

RESSALVA

Atendendo solicitação do(a) autor(a), o texto completo desta dissertação será disponibilizado somente a partir de 08/04/2023.

**DESACETILASES DE HISTONAS NO CÂNCER COLORRETAL:
EXPRESSÃO GÊNICA, INDICADORES PROGNÓSTICOS
E PERSPECTIVAS PARA A TERAPIA EPIGENÉTICA**

DIOGO PESSUTO CALLEGARI

Botucatu-SP
Março/2021

UNIVERSIDADE ESTADUAL PAULISTA "JULIO DE MESQUITA FILHO"
INSTITUTO DE BIOCIÊNCIAS DE BOTUCATU
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS BIOLÓGICAS - GENÉTICA

DESACETILASES DE HISTONAS NO CÂNCER COLORRETAL:
EXPRESSÃO GÊNICA, INDICADORES PROGNÓSTICOS
E PERSPECTIVAS PARA A TERAPIA EPIGENÉTICA

Diogo Pessuto Callegari
Mestrando

Profa. Dra. Cláudia Aparecida Rainho
Orientadora

Dissertação apresentada ao Instituto de Biociências, Campus de Botucatu, UNESP, como parte dos pré-requisitos necessários para obtenção do título de Mestre no Programa de Pós-Graduação em Ciências Biológicas - Genética.

Botucatu-SP
Março/2021

FICHA CATALOGRÁFICA ELABORADA PELA SEÇÃO TÉC. AQUIS. TRATAMENTO DA INFORM.
DIVISÃO TÉCNICA DE BIBLIOTECA E DOCUMENTAÇÃO - CÂMPUS DE BOTUCATU - UNESP
BIBLIOTECÁRIA RESPONSÁVEL: ROSEMEIRE APARECIDA VICENTE-CRB 8/5651

Callegari, Diogo Pessuto.

Desacetilases de histonas no câncer colorretal :
expressão gênica, indicadores prognósticos e perspectivas
para a terapia epigenética / Diogo Pessuto Callegari. -
Botucatu, 2021

Dissertação (mestrado) - Universidade Estadual Paulista
"Júlio de Mesquita Filho", Instituto de Biociências de
Botucatu

Orientador: Cláudia Aparecida Rainho
Capes: 20205007

1. Cólon (Anatomia) - Câncer. 2. Expressão gênica.
3. Epigenética. 4. Histona desacetilases. 5. Sirtuínas.

Palavras-chave: Alvos terapeuticos; HDACs clássicas;
Marcas epigenéticas; Modificadores de histonas; Sirtuínas.

Agradecimentos

Aos meus pais e ao meu irmão que sempre me auxiliaram e apoiaram, cada qual com seu jeito único.

Aos meus amigos de laboratório *João, Fernanda, Barbara e Naiade* que desde o começo me ensinam sobre ciência, trabalho em grupo e principalmente sobre a vida. Agradeço por tê-los conhecido.

Agradeço imensamente a minha orientadora *Profa. Cláudia Rainho* por participar como chefe da banca da defesa e confiar em meu potencial mesmo sabendo das minhas limitações, sempre buscando a melhor maneira de ensinar e passar seu conhecimento adiante não só para mim, como também para as pessoas em seu entorno. Continuarei agradecendo pela oportunidade de trabalhar com você e conhecer todo esse mundo que você me apresenta ao longo do tempo.

Aos membros da banca de defesa *Dra. Luisa Matos do Canto Alvim e Dr. Danilo Moretti-Ferreira*. Muito obrigado por estarem presentes neste momento tão importante de minha vida e contribuírem com o conhecimento tão valioso de vocês.

Ao *Dr. José Roberto Fígaro Caldeira* e a *Dra Ana Maria de Andrade Caldeira* por iluminarem as pedras que enfrentei como também as que irei enfrentar pelos caminhos da vida, somado a necessidade de encarar isso de peito aberto, muito obrigado por isso e por todo o carinho.

Agradeço também a *Lethícia Valencise* que me acompanha nesse final da jornada, sempre me ensinando calma, paciência, e principalmente o real significado do que é o companheirismo.

Aos docentes e funcionários do Departamento de Genética (atualmente, Departamento de Ciências Químicas e Biológicas – Setor Genética) e do Programa de Pós-graduação em Ciências Biológicas – Genética.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – CAPES, pela bolsa concedida.

“Mas, com o tempo, como ele se recusasse a tornar a se agarrar, a corrente o levantou, livrando-o do fundo, e ele não se machucou nem se magoou mais.”

“E as criaturas mais abaixo no rio, para quem ele era um estranho, exclamaram: ‘Vejam, um milagre! Uma criatura como nós, e no entanto, voa! Vejam, é o Messias que chegou para nos salvar!’

“E aquele que foi carregado pela corrente disse: Não sou mais Messias do que vocês. O rio tem prazer em nos erguer à liberdade, se ousamos nos soltar. O nosso verdadeiro trabalho é essa viagem, essa aventura.”

- Richard Bach. *Illusions* (1977).

Sumário

Lista de Abreviaturas e Siglas.....	7
Resumo.....	13
Abstract.....	15
Capítulo I.....	16
1. Introdução.....	17
1.1. Câncer de intestino: epidemiologia, fatores de risco, classificação.....	17
1.2. Câncer de intestino: fatores epigenéticos e patologia molecular.....	20
1.3. Desacetilases de histonas e o câncer colorretal.....	25
1.4. Efeito das modificações de histonas na resposta à terapia: HDACs e a resistência ao irinotecano em CRCs.....	30
2. Objetivos.....	33
2.1. Objetivo Geral.....	33
2.2. Objetivos específicos.....	33
Capítulo II.....	40
Abstract.....	41
Introduction.....	41
Method and Materials.....	45
Results and Discussions.....	47
Conclusions.....	61
Acknowledgments.....	62
References.....	63
Appendix.....	74
Conclusão.....	104
Apêndices.....	104

Lista de Abreviaturas e Siglas

<i>5-FU</i>	<i>5-fluorouracila</i>
<i>TP53BP1</i>	<i>tumor protein p53 binding protein 1</i>
<i>AC</i>	<i>Acetil Group</i>
<i>APC</i>	<i>Adenomatous polyposis coli</i>
<i>AR</i>	<i>androgen receptor</i>
<i>ATM</i>	<i>ATM serine/threonine kinase</i>
<i>Bcl-2</i>	<i>BCL2 apoptosis regulator</i>
<i>BMF</i>	<i>Bcl2 modifying factor</i>
<i>BRAF</i>	<i>B-Raf proto-oncogene, serine/threonine kinase</i>
<i>BRCA1</i>	<i>BRCA1 DNA repair associated</i>
<i>CIMP</i>	<i>CpG island methylator phenotype</i>
<i>CIMP-H</i>	<i>High CpG island methylator phenotype</i>
<i>CIMP-L</i>	<i>Low CpG island methylator phenotype</i>
<i>CIN</i>	<i>chromosomal instability</i>
<i>CNVs</i>	<i>copy number variations</i>
<i>Co115</i>	<i>cell line (Co-115) from a human colon carcinoma transplanted into nude mice.</i>
<i>COAD</i>	<i>colon adenocarcinoma</i>
<i>CPE</i>	<i>Consensus Purity Estimate</i>
<i>CPT-11</i>	<i>irinotecan</i>
<i>CRC</i>	<i>colorectal cancer</i>
<i>CSL</i>	<i>recombination signal binding protein for immunoglobulin kappa J region</i>
<i>DACT1</i>	<i>dishevelled binding antagonist of beta catenin 1</i>
<i>DCC</i>	<i>DCC netrin 1 receptor</i>

<i>DDR</i>	<i>DNA Damage Repair</i>
<i>DEGs</i>	<i>differential expressed genes</i>
<i>DLD1</i>	<i>human colon carcinoma cells</i>
<i>DNA</i>	<i>deoxyribonucleic acid</i>
<i>DNMT1</i>	<i>DNA-metiltransferase 1</i>
<i>DNMT3A</i>	<i>DNA methyltransferase 3 alpha</i>
<i>DNMT3B</i>	<i>DNA methyltransferase 3 beta</i>
<i>DNMTi</i>	<i>DNA-metiltransferase inhibitor</i>
<i>DNMTs</i>	<i>DNA methyltransferase</i>
<i>DSB</i>	<i>DNA double-strand breaks</i>
<i>FC</i>	<i>Fold-change</i>
<i>FDA</i>	<i>Food and Drug Administration</i>
<i>FOLFIRI</i>	<i>chemotherapy regimen for treatment of colorectal cancer with folinic acid, fluorouracil, and irinotecan</i>
<i>FOLFOX</i>	<i>chemotherapy regimen for treatment of colorectal cancer with folinic acid, fluorouracil, and oxaliplatin</i>
<i>FOXO</i>	<i>forkhead box O</i>
<i>GDC</i>	<i>Genomic Data Commons</i>
<i>GTEX</i>	<i>Genotype-Tissue Expression</i>
<i>H2A</i>	<i>histone H2A</i>
<i>H2B</i>	<i>histone H2B</i>
<i>H3</i>	<i>histone H3</i>
<i>H3K9</i>	<i>ninth lysine residue of histone H3 protein</i>
<i>H4</i>	<i>histone H4</i>
<i>H4K16ac</i>	<i>acetylation in the sixteenth lysine residue of the histone H4 protein</i>
<i>H4K20me2</i>	<i>dimethylation in the twentieth lysine residue of the histone H4 protein.</i>
<i>HATs</i>	<i>histone acetyltransferase</i>

<i>HCT116</i>	<i>human colon carcinoma cells</i>
<i>HDAC</i>	<i>histone deacetylase</i>
<i>HDAC1</i>	<i>histone deacetylase 1</i>
<i>HDAC2</i>	<i>histone deacetylase 2</i>
<i>HDAC3</i>	<i>histone deacetylase 3</i>
<i>HDAC4</i>	<i>histone deacetylase 4</i>
<i>HDAC5</i>	<i>histone deacetylase 5</i>
<i>HDAC6</i>	<i>histone deacetylase 6</i>
<i>HDAC7</i>	<i>histone deacetylase 7</i>
<i>HDAC8</i>	<i>histone deacetylase 8</i>
<i>HDAC9</i>	<i>histone deacetylase 9</i>
<i>HDAC10</i>	<i>histone deacetylase 10</i>
<i>HDAC11</i>	<i>histone deacetylase 11</i>
<i>HDACi</i>	<i>histone deacetylase inhibitor</i>
<i>HDMs</i>	<i>histone desmethylase</i>
<i>HMTs</i>	<i>histone methylase</i>
<i>HNPCC</i>	<i>Hereditary Non Polyposis Colorectal Cancer</i>
<i>HOXC6</i>	<i>homeobox C6</i>
<i>HSP70</i>	<i>heat shock protein family A</i>
<i>HSP90</i>	<i>heat shock protein 90</i>
<i>HT-29</i>	<i>human colon carcinoma cell line</i>
<i>ING2</i>	<i>inhibitor of growth family member 2</i>
<i>KIF26B</i>	<i>kinesin family member 26B</i>
<i>KRAS</i>	<i>KRAS proto-oncogene, GTPase</i>
<i>LBH-589</i>	<i>Panobinostat</i>

<i>LDHB</i>	<i>lactate dehydrogenase B</i>
<i>lncRNAs</i>	<i>Long non-coding RNA</i>
<i>LoVo</i>	<i>human colon carcinoma cell line</i>
<i>LV</i>	<i>Leucovorin - folinic acid</i>
<i>MAP syndrome</i>	<i>MUTYH-associated adenomatous polyposis</i>
<i>MAPK</i>	<i>mitogen-activated protein kinase</i>
<i>MECP2</i>	<i>methyl-CpG binding protein 2</i>
<i>MEF2</i>	<i>Myocyte enhancer factor 2</i>
<i>MGMT</i>	<i>O-6-methylguanine-DNA methyltransferase</i>
<i>MLH1</i>	<i>mutL homolog 1</i>
<i>MMP13</i>	<i>matrix metalloproteinase 13</i>
<i>MMR</i>	<i>mismatch repair</i>
<i>mRNA</i>	<i>Messenger Ribonucleic acid</i>
<i>MSH2</i>	<i>mutS homolog 2</i>
<i>MSH6</i>	<i>mutS homolog 6</i>
<i>MSI</i>	<i>microsatellite instability</i>
<i>MSI-H</i>	<i>High microsatellite instability</i>
<i>MSI-L</i>	<i>Low microsatellite instability</i>
<i>MSS</i>	<i>microsatellite stability</i>
<i>MUTYH</i>	<i>mutY DNA glycosylase</i>
<i>MyoD</i>	<i>myogenic differentiation</i>
<i>NAD⁺</i>	<i>nicotinamide adenine dinucleotide</i>
<i>Oxa</i>	<i>Oxaliplatin</i>
<i>P300</i>	<i>histone acetyltransferase</i>
<i>p53</i>	<i>tumor protein 53</i>

<i>PCA</i>	<i>Principal component analysis</i>
<i>PI3K</i>	<i>phosphatidylinositol-4,5-bisphosphate 3-kinase</i>
<i>PMS2</i>	<i>PMS1 homolog 2, mismatch repair system component</i>
<i>PPAR-Y</i>	<i>Peroxisome Proliferator-Activated Receptor Gamma</i>
<i>pRb</i>	<i>RB transcriptional corepressor 1</i>
<i>PTMs</i>	<i>post-translational modification</i>
<i>RKO</i>	<i>human colon carcinoma cell line</i>
<i>RNA</i>	<i>Ribonucleic acid</i>
<i>RT-PCR</i>	<i>Real time Polimerase Chain Reaction</i>
<i>S100A2</i>	<i>S100 calcium binding protein A2</i>
<i>SET1</i>	<i>SET domain containing 1</i>
<i>SFN</i>	<i>sulforaphane</i>
<i>SHMT2</i>	<i>serine hydroxymethyltransferase 2</i>
<i>SHP</i>	<i>nuclear receptor subfamily 0 group B member 2</i>
<i>shRNA</i>	<i>Short hairpin RNA</i>
<i>Sir2</i>	<i>NAD-dependent histone deacetylase SIR2</i>
<i>SIRT</i>	<i>sirtuins</i>
<i>SIRT1</i>	<i>Sirtuin 1</i>
<i>SIRT2</i>	<i>Sirtuin 2</i>
<i>SIRT3</i>	<i>Sirtuin 3</i>
<i>SIRT4</i>	<i>Sirtuin 4</i>
<i>SIRT5</i>	<i>Sirtuin 5</i>
<i>SIRT6</i>	<i>Sirtuin 6</i>
<i>SIRT7</i>	<i>Sirtuin 7</i>
<i>SMAD3</i>	<i>SMAD family member 3</i>

<i>SMAD4</i>	<i>SMAD family member 4</i>
<i>SMRT/N-CoR</i>	<i>nuclear receptor corepressor 2</i>
<i>SN-38</i>	<i>7-etil-10-hidroxicamptotecina, active metabolite of irinotecan / irinotecan analog</i>
<i>SW480</i>	<i>human colon carcinoma cell line</i>
<i>TCF</i>	<i>hepatocyte nuclear factor 4 alpha</i>
<i>TCGA</i>	<i>The Cancer Genome Atlas</i>
<i>TNM</i>	<i>TNM staging system</i>
<i>TOPO I</i>	<i>Topoisomerase 1</i>
<i>TP53</i>	<i>tumor protein p53</i>
<i>TSA</i>	<i>tricostatin A</i>
<i>VIP</i>	<i>vasoactive intestinal peptide</i>
<i>WNT</i>	<i>Wnt signaling pathway</i>

Resumo

O termo câncer colorretal (CRC) inclui os tumores localizados no intestino grosso separado em cólon e reto. Globalmente, o CRC apresenta alta incidência e taxa de mortalidade. A doença é heterogênea, possui etiologia complexa, desenvolvendo-se em múltiplas etapas em decorrência do acúmulo de alterações genéticas e epigenéticas. Estudos recentes identificaram perfis moleculares específicos caracterizados por alterações genéticas e epigenéticas: i) instabilidade cromossômica (CIN), ii) instabilidade de microssatélites (MSI), iii) fenótipo metilador de ilha CpG (CIMP) e iv) hipometilação global do DNA. Alterações epigenéticas, incluindo mudanças no perfil de metilação do DNA e de modificações pós-traducionais das histonas, são detectadas em todas as etapas do processo de carcinogênese. Neste contexto, a perda da acetilação da cromatina mediada pelas enzimas denominadas desacetilases de histonas (HDACs) têm sido implicadas na progressão do CRC e na resistência à quimioterapia. Este estudo teve como objetivo principal avaliar o padrão de expressão de genes codificadores de desacetilases de histonas em CRC e sua correlação com parâmetros clínico-patológicos. O trabalho encontra-se apresentado em duas partes: o primeiro capítulo apresenta os principais fatores epidemiológicos do câncer colorretal, sua classificação, fatores epigenéticos que acompanham seu desenvolvimento e sua relação com as HDACs, assim como o efeito das modificações de histonas na resposta à terapia. O segundo capítulo apresenta a análise do padrão de expressão dos 18 genes codificadores de desacetilases de histonas em CRC comparado a amostras de intestino normal a partir de dados de RNA-Seq disponíveis nos bancos de dados públicos *The Cancer Genome Atlas* (TCGA) e *Genotype-Tissue Expression* (GTEx) bem como possíveis correlações entre os genes diferencialmente expressos com parâmetros prognósticos. Dentre os dados disponíveis, foram selecionados aqueles obtidos a partir de amostras de tumores com alto nível de pureza de acordo com o valor de CPE (*Consensus Purity Estimate*). As análises foram conduzidas com o pacote DESeq2 para linguagem R (R Bioconductor v 3.8). Foram detectados seis genes diferencialmente expressos ($\log_2\text{FoldChange} > +1$ ou < -1 e $p\text{-value} < 0,0001$; student t-test): os genes *HDAC1*, *2* e *8* apresentaram altos níveis de expressão enquanto os

genes *HDAC 4*, *9* e *10* apresentaram níveis menores nos tumores quando comparados tecido normal. Em conjunto, os genes diferencialmente expressos mostraram padrão heterogêneo entre as amostras tumorais e foram associados com a ocorrência de mutações no oncogene *BRAF* (genes *HDAC1* e *HDAC10*), instabilidade de microssatélites (gene *HDAC2*) e o fenótipo CIMP (gene *HDAC10*). Considerando-se as interações entre os diferentes membros da família das HDACs e a participação dessas proteínas em complexos repressores da transcrição, novos estudos são necessários para se investigar o impacto das alterações descritas na biologia tumoral e suas implicações para as estratégias de terapia epigenética.

Palavras chave: alvos terapêuticos; HDACs clássicas; marcas epigenéticas; modificadores de histonas; sirtuínas

Abstract

The term colorectal cancer (CRC) includes tumors located in the large intestine composed by colon and rectum. Globally, CRC has high incidence and mortality rates. This disease is heterogeneous, presents a complex etiology and develops in a multistep process, due to the accumulation of genetic and epigenetic changes. Recent advances have identified specific molecular profiles characterized by genetic and epigenetic alterations: *i*) chromosomal instability (CIN), *ii*) microsatellite instability (MSI), *iii*) CpG island-methylating phenotype (CIMP), and *iv*) global DNA hypomethylation. Epigenetic alterations can be detected during all stages of carcinogenesis, including DNA methylation and post-translational modifications of histones. In this context, the loss of chromatin acetylation mediated by histone deacetylases (HDACs) has been implicated in the progression and resistance to chemotherapy of CRC. This study evaluated the expression pattern of the encoding genes of histone deacetylases in CRC and its correlation with clinical-pathological parameters. Here, this work is presented in two parts: the first chapter summarizes the main epidemiological features of CRC, its classification, epigenetic factors that contribute to its development, and its relationship with HDACs, as well as the effect of histone modifications on the response to therapy. The second chapter presents the analysis of the expression pattern of the 18 histone deacetylase genes of CRC, compared to normal intestine samples using RNA-Seq data available in the public databases The Cancer Genome Atlas (TCGA) and Genotype-Tissue Expression (GTEx), respectively. We also investigate the possible correlations between differentially expressed genes with prognostic parameters. Among the available data, those samples obtained from tumors with a high level of purity were selected, according to the CPE value (Consensus Purity Estimate). The analysis was conducted with the DESeq2 package for R language (R Bioconductor v 3.8). Six differentially expressed genes were detected ($\log_2\text{FoldChange} > +1$ or < -1 and $p\text{-value} < 0.0001$; student t-test). The *HDAC1*, *2*, and *8* genes showed high levels of expression while *HDAC4*, *9*, and *10* presented lower levels in the tumors compared to normal

tissue. The differentially expressed genes showed a heterogeneous pattern among tumor samples and were associated with the occurrence of mutations in the *BRAF* oncogene (*HDAC1* and *HDAC10*), microsatellite instability (*HDAC2*), and the CIMP phenotype (*HDAC10*). Based on the protein-protein interactions of members of the HDAC family and their participation in repressor complexes of transcription, further studies are needed to investigate the impact of these changes in tumor biology features, and their implications for epigenetic therapy strategies.

Key words: therapeutic targets; classic HDACs; epigenetic marks; histone modifiers; sirtuins.

1. Introdução

1.1. Câncer de intestino: epidemiologia, fatores de risco, classificação

O termo câncer de intestino inclui os tumores que se desenvolvem no intestino grosso caracterizado pelas regiões do cólon ascendente próximo ao apêndice, transverso, descendente seguido da região do sigmoide e por fim os tumores encontrados no reto (final do intestino) (Figura 1), sendo também denominados como câncer de cólon e reto ou colorretal (CRC, do inglês *colorectal cancer*). Globalmente, esses tumores representam o terceiro tipo de câncer mais comum e a segunda causa de morte por câncer; estimativas recentes indicam até 1,9 milhões de novos casos de CRC na população mundial e 935.173 mortes por ano [1]. No Brasil, foram estimados 41.010 novos casos de CRC em 2020 com 20.540 acometendo homens e 20.470 diagnosticados em mulheres [2]. As opções de tratamento para o câncer de intestino incluem a cirurgia, quimioterapia, radioterapia e imunoterapia. A escolha da modalidade terapêutica depende da localização, tamanho e estágio de progressão da doença. Quando diagnosticados precocemente, os protocolos terapêuticos são mais efetivos, sendo frequentemente relacionados à cura em estágios iniciais, o que destaca a importância da detecção precoce no manejo do CRC [3].

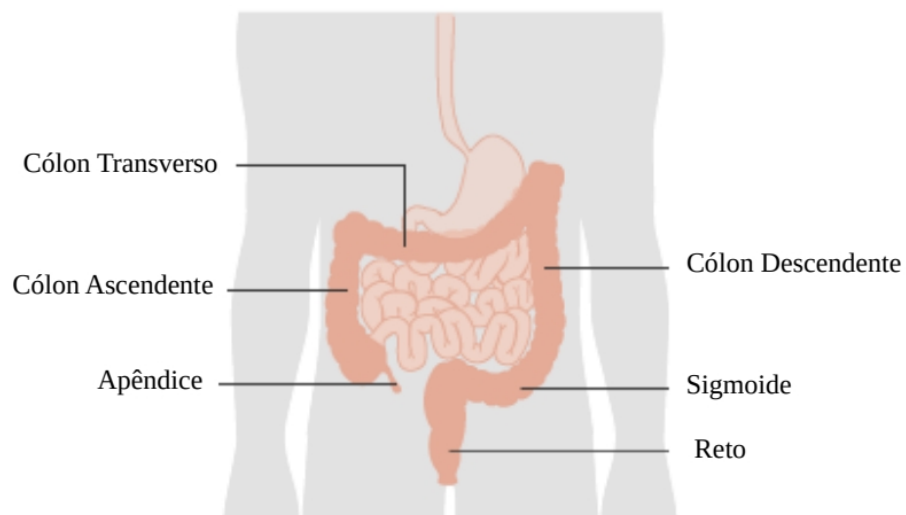


Figura 1. Divisão do intestino grosso representado pelas regiões do cólon ascendente, transverso, sigmoide e reto. A imagem também retrata o apêndice localizado no início do intestino grosso. Imagem retirada do site <https://www.cancerresearchuk.org/about-cancer/bowel-cancer/about-bowel-cancer> (com modificações).

Não há um fator de risco único associado ao desenvolvimento da maioria dos casos de CRC esporádicos. Além da idade e sexo, diferentes estudos epidemiológicos identificaram fatores de risco ambientais e genéticos [4]. Os fatores ambientais incluem consumo exagerado de carne vermelha, carne processada, uso de tabaco e bebidas alcoólicas [5]. Já a ingestão de frutas e vegetais ricos em fibras, acompanhada de atividades físicas regulares e controle do peso corporal foram indicados como fatores protetores em alguns estudos [6,7].

Entre os fatores genéticos, o histórico familiar foi associado ao aumento do risco relativo em aproximadamente duas vezes entre parentes de primeiro grau do paciente, especialmente em jovens. Os indivíduos considerados como de alto risco para o desenvolvimento da doença incluem os portadores de mutações em genes específicos associados às síndromes de cânceres hereditários. Alguns exemplos envolvem: a polipose adenomatosa familiar (gene *APC*), a síndrome de Lynch/*HNPCC* (genes *MLH1*, *MSH2*, *MSH6* e *PMS2*) e as síndromes com pólipos hamartomatosos (síndrome de Peutz-Jeghers, síndrome da polipose juvenil e síndrome de Cowden). No conjunto, estima-se que essas síndromes são responsáveis por aproximadamente 5% dos casos de câncer colorretal [4].

Clinicamente, os CRCs são classificados de acordo com a extensão da invasão local (T), comprometimento de linfonodos regionais (N) e a presença de metástases (M). Esse sistema de estadiamento combinado é utilizado para definir a extensão da doença ao diagnóstico (classificação TNM) e proporciona a base para as decisões terapêuticas (Figura 2) [8].

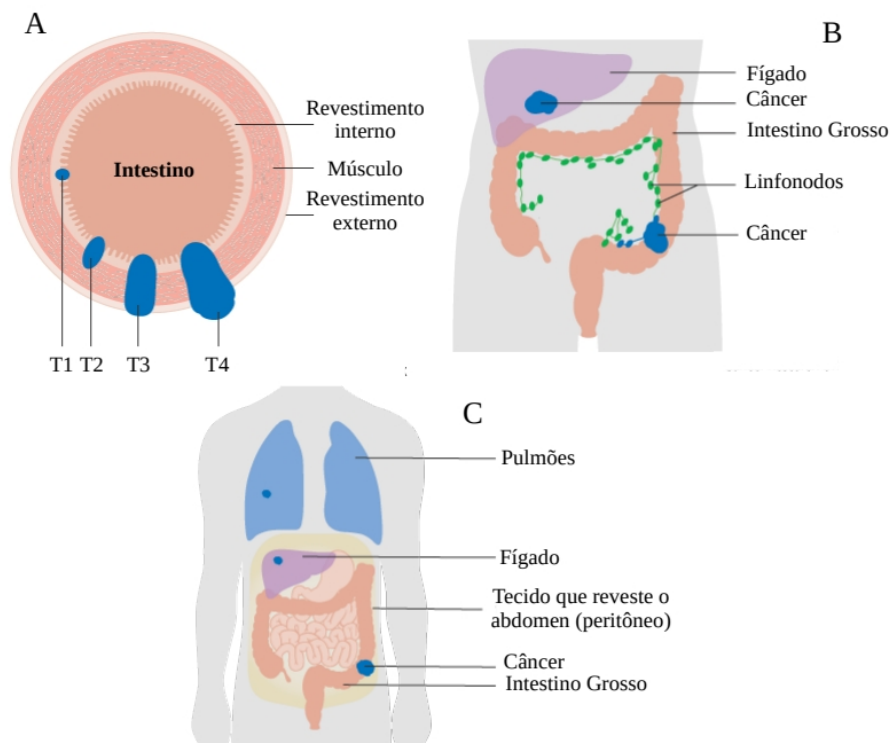


Figura 2. O estadiamento baseado na classificação tumor, linfonodo e metástase (TNM) é um dos sistemas usados para o estadiamento do câncer colorretal. Existem 4 estágios de tamanho do tumor no câncer de intestino (A) T1: tecido tumoral está contido apenas na camada interna do intestino. T2: presença do tecido tumoral na camada muscular da parede intestinal. T3: o tumor inicia seu crescimento para o revestimento externo da parede intestinal, mas não através dele. T4: nesse ponto a metástase acontece, o tumor cresce através do revestimento externo da parede do intestino e se espalha para outras regiões. Nos últimos estágios (T3 e T4), a invasão do câncer nos linfonodos acontece (B), seguida da metástase para órgãos próximos (C). Imagem retirada do site <https://www.cancerresearchuk.org/about-cancer/bowel-cancer/stages-types-and-grades/TNM-staging> (com modificações).

1.2. Câncer de intestino: fatores epigenéticos e patologia molecular

A epigenética dedica-se ao estudo das mudanças na expressão gênica que não são atribuídas às alterações na sequência de bases nitrogenadas do DNA. Globalmente, os mecanismos epigenéticos são mediados pela metilação do DNA, modificações pós-traducionais das histonas e pelos RNAs não codificadores, sendo fundamentais no remodelamento da conformação da cromatina e na regulação da expressão gênica [9]. Os mecanismos epigenéticos são essenciais para a determinação e diferenciação de células-tronco para a reprogramação de células somáticas durante o desenvolvimento normal e apresentam perfis de distribuição distintos em diferentes tipos celulares e tecidos, determinando diferentes epigenomas [10].

A metilação do DNA é a modificação epigenética melhor estudada em doenças humanas. Em geral, a metilação do DNA refere-se principalmente à ligação covalente de um grupo metil no carbono da posição 5 da citosina que é imediatamente seguida por uma guanina, formando um

dinucleotídeo 5'-CpG-3'. Embora a maioria dos sítios CpG isolados no genoma sejam metilados, as ilhas CpGs associadas a promotores gênicos encontram-se não metiladas [11]. Alterações epigenéticas estão implicadas na carcinogênese: a perda dos perfis normais de metilação do DNA, tais como a hipo ou a hipermetilação, acompanham o avanço natural da doença [12]. Sequências repetitivas e promotores de genes regulados por *imprinting* se tornam hipometilados, resultando em instabilidade genômica ou ativação gênica anormal, respectivamente [13]. Adicionalmente, a hipermetilação de promotores específicos de genes supressores tumorais, levando ao seu silenciamento, tem sido implicada na origem e progressão de diversos tipos de cânceres [14].

O câncer é fundamentalmente uma doença decorrente do acúmulo de alterações genéticas e epigenéticas [15]. Nas últimas décadas, houve uma grande expansão do conhecimento sobre as bases moleculares do processo de carcinogênese. Os avanços no entendimento dos mecanismos moleculares fundamentais para o desenvolvimento do CRC levaram à identificação de marcadores prognósticos e preditivos de resposta ao tratamento. Por exemplo, as mutações ativadoras do oncogene *KRAS* e as mutações de perda de função dos genes de reparo de mau pareamento são atualmente utilizadas na prática oncológica: esses marcadores fornecem informações para a estratificação de risco e para a escolha das melhores opções de tratamento em CRC avançados. No entanto, atualmente não há marcadores prognósticos confiáveis para pacientes portadores de CRC em estágios iniciais que poderiam se beneficiar da quimioterapia adjuvante [16].

No conjunto, a doença caracteriza-se por heterogeneidade intra e inter-tumoral, podendo progredir por vias moleculares distintas: a via adenoma-carcinoma, considerada o modelo convencional inclui a maioria dos casos e a via alternativa denominada “*serrated pathway*” ou via serrilhada. O perfil de alterações moleculares se relaciona com as diferenças observadas no prognóstico e na resposta aos tratamentos nos casos avançados de CRC [17,18].

Acredita-se que uma fração importante dos tumores colorretais origina-se de neoplasias benignas denominadas pólipos adenomatosos, podendo evoluir para adenocarcinomas ao longo do tempo (~60% dos casos). Dependendo da arquitetura glandular, pleomorfismo celular e padrão da

secreção de muco, os adenocarcinomas podem ser categorizados em três graus de diferenciação: bem diferenciado (grau I), moderadamente diferenciado (grau II) e indiferenciado (grau III) [19].

A via de progressão adenoma-carcinoma é acompanhada pelo acúmulo de mutações somáticas adquiridas, incluindo a ativação de oncogenes (por exemplo, *KRAS*) e da inativação em genes supressores de tumor, como o gene *TP53*. A progressão em múltiplas etapas também inclui a instabilidade cromossômica, evidenciada pelo acúmulo de alterações cromossômicas numéricas e estruturais [17]. Essa via é considerada convencional e caracteriza-se morfológicamente pela presença de adenomas tubulosos como lesões precursoras. A progressão dessas lesões é acompanhada pelo acúmulo de múltiplas mutações adquiridas. Em nível molecular, essa via caracteriza-se dois mecanismos principais [20]:

- Instabilidade Cromossômica (CIN): evidenciada pelo acúmulo de alterações cromossômicas numéricas e estruturais, mutações somáticas que afetam vias controladas por genes supressores de tumor, tais como *DCC*, *APC* (gene *APC* regulador da via de sinalização WNT), *SMAD4* e *TP53*, e também mutações ativadoras em oncogenes (por exemplo, *KRAS*). As vias de sinalização Wnt/ β -catenin, MAPK, PI3K e TGF- β também estão comprometidas. Correspondem a 60-80% dos CRCs esporádicos.

- Instabilidade de Microssatélites (MSI): Caracterizado pela baixa expressão dos genes envolvidos no sistema reparo do tipo MMR (*Mismatch Repair*), tais como os genes *MLH1* (*mutL homolog 1*), *MSH2* (*mutS homolog 2*), *MSH6* (*mutS homolog 6*) e *PMS2* (*PMS1 homolog 2*), alta frequência de mutação nos genes *BAX*, *TGF β R2* e *BRAF*, e baixa frequência de mutações em *TP53*, *SMAD4* e *KRAS*. A deficiência no processo de reparo do DNA de mau pareamento leva ao acúmulo de pequenas inserções e de deleções nas repetições de nucleotídeos denominadas “*short tandem repeats*” ou microssatélites. As mutações nos genes MMR e a instabilidade de microssatélites (MSI, do inglês *Microsatellite Instability*) são observadas em 15 a 20% dos CRCs esporádicos, sendo também utilizados como marcadores associados a síndrome de predisposição HNPCC (3% dos CRCs) [21,22].

A presença de MSI representa a evidência fenotípica de que o MMR não está funcionando normalmente e a análise de um painel de microssatélites é utilizada para classificar os CRCs em alta instabilidade de microssatélites (MSI-H), baixa instabilidade de microssatélites (MSI-L) e estabilidade de microssatélites (MSS). Estudos iniciais, indicaram que os tumores classificados como MSI-L não diferiram dos considerados MSS, demonstrando praticamente nenhuma diferença clínico-patológica [23]. Já os tumores com MSI-H possuem como principal característica a presença de hipermetilação com consequente silenciamento e inativação do gene *MLH1* (presente em 80% dos CRCs com MSI) [21,22]. O gene *MLH1* codifica uma proteína com função no reconhecimento de eventos do tipo mau pareamento de bases, fundamental no processo de MMR [24].

A segunda via ou via alternativa denominada “*serrated pathway*”, destaca-se pela presença de adenomas ou pólipos serrilhados como lesões precursoras (10 a 30% dos casos de CRC esporádicos). Os pólipos serrilhados são caracterizados por criptas luminiais com arquitetura glandular serrilhada [24] e podem ser classificados em três subtipos principais: pólipos hiperplásicos; adenoma serrilhado séssil com ou sem displasia e adenoma serrilhado tradicional [25].

Em nível molecular a via serrilhada é caracterizada por mutações somáticas adquiridas nos oncogenes *BRAF* e *KRAS* e alta frequência de metilação do DNA em dinucleotídeos CpGs caracterizando o fenótipo CIMP (*CpG Island Methylator Phenotype*) [26,27]. Em adição, a maioria dos CRCs que se origina a partir dessas lesões precursoras também apresenta MSI-H decorrente da inativação do gene *MLH1* por hipermetilação da sua região promotora e deficiência de reparo do DNA do tipo excisão de base associada ao gene *MUTYH* (*MAP syndrome, MUTYH associated polyposis syndrome*) [28].

O fenótipo hipermetilador de Ilhas CpG (CIMP) caracteriza-se pelo silenciamento de múltiplos genes supressores de tumor por mecanismos epigenéticos, podendo ou não apresentar mutações no oncogene *BRAF* [29]. O fenótipo CIMP pode ser classificado em três tipos de acordo com a presença de MSI e do nível de metilação encontrado nos genes *MLH1* e *MGMT* [21]:

- CIMP-High com MSI: MSI devido metilação no gene *MLH1* e mutação em *BRAF*;
- CIMP-High: gene *MLH1* parcialmente metilado, com presença de MSS e mutações em *BRAF*.
- CIMP-Low: gene *MGMT* metilado, com presença de MSS e mutações em *KRAS*.

Atualmente diversos programas e recursos computacionais estão sendo explorados para a identificação das vias moleculares e biomarcadores para os diferentes tipos de cânceres. O uso de ferramentas computacionais para as análises de expressão genica *in silico* permite identificar genes diferencialmente expressos comparados ao tecido normal, auxiliando na pesquisa de possíveis marcadores da doença. A identificação de genes diferencialmente expressos no câncer é uma etapa importante para a detecção de genes alvos, elegíveis para análises funcionais *in vivo* ou *in vitro* subsequentes. Alguns recursos de exploração visual como o *UCSC Xena browser* [30] permitem ao usuário avaliar a relação entre *datasets* genômicos e/ou das variáveis fenotípicas, incluindo 127 *cohorts* e 1559 *datasets* provenientes dos mais variados *hubs* como *The Cancer Genome Atlas* (TCGA) ou *Genomic Data Commons* (GDC) [31,32]. Por exemplo, estudos como o realizado por Alajez e colaboradores (2016) utilizaram as ferramentas computacionais citadas acima para o entendimento da heterogeneidade intra e inter-tumoral encontrada no CRC. Os autores conduziram análises de expressão gênica baseada em dados normalizados de microarranjos, utilizando valores de 1.5 de *fold-change* (FC) e $p < 0.05$ para filtragem em três *datasets* diferentes para CRC e identificaram cinco genes altamente expressos (*S100A2*, *VIP*, *HOXC6*, *DACT1*, *KIF26B*) associados a uma baixa sobrevida dos pacientes [33].

1.3. Desacetilases de histonas e o câncer colorretal

As modificações das histonas representam um segundo nível de regulação epigenética. O nucleossomo é a unidade básica estrutural e funcional da cromatina. É formado por um octâmero de histonas envolvido pela molécula de DNA (~146pb). As histonas H2A, H2B, H3 e H4 detêm informações epigenéticas determinadas pelas modificações pós-traducionais em aminoácidos específicos, geralmente localizados na região NH₂ terminal dessas proteínas. Enquanto as modificações epigenéticas do DNA são moduladas pelas enzimas denominadas DNA metiltransferases (DNMT1, DNMT3A e DNMT3B) e desmetilases (TETs), proteínas específicas também são responsáveis pela adição e remoção de grupos químicos como os radicais acetil e metil em aminoácidos específicos nas histonas, levando a mudanças no nível de compactação da cromatina [34]. Por exemplo, a acetilação de resíduos de lisina das histonas neutraliza a carga positiva desses aminoácidos acarretando na diminuição da interação eletrostática entre a histona e o DNA, carregado negativamente [35]. Consequentemente, observa-se a maior relaxamento na conformação da eucromatina quando a acetilação das lisinas coincide com regiões promotoras, favorecendo a transcrição (Figura 3).

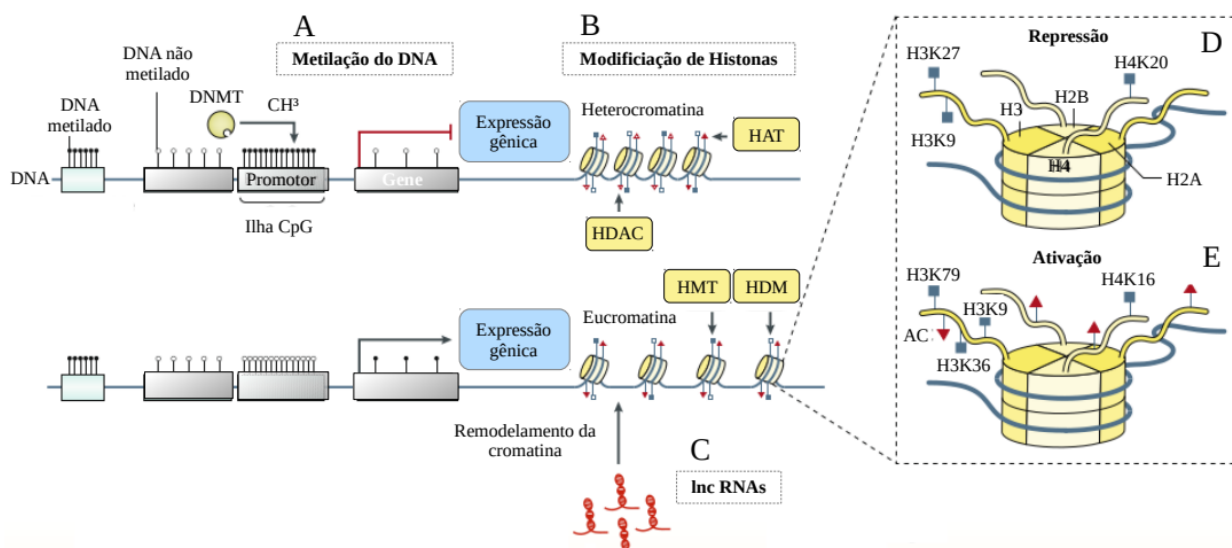


Figura 3. As principais modificações epigenéticas no câncer colorretal incluem: a hipermetilação de ilhas CpGs em regiões promotoras de genes supressores de tumores (A) por DNA metiltransferases (DNMTs) [37,38], inibindo a expressão gênica e favorecendo a tumorigênese e a expressão de proto-oncogenes [39]; modificações de histonas pela inclusão e retirada de grupos acetil (AC, identificados pelo triângulo vermelho) realizadas pelas histonas acetiltransferases (HATs) e desacetilases de histonas (HDACs), respectivamente, geralmente enfraquecendo o estado de compactação da cromatina e tornando o DNA acessível aos fatores de transcrição [40]. A metilação das histonas é regulada por histonas metiltransferases (HMTs) e histonas desmetilases (HDMs) que criam locais de encaixe para proteínas específicas, podendo reprimir ou aumentar a expressão gênica (B, D e E) [41]; Remodelamento da cromatina por intermédio dos RNAs longos não codificantes (lncRNAs) influenciando a expressão gênica (C) [42]. Fonte: Ref [36] (com modificações).

As metiltransferases (HMTs), desmetilases (HDMs), acetiltransferases (HATs) e a superfamília das desacetilases (HDACs) de histonas são exemplos de efetores epigenéticos [43]. A remoção do radical acetil das lisinas pelas HDACs está associada à compactação da cromatina e à ausência de transcrição [44]. Em particular, a acetilação de resíduos de lisina nas histonas H3 e H4 em CRC vem sendo investigada em diversos estudos que buscam associar as modificações de histonas com a progressão do CRC [45-47]. Outros tipos de modificações de histonas também ocorrem como a sumoilação, ubiquitinação de lisinas, fosforilação de serinas e treoninas [48].

Nas células humanas foram identificadas 18 representantes na família das HDACs, tipicamente divididas em quatro classes: as classes I, II (IIa/IIb) e IV são consideradas desacetilases clássicas e são dependentes da ligação de uma molécula de zinco para atividade de desacetilação. Já as desacetilases denominadas sirtuínas, por diferirem no seu mecanismo de ação dependente de uma molécula de NAD⁺, são agrupadas como classe III, sendo homólogas a família Sir2 de *Saccharomyces cerevisiae* [49]. As HDACs clássicas geralmente localizam-se no núcleo e possuem expressão ubíqua; no entanto, a localização citoplasmática e a expressão restrita a tecidos específicos também foram detectadas (por exemplo, expressão tecido específica da HDAC3) [50].

Além da atividade sobre as histonas, as HDACs apresentam vários substratos alternativos como alvo, que participam de diferentes funções nos processos celulares envolvidos no câncer [46]. As HDACs de classe I, apresentam no geral substratos alternativos relacionados a proliferação e sobrevivência celular, como exemplos as proteínas: pRb, SHP, BRCA1, MECP2, ATM, MEF2, MyoD, p53, NF- κ B, AR, DNMT1, HSP70. Já as HDACs de classe II participam nas modificações das proteínas 14-3-3, SHP, HSP90 e α -tubulina, relacionadas a regulação do citoesqueleto e mobilidade, assim como funções celulares endoteliais de gliconeogênese, crescimento dos miócitos cardíacos, diferenciação dos timócitos e na recombinação homóloga durante o reparo de DNA. Outros exemplos de processos celulares como: autoimunidade, envelhecimento, sobrevivência, migração e invasão celular, foram relacionados as modificações nos substratos não-histônicos PPAR- γ , p53, P300, α -tubulina e FOXO por intermédio das HDACs de classe III (SIRT1 e SIRT2) [51].

As HDACs de classe II (HDAC4, HDAC6, HDAC9 e HDAC10) foram implicadas nos mecanismos de resposta aos danos causados ao DNA (DDR, DNA Damage Repair), como o reparo das quebras de fita dupla (DSB, DNA *double-strand breaks*). Estudos iniciais demonstraram que a depleção da HDAC4 reduziu a expressão da proteína 53BP1, não possibilitando a regeneração das DSBs na fase G2 [52]. A importância das HDACs no reparo de quebras de DNA por recombinação homóloga foi indicada pelos estudos de Kotian et al. (2011). Esses autores demonstraram que, além do *knockdown* da HDAC9 ou da HDAC10, o tratamento das células HeLa-DR com inibidores de HDACs resultou em maior sensibilidade à mitomicina C [53].

As HDACs também foram implicadas em outras vias de reparo a danos do DNA. O gene *MSH2*, como já citado anteriormente, atua no sistema de MMR. Esse gene é diretamente regulado por HDACs de classe IIb, especificamente a HDAC6. A desacetilação, seguida da ubiquitinação de *MSH2*, foi associada à tolerância aos danos do DNA [54]. Por outro lado, quando a desacetilação de *MSH2* é mediada pela HDAC10, ocorre um aumento na atividade do reparo de DNA [55]. As HDACs de classe III também foram relacionadas ao reparo do DNA, como a SIRT1 e SIRT6 envolvidas diretamente na resposta a DDR [56-59].

O aumento da expressão das HDACs, principalmente das enzimas de classe I, foi detectado em CRC: as HDACs 1, 2 e 3 além de promover a proliferação em linhagens celulares de CRC (HT-29 e SW480), também foram relacionados com a diminuição da sobrevida dos pacientes [60,61]. Altos níveis de expressão do gene codificador da HDAC2 também foram relacionados tanto a pacientes com câncer retal não responsivos ao tratamento neoadjuvante com 5-fluorouracil [62], como também aos eventos iniciais do desenvolvimento do CRC [63]. A expressão da desacetilase SIRT1 foi associada a marcadores prognósticos conhecidos, como a MSI e o fenótipo CIMP em CRCs. Nosho et al. (2009) detectaram alta expressão da SIRT1 por imuno-histoquímica associada ao fenótipo CIMP-H com presença de MSI. Porém, sem associações significantes para eventos moleculares e características clínicas incluindo sobrevida dos pacientes e índice de massa corporal [64].

As terapias epigenéticas atuais buscam por benefícios no tratamento de pacientes com câncer pela inibição das enzimas envolvidas no silenciamento epigenético, como o uso de inibidores das DNMTs (DNMTi) e das HDACs (HDACi). A diversidade de substratos sujeitos à ação das HDACs indica que essas enzimas podem interferir em vários processos e vias envolvidas no início e progressão do câncer, justificando o grande interesse no uso dessas proteínas como novos alvos terapêuticos [65]. Alguns estudos também descreveram que o desenvolvimento de resistência a drogas em células tumorais está relacionado diretamente à regulação epigenética [66], resultando na expressão anormal de genes supressores de tumores e genes envolvidos nas vias específicas dos agentes quimioterápicos [67]. Em adição, demonstrou-se que baixas doses de HDACi foram mais efetivas quando comparadas com os efeitos epigenéticos dos inibidores de DNMTs em células tumorais [68].

Estudos *in vitro* também demonstram a associação entre HDACs com a sensibilidade à quimioterápicos. No estudo de Alzoubi et al. (2016), o nível de expressão da HDAC2 foi menor em linhagens apresentando mutações no gene TP53 (SW480 e HT29) em comparação com a linhagem de câncer de cólon HCT116 (TP53 +/+). Em adição, também foi demonstrado que a expressão aumentada de HDAC2 correlacionou-se com resistência a drogas enquanto a sua depleção por shRNA sensibilizou a linhagem celular HT-29 que apresenta resistência a múltiplos quimioterápicos, como a 5-fluorouracila (5-FU) e oxaliplatina (Oxa) [69]. O *knockdown* da HDAC8 na linhagem celular HCT116 levou à inibição da proliferação celular dessas células, assim como em linhagens derivadas de câncer cervical e pulmão [70].

A importância da inibição de HDACs no CRC também foi indiretamente sugerida pelo sulforafano (SFN), um isotiocianato encontrado em vegetais crucíferos como brócolis [71]. No estudo de Dashwood e Ho (2007), os autores demonstraram que o SFN inibiu o crescimento de pólipos intestinais espontâneos em ratos, fornecendo evidências de que o mecanismo de quimioprevenção mediado pelo SFN ocorre por vias epigenéticas associadas a inibição de HDACs. Os autores destacaram que as alterações nos níveis de acetilação de histonas e das atividades de HDACs *in vivo*, indicam que as HDACs podem ser moduladas pelos produtos de origem natural.

1.4. Efeito das modificações de histonas na resposta à terapia: HDACs e a resistência ao irinotecano em CRCs

Por mais de 40 anos, o quimioterápico 5-FU associado à leucovorina (LV, ou ácido folínico) foi utilizado no tratamento de pacientes com CRC metastático, porém as taxas de resposta ao tratamento eram pobres, de aproximadamente 20%, sem melhora significativa nas taxas de sobrevida. No entanto, com o desenvolvimento de novos agentes quimioterápicos e de novas técnicas de administração das medicações, observou-se melhora na eficiência dos tratamentos e a conversão de lesões antes consideradas não operáveis em lesões operáveis após tratamento neoadjuvante [72]. Com a introdução de dois novos quimioterápicos: o irinotecano e a oxaliplatina, as taxas de resposta terapêutica e de conversão para lesões operáveis aumentaram [73]. Esses quimioterápicos foram incluídos nos protocolos adjuvantes: FOLFIRI (5-FU + LV + irinotecano) e FOLFOX (5-FU + LV + oxaliplatina), demonstrando taxas de respostas praticamente idênticas de aproximadamente 55% para ambos os regimes de tratamento [74]. Em adição, novos estudos demonstraram que existem diferenças da resposta a esses protocolos terapêuticos dependendo da posição anatômica do tumor (direita ou esquerda) [75]. No entanto, os mecanismos associados à resistência constitutiva ou adquirida a esses quimioterápicos em CRC ainda não foram elucidados [76].

Em 1996, o irinotecano (CPT-11) foi autorizado pela agência federal dos Estados Unidos *Food and Drug Administration* (FDA) para tratamento do CRC [76]. Essa droga é um composto semissintético derivado da camptotecina e inibe seletivamente a topoisomerase I (TOPO I), aprisionando-a e todo seu complexo ao DNA. Como resultado da inibição da TOPO I, as quebras da dupla fita do DNA (DSB) acumulam-se levando as células à apoptose [77]. Entre 2-8% de CPT-11 é convertido intracelularmente em seu metabólito ativo 7-etil-10-hidroxicamptotecina (SN-38), que possui maior atividade (100 a 1000 vezes) do que o CPT-11 [78,79].

Em mamíferos, o reparo de DSBs é realizado por mecanismos complexos que são sinalizados por modificações nas histonas, envolvendo principalmente a acetilação da lisina 16 (H4K16ac) e a bimetilação da lisina 20 (H4K20me2) da histona H4. Por exemplo, a interação entre fatores de reparo como a proteína 53BP1 com a H4K20me2, modifica a cromatina e atua no reparo de quebras do DNA [80]. Em um estudo recente, Meisenberg et al (2017) desenvolveram modelos *in vitro* de resistência ao irinotecano nas linhagens de câncer colorretal RKO e DLD1, demonstrando que a resistência ao quimioterápico estava associada com a perda da acetilação da lisina 16 da histona H4 (H4K16ac). O mesmo estudo demonstrou um efeito parcial na reversão da resistência ao CPT-11 quando células tumorais foram tratadas com os inibidores de HDAC tricostatina A (TSA) e panobinostat (LBH-589) [81].

Conclusões

- Os resultados das análises para a identificação dos genes diferencialmente expressos, entre os que codificam as desacetilases de histonas, em adenocarcinoma de cólon foram dependentes da localização anatômica do tecido normal usado como referência (cólon sigmoide ou transverso).
- Seis genes codificadores de HDACs clássicas foram detectados como diferencialmente expressos independente da localização do tecido normal: três considerados altamente expressos são pertencentes à classe I (*HDAC1*, *HDAC2* e *HDAC8*); os demais apresentaram regulação negativa e pertencem às classes IIa e IIb (*HDAC4*, *HDAC9* e *HDAC10*).
- A comparação entre amostras pareadas tumoral e normal adjacente confirmou os baixos níveis de expressão da *HDAC9* em adenocarcinomas de cólon.
- O perfil dos genes diferencialmente expressos foi heterogêneo, com a identificação clara de dois subgrupos tumorais. Porém, não foi observado uma diferença nítida entre os mesmos quando correlacionados com os parâmetros clínicos e patológicos.
- Os níveis de expressão dos genes diferencialmente expressos nas amostras tumorais foram correlacionados com importantes parâmetros moleculares, incluindo a presença de mutações no oncogene *BRAF* (*HDAC1* e *HDAC10*), instabilidade de microssatélite (*HDAC2*) e fenotipo CIMP (*HDAC10*).