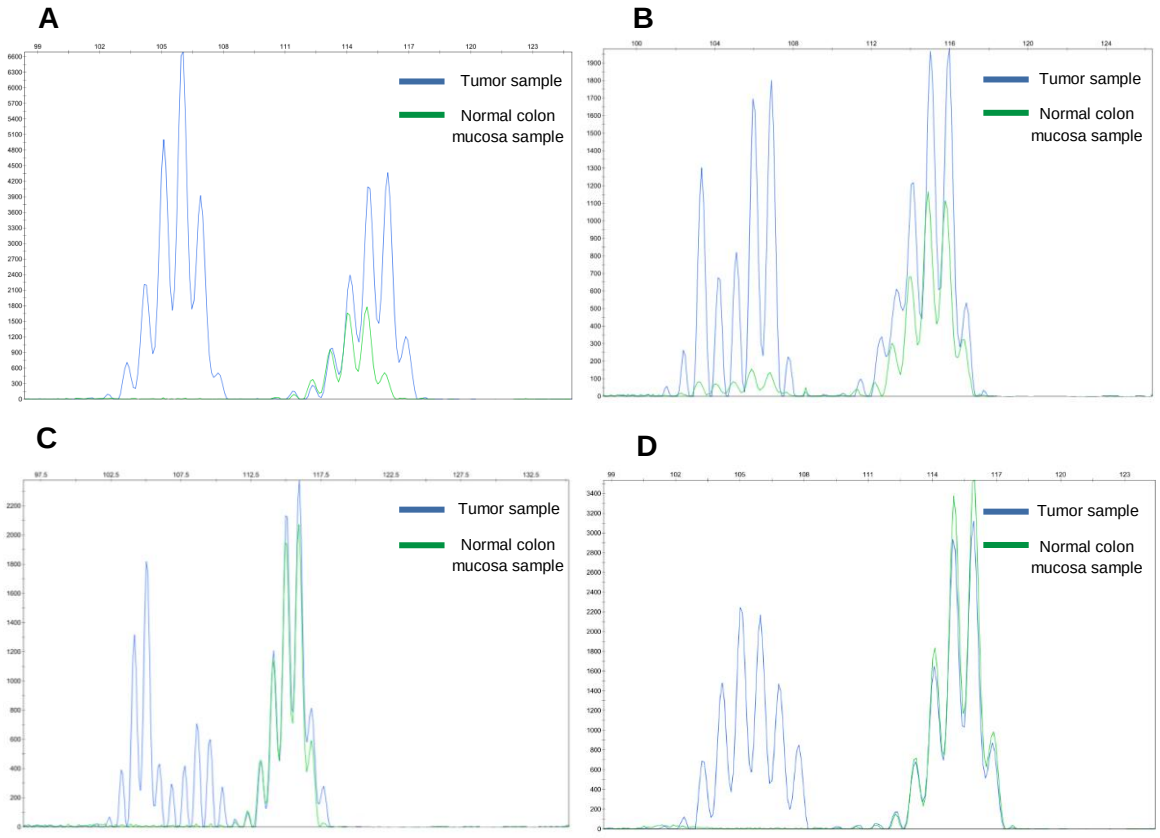


Figure 17 - Fragment analysis showing microsatellite sequences of mononucleotide BAT26 in cases #3 (A), #10 (B), #27 (C), and #38 (D). This analysis was performed in tumors (blue) and normal colon mucosa (green) samples in order to compare both tissues.

V. DISCUSSION

V. DISCUSSION

It is nowadays recognized that epimutations are often associated with inactivation of tumor-suppressor genes and activation of proto-oncogenes that play a central role in carcinogenesis (Herceg and Vaissière, 2011; Sharma *et al*., 2009). The epigenetic disruption can predispose to mutation events, emphasizing the importance of both genetic and epigenetic alterations in carcinogenesis (Dobrovic and Kristensen, 2009; Sharma *et al*., 2009).

Lynch syndrome is caused by germline mutations in the MMR genes, mainly affecting *MLH1* and *MSH2.* Besides genetic mutations, constitutional epigenetic silencing of the MMR genes *MLH1* and *MSH2* has been recently reported as another possible cause of LS (Crépin *et al*., 2012; Hitchins, 2013; Lynch *et al*., 2015; Sehgal *et al*., 2014). Recently, constitutional *MLH1* methylation has been identified in a subset of cases with clinical criteria for LS and with loss of the MLH1 protein in the tumor, but without identified germline mutations in the MMR genes (Crépin *et al*., 2012; Hitchins, 2013; Hitchins and Ward, 2009; Morak *et al*., 2008; Pineda *et al*., 2012; Ward *et al*., 2012). Constitutional methylation of *MLH1* is characterized by dense CpG island promoter methylation of a single allele throughout the somatic tissues, accompanied by transcriptional inactivation of the affected allele (Coppedè *et al*., 2014; Hitchins and Ward, 2009).

1. *MLH1* constitutional methylation

In this study we aimed to investigate the prevalence of *MLH1* constitutional methylation in patients who met clinical criteria for LS, as well as to deduce the pattern of constitutional methylation inheritance. For this purpose we screened 38 probands from families that fulfilled clinical criteria for LS, without germline mutations in *MLH1* and presenting loss of MLH1 protein expression by IHC. We identified four patients with constitutional methylation of *MLH1*. The frequency of constitutional *MLH1* methylation was 10.5% (4/38), indicating that *MLH1* methylation may account for a non-negligible proportion of LS. To our knowledge and to date, there are at least 55 index cases without a conventional pathogenic germline MMR mutation that are carriers of constitutional *MLH1* methylation (Castillejo *et al*., 2015; Hitchins, 2013). Here, we report four additional patients presenting constitutional *MLH1* methylation.

In accordance with previous reports, on the clinical and pathological characteristics of the patients with constitutional *MLH1* methylation, two of the four cases identified in this study had developed multiple LS tumors at an early age (Castillejo *et al*., 2015; Hitchins *et al*., 2007; Hitchins and Ward, 2009; Hitchins, 2013; Pineda *et al*., 2012; Ward *et al*., 2012), one having metachronous (case #3) and the other synchronous tumors (case #38). Regarding family history, 3/4 of the patients with constitutional methylation presented relatives with LS-associated cancers, but their ages at diagnosis are compatible with sporadic origins. The fourth patient with constitutional *MLH1* methylation did not present family history of cancer, which is quite common in this pathogenic mechanism (Castillejo *et al*., 2015; Crépin *et al*., 2012; Hitchins *et al*., 2007; Hitchins and Ward, 2009; Hitchins, 2013; Ward *et al*., 2012).

On all cases, methylation of the *MLH1* promoter was studied in the five regions analysed, including the C- and D-Deng regions, being the C-Deng region the most directly correlated with transcriptional silencing and resulting loss of MLH1 protein expression (Deng *et al*., 2002). The methylation level detected in PBL at the C-Deng region was an average of 28% (table 6), indicating that this alteration is present in mosaic (the affected allele is not methylated in all cells), as described in other studies (Castillejo *et al*., 2015; Crépin *et al*., 2012; Pineda *et al*., 2012; Sloane *et al*., 2015). In these patients, methylation of the *MLH1* gene was also detected in other tissues, like colorectal tumors and normal colon mucosa (endoderm), oral mucosa (ectoderm) and muscle (mesoderm), representing the three germ layers, and with methylation frequencies that are not significantly different from that found in PBL (mesoderm). The normal developmental process of epigenetic reprogramming in mammals has two major waves of reprogramming reset the epigenome. The first wave, the demethylation, occurs following fertilization in the early embryo and the second, *de novo* methylation, takes place in primordial germ cells (Fleming *et al*., 2008). Thus, since the methylation is present in equivalent levels in tissues derived from the three germ layers, we can infer that in these cases the methylation of *MLH1* gene occurred early during embryogenesis, as previously reported in other studies (Hitchins, 2013; Ward *et al*., 2012). In fact, we were able to study three relatives of patient #27 (a man with early onset CRC who showed methylation of the *MLH1* promoter of around 30% in PBL samples), namely the parents and the sister, and none of them presented constitutional *MLH1* methylation, suggesting that in this patient *MLH1* methylation arose *de novo*, similarly to most cases reported to date (Hitchins and Ward, 2009; Hitchins, 2013).

The pattern of transmission of the different forms of epimutations in the *MLH1* gene depends on its origin, and may be divided into two categories. First, the primary

epimutations usually arise spontaneously *de novo* and are reversible (but may occasionally be transmitted to the next generation), not following a Mendelian pattern of transmission (Hitchins, 2013). On the other hand, secondary epimutations occur as the result of a genetic mutation in *cis*, and therefore may follow the classic pattern of autosomal dominant transmission (Castillejo *et al.*, 2015; Hitchins, 2013). A dominantly transmitted constitutional *MLH1* methylation has been linked to a *MLH1* haplotype bearing two single-nucleotide variants: c.-27C>A and c.85G>T (p.Ala29Ser) (Kwok *et al*., 2014; Pineda *et al*., 2012). Studies of the c.-27C>A variant offer the most compelling evidence that *MLH1* promoter variants can directly affect the regulation of *MLH1* (Hitchins, *et al*., 2011; Kwok, *et al*., 2014; Ward, *et al*., 2012) and has been associated with reduced transcriptional activity and the dominant inheritance of a mosaic constitutional *MLH1* methylation (Hitchins, *et al*., 2011; Hesson *et al*., 2014). In our series this variant was not found and, we excluded the hypothesis of secondary epimutations as the mechanism for the cases we here present. Sanger sequencing of the whole *MLH1* promoter (from c.-1469 to intron 1) was performed in PBL from all probands (n=38) and in the available relatives (n=3). No variants affecting the binding of MS-MLPA probes or *HhaI* restriction sites were found, making it more likely that the patients here reported have primary constitutional methylation. The fact that in our cases the methylation was present in mosaic makes it more likely to be a primary epimutation. However, we cannot exclude the remote possibility that there is a mutation in *cis* (outside the regions studied), causing a secondary methylation. For a more definitive classification of these epimutations, it would be necessary to perform additional studies, including offspring haplotypes studies. To know the precise mechanism would allow better risk predictions for the relatives concerning the development of neoplasms associated with LS and more tailored screening and prophylactic measures.

2. *MLH1* expression pattern

We have also proposed to verify if *MLH1* RNA expression was altered in PBL samples, when compared to selected controls and using *B2M* as internal control. We concluded that there was significant loss of expression of *MLH1* in the four positive cases with constitutional methylation of the *MLH1* gene when compared to controls. Our results are in agreement with other studies showing that constitutional methylation causes a transcriptional silencing of *MLH1* gene and indicating that these events are correlated (Pineda *et al*., 2012; Ward *et al*., 2012).

Two of the probands positive for constitutional methylation of the *MLH1* promoter (patient #10 and #27) were heterozygous for the coding polymorphism c.655A>G, p.Ile219Val (rs1799977) within *MLH1* exon 8. We took advantage of this heterozygous polymorphism to determine the effect of the methylation in the *MLH1* transcriptional activity. In order to evaluate if the *MLH1* promoter methylation were monoallelic, we sequenced the cDNA of these patients. In both cases, both alleles were present, but one of them was less expressed when compared with controls. These results we confirmed by SNuPE analysis performed on gDNA and cDNA. Allele-specific expression values obtained revealed a decrease in the A allele of 45% and of 27% (relatively to the controls) in cases #10 and #27, respectively. Therefore, it is assumed that the allele methylated is the A allele showing lower expression. In summary, the RNA studies demonstrated a decrease of the *MLH1* gene expression in cases with constitutional methylation, indicating that these events are correlated. Recent studies have also shown that *MLH1* methylation can present itself as a constitutional alteration that results in the silencing of the affected allele (Gazzoli *et al*., 2002; Goel *et al*., 2011; Hitchins *et al*., 2005; Morak *et al*., 2008; Mrkonjic *et al*., 2010; Sloane *et al*., 2015).

3. Molecular characterization of the tumors

In patients with LS, the somatic inactivation, or “second hit”, of the wild-type allele of the affected MMR gene leads to the abnormal function of the MMR gene. In turn, this leads to the accumulation of errors during DNA replication, especially in repetitive sequences known as microsatellites (Boland and Goel, 2010). Consequently, tumors from patients with LS characteristically demonstrate MMR deficiency, defined as the presence of microsatellite instability (MSI) and/or the loss of MMR protein expression, which are the hallmarks of this disorder (Cicek *et al*., 2011; Kawakami *et al*., 2015). This feature is present in more than 90% of LS-associated colorectal tumors, in general and also in those associated with somatic or constitutional epigenetic silencing of the *MLH1* gene. Our results are in agreement with all of the published data, since all four tumors analyzed were MSI-H and present loss of expression of MLH1 both at the mRNA and protein level (Bouzourene *et al*., 2010; Cunningham *et al*., 2010; Deng *et al*., 2002; Domingo *et al*., 2004).

The *BRAF* V600E mutation has been detected predominantly in sporadic CRC. Consequently, the presence of a *BRAF* mutation is usually, evidence against the

presence of LS (Giardiello *et al*., 2014). The evidence gathered so far strongly indicates that the presence of a *BRAF* mutation may exclude the occurrence of *MLH1*

constitutional methylation (Pérez-Carbonell *et al*., 2010). The p.Val600Glu *BRAF* mutation was absent in the tumors of the all four carriers of constitutional *MLH1* methylation here reported, as it is the case for the majority of the reported cases in the literature (Crépin *et al*., 2012; Goel *et al*., 2011; Pineda *et al*., 2010). However, van Roon *et al*. (2010), Parsons *et al*. (2012), and Crépin *et al*. (2012) have reported LS cases with the p.Val600Glu *BRAF* mutation, showing that it is possible the co-existence of *BRAF* mutation with *MLH1* constitutional methylation (van Roon *et al*., 2010; Crépin *et al*., 2012; Parsons *et al*., 2012), contrarily to what happens in LS due to germline mutations.

VI. CONCLUSION

VI. CONCLUSION

It is nowadays recognized that analysis of DNA methylation patterns offers an alternative strategy for the study of inactivation of tumor suppressor genes, when patients present loss of MLH1 protein expression and have no *MLH1* pathogenic germline mutation. After completion of this study we conclude:

- Constitutional *MLH1* promoter methylation is present in about 10.5% of the patients without germline mutations and who present loss of protein MLH1 expression in their tumors;
- Constitutional methylation is present usually in mosaicism in tissues derived from all the germ layers, indicating that this phenomenon occurs in the early embryo;
- Constitutional methylation is monoallelic and is associated with a significant decrease in *MLH1* expression when compared with controls, indicating that these events are correlated;
- Patients with constitutional methylation exhibit MSI-H and do not (or only rarely) present the p.Val600Glu *BRAF* mutation;
- The fact that mosaic constitutional methylation was observed in all cases, together with its absence in the parents of the patient from whom relatives could be evaluated, makes it likely that the most common mechanism of constitutional methylation is primary epimutation.

VII. FUTURE PERSPECTIVES

VII. FUTURE PERSPECTIVES

This study may benefit from further analysis to support our conclusions and to allow a more specific evaluation of the clinical implications for the relatives of the patients with constitutional *MLH1* methylation. Thus, we plan:

- To quantify the methylation of the *MLH1* gene by another technique, for example by pyrosequencing, in order to be able to confirm the results obtained by the MS-MLPA technique;
- To analyze all available clinical data of each patient with constitutional methylation and their relatives;
- To construct haplotypes in order to comprehend the transmission pattern in each family and the origin of methylation;
- To increase our series in order to identify more index cases showing this epigenetic mechanism of LS.

VIII. REFERENCES

VIII. REFERENCES

Al-Sohaily, S., Biankin, A., Leong, R., Kohonen-Corish, M., and Warusavitarne, J. (2012). Molecular pathways in colorectal cancer. Jounal Gastroenteroly Hepatology. *27(9)*, 1423–1431.

Armaghany, T., Wilson, J.D., Chu, Q., and Mills, G. (2012). Genetic alterations in colorectal cancer. Gastrointestinal Cancer Research. *5(1)*, 19–27.

Astin, M., Griffin, T., Neal, R.D., Rose, P., and Hamilton, W. (2011). The diagnostic value of symptoms for colorectal cancer in primary care. British Journal of General Practice. *61(586)*, 231–243.

Arends, M. J. (2013). Pathways of colorectal carcinogenesis. Applied Immunohistochemistry & Molecular Morphology. *21(2)*, 97–102.

Balmaña, J., Balaguer, F., Cervantes, A., and Arnold, D. (2013). Familial risk-colorectal cancer: ESMO clinical practice guidelines. Annals of Oncology. *24*, *suppl 6,* 73-80*.*

Banno, K., Kisu, I., Yanokura, M., Tsuji, K., Masuda, K., Ueki, A., *et al*. (2012). Epimutation and cancer: a new carcinogenic mechanism of Lynch syndrome (Review). International Journal of Oncology. *41*, 793–797.

Berg, M., and Søreide, K. (2011). Genetic and epigenetic traits as biomarkers in colorectal cancer. International Journal of Molecular Sciences. *12(2)*, 9426–9439.

Bogaert, J., and Prenen, H. (2014). Molecular genetics of colorectal cancer. Annals of Gastroenterology. *27(1)*, 9–14.

Boland, R., Thibodeau, S., Halmiton, S.,Sidransky, D., Eshleman, JR., Burt, RW., *et al.* (1998). A national cancer institute workshop on microsatellite instability for cancer detection and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer. Cancer Research. *58(22)*, 5248-57.

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

Bosetti, C., Rosato, V., Gallus, S., Cuzick, J., and La Vecchia, C. (2012). Aspirin and cancer risk: A quantitative review to 2011. Annals of Oncology. *23(6)*, 1403–1415.

Bouzourene, H., Hutter, P., Losi, L., Martin, P., and Benhattar, J. (2010). Selection of patients with germline *MLH1* mutated Lynch syndrome by determination of *MLH1* methylation and *BRAF* mutation. Familial Cancer. *9(2)*, 167–172.

Brown, J.R., and DuBois, R.N. (2005). COX-2: a molecular target for colorectal cancer prevention. Journal of Clinical Oncology. *23(12)*, 2840–2855.

Bujanda, L., Ph, D., Cubiella, J., Salas, D., Hernández, C., Sc, B., Jover, R., Montalvo, I., Ponce, M., Andrés, M., *et al*. (2012). Colonoscopy versus fecal immunochemical testing in colorectal-cancer screening. New England Journal of Medicine. *366(8)*, 697–706.

Calistri, D., Rengucci, C., Seymour, I., Lattuneddu, A., Polifemo, A.M., Monti, F., *et al*. (2005). Mutation analysis of p53, K-ras, and *BRAF* genes in colorectal cancer progression. Journal of Cellular Physiology. *204(2)*, 484–488.

Cappell, M.S. (2008). Pathophysiology, clinical presentation, and management of colon cancer. Gastroenterology Clinics of North America. *37(1)*, 1–24.

Castillejo, A., Hernandez-Illan, E., Rodriguez-Soler, M., Perez-Carbonell, L., Egoavil, C., Barbera, V.M., *et al*. (2015). Prevalence of *MLH1* constitutional epimutations as a cause of Lynch syndrome in unselected versus selected consecutive series of patients with colorectal cancer. Journal of Medical Genetics. *52(7)*, 498–502.

Centelles, J. J. (2012). General aspects of colorectal cancer. ISRN oncology. *2012*, 139–268.

Chomczynski, P. (1993). A reagent for the single-step simultaneous isolation of RNA, DNA and proteins from cell and tissue samples. Biotechniques 15, 532–537.

Chu, E., Alberts, S., Cohen, D., Van Cutsem, E. (2008). New treatment strategies formetastatic colorectal cancer (Connecticut, CMPMedica).

Cicek, M.S., Lindor, N.M., Gallinger, S., Bapat, B., Hopper, J.L., Jenkins, M. A., *et al*. (2011). Quality assessment and correlation of microsatellite instability and immunohistochemical markers among population- And clinic-based colorectal tumors: Results from the colon cancer family registry. Journal of Molecular Diagnostics. *13(3)*, 271–283.

Colussi, D., Brandi, G., Bazzoli, F., and Ricciardiello, L. (2013). Molecular pathways involved in colorectal cancer: implications for disease behavior and prevention. International Journal of Molecular Sciences. *14(8)*, 16365–16385.

Coppedè, F., Lopomo, A., Spisni, R., and Migliore, L. (2014). Genetic and epigenetic biomarkers for diagnosis, prognosis and treatment of colorectal cancer. World Journal of Gastroenterology. *20(4)*, 943–956.

Crépin, M., Dieu, M.-C., Lejeune, S., Escande, F., Boidin, D., Porchet, N., Morin, G., *et al*. (2012). Evidence of constitutional *MLH1* epimutation associated to transgenerational inheritance of cancer susceptibility. Human Mutation. *33(1)*, 180–188.

Cunningham, D., Atkin, W., Lenz, H.-J., Lynch, H.T., Minsky, B., Nordlinger, B., and Starling, N. (2010). Colorectal cancer. Lancet. *375 (9719)*, 1030–1047.

Dawson, M. A., and Kouzarides, T. (2012). Cancer epigenetics: From mechanism to therapy. The journal Cell. *150(1)*, 12–27.

Deng, G., Chen, A., Hong, J., Chae, H.S., and Kim, Y.S. (1999). Methylation of CpG in a small region of the h*MLH1* promoter invariably correlates with the absence of gene expression. Cancer Research. *59(1)*, 2029–2033.

Deng, G., Peng, E., Gum, J., Terdiman, J., Sleisenger, M., and Kim, Y.S. (2002). Methylation of h*MLH1* promoter correlates with the gene silencing with a region-specific manner in colorectal cancer. British Journal of Cancer. *86(4)*, 574–579.

Deschoolmeester, V., Boeckx, C., Baay, M., Weyler, J., Wuyts, W., Van Marck, E., *et al*. (2010). KRAS mutation detection and prognostic potential in sporadic colorectal cancer using high-resolution melting analysis. British Journal of Cancer. *103(10)*, 1627–1636.

DeVita, V., Hellman, S., Rosenberg, S.(2011). DeVita, Hellman and Rosenberg's cancer: Principles & Practice of Oncology. Lippincott Williams & Wilkins, 9th edition.

Dobrovic, A., and Kristensen, L.S. (2009). DNA methylation, epimutations and cancer predisposition. The International Journal of Biochemistry & Cell Biology. *41(1)*, 34–39.

Domingo, E., Laiho, P., Ollikainen, M., Pinto, M., Wang, L., French, A. J., *et al*. (2004). *BRAF* screening as a low-cost effective strategy for simplifying HNPCC genetic testing. Journal of Medical Genetics. *41(9)*, 664–668.

Douillard, J.Y., Siena, S., Cassidy, J., Tabernero, J., Burkes, R., Barugel, M., *et al*. (2010). Randomized, Phase III trial of panitumumab with infusional fluorouracil, leucovorin, and oxaliplatin (FOLFOX4) Versus FOLFOX4 alone as first-line treatment in patients with previously untreated metastatic colorectal cancer: The PRIME study. Journal of Clinical Oncology. *28(31)*, 4697–4705.

Dunlop, M.G., Tenesa, A., Farrington, S.M., Ballereau, S., Brewster, D.H., Koessler, T., *et al*. (2013). Cumulative impact of common genetic variants and other risk factors on colorectal cancer risk in 42,103 individuals. Gut BMJ Journals. *62(6)*, 871–881.

Edge, S. B., and Compton, C.C. (2010). The American Joint Committee on Cancer: the 7th edition of the AJCC cancer staging manual and the future of TNM. Annals of Surgical Oncology. *17(6)*, 1471–1474.

Fearon, E.R. (2011). Molecular genetics of colorectal cancer. Annual Review of Pathology. *6*, 479–507.

Fearon, E.R., and Vogelstein, B. (1990). A genetic model for colorectal tumorigenesis. The journal Cell. *61(5)*, 759–767.

Ferlay, J., Steliarova-Foucher, E., Lortet-Tieulent, J., Rosso, S., Coebergh, J.W.W., Comber, H., *et al*. (2013). Cancer incidence and mortality patterns in Europe: estimates for 40 countries in 2012. European Journal of Cancer. *49(6)*, 1374–1403.

Fleming, J.L., Huang, T.H.M., and Toland, A.E. (2008). The role of parental and grandparental epigenetic alterations in familial cancer risk. Cancer Research. *68(2)*, 9116–9121.

Gausachs, M., Mur, P., Corral, J., Pineda, M., González, S., Benito, L., *et al*. (2012). *MLH1* promoter hypermethylation in the analytical algorithm of Lynch syndrome: a cost-effectiveness study. European Journal of Human Genetics. *20(7)*, 762–768.

Gazzoli, I., Loda, M., Garber, J., Tumor, M.I., Syngal, S., and Kolodner, R.D. (2002). A hereditary nonpolyposis colorectal carcinoma case associated with hypermethylation of the *MLH1* gene in normal tissue and loss of heterozygosity of the unmethylated allele in the resulting microsatellite instability-high tumor. Cancer Research. *62(14)*, 3925–3928.

Giardiello, F.M., Allen, J.I., Axilbund, J.E., Boland, C.R., Burke, C. A, Burt, R.W., *et al*. (2014). Guidelines on genetic evaluation and management of Lynch syndrome: a consensus statement by the US Multi-Society Task Force on colorectal cancer. Gastroenterology. *147(2)*, 502–526.

Giovannucci, E. (2002). Modifiable risk factors for colon cancer. Gastroenterology Clinics of North America. *31(4)*, 925-943.

Goel, A., Nguyen, T.-P., Leung, H.-C.E., Nagasaka, T., Rhees, J., Hotchkiss, *et al*. (2011). *De novo* constitutional *MLH1* epimutations confer early-onset colorectal cancer in two new sporadic Lynch syndrome cases, with derivation of the epimutation on the paternal allele in one. International Journal of Cancer. *128(4)*, 869–878.

Gryfe, R. (2009). Inherited colorectal cancer syndromes. Clinics in Colon and Rectal Surgery. *22(4)*, 198–208.

Haggar, F. A, and Boushey, R.P. (2009). Colorectal cancer epidemiology: incidence, mortality, survival, and risk factors. Clinics in Colon and Rectal Surgery. *22(4)*, 191–197.

Hamilton, S.R. and Aaltonen, L. (2000). World Health Organization Classification of tumours, pathology and genetics of tumours of the digestive system. Lyon, IARC press Lyon.

Hassan, C., Zullo, A., Risio, M., Rossini, F.P., and Morini, S. (2005). Histologic risk factors and clinical outcome in colorectal malignant polyp: A pooled-data analysis. Diseases of the Colon & Rectum. *48(8)*, 1588–1596.

He J., and Efron, J. E. (2011). Screening for colorectal cancer. Advances in Surgery. *45*, 31–44.

Heitman, S.J., Hilsden, R.J., Au, F., Dowden, S., and Manns, B.J. (2010). Colorectal cancer screening for average-risk North Americans: an economic evaluation. PLoS Med. *7(11),* e1000370.

Herceg, Z., and Vaissière, T. (2011). Epigenetic mechanisms and cancer: an interface between the environment and the genome. Epigenetics. *6(7)*, 804–819.

Hesson, L.B., Packham, D., Kwok, C., Nunez, A.C., Schmidt, C., Fields, M., *et al*. (2014). Lynch syndrome associated with two *MLH1* promoter variants and allelic imbalance of *MLH1* expression. Human Mutation. *36(6)*, 622–630.

Hitchins, M.P. (2010). Inheritance of epigenetic aberrations (constitutional epimutations) in cancer susceptibility. Advances in Genetics. *70*, 201–243.

Hitchins, M.P. (2013). The role of epigenetics in Lynch syndrome. Familial Cancer. *12(2)*, 189–205.

Hitchins, M.P., and Ward, R.L. (2009). Constitutional (germline) *MLH1* epimutation as an aetiological mechanism for hereditary non-polyposis colorectal cancer. Journal of Medical Genetics. *46(12)*, 793–802.

Hitchins M, Owens S, Kwok CT, Godsmark G, Algar U, Ramesar R. (2011). Identification of new cases of early-onset colorectal cancer with an MLH1 epimutation in an ethnically diverse South African cohort. Clin Genet. 80, 428–434.

Hitchins, M., Williams, R., Cheong, K., Halani, N., Lin, V. P., Packham, D., Ku, S., Buckle, A., Hawkins, N., Burn, J., *et al*. (2005). *MLH1* germline epimutations as a factor in hereditary nonpolyposis colorectal cancer. Gastroenterology. *129(5)*, 1392–1399.

Hitchins, M.P., Wong, J.J., Suthers, G., Suter, C.M., Martin, D.I., Hawkins, N.J., Ward, R.L. (2007). Inheritance of a cancer-associated. New England Journal of Medicine*. 356(7),* 697–705.

Hugen, N., Van de Velde, C.J.H., De Wilt, J.H.W., and Nagtegaal, I.D. (2014). Metastatic pattern in colorectal cancer is strongly influenced by histological subtype. Annals of Oncology. *25(3)*, 651–657.

Hughes, M.R., and Huang, E.H. (2011). Molecular basis of hereditary colorectal cancer. Semin. Colon and Rectal Surgery. *22(2)*, 65–70.

Huxley, R.R., Ansary-Moghaddam, A., Clifton, P., Czernichow, S., Parr, C.L., and Woodward, M. (2009). The impact of dietary and lifestyle risk factors on risk of colorectal

cancer: A quantitative overview of the epidemiological evidence. International Journal of Cancer. *125(1)*, 171–180.

Iacopetta, B. (2003). TP53 mutation in colorectal cancer. Human Mutation. *21(3)*, 271–276.

Ilyas, M., Straub, J., Tomlinson, I.P., and Bodmer, W.F. (1999). Genetic pathways in colorectal and other cancers. European Journal of Cancer. *35(14)*, 1986–2002.

Jasperson, K.W., Tuohy, T.M., Neklason, D.W., and Burt, R.W. (2010). Hereditary and familial colon cancer. Gastroenterology. *138(6)*, 2044–2058.

Jemal, A., Bray, F., and Ferlay, J. (2011). Global Cancer Statistics. Cancer Journal for Clinicians. *61(2)*, 69–90.

Jiricny, J. (2006). MutLalpha: at the cutting edge of mismatch repair. The journal Cell.*126(2)*, 239–241.

Joensuu, E.I., Abdel-Rahman, W.M., Ollikainen, M., Ruosaari, S., Knuutila, S., and Peltomäki, P. (2008). Epigenetic signatures of familial cancer are characteristic of tumor type and family category. Cancer Research. *68(12)*, 4597–4605.

Jun, S.H., Kim, T. G., Ban, C. (2006). DNA mismatch repair system. Classical and fresh roles. FEBS Journal. *273(8)*, 1609–1619.

Kanthan, R., Senger, J.-L., and Kanthan, S.C. (2012). Molecular events in primary and metastatic colorectal carcinoma: a review. Pathology Research International. *2012*, 497–597.

Kastrinos, F., Syngal, S. (2011). Inherited colorectal cancer syndromes. Cancer Journal. *17(6)*, 405–415.

Kawakami, H., Zaanan, A., and Sinicrope, F.A. (2015). Implications of mismatch repair-deficient status on management of early stage colorectal cancer. Journal of Gastrointestinal Oncology. *16(7)*, 30.

Kheirelseid, E. a H., Miller, N., and Kerin, M.J. (2013). Molecular biology of colorectal cancer : Review of the literature. American Journal of Molecular Biology. *3*, 72–80.

Kwok, C.T., Vogelaar, I.P., van Zelst-Stams, W.A., Mensenkamp, A.R., Ligtenberg, M.J., Rapkins, R.W., *et al.* (2014). 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. *22(5)*, 617–624.

Labianca, R., Nordlinger, B., Beretta, G.D., Brouquet, A., and Cervantes, A. (2010). Primary colon cancer: ESMO clinical practice guidelines for diagnosis, adjuvant treatment and follow-up. Annals of Oncology. *21 Suppl 5*, 70–77.

Labianca, R., Nordlinger, B., Beretta, G.D., Mosconi, S., Mandalà, M., Cervantes, A., and Arnold, D. (2013). Early colon cancer: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. Annals of Oncology. *24 Suppl 6*, 64–72.

Lanza, G., Messerini, L., Gafà, R., and Risio, M. (2011). Colorectal tumors: the histology report. Digestive and Liver Disease. *43 Suppl 4*, 344–355.

Levin, B., Lieberman, D. a., McFarland, B., Andrews, K.S., Brooks, D., Bond, J., *et al*. (2008). Screening and surveillance for the early detection of colorectal cancer and adenomatous polyps, 2008: A joint guideline from the American Cancer Society, the US Multi-Society Task Force on colorectal cancer, and the American College of Radiology. Gastroenterology. *134(5)*, 1570–1595.

Ligtenberg, M.J., Kuiper, R.P., Chan, T.L., Goossens, M., Hebeda, K.M., Voorendt, M., *et al.* (2009). Heritable somatic methylation and inactivation of *MSH2* in families with Lynch syndrome due to deletion of the 3' exons of TACSTD1. Nature Genetics. *41(1)*, 112–117.

Lynch, H.T., and Shaw, T.G. (2013). Practical genetics of colorectal cancer. Chinese Clinical Oncology. *2(2)*, 12.

Lynch, H.T., Snyder, C.L., Shaw, T.G., Heinen, C.D., and Hitchins, M.P. (2015). Milestones of Lynch syndrome: 1895–2015. Nature Reviews Cancer. *15(3)*,181–194.

Markowitz, S.D., and Bertagnolli, M.M. (2009). Molecular origins of cancer: Molecular basis of colorectal cancer. New England Journal of Medicine. *361(25)*, 2449–2460.

Martín-López, J. V, and Fishel, R. (2013). The mechanism of mismatch repair and the functional analysis of mismatch repair defects in Lynch syndrome. Familial. Cancer. *12(2)*, 159–168.

Marzouk, O., and Schofield, J. (2011). Review of histopathological and molecular prognostic features in colorectal cancer. Cancers (Basel). *3(2)*, 2767–2810.

Mills, A. a. (2005). p53: Link to the past, bridge to the future. Genes & Development. *19(18)*, 2091–2099.

Morak, M., Schackert, H.K., Rahner, N., Betz, B., Ebert, M., Walldorf, C., *et al*. (2008). Further evidence for heritability of an epimutation in one of 12 cases with *MLH1* promoter methylation in blood cells clinically displaying HNPCC. European Journal of Human Genetics. *16(7)*, 804–811.

Mrkonjic, M., Roslin, N.M., Greenwood, C.M., Raptis, S., Pollett, A., Laird, P.W., Pethe, V., *et al*. (2010). Specific variants in the *MLH1* gene region may drive DNA methylation, loss of protein expression, and MSI-H colorectal cancer. PLoS One. *5(10)*, 1–10.

Mundade, R., Imperiale, T.F., Prabhu, L., and Loehrer, P.J. (2014). Genetic pathways , prevention, and treatment of sporadic colorectal cancer. Oncoscience*. 1(6),* 400–406.

Naccarati, A., Polakova, V., Pardini, B., Vodickova, L., Hemminki, K., Kumar, R., and Vodicka, P. (2012). Mutations and polymorphisms in TP53 gene - An overview on the role in colorectal cancer. Mutagenesis. *27(2)*, 211–218.

Nacional Comprehensive Cancer Network (NCCN) Guidelines.Genetic/Familial High-Risk. Assessment: Colorectal. (2015). Version 1. http://www.nccn.org/.

Nallapareddy, S. (2011). Early diagnosis and treatment of cancer series: colorectal cancer. (Elsevier). *Chapter 15*. 167–183.

Nygren, a. O.H., Ameziane, N., Duarte, H.M.B., Vijzelaar, R.N.C.P., Waisfisz, Q., Hess, C.J., *et al*. (2005). Methylation-Specific MLPA (MS-MLPA): Simultaneous detection of CpG methylation and copy number changes of up to 40 sequences. Nucleic Acids Research. *33(14)*, 1–9.

Pande, M., Wei, C., Chen, J., Amos, C.I., Lynch, P.M., Lu, K.H., *et al*. (2012). Cancer spectrum in DNA mismatch repair gene mutation carriers: results from a hospital based Lynch syndrome registry. Familial. Cancer. *11(3)*, 441–447.

Parsons, M.T., Buchanan, D.D., Thompson, B., Young, J.P., and Spurdle, A. B. (2012). Correlation of tumour *BRAF* mutations and *MLH1* methylation with germline mismatch repair (MMR) gene mutation status: a literature review assessing utility of tumour features for MMR variant classification. Journal of Medical Genetics. *49(3)*, 151–157.

Perea, J., Lomas, M., and Hidalgo, M. (2011). Molecular basis of colorrectal cancer: towards an individualized management?. Revista Española de Enfermedades Digestivas. *103(1)*, 29–35.

Pérez-Carbonell, L., Alenda, C., Payá, A., Castillejo, A., Barberá, V.M., Guillén, C., *et al*. (2010). Methylation analysis of *MLH1* improves the selection of patients for genetic testing in Lynch syndrome. Journal of Molecular Diagnostics. *12(4)*, 498–504.

Pineda, M., González, S., Lázaro, C., Blanco, I., and Capellá, G. (2010). Detection of genetic alterations in hereditary colorectal cancer screening. Mutation Research. *693(1-2)*, 19–31.

Pineda, M., Mur, P., Iniesta, M.D., Borràs, E., Campos, O., Vargas, G., *et al*. (2012). *MLH1* methylation screening is effective in identifying epimutation carriers. European Journal of Human Genetics. *20(12)*, 1256–1264.

Pino, M.S., and Chung, D.C. (2010). The chromosomal instability pathway in colon cancer. Gastroenterology. *138(6)*, 2059–2072.

Platz, E. A., Willett, W.C., Colditz, G. A., Rimm, E.B., Spiegelman, D., and Giovannucci, E. (2000). Proportion of colon cancer risk that might be preventable in a cohort of middle-aged US men. Cancer Causes Control. *11(7)*, 579–588.

Poston, G.J., Figueras, J., Giuliante, F., Nuzzo, G., Sobrero, A.F., Gigot, J.-F., *et al*. (2008). Urgent need for a new staging system in advanced colorectal cancer. Journal of Clinical Oncology. *26(29)*, 4828–4833.

Purim, O., Gordon, N., and Brenner, B. (2013). Cancer of the colon and rectum: potential effects of sex-age interactions on incidence and outcome. Medical Science Monitor. *19*, 203–209.

Quirke, P., Risio, M., Lambert, R., Von Karsa, L., and Vieth, M. (2011). Quality assurance in pathology in colorectal cancer screening and diagnosis-European recommendations. Virchows Archiv. *458(1)*, 1–19.

Rasmussen, L.J., Heinen, C.D., Royer-Pokora, B., Drost, M., Tavtigian, S., Hofstra, R.M.W., and de Wind, N. (2012). Pathological assessment of mismatch repair gene variants in Lynch syndrome: past, present, and future. Human Mutation. *33(12)*, 1617–1625.

Rustgi, A.K. (2007). The genetics of hereditary colon cancer. Genes & Development. *21(20)*, 2525–2538.

Saltz, L.B., Clarke, S., Díaz-Rubio, E., Scheithauer, W., Figer, A., Wong, R., *et al*. (2008). Bevacizumab in combination with oxaliplatin-based chemotherapy as first-line therapy in metastatic colorectal cancer: A randomized phase III study. Journal of Clinical Oncology. *26(12)*, 2013–2019.

Seisenberger, S., Andrews, S., Krueger, F., Arand, J., Walter, J., Santos, F., *et al*. (2012). The dynamics of genome-wide DNA methylation reprogramming in mouse primordial germ cells. Molecular Cell. *48(6)*, 849–862.

Schmoll, H.J., Van Cutsem, E., Stein, a, Valentini, V., Glimelius, B., Haustermans, K., *et al*. (2012). ESMO consensus guidelines for management of patients with colon and rectal cancer. a personalized approach to clinical decision making. Annals of Oncology. *23(10)*, 2479–2516.

Sehgal, R., Sheahan, K., O'Connell, P.R., Hanly, A.M., Martin, S.T., and Winter, D.C. (2014). Lynch Syndrome: An updated review. Genes (Basel). *5(3)*, 497–507.

Sharma, S., Kelly, T.K., and Jones, P. A. (2009). Epigenetics in cancer. Carcinogenesis. *31(1)*, 27–36.

Sheffer, M., Bacolod, M.D., Zuk, O., Giardina, S.F., Pincas, H., Barany, F., *et al*. (2009). Association of survival and disease progression with chromosomal instability: a

genomic exploration of colorectal cancer. Proceedings of the National Academy of Sciences. U. S. A. *106(17)*, 7131–7136.

Shia, J., Klimstra, D.S., Bagci, P., Basturk, O., and Adsay, N.V. (2012). TNM staging of colorectal carcinoma: issues and caveats. Seminars in Diagnostic Pathology. *29(3)*, 142–153.

Siegel, R., DeSantis, C., Virgo, K., Stein, K., Mariotto, A., Smith, T., *et al*. (2013). Cancer treatment and survivorship statistics, 2012. CA. Cancer Journal for Clinicians. *62(4)*, 220–241.

Siegel, R., Ma, J., Zou, Z., and Jemal, A. (2014). Cancer statistics, 2014. CA. Cancer Journal for Clinicians. *64(1)*, 9–29.

Sloane, M. A., Nunez, A.C., Packham, D., Kwok, C.-T., Suthers, G., Hesson, L.B., and Ward, R.L. (2015). Mosaic epigenetic inheritance as a cause of early-onset colorectal cancer. JAMA Oncology. *141(5)*, 428.

Søreide, K., Nedrebø, B.S., Knapp, J.C., Glomsaker, T.B., Søreide, J.A., and Kørner, H. (2009). Evolving molecular classification by genomic and proteomic biomarkers in colorectal cancer: Potential implications for the surgical oncologist. Surgical Oncology. *18(1)*, 31–50.

Stegeman, I., de Wijkerslooth, T.R., Stoop, E.M., van Leerdam, M.E., Dekker, E., van Ballegooijen, M., *et al*. (2013). Colorectal cancer risk factors in the detection of advanced adenoma and colorectal cancer. Cancer Epidemiology. *37(3)*, 278–283.

Takayama, T., Miyanishi, K., Hayashi, T., Sato, Y., and Niitsu, Y. (2006). Colorectal cancer: genetics of development and metastasis. Journal of Gastroenterology. *41(3)*, 185–192.

Theodoratou, E., Farrington, S.M., Tenesa, A., McNeill, G., Cetnarskyj, R., Korakakis, E., *et al*. (2014). Associations between dietary and lifestyle risk factors and colorectal cancer in the Scottish population. European Journal of Cancer Prevention. *23(1)*, 8–17.

Tortora, G. J. and B. Derrickson (2012). Principles of Anatomy and Physiology*, 13th Edition*, Wiley**:** 1222.

Townsend, C. M. (2007). Sabiston Textbook of Surgery: Expert Consult: Online & Print, Saunders.

Umar, A., Boland, C.R., Terdiman, J.P., Syngal, S., Chapelle, a. D.L., Ruschoff, J., *et al*. (2004). Revised Bethesda guidelines for hereditary nonpolyposis colorectal cancer (Lynch syndrome) and microsatellite instability. JNCI, Journal of the National Cancer Institute. *96(4)*, 261–268.

Van Cutsem, E., Nordlinger, B., and Cervantes, A. (2010). Advanced colorectal cancer: ESMO clinical practice guidelines for treatment. Annals of Oncology. *21 Suppl 5*, 93–97.

Vargas, A.J., and Thompson, P. (2012). Diet and nutrient factors in colorectal cancer risk. Nutrition in Clinical Practice. *27(5)*, 613–623.

Vasen, H.F. A., Möslein, G., Alonso, A, Bernstein, I., Bertario, L., Blanco, I., *et al*. (2007). Guidelines for the clinical management of Lynch syndrome (hereditary non-polyposis cancer). Journal of Medical Genetics. *44(6)*, 353–362.

Vasen, H.F.A., Watson, P., Mecklin, J.P., and Lynch, H.T. (1999). New clinical criteria for hereditary nonpolyposis colorectal definition of HNPCC. Gastroenterology. *116(6)*. 1453–1456.

Ward, R.L., Dobbins, T., Lindor, N.M., Rapkins, R.W., and Hitchins, M.P. (2012). Identification of constitutional *MLH1* epimutations and promoter variants in colorectal cancer patients from the Colon Cancer Family Registry. Genetics in Medicine. *15(1).* 25–35.

Weisenberger, D.J., Siegmund, K.D., Campan, M., Young, J., Long, T.I., Faasse, M. A., *et al*. (2006). CpG island methylator phenotype underlies sporadic microsatellite instability and is tightly associated with *BRAF* mutation in colorectal cancer. Nature Genetics. *38(7)*, 787–793.

Widmaier, E., Raff, H., Strang, K, Vander'sHuman physiology: The mechanisms of body function, McGraw-Hill Education, 13th edition.

Winawer, S.J., Zauber, A.G., Fletcher, R.H., Stillman, J.S., O'Brien, M.J., Levin, B., Smith, R. A., *et al*. (2006). Guidelines for colonoscopy surveillance after polypectomy: a

consensus update by the US Multi-Society Task Force on colorectal cancer and the American Cancer Society. Gastroenterology. *130(6)*, 1872–1885.

Wolpin, B.M., and Mayer, R.J. (2008). Systemic treatment of colorectal cancer. Gastroenterology. *134(5)*, 1296–1310.

Wong, J.J.L., Hawkins, N.J., and Ward, R.L. (2007). Colorectal cancer: a model for epigenetic tumorigenesis. Gut BMJ Journals. *56(1)*, 140–148.

Worthley, D.L., Whitehall, V.L., Spring, K.J., and Leggett, B. A. (2007). Colorectal carcinogenesis: Road maps to cancer. World Journal of Gastroenterology. *13(28)*, 3784–3791.

Zbuk, K.M., and Eng, C. (2007). Hamartomatous polyposis syndromes. Nature Clinical Practice Gastroenterology Hepatology. *4(9)*, 492–502.