UNIVERSIDADE ESTADUAL PAULISTA
"JÚLIO DE MESQUITA FILHO"
Campus de Botucatu

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

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

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

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

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

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

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

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 (log2FoldChange >+1 ou < –1 e p-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 (log2FoldChange >+1 or <–1 and p-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 potein-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.

Capítulo I

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].

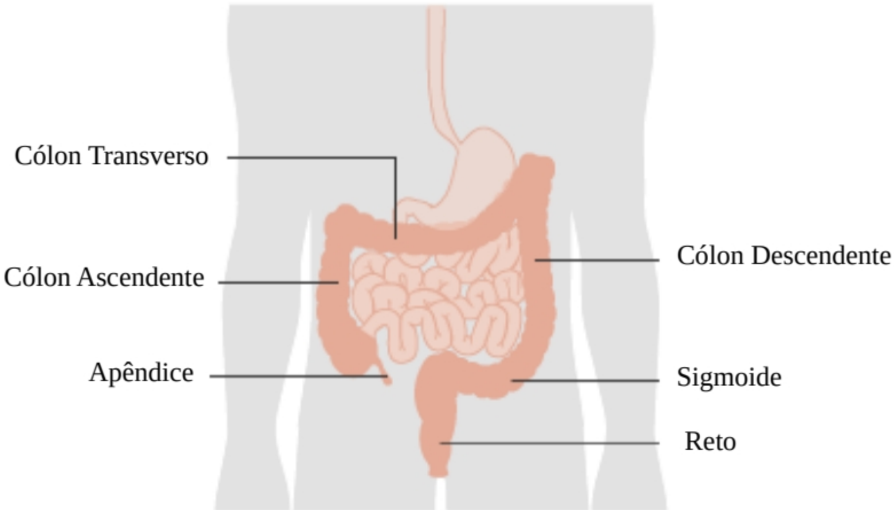

Figura 1. Divisão do intestino grosso representado pelas regiões do cólon ascendente, transverso, sigmoide e reto. A imagem também retrata o apêndice localizado no inicio do intestino grosso. Imagem retirada do site https://www.cancerresearchuk.org/about-cancer/bowel-cancer/about-bowel-cancer (com modificações).

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].

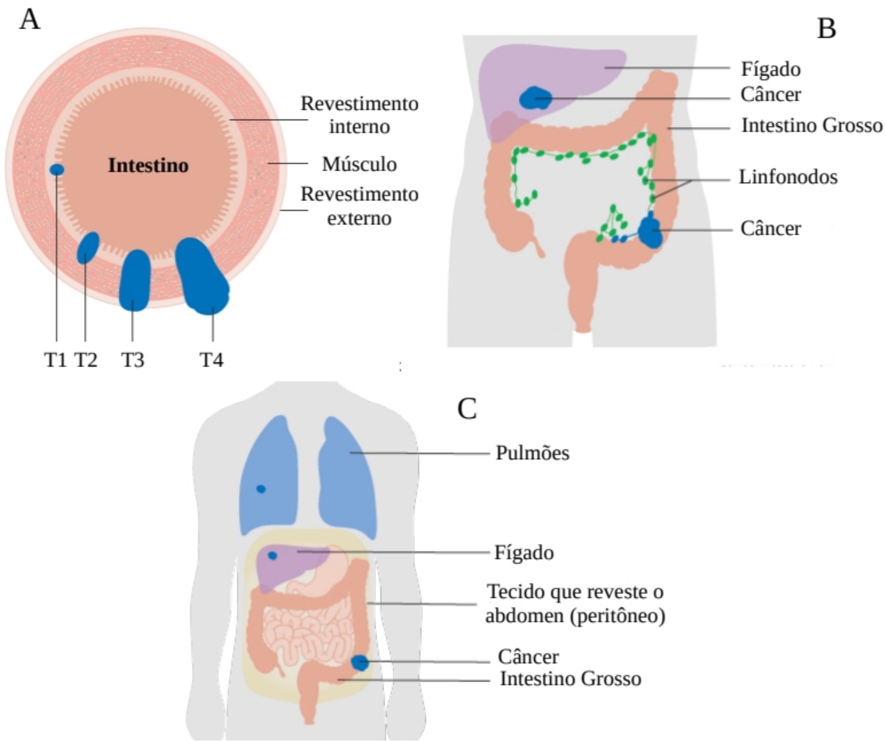

Figura 2. O estadiamento baseado na classificação tumor, linfonodo e metástase (TNM) é um dos sistemas usados para o estadiamento do câncer colorretal. Existem 4 estágios de tamanho do tumor no câncer de intestino (A) T1: tecido tumoral está contido apenas na camada interna do intestino. T2: presença 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 morfologicamente 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*, *TGFβRII* 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 clinico-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 luminais 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 NH_2 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).

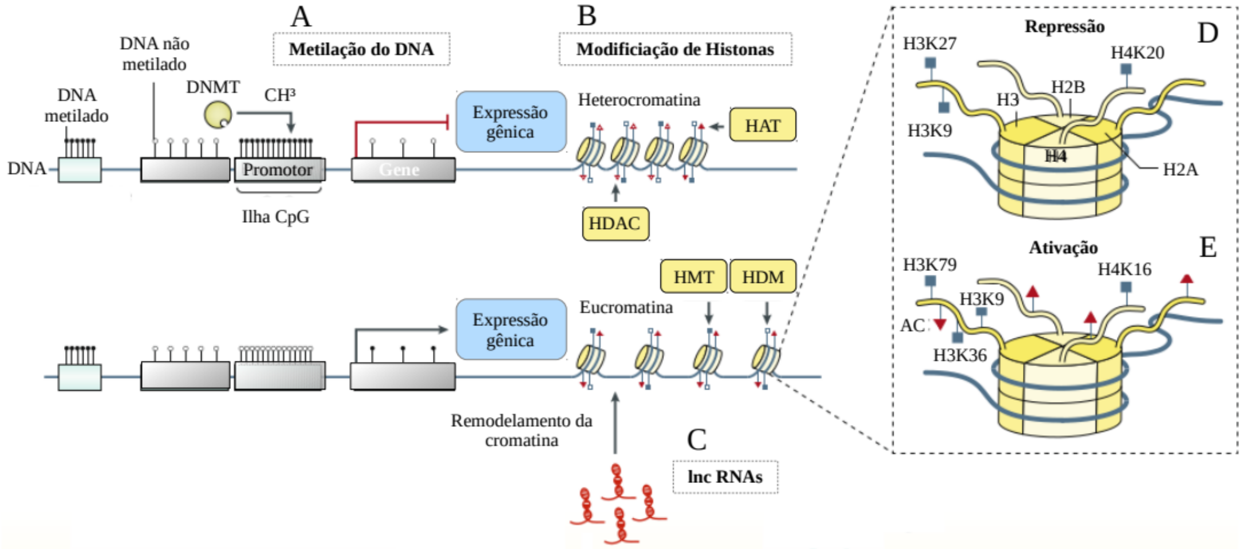

Figura 3. As principais modificações epigenéticas no câncer colorretal incluem: a hipermetilação de ilhas CpGs em regiões promotoras de genes supressores de tumores (**A**) por DNA metiltransferases (DNMTs) [37,38], inibindo a expressão gênica e favorecendo a tumorigênese e a expressão de proto-oncogenes [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 especificas, 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 especifica 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-kβ, 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-Y, 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 clinicas 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 especificas 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 evidencias 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 + oxailiplatina), 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 agencia 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 iritonecano 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 panonbinostat (LBH-589) [81].

2. Objetivos

2.1. Objetivo Geral

O objetivo desse estudo foi caracterizar o padrão de expressão dos genes codificadores de desacetilases de histonas e seus potenciais impactos no prognóstico do câncer colorretal.

2.2. Objetivos específicos

- Recuperação dos dados clínicos e de expressão gênica obtidos por RNA-Seq de tecidos tumorais e normais referentes a pacientes com câncer colorretal do projeto TCGA, assim como dados de tecido normal de cólon do projeto GTEx;

- Criação de *datasets* filtrando valores de expressão dos dezoito genes codificadores das HDACs humanas conhecidas, classificando as amostras tumorais pelo critério de pureza CPE, localização anatômica e tipo de tecido normal;

- Identificação dos genes diferencialmente expressos a partir do painel dos dezoito genes que compõem a família das desacetilases de histonas e análise do seu perfil de expressão em CRC;

- Correlacionar os genes das desacetilases de histonas diferencialmente expressos com parâmetros clínicos e patológicos bem como também com a sobrevida global dos pacientes.

3. Referências

[1] Cancer Tomorrow. 2021. Available at: https://gco.iarc.fr/tomorrow/home. Accessed March 15, 2021.

[2] Tipos de Câncer. INCA - Instituto Nacional de Câncer 2019. Available at: https://www.inca.gov.br/estimativa/sintese-de-resultados-e-comentarios. Accessed March 15, 2021.

[3] Cordeiro F. Diretrizes para diagnóstico, estadiamento e tratamento cirúrgico e multidisciplinar do câncer colorretal. Revista da Associação Médica Brasileira 2004; 50(1): 10–11. doi:10.1590/S0104-42302004000100017.

[4] Brenner H et al. Colorectal cancer. Lancet 2014; 383(9927): 1490–1502. doi:10.1016/S0140-6736(13)61649-9.

[5] Bultman SJ. Interplay between diet, gut microbiota, epigenetic events, and colorectal cancer. Mol Nutr Food Res 2017; 61(1). doi:10.1002/mnfr.201500902.

[6] Fleischauer AT et al. Garlic consumption and cancer prevention: meta-analyses of colorectal and stomach cancers. Am J Clin Nutr 2000; 72(4): 1047–1052. doi:10.1093/ajcn/72.4.1047.

[7] Matsubara K et al. Vitamin B6-mediated suppression of colon tumorigenesis, cell proliferation, and angiogenesis (review). J Nutr Biochem 2003; 14(5): 246–250. doi:10.1016/s0955-2863(03)00009-3.

[8] O'Sullivan B, Brierley J. UICC Manual of Clinical Oncology, 9th Edition | Wiley.

[9] Flavahan WA et al. Epigenetic plasticity and the hallmarks of cancer. Science 2017; 357(6348). doi:10.1126/science.aal2380.

[10] Wang X. Stem cells in tissues, organoids, and cancers. Cell Mol Life Sci 2019; 76(20): 4043–4070. doi:10.1007/s00018-019-03199-x.

[11] Jones MJ et al. DNA methylation and healthy human aging. Aging Cell 2015; 14(6): 924–932. doi:10.1111/acel.12349.

[12] Hammoud SS et al. Epigenetic regulation of colon cancer and intestinal stem cells. Current Opinion in Cell Biology 2013; 25(2): 177–183. doi:10.1016/j.ceb.2013.01.007.

[13] Jones PA, Laird PW. Cancer epigenetics comes of age. Nat Genet 1999; 21(2): 163–167. doi:10.1038/5947.

[14] Toh TB et al. Epigenetics in cancer stem cells. Mol Cancer 2017; 16. doi:10.1186/s12943-017-0596-9.

[15] Como surge o câncer? INCA - Instituto Nacional de Câncer 2018. Available at: https://www.inca.gov.br/como-surge-o-cancer. Accessed March 15, 2021.

[16] Koncina E et al. Prognostic and Predictive Molecular Biomarkers for Colorectal Cancer: Updates and Challenges. Cancers (Basel) 2020; 12(2). doi:10.3390/cancers12020319.

[17] Ferreira CG, Rocha JCC da. Oncologia molecular., 2nd ed. Atheneu, 2010.

[18] Snover DC. Update on the serrated pathway to colorectal carcinoma. Hum Pathol 2011; 42(1): 1–10. doi:10.1016/j.humpath.2010.06.002.

[19] Aaltonen LA et al., eds. Pathology and genetics of tumours of the digestive system. Lyon : Oxford: IARC Press ; Oxford University Press (distributor,), 2000.

[20] De Palma FDE et al. The Molecular Hallmarks of the Serrated Pathway in Colorectal Cancer. Cancers (Basel) 2019; 11(7). doi:10.3390/cancers11071017.

[21] Jass JR. Classification of colorectal cancer based on correlation of clinical, morphological and molecular features. Histopathology 2007; 50(1): 113–130. doi:10.1111/j.1365-2559.2006.02549.x.

[22] Yamagishi H et al. Molecular pathogenesis of sporadic colorectal cancers. Chin J Cancer 2016; 35. doi:10.1186/s40880-015-0066-y.

[23] Ogino S et al. CpG island methylator phenotype, microsatellite instability, BRAF mutation and clinical outcome in colon cancer. Gut 2009; 58(1): 90–96. doi:10.1136/gut.2008.155473.

[24] Szylberg Ł et al. Serrated polyps and their alternative pathway to the colorectal cancer: a systematic review. Gastroenterol Res Pract 2015; 2015: 573814. doi:10.1155/2015/573814.

[25] Rex DK et al. Serrated lesions of the colorectum: review and recommendations from an expert panel. Am J Gastroenterol 2012; 107(9): 1315–1329; quiz 1314, 1330. doi:10.1038/ajg.2012.161.

[26] Barault L et al. Hypermethylator phenotype in sporadic colon cancer: study on a population-based series of 582 cases. Cancer Res 2008; 68(20): 8541–8546. doi:10.1158/0008-5472.CAN-08-1171.

[27] Mojarad EN et al. The CpG island methylator phenotype (CIMP) in colorectal cancer. Gastroenterol Hepatol Bed Bench 2013; 6(3): 120–128.

[28] Moussata D et al. Endoscopic and histologic characteristics of serrated lesions. World J Gastroenterol 2015; 21(10): 2896–2904. doi:10.3748/wjg.v21.i10.2896.

[29] Abdelfatah E et al. Epigenetic therapy in gastrointestinal cancer: the right combination. Therap Adv Gastroenterol 2016; 9(4): 560–579. doi:10.1177/1756283X16644247.

[30] Goldman M et al. The UCSC Xena platform for public and private cancer genomics data visualization and interpretation. bioRxiv 2019: 326470. doi:10.1101/326470.

[31] Tomczak K et al. The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. Contemp Oncol (Pozn) 2015; 19(1A): A68-77. doi:10.5114/wo.2014.47136.

[32] Grossman RL et al. Toward a Shared Vision for Cancer Genomic Data. N Engl J Med 2016; 375(12): 1109–1112. doi:10.1056/NEJMp1607591.

[33] Alajez NM. Large-Scale Analysis of Gene Expression Data Reveals a Novel Gene Expression Signature Associated with Colorectal Cancer Distant Recurrence. PLoS One 2016; 11(12): e0167455. doi:10.1371/journal.pone.0167455.

[34] Jones PA, Baylin SB. The fundamental role of epigenetic events in cancer. Nat Rev Genet 2002; 3(6): 415–428. doi:10.1038/nrg816.

[35] Dawson MA, Kouzarides T. Cancer epigenetics: from mechanism to therapy. Cell 2012; 150(1): 12–27. doi:10.1016/j.cell.2012.06.013.

[36] Goel A, Boland CR. Epigenetics of colorectal cancer. Gastroenterology 2012; 143(6): 1442-1460.e1. doi:10.1053/j.gastro.2012.09.032.

[37] Kanai Y, Hirohashi S. Alterations of DNA methylation associated with abnormalities of DNA methyltransferases in human cancers during transition from a precancerous to a malignant state. Carcinogenesis 2007; 28(12): 2434–2442. doi:10.1093/carcin/bgm206.

[38] Ng JM-K, Yu J. Promoter hypermethylation of tumour suppressor genes as potential biomarkers in colorectal cancer. Int J Mol Sci 2015; 16(2): 2472–2496. doi:10.3390/ijms16022472.

[39] Luo J et al. S-adenosylmethionine inhibits the growth of cancer cells by reversing the hypomethylation status of c-myc and H-ras in human gastric cancer and colon cancer. Int J Biol Sci 2010; 6(7): 784–795. doi:10.7150/ijbs.6.784.

[40] Struhl K. Histone acetylation and transcriptional regulatory mechanisms. Genes Dev 1998; 12(5): 599–606. doi:10.1101/gad.12.5.599.

[41] Huang T et al. Targeting histone methylation for colorectal cancer. Therap Adv Gastroenterol 2017; 10(1): 114–131. doi:10.1177/1756283X16671287.

[42] Guttman M, Rinn JL. Modular regulatory principles of large non–coding RNAs. Nature 2012; 482(7385): 339–346. doi:10.1038/nature10887.

[43] Sharma S et al. Epigenetics in cancer. Carcinogenesis 2010; 31(1): 27–36. doi:10.1093/carcin/bgp220.

[44] Muntean AG, Hess JL. Epigenetic dysregulation in cancer. Am J Pathol 2009; 175(4): 1353–1361. doi:10.2353/ajpath.2009.081142.

[45] Glozak MA, Seto E. Histone deacetylases and cancer. Oncogene 2007; 26(37): 5420–5432. doi:10.1038/sj.onc.1210610.

[46] Li Y, Seto E. HDACs and HDAC Inhibitors in Cancer Development and Therapy. Cold Spring Harb Perspect Med 2016; 6(10). doi:10.1101/cshperspect.a026831.

[47] Patnaik S, Anupriya null. Drugs Targeting Epigenetic Modifications and Plausible Therapeutic Strategies Against Colorectal Cancer. Front Pharmacol 2019; 10: 588. doi:10.3389/fphar.2019.00588.

[48] Patel DJ, Wang Z. Readout of epigenetic modifications. Annu Rev Biochem 2013; 82: 81–118. doi:10.1146/annurev-biochem-072711-165700.

[49] Xu WS et al. Histone deacetylase inhibitors: molecular mechanisms of action. Oncogene 2007; 26(37): 5541–5552. doi:10.1038/sj.onc.1210620.

[50] Seto E, Yoshida M. Erasers of histone acetylation: the histone deacetylase enzymes. Cold Spring Harb Perspect Biol 2014; 6(4): a018713. doi:10.1101/cshperspect.a018713.

[51] Singh AK et al. Targeting Histone Deacetylases with Natural and Synthetic Agents: An Emerging Anticancer Strategy. Nutrients 2018; 10(6). doi:10.3390/nu10060731.

[52] Kao GD et al. Histone deacetylase 4 interacts with 53BP1 to mediate the DNA damage response. J Cell Biol 2003; 160(7): 1017–1027. doi:10.1083/jcb.200209065.

[53] Kotian S et al. Histone deacetylases 9 and 10 are required for homologous recombination. J Biol Chem 2011; 286(10): 7722–7726. doi:10.1074/jbc.C110.194233.

[54] Zhang M et al. HDAC6 deacetylates and ubiquitinates MSH2 to maintain proper levels of MutSα. Mol Cell 2014; 55(1): 31–46. doi:10.1016/j.molcel.2014.04.028.

[55] Radhakrishnan R et al. Histone deacetylase 10 regulates DNA mismatch repair and may involve the deacetylation of MutS homolog 2. J Biol Chem 2015; 290(37): 22795–22804. doi:10.1074/jbc.M114.612945.

[56] Gorospe M, de Cabo R. AsSIRTing the DNA damage response. Trends Cell Biol 2008; 18(2): 77–83. doi:10.1016/j.tcb.2007.11.007.

[57] Mostoslavsky R et al. Genomic instability and aging-like phenotype in the absence of mammalian SIRT6. Cell 2006; 124(2): 315–329. doi:10.1016/j.cell.2005.11.044.

[58] Palacios JA et al. SIRT1 contributes to telomere maintenance and augments global homologous recombination. J Cell Biol 2010; 191(7): 1299–1313. doi:10.1083/jcb.201005160.

[59] Toiber D et al. SIRT6 recruits SNF2H to DNA break sites, preventing genomic instability through chromatin remodeling. Mol Cell 2013; 51(4): 454–468. doi:10.1016/j.molcel.2013.06.018.

[60] Spurling CC et al. HDAC3 overexpression and colon cancer cell proliferation and differentiation. Mol Carcinog 2008; 47(2): 137–147. doi:10.1002/mc.20373.

[61] Weichert W et al. Class I histone deacetylase expression has independent prognostic impact in human colorectal cancer: specific role of class I histone deacetylases in vitro and in vivo. Clin Cancer Res 2008; 14(6): 1669–1677. doi:10.1158/1078-0432.CCR-07-0990.

[62] Canto LM do et al. Locally advanced rectal cancer transcriptomic-based secretome analysis reveals novel biomarkers useful to identify patients according to neoadjuvant

chemoradiotherapy response. Scientific Reports 2019; 9(1): 8702. doi:10.1038/s41598-019-45151-w.

[63] Stypula-Cyrus Y et al. HDAC up-regulation in early colon field carcinogenesis is involved in cell tumorigenicity through regulation of chromatin structure. PLoS One 2013; 8(5): e64600. doi:10.1371/journal.pone.0064600.

[64] Nosho K et al. SIRT1 histone deacetylase expression is associated with microsatellite instability and CpG island methylator phenotype in colorectal cancer. Mod Pathol 2009; 22(7): 922–932. doi:10.1038/modpathol.2009.49.

[65] Marks PA, Xu W-S. Histone deacetylase inhibitors: Potential in cancer therapy. J Cell Biochem 2009; 107(4): 600–608. doi:10.1002/jcb.22185.

[66] Steele N et al. Combined inhibition of DNA methylation and histone acetylation enhances gene re-expression and drug sensitivity in vivo. Br J Cancer 2009; 100(5): 758–763. doi:10.1038/sj.bjc.6604932.

[67] Zeller C et al. Candidate DNA methylation drivers of acquired cisplatin resistance in ovarian cancer identified by methylome and expression profiling. Oncogene 2012; 31(42): 4567–4576. doi:10.1038/onc.2011.611.

[68] Robert C, Rassool FV. HDAC inhibitors: roles of DNA damage and repair. Adv Cancer Res 2012; 116: 87–129. doi:10.1016/B978-0-12-394387-3.00003-3.

[69] Alzoubi S et al. Synergy between histone deacetylase inhibitors and DNA-damaging agents is mediated by histone deacetylase 2 in colorectal cancer. Oncotarget 2016; 7(28): 44505–44521. doi:10.18632/oncotarget.9887.

[70] Vannini A et al. Crystal structure of a eukaryotic zinc-dependent histone deacetylase, human HDAC8, complexed with a hydroxamic acid inhibitor. Proc Natl Acad Sci U S A 2004; 101(42): 15064–15069. doi:10.1073/pnas.0404603101.

[71] Dashwood RH, Ho E. Dietary histone deacetylase inhibitors: from cells to mice to man. Semin Cancer Biol 2007; 17(5): 363–369. doi:10.1016/j.semcancer.2007.04.001.

[72] Giacchetti S et al. Long-term survival of patients with unresectable colorectal cancer liver metastases following infusional chemotherapy with 5-fluorouracil, leucovorin, oxaliplatin and surgery. Ann Oncol 1999; 10(6): 663–669. doi:10.1023/a:1008347829017.

[73] Adam R et al. Rescue surgery for unresectable colorectal liver metastases downstaged by chemotherapy: a model to predict long-term survival. Ann Surg 2004; 240(4): 644–657; discussion 657-658. doi:10.1097/01.sla.0000141198.92114.f6.

[74] Secretaria de Ciência, Tecnologia e Insumos Estratégicos. Cetuximabe em combinação com FOLFIRI ou FOLFOX no tratamento do câncer colorretal metastático KRAS selvagem com

metástases hepáticas exclusivas. 2013. Available at: https://docs.bvsalud.org/biblioref/2017/06/836900/cetuximabe-colorretal-final.pdf.

[75] Shen H et al. Different treatment strategies and molecular features between right-sided and left-sided colon cancers. World J Gastroenterol 2015; 21(21): 6470–6478. doi:10.3748/wjg.v21.i21.6470.

[76] Van der Jeught K et al. Drug resistance and new therapies in colorectal cancer. World J Gastroenterol 2018; 24(34): 3834–3848. doi:10.3748/wjg.v24.i34.3834.

[77] Ashour ME et al. Topoisomerase-mediated chromosomal break repair: an emerging player in many games. Nature Reviews Cancer 2015; 15(3): 137–151. doi:10.1038/nrc3892.

[78] Liew ST, Yang L-X. Design, synthesis and development of novel camptothecin drugs. Curr Pharm Des 2008; 14(11): 1078–1097. doi:10.2174/138161208784246180.

[79] Hicks LD et al. Improved, selective, human intestinal carboxylesterase inhibitors designed to modulate 7-ethyl-10-[4-(1-piperidino)-1-piperidino]carbonyloxycamptothecin (irinotecan; CPT-11) toxicity. J Med Chem 2009; 52(12): 3742–3752. doi:10.1021/jm9001296.

[80] Price BD, D’Andrea AD. Chromatin remodeling at DNA double-strand breaks. Cell 2013; 152(6): 1344–1354. doi:10.1016/j.cell.2013.02.011.

[81] Meisenberg C et al. Epigenetic changes in histone acetylation underpin resistance to the topoisomerase I inhibitor irinotecan. Nucleic Acids Res 2017; 45(3): 1159–1176. doi:10.1093/nar/gkw1026.

Capítulo II

The prognostic value of histone deacetylase family members expression in colorectal cancer

Diogo Pessuto Callegari[1], Fernanda Aparecida dos Santos France [1], João Henrique Maia Assumpção [1], Marcus Vinicius Niz Alvarez [1,2] and Claudia Aparecida Rainho[1]*

1. Institute of Biosciences of Botucatu, Sao Paulo State University, Botucatu, São Paulo, Brazil

2. Institute of Biotechnology, Sao Paulo State University, Botucatu, São Paulo, Brazil

Abstract

Histone deacetylases catalyze the removal of acetyl groups of histones and non-histone proteins: the target subtracts regulate chromatin structure and transcription, as well as control several cancer-related processes. In humans, the histone deacetylase family contains 18 enzymes that use either zinc or NAD (+)-dependent mechanisms. Here, RNA-Seq expression data from The Cancer Genome Atlas (TCGA) and Genotype-Tissue Expression (GTEx) were used to investigate the expression patterns of genes encoding histone deacetylases in colorectal cancer (CRC), including colon and rectal adenocarcinomas (TCGA-COAD and TCGA-READ, respectively). Tumor samples were selected according to tissue purity and compared to normal specimens. We detected six differentially expressed genes (DEGs) in high-purity COAD: *HDAC1, HDAC2,* and *HDAC8* were up-regulated, and *HDAC4, HDAC9*, and *HDAC10* were down-regulated. Further, exploratory clustering analysis revealed a heterogeneous expression pattern with two subgroups: A (n=153 cases) and B (n=30 cases). The expression levels of *HDAC1* and *HDAC10* were correlated with the occurrence of mutations in the *BRAF* oncogene; lower expression levels of *HDAC2* and *HDAC10* were correlated with MSI-high and CIMP-low tumors, respectively. Considering the protein-protein interactions among histone deacetylases and other important regulator complexes, future investigations are necessary to confirm the use of HDAC as biomarkers as well the implications of *HDAC* differential expression patterns for the epigenetic therapy.

Keywords: epigenetic marks, histone modifications, classic HDACs, sirtuins, therapeutic target, colorectal cancer prognosis

1. **Introduction**

Colorectal cancer (CRC), including adenocarcinomas arising from the colon (COAD) and rectal regions (READ), represents the third most common type of cancer in the world and is the second leading cause of death by cancer. Recent estimates indicate approximately 1.9 million new cases and 935,173 deaths per year worldwide [1]. CRC is a disease with complexetiology that emerges due to the stepwise accumulation of genetic and epigenetic alterations [2,3]. Advances in understanding the fundamental molecular mechanisms implicated in CRC development have led to the identification of prognostic and predictive markers. The activating driver mutations in *BRAF* (B-Raf proto-oncogene, serine/threonine kinase) and *KRAS* (KRAS proto-oncogene, GTPase) are critical events in CRC biology by disrupting the RAS/RAF/MEK/MAP-kinase signaling pathway [4]. Besides, many tumor suppressor genes are inactivated by genetic and epigenetic changes in CRC. For example, loss of function of *MLH1* (mutL homolog 1) gene is a key event. Epigenetic silencing of this gene due to promoter methylation is a useful biomarker for risk assessment, diagnosis as well as prediction of therapeutic response [5]. The loss of function of *MLH1* is linked to mismatch repair (MMR) deficiency leading to microsatellite instability (MSI). Mutations in MMR genes and MSI are observed in 15 to 20% of sporadic CRC [6,7]. Importantly, MMR deficiency is also associated with the CIMP (CpG island methylator phenotype) and is relevant for therapy response predictions [8]. A recent pan-cancer study comprising >6,000 human tumors identified DNA methylation changes at genes that converge to few pathways. Further, concomitant genetic and epigenomic changes converged on the same pathways: *i*) mutations in WNT pathway components and *ii*) chromatin modifiers associated with DNA methylation instability [9]. Nevertheless, there are no prognostic nor reliable markers to predict whether early-stage CRC patients could benefit from adjuvant chemotherapy [10].

Epigenetic changes, including aberrant DNA methylation and post-translational histone modifications, have been implicated in the origin and progression of human cancer, as well as in drug resistance mechanisms [11–13]. In this context, epigenetic modifications can contribute as biomarkers to improve diagnosis, prognosis, and prediction of treatment responses [18]. The reversible nature of the epigenetic modifications makes them potential targets for epigenetic therapy for cancer. Currently, different strategies have been employed for the use of epigenetic therapy as monotherapy or along with other modalities of treatment, such as radiotherapy and chemotherapy [14,15]. For instance, inhibitors of DNA methyltransferase (DNMTi), histone deacetylases (HDACi), and histone methylases (HMTi) were approved for cancer treatment by the United States Food and Drug Administration (US-FDA). Several clinical trials have been conducted to explore new combinations and small molecules termed epi-drugs for the treatment of solid tumors, namely those showing aggressive phenotype and resistance to traditional therapeutic approaches [15–18].

The components of the epigenetic regulation systems, also known as proteins of the epigenetic machinery, have been widely studied as promising druggable targets [19]. These proteins modulate chromatin dynamics and transcriptional regulation by adding or erasing post-translational modifications (PTMs) in specific amino acid residues located at the histone tails of H2A, H2B, H3, and H4 [20]. Within this framework, histone acetylation/deacetylation has been believed to be a key event. These histone PTMs are regulated by opposing enzymatic functions of histone acetyltransferases (HATs) and histone deacetylase superfamilies [21,22]. The eighteen members of the histone deacetylase family are grouped into four classes: classic deacetylases (HDACs), sharing the same catalytic mechanism and conserved zinc-dependent deacetylation domain, are grouped in class I (HDACs 1, 2, 3, and 8), class IIa (HDACs 4, 5, 7 and 9); IIb (HDACs 6 and 10), and IV (HDAC11) while sirtuins (SIRT 1-7) are denominated class III, due to their mechanism of action dependent on the cofactor NAD (+) [23]. Class I HDACs are nuclear and ubiquitously expressed, while classes II, III, and IV may be nuclear and/or cytoplasmic expressed in tissue-specific patterns [24].

In addition to histones, the histone deacetylases have alternative non-histone substrates, which regulate different functions of the cellular processes involved in cancer, such as cell proliferation and survival (class I HDACs), cytoskeleton regulation, mobility, endothelial cell functions of gluconeogenesis, growth of cardiac myocytes, differentiation of thymocytes and homologous recombination during DNA repair [25]. Other cellular processes, such as autoimmunity, aging, survival, migration, and cell invasion are modulated by class III HDACs [26]. The diversity of substrates indicates that these enzymes can interfere in several processes and pathways involved in the onset and progression of cancer, justifying the great interest in the usage of these proteins as new therapeutic targets [27]. Evidence supports that these enzymes are involved in gene silencing during drug resistance development. In particular, acetylation and deacetylation of lysines of histones H3 and H4 have been associated with the progression of CRC [25,28,29].

CRC consists of a heterogeneous disease showing several epigenetic alterations, frequently associated with distinct molecular features, traditionally classified into chromosomal instability (CIN), MSI (High, Low, Stable), and CIMP (High and Low) [18]. The current knowledge regarding histone modifiers and histone marks in CRC is yet sparse. Previous studies suggested the inter-and intra-tumoral heterogeneity of histone marks and HDAC expression in this cancer type [30]. Therefore, the present study focused on the evaluation of the gene expression patterns and the prognostic value of the panel of genes encoding histone deacetylases, using public resources such as The Cancer Genome Atlas (TGCA) and The Genotype-Tissue Expression (GTEx) projects [31]. The clinical, and expression data based on RNA-Seq was retrieved to identify the differentially expressed members of the histone deacetylase family as promising biomarkers for CRC diagnosis, prognosis, and therapeutic targets as well as to investigate its potential correlations with relevant clinical and molecular parameters.

2. Method and Materials

2.1. Data collection, pre-processing, and tumor sample selection according to purity score of the tumor samples

Transcriptional profile based on RNA-Seq from The Cancer Genome Atlas (https://portal.gdc.cancer.gov/), including GDC TCGA Colon Cancer (TCGA-COAD, n= 471 samples) and GDC TCGA Rectum Cancer (TCGA-READ, n=167 samples) datasets, were retrieved from the Genomic Data Commons (GDC) portal (https://portal.gdc.cancer.gov/). The tumor samples were selected according to tissue purity score as previously reported [32]. This previous study analyzed 9,364 tumor samples and determined a purity score termed CPE (Consensus Purity Estimate). The CPE score is the median of four existing methods to estimate the purity of tumor tissues: ABSOLUTE [33], ESTIMATE [34], unmethylated leukocytes, and immuno-histochemistry. Considering that the median CPE value was 0.7805 and 0.7913 for COAD and READ samples, respectively, a CPE greater than 0.8 was considered as high-purity tumor samples and selected for further analysis.

Gene expression profiles from normal intestine colon tissue were downloaded from The Genotype Tissue Expression (GTEx) project using the UCSC data Xena Explorer (http://xena.ucsc.edu/) [35]. Normal tissue samples were classified into two groups: matched normal tissue adjacent to tumor samples (Table 1) and 308 normal tissue from GTEx project. The clinic-pathological data were used to investigate the potential correlations with the differential expression of the histone deacetylase family members.

2.2. Differential gene expression analysis

The RNA-Seq quantification data for gene expression analysis from TCGA and GTEx projects are determined as HTSeq and RSEM-expected counts, respectively. Thus, we used the free and open web based tool UCSC Xena Explorer (docs.gdc.cancer.gov/Data/Bioinformatics_Pipelines/Expression_mRNA_Pipeline) to obtain uniformly processed and unified data. Differential expression patterns of the 18 members of the gene family encoding known human histone deacetylase enzymes were determined between CRC (COAD and READ) and normal tissues. The differentially expressed genes (DEGs) were defined considering log2FoldChange > +1 or < – 1 and adjusted p-value of <0. 0001 (student t-test). All statistical analysis and data normalization were performed following the manual of the DESeq2 package for R language [36], available on R Bioconductor (v 3.8) (bioconductor.org/packages/release/bioc/html/DESeq2.html). Euclidean distance heatmaps and principal component analysis (PCA) were created using *R* pheatmap available in the DESeq2 package. Boxplot and distribution graphics were plotted in the same package. Venn diagrams were designed using the LibreOffice Draw software.

2.3. Correlation between DEGs with clinical-pathological parameters and molecular features

To identify the possible correlations between DEGs with classic prognostic parameters and known biomarkers of CRC, the following clinical-pathological parameters were analyzed: gender, age, body mass index (Normal, Overweight, Obesity I, II, and III) and TNM staging (stages I-IV). Tumor molecular features, when available, were also evaluated: the presence of mutations in *KRAS* and *BRAF* oncogenes, microsatellite instability classified in high levels (MSI-H), low levels (MSI-L) or microsatellite stability (MSS) and CIMP phenotype grouped as high methylation levels (CIMP-H) or low (CIMP-L). Statistical analysis was conducted using the Wilcoxon and Kruskal-Wallis nonparametric tests using Rstats [37]. The survival analysis was based on Kaplan–Meier curve, Z-score values for each DEG, and quartiles as cutoff values for high and low expression groups (75% and 25%, respectively) on R survminer [38]. The $p \leq 0.05$ was considered significant.

3. Results and Discussions

3.1. Transcriptional expression pattern of histone deacetylase genes in patients with CRC

The erasers of histone acetylation are key modulators of chromatin structure and transcription as well as regulators of non-histone substrates, including several cancer-related proteins. Notwithstanding, proteins of the epigenetic machinery are involved in tumor genetic and epigenetic intra-tumoral heterogeneity, phenotypic adaptation, tumor progression, and immune evasion [39], as proposed for CRC evolution [40-41]. Indeed, CRC is considered a heterogeneous disease at both genetic and epigenetic levels. This study evaluated the transcriptional expression patterns and the prognostic value of the histone deacetylase family members in human CRC primary tumor samples.

Class I HDACs, including the co-expressed HDAC1 and 2 in adult tissues, are ubiquitously expressed, while other HDCAs from Class IIa and IIb demonstrated a tissue-specific or variable pattern of expression among healthy tissues, respectively [42]. In addition, some studies indicate that tumor purity is an important confounder in evaluating gene expression profiles and its correlation with clinicopathological parameters and biological features such as molecular subtypes of human cancers [43]. Considering that tumor specimens contain non-cancerous cells, including immune cells and fibroblasts from stroma, we first selected CRC samples showing high CPE tumor purity levels [32]. Among the cohort of 638 CRCs recovered in the COAD (471 tumor samples) and READ (167 tumor samples) datasets, 9 were excluded due to the absence of tumor purity data. The remaining 629 tumor samples were classified according to tumor purity. The COAD and READ projects presented 185 and 74 samples, respectively, classified as high-purity tumors. In parallel, four of 308 samples of normal colon tissue from the GTEx project were removed from subsequent analysis because they did not have gene expression data based on the RNA-Seq technique. The remaining 304 normal samples were divided into groups regarding their anatomical portion to perform differential gene expression analysis (normal sigmoid- or transverse-colon). These data are summarized in Table 1.

Table 1. Distribution tumor and normal adjacent samples from National Cancer Institute - GDC data portal repository (The Cancer Genome Atlas project) and normal samples according to with The Genotype-Tissue Expression project (GTEx). CRC was classified according to tumor purity.

Project	**Primary Site**	**Number of tumor samples**	**Tumor purity (CPE)***			**Number of normal samples**
		All samples	CPE ≤ 0.8	**High-purity tumors (CPE > 0.8)**	Normal adjacent (TCGA)	GTEx
GDC TCGA Colon Cancer (COAD)	Colon	471	278	**185**	41	308
GDC TCGA Rectum Cancer (READ)	Rectum	167	92	**74**	10	0

* CPE = Consensus Purity Estimate system

Eight genes were identified as differentially expressed in high-purity COAD samples, with three of them characterized with high expression levels (*HDAC1*, *HDAC2*, and *HDAC8*) and five showing low expression levels (*HDAC4*, *HDAC7*, *HDAC9*, *HDAC10*, and *SIRT4*) (Fig 1A). This pattern was not reproduced in the group of COAD showing lower-purity (CPE ≤ 0.8), since neither *HDAC8* was observed as over-expressed nor *HDAC9* under-expressed (appendix 1). The overall structure of the analyzed dataset showed the distribution of tumors and normal tissue samples in distinct clusters after PCA. Although with some overlap, the normal tissue trends to group according to the anatomical region, *i.e.*, the sigmoid and transverse colon (Fig 1B). Thus, we also compared the expression levels of COAD samples with normal tissues grouped according to these two anatomical regions (COAD *versus* normal sigmoid-colon and COAD *versus* normal transverse-colon) (Appendix 1). Separately, the comparison between COAD samples with normal transverse-colon region showed four additional DEGs: the up-regulation of the *SIRT5* and *SIRT6* genes as well as low expression of *HDAC7* and *SIRT4* (Fig.1C). The comparison between COAD and normal sigmoid-colon tissue included only one aditional DEG (high expression levels of *HDAC3* gene).

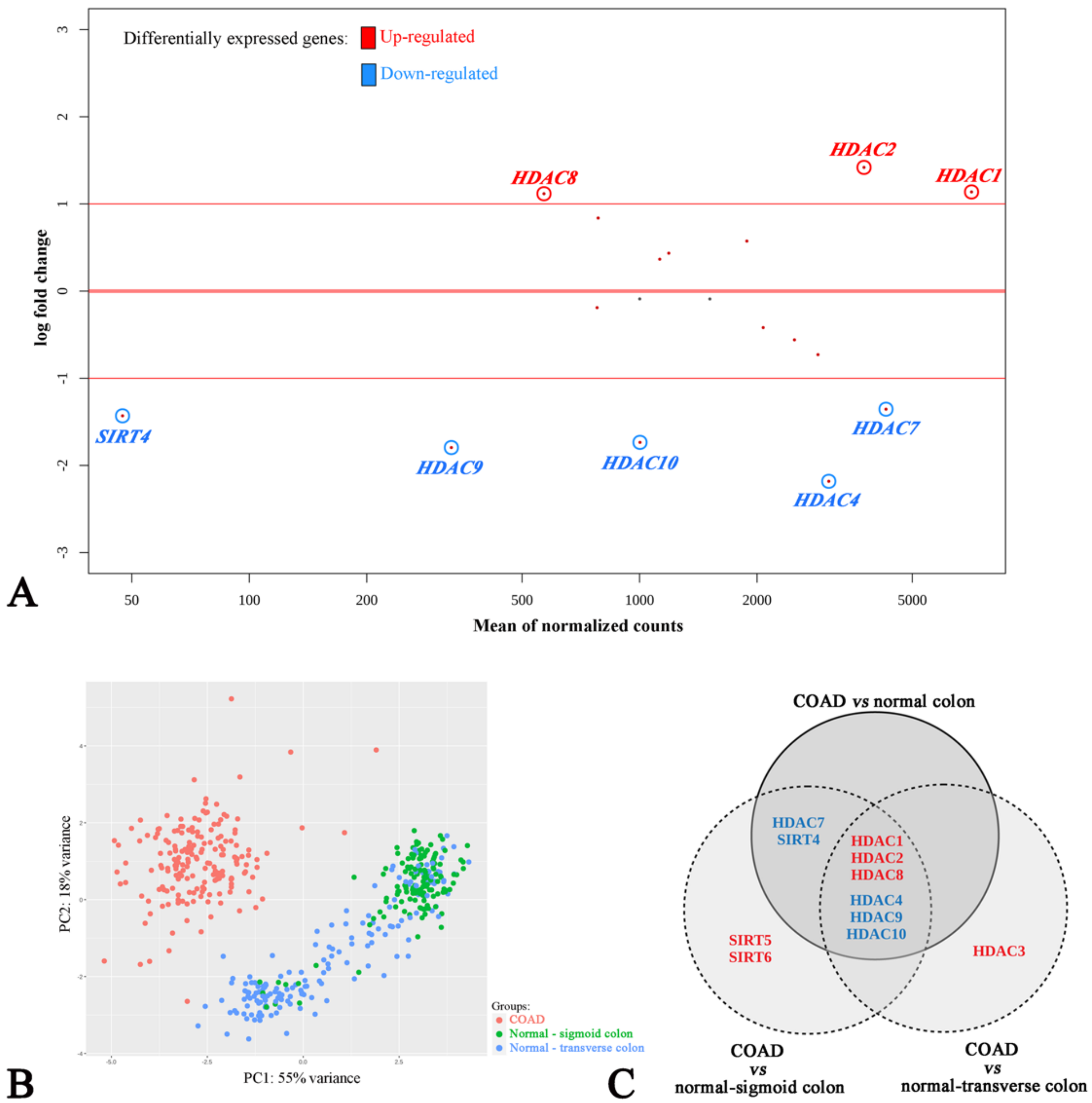

Figure 1. Gene expression analysis in COAD **A**) Differentially expressed genes (DEGs) in tumor samples compared to normal colon tissues. Up- and down-regulated genes are indicated in red and blue, respectively. **B**) Principal Component Analysis (PCA) applied to the datasets indicating the separation of tumor and normal tissues, besides a trend to split the normal samples according to anatomic regions (sigmoid- and transverse-colon). **C**) Venn diagram of DEGs in different grouped into three comparison groups represented by three circles. The six common DEGs to these comparison groups are detailed in the overlapping portions of the different circles

Previous studies have reported that CRC located at different anatomic locations varies regarding microbiome constitution, risk factors, incidence, pathogenesis, molecular pathways, and patient outcomes [44]. Accordingly, the CRC sublocations are right-sided (cecum, ascending, and two-thirds of the transverse colon) and left-sided (one-third of transverse, descending, and sigmoid colon), separated by the splenic flexure. Left-sided tumors typically present lower levels of *KRAS* and *BRAF* mutations, CIMP-L and MSS, and better overall survival compared to right-sided tumors. The latter group is reported to show poor differentiation, higher TNM staging, and higher risk of peritoneal tumor seeding [45]. In this study, we detected distinct lists of DEGs depending on the anatomical region of the normal tissue samples used as reference. This finding may reflect the distinct embryological origins of the intestine segments [45]. The limitation here is that TCGA data do not establish the original site of CRC samples, so we could not filter the tumor samples according to this parameter. However, future studies could investigate DEGs in matched tumor and normal samples according to anatomical sites, especially to confirm the up-regulation of *SIRT5* and *SIRT6* and downregulation of *HDAC7* and *SIRT4* observed after the analysis using transverse-colon (right-side) as reference. We also observed high expression levels of *HDAC3* after comparing tumor *versus* normal sigmoid-colon (left-side) samples.

Figure 2A illustrates the heat map of the six DEGs in COAD and normal tissue samples (GTEx). Based on the above-mentioned findings demonstrating that DEGs depend on the specific choice of normal samples in the dataset, we set out to investigate whether we can detect biologically relevant information in CRC. The clustering analysis showed a heterogeneous expression pattern among COAD samples, which formed two clusters: the larger and small tumor groups were termed subgroup A and B (Fig 1B), respectively, and partially replicate the separation of CRC samples in PCA (data not shown). A total of 153 and 30 COAD patients were classified in subgroups A and B, respectively. The subgroup B showed proximity to normal colon samples than the first group but was separated in another clade (Fig 2B). The distribution of clinic-pathological data between the COAD subgroups A and B were not different but were the basis for the search for possible correlations between the expression levels of the six DEGs with prognostic parameters and patient survival (Table 2).

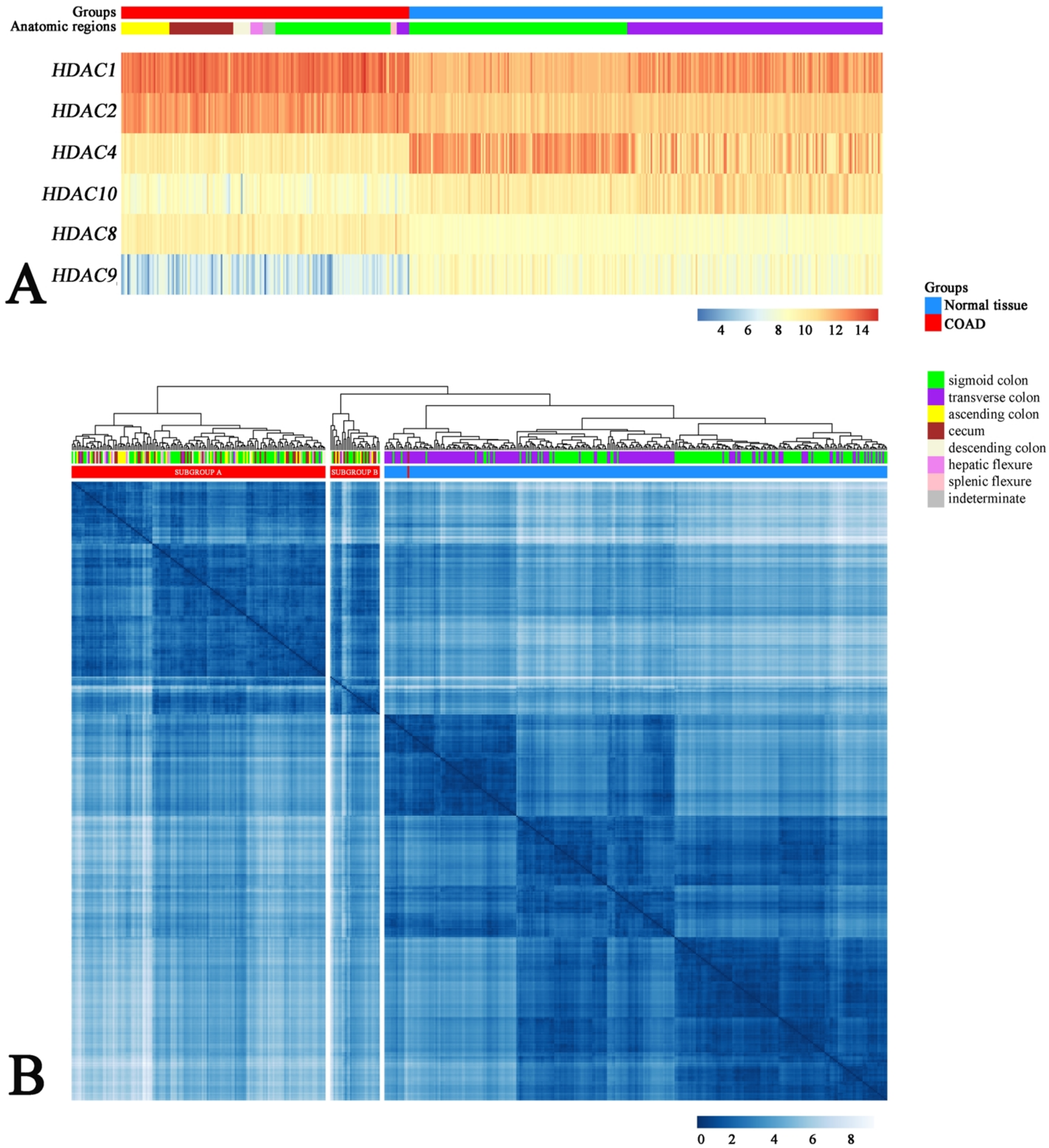

Figure 2. A) Heat map of the expression levels of the six DEGs in COAD distributed according to anatomic regions. **B**) Clustering of tissue-specific expression patterns derived from DEG data. CRC samples grouped into two subsets termed A and B. Further, the normal tissue samples were grouped in subsets replicating the separation observed in PCA.

Table 2. Distribution of clinical variables and molecular features of COAD according to subtypes A and B.

Variables	**COAD**		
	Subgroup A (Number of cases)	**Subgroup B (Number of cases)**	***p* value (*)**
Gender	**152**	**27**	>0.9999
Female	68	12	
Male	84	15	
Age at diagnosis	**152**	**27**	0.7382
≥50	137	24	
<50	15	3	
Body Mass Index (N/Over *vs* Obesity)	**67**	**12**	0.0192
Normal	27	1	
Overweight	23	5	
Obesity I	7	2	
Obesity II	7	3	
Obesity III	3	1	
Stage TNM (I/II vs III/IV)	**150**	**27**	0.6810
Stage I	25	4	
Stage II	58	10	
Stage III	39	8	
Stage IV	28	5	
KRAS Mutation	**22**	**3**	0.2300
Presence	11	3	
Absence	11	n.d.	
BRAF Mutation	**17**	**1**	>0.9999
Presence	2	n.d.	
Absence	15	1	
Microsatellite Instability	**152**	**26**	0.0764
High	21	2	
Low	29	1	
MSS	102	23	
CIMP	**21**	**3**	>0.9999
High	10	2	
Low	11	1	

(*) p value obtained from Chi square or Fisher's exact test. Bonferroni correction was applied to statistical adjustment of each p-value (statistically significant p < 0,00625). (n.d.) = not detected.

To investigate the subjacent molecular mechanisms to the expression changes of the six DEGs in COAD, we investigated the occurrence of genetic changes including copy number variations (CNVs) and gene mutations in the public databases (appendix 2). Considering the ubiquity of class I HDACs and the expected higher count values, we hypothesized that when there are slight CNV alterations, they may reflect a surge of transcriptional activity, thus a possible explanation for their up-regulation. However, there was no clear correlation between increased/reduced gene expression and CNV alterations (gains/losses). A noteworthy example is *HDAC9*, which is reported to have high frequencies of CNV gains, but at the transcriptional level, was detected as down-regulated in COAD. Assuming the possible discrepancy of genomic and transcriptomic data, we highlight that further proteomic analysis should be performed to better characterize the histone deacetylase landscape in CRC. Unfortunately, there are no proteomic data for histone deacetylases in TCGA.

Abnormal expression levels of histone deacetylases could impact several multi-protein complexes and cellular processes. Class I HDACs participate in multiprotein complexes responsible for the repression of gene transcription: Sin3, NuRD, CoREST, and NCoR/SMRT complexes. These enzymes are also found in more specialized complexes with specific DNA-binding proteins, such as CSL and TCF transcription factors [46], while HDAC8 functions alone or along with SMAD3/4 heterotrimer [47]. For example, the expression of *HDAC3* is considered significantly higher in CRC (primary tissues and cell lines), and has been correlated with clinical features: high levels were described in poorly differentiated tumors, higher TNM staging and associated with worse survival rates [48-51]. Although overexpressed only when comparing COAD versus normal transverse-colon, our data suggest that *HDAC3* could impact CRC biology. Class IIa HDACs participate in the SMRT/N-CoR [52] and may form a bridge between the SMRT–HDAC3 complex and MEF2 transcription factors due to its N-terminal adaptor domain [53]. The functions of class IIb HDACs are poorly understood, as are typically found in the cytoplasm. HDAC6 is involved in the deacetylation of α-tubulin, cortactin, chaperones, and IFNαR [53], while HDAC10 is believed to depend on HDAC3 to function, and could repress transcription via other mechanisms [54].

The sirtuins can be found in the nucleus, cytoplasm or mitochondria, according to their respective functions. The nuclear ones (SIRT1, SIRT2, SIRT6, and SIRT7) regulate chromatin, transcription, cell cycle, and metabolism, but studies have shown that there are ambiguous functions for sirtuins in human cancers [55]. For instance, *SIRT6* is thought to be a tumor suppressor gene, usually silenced or deleted in colon cancers (both primary tumors and cell line samples) [56-58], that could modulate gene expression by H3K9 deacetylation or protein deacetylation [59,60]. However, herein *SIRT6* is up-regulated in COAD compared to normal sigmoid-colon samples. Considering that the analysis considering normal transverse-colon samples showed different results, we cannot exclude a possible bias due to the differences in normal cell programs from different anatomical regions, not necessarily related to CRC tumorigenesis.

The *SIRT4* gene is suggested to be a tumor suppressor (studies in other cancer types [55]. *SIRT4* expression is notably reduced in CRC when compared to normal tissues. The correlation between *SIRT4* lower expression levels, less differentiated tumors, higher TNM staging, higher recurrence rates, and worse survival has been previously reported. Moreover, *SIRT4* is down-regulated from the initial steps of tumorigenesis until its advanced states. Besides, it was also found that the overexpression of this gene inhibited the proliferation, invasion, and migration of CRC cell lines [61]. The upregulation of the *SIRT5* mRNA and protein levels was associated with larger tumor sizes, higher tumor stage and TNM staging, and higher risk of mortality, thus it correlates with poor prognosis of CRC [62]. Additionally, this enzyme is assumed to regulate lactate dehydrogenase B (LDHB), T cell activity, and to desuccinylate the mitochondrial serine hydroxymethyltransferase SHMT2 [63-65]. In CRC, the gene product and protein levels of SIRT7 were found to be higher in CRC than in normal samples, and the protein levels were significantly correlated with poor survival [66]. Deacetylation of H3K18Ac by SIRT7 is necessary for maintaining essential features of human cancer cells, including anchorage-independent growth and escape from contact inhibition [67], although SIRT7 may require to be modulated by other factors, such as Dicer [68].

Finally, six DEGs (up-regulated *HDAC1*, *HDAC2*, and *HDAC8*, and down-regulated *HDAC4*, *HDAC9*, and *HDAC10* gene) were selected for further analysis. Given the fact that normal tissue origin did not interfere with their significance, they can be more robust and more likely implicated in CRC development and progression (as driver or passenger alterations in gene expression).

3.2. mRNA expression levels of the HDAC family members are significantly associated with clinicopathological parameters of patients with COAD

High expression levels of the *HDAC1* gene were correlated with the presence of a mutation in the *BRAF* gene (p=0.01471) while a trend was observed between the lower expression level of the gene *HDAC10* and mutation in this oncogene (p=0.05882). Lower expression levels of the gene encoding the HDAC10 were also correlated with the CIMP-Low phenotype (p=0.006192). Besides, the expression of the *HDAC2* gene was correlated with high microsatellite instability (MSI-H) (p=0.04468) compared with the groups of MSI-L and MSS (low microsatellite instability and microsatellite stability, respectively). The MSI-L and MSS groups do not show significant value in multiple pairwise-comparison between groups but have demonstrated a correlation with high expression of *HDAC2*. In subgroup B tumors, lower expression levels of the *HDAC4* gene were correlated with younger patients (age smaller than 50 years) (Fig. 3). The survival analysis of all DEGs did not show significant results, demonstrating that the curves of both high and low expression were similar (appendix 3).

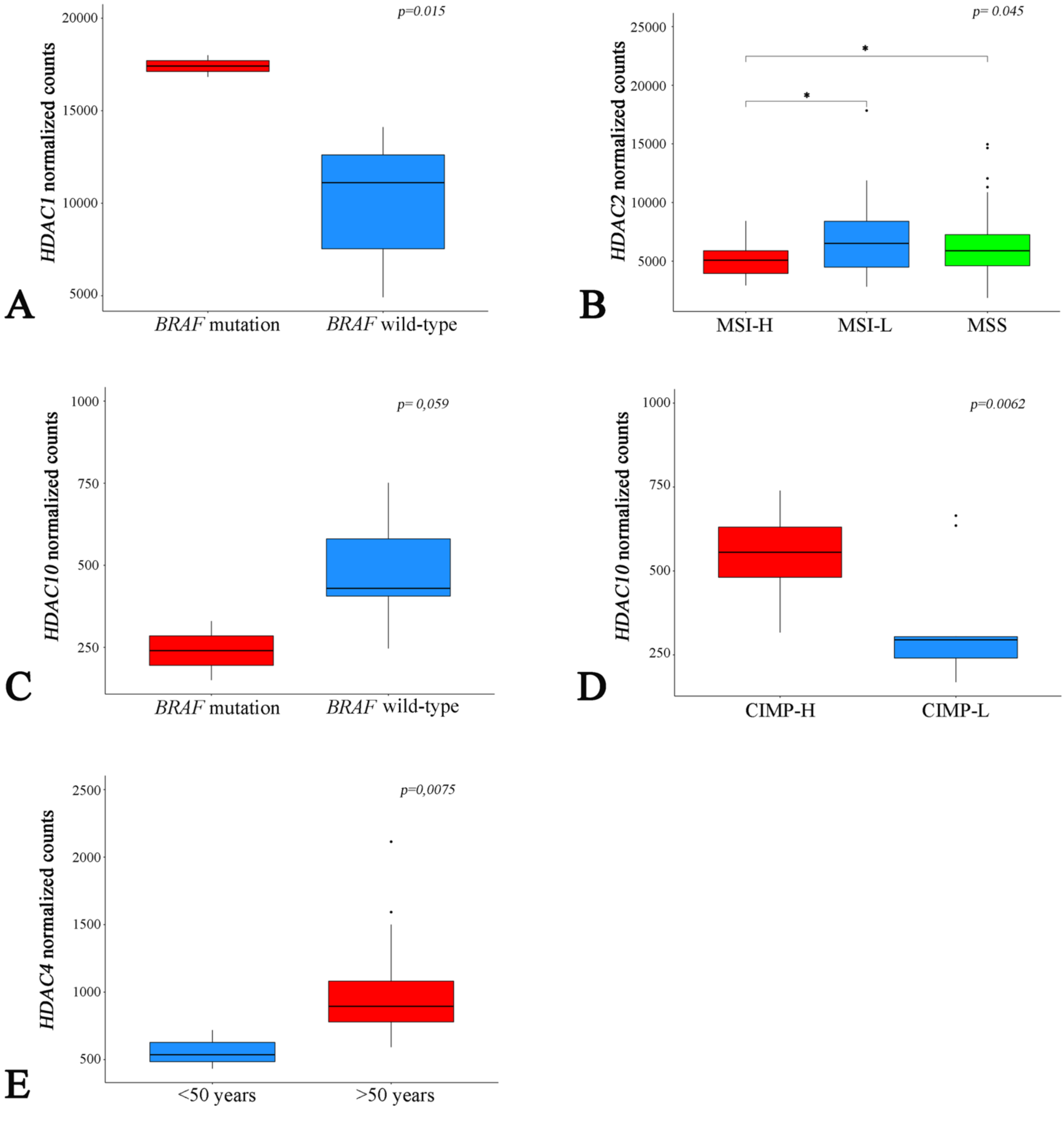

Figure 3. Normalized mRNA expression levels of differentially expressed *HDAC* genes and their correlation with clinic-pathological parameters in the two COAD subgroups A and B. There are three significant associations (**A**, **B**, and **D**), and one trend (**C**) for the subgroup A. For the subgroup B, one significant association was found with age at diagnosis (**E**). **A**) High *HDAC1* expression levels correlated with the presence of *BRAF* mutations. **B**) Low expression levels of *HDAC2* were correlated with high microsatellite instability (MSI-H) compared to low microsatellite instability (MSI-L) and microsatellite stability (MSS). **C**) **HDAC10** counts showed a trend to correlate with the occurrence of *BRAF* mutations. **D**) Increased levels of *HDAC10* correlated with the CpG island methylator phenotype (CIMP) [high methylation (CIMP-H) *versus* low methylation (CIMP-L)]. **E**) *HDAC4* expression levels correlated with patients' age. p-value ≤0.05 was considered significant.

The six HDACs differentially expressed in COAD are grouped in three up-regulated and encoding class I HDACs (HDAC1, HDAC2, and HDAC8) while three were down-regulated genes (HDAC4, HDAC9, and HDAC10) and classified as class IIa/b HDACs (Fig 1C). These HDACs are related to important biological functions during colon cancer carcinogenesis, but also in other types of cancers [69]. In the literature, class I HDAC isoforms are founded highly expressed in a subset of CRCs with significant results for HDAC1 (36.4%), HDAC2 (57.9%), and HDAC3 (72.9%) after immunohistochemistry. The expression levels of these proteins were significantly enhanced in strongly proliferating and poorly differentiated tumors. Also, class I HDACs expression levels were associated with reduced patient survival, with HDAC2 expression being an independent survival indicator [51].

The high expression levels of *HDAC1* were correlated with mutations in the *BRAF* gene (p=0.01471) (Fig 3A). In CRC, high expression of *HDAC1* at mRNA and protein expression levels were greater in tumor tissue [70-73]. Additionally, these studies demonstrated a correlation between the expression of HDAC1 in patients with poor prognostic parameters, including depth tumor invasion, distant metastasis, advanced stage, and low survival rate [74]. The combination of *ING2* and *HDAC1* overexpression remarkably enhances MMP13 expression, when compared to only ING2 overexpression [75]. Some studies have related to the higher expression of HDAC1 and HDAC7 (a class IIa HDAC) in colorectal cancer [76] and their correlation with stages III and IV as well poorly differentiated tumors in comparison to moderately differentiated ones [77].

In primary tumor and colon cancer cell lines, the up-regulation of *HDAC1* and *HDAC2* were linked to the expression of the Sp family of transcription factors, responsible for regulating the transcriptional activity of genes implicated in several cellular processes. These HDACs up-regulate Sp1 and Sp3 and other epigenetic modulators, such as SET1 and p300, forming an interaction network regulating the expression of HDAC1 and HDAC2 in colon cancer cells [78]. These findings demonstrate diverse and important crosstalk on histone modifications and alterations of histone modifications in colon cancer. Our data corroborate previous studies that reported the up-regulation of the *HDAC1* gene at both mRNA and protein levels, in CRC samples [79,80], without significant correlations with the variables investigated. Evidence indicates that HDAC1 interacts with other modulators to contribute to CRC biology [81,82], cooperating with HDAC8 to repress BMF, a gene involved in Bcl-2 related apoptotic mechanisms and resistance to treatment with HDACis [83]. In addition, a truncating mutation of *HDAC2* gene was detected in CRC showing microsatellite instability. This mutation leads to the loss of enzymatic activity [84] and renders cells resistant to the treatment with the classical HDACi inhibitor trichostatin A (TSA) [84]. Furthermore, ectopic expression of *HDAC2* in mutant cancer cells induces a reduction of colony formation and inhibition of tumor growth in xenograft nude mice [84], suggesting a putative tumor-suppressor role for HDAC2 in this cellular setting. In support of this idea, the tumor-suppressor Rb represses gene transcription by recruiting class I HDACs [85], indicating that HDAC2 mutations could induce the de-repression of Rb-silenced genes. Interestingly, in the context of microsatellite stable colorectal tumors, elevated expression of *HDAC2* is observed [84,86].

Besides the high expression of the *HDAC2* gene in the COAD group (Fig 1C), we also observed that the higher expression values in subgroup A were correlated with microsatellite stability (MSS) and low microsatellite instability (MSI-L) (p=0.04468) compared with the group showing high microsatellite instability (MSI-H) (Fig 3B). Elevated *HDAC2* mRNA and protein expression were demonstrated in tumor tissue compared with normal tissue [70,87-92], as well as their nuclear expression was detected at high levels in CRC (81.0%), adenomas (62.1%), and normal tissue (53.1%) [93-95] or differentially expressed in aneuploid than in diploid carcinomas [96]. However, higher expression levels of *HDAC2* gene in normal tissue was found by RT-PCR compared with adenomatous polyp or colon carcinoma tissues [97].

The relationship between HDAC2 and MSI phenotype has been previously addressed. *In vitro* studies using two colorectal cancer cell lines (RKO and Co115) characterized by a frameshift mutation in the *HDAC2* gene leading to loss of its activity, and three cell lines with the wild-type gene (SW48, LoVo, and HCT −116), demonstrated the relationship between the MSI phenotype and resistance to anti-proliferative and pro-apoptotic effects of known HDACis, such as TSA, but not to other inhibitors such as butanoic and valproic acid. The authors have analyzed the acetylation level of histones H3 and H4 and suggested that, although butanoic and valproic acids induce hyperacetylation of histones H3 and H4 in all cell lines, the HDACi TSA was unable to induce hyperacetylation in HDAC2-deficient cells, only in those carrying the wild-type gene. Importantly, the knockdown of the *HDAC2* gene in cell lines carrying the wild-type gene led to the resistance to TSA. Taken together, these findings indicated that resistance to TSA in sporadic CRC samples with the presence of MSI can be mediated by the status of the *HDAC2* gene and that it could be considered a predictive marker of therapeutic response [84].

Among the down-regulated DEGs, lower levels of *HDAC4* gene were correlated with younger patients (aged under 50 years) in subgroup B (p=0.007521) (Fig. 3E), and HDAC10 lower expression levels were correlated with the CIMP-low phenotype (p=0.006192) in subgroup A. Lower expression of *HDAC4* and Tip60 genes were described in the colon and lung. The Tip60 encodes the catalytic subunit of the Tip60 chromatin-remodeling complex involved in DNA repair [98].

6. Conclusions

In this study, we characterized the expression profile of the histone deacetylase gene family in CRC. Of the six DEGs detected in high-purity CRC samples, three are consistent with previous studies (up-regulation of *HDAC1*, *2*, and downregulation of *HDAC10*). Additionally, we report the up-regulation of *HDAC8* and down-regulation of *HDAC9* and *HDAC10*, for which literature is scarce. We highlight the importance of *HDAC1*, *HDAC2*, *HDAC4*, and *HDAC10* for CRC biology, given the association with molecular features such as *BRAF* mutations, MSI, and CIMP status. Interestingly, the differential gene expression results depended on the anatomical region of normal tissue samples used as a reference, possibly a reflection of different embryonic origins and anatomic-related specialized cell programs. We have detected intratumor heterogeneity, which reflects the aggregation of two subgroups of histone deacetylase expression, although their biological profiles have not yet been detailed. The role of HDACs in CRC therapeutic response depends on several variables. These enzymes interact with multiprotein complexes involving other epigenetic modulators of gene expression. Further studies are necessary to elucidate the protein-protein interactions and protein networks of histone deacetylase family members. In this perspective, histone deacetylase expression is a promising target for biomarker and pharmacological targeting for epigenetic therapy of CRC.

Acknowledgments

This study was supported by the research grant from the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior-Brasil (CAPES) - Finance Code 001.

References

[1]Cancer Tomorrow. 2021. Available at: https://gco.iarc.fr/tomorrow/home. Accessed March 15, 2021.

[2]You JS, Jones PA. Cancer genetics and epigenetics: two sides of the same coin? *Cancer Cell* 2012; 22(1): 9–20. doi:10.1016/j.ccr.2012.06.008.

[3]Takeshima H, Ushijima T. Accumulation of genetic and epigenetic alterations in normal cells and cancer risk. *npj Precision Oncology* 2019; 3(1): 1–8. doi:10.1038/s41698-019-0079-0.

[4]Das V *et al.* Predictive and prognostic biomarkers in colorectal cancer: A systematic review of recent advances and challenges. *Biomed Pharmacother* 2017; 87: 8–19. doi:10.1016/j.biopha.2016.12.064.

[5]Werner RJ *et al.* Epigenetics and Precision Oncology. *Cancer J* 2017; 23(5): 262–269. doi:10.1097/PPO.0000000000000281.

[6]Jass JR. Classification of colorectal cancer based on correlation of clinical, morphological and molecular features. *Histopathology* 2007; 50(1): 113–130. doi:10.1111/j.1365-2559.2006.02549.x.

[7]Yamagishi H *et al.* Molecular pathogenesis of sporadic colorectal cancers. *Chin J Cancer* 2016; 35. doi:10.1186/s40880-015-0066-y.

[8]Etienne-Grimaldi M-C *et al.* Molecular patterns in deficient mismatch repair colorectal tumours: results from a French prospective multicentric biological and genetic study. *British Journal of Cancer* 2014; 110(11): 2728–2737. doi:10.1038/bjc.2014.213.

[9]Saghafinia S *et al.* Pan-Cancer Landscape of Aberrant DNA Methylation across Human Tumors. *Cell Rep* 2018; 25(4): 1066-1080.e8. doi:10.1016/j.celrep.2018.09.082.

[10]Koncina E *et al.* Prognostic and Predictive Molecular Biomarkers for Colorectal Cancer: Updates and Challenges. *Cancers (Basel)* 2020; 12(2). doi:10.3390/cancers12020319.

[11]Guo L *et al.* Targeting epigenetic regulatory machinery to overcome cancer therapy resistance. *Seminars in Cancer Biology* 2021. doi:10.1016/j.semcancer.2020.12.022.

[12]Flavahan WA *et al.* Epigenetic plasticity and the hallmarks of cancer. *Science* 2017; 357(6348). doi:10.1126/science.aal2380.

[13]Sharma SV *et al.* A chromatin-mediated reversible drug-tolerant state in cancer cell subpopulations. *Cell* 2010; 141(1): 69–80. doi:10.1016/j.cell.2010.02.027.

[14]Mohammad HP *et al.* Targeting epigenetic modifications in cancer therapy: erasing the roadmap to cancer. *Nature Medicine* 2019; 25(3): 403–418. doi:10.1038/s41591-019-0376-8.

[15]Morel D *et al.* Combining epigenetic drugs with other therapies for solid tumours — past lessons and future promise. *Nature Reviews Clinical Oncology* 2020; 17(2): 91–107. doi:10.1038/s41571-019-0267-4.

[16]Cheng Y *et al.* Targeting epigenetic regulators for cancer therapy: mechanisms and advances in clinical trials. *Signal Transduction and Targeted Therapy* 2019; 4(1): 1–39. doi:10.1038/s41392-019-0095-0.

[17]Ronnekleiv-Kelly SM *et al.* Epigenetic therapy and chemosensitization in solid malignancy. *Cancer Treat Rev* 2017; 55: 200–208. doi:10.1016/j.ctrv.2017.03.008.

[18]Jung G *et al.* Epigenetics of colorectal cancer: biomarker and therapeutic potential. *Nat Rev Gastroenterol Hepatol* 2020; 17(2): 111–130. doi:10.1038/s41575-019-0230-y.

[19]Moreira-Silva F *et al.* Repurposing Old Drugs into New Epigenetic Inhibitors: Promising Candidates for Cancer Treatment? *Pharmaceutics* 2020; 12(5). doi:10.3390/pharmaceutics12050410.

[20]Park S-Y, Kim J-S. A short guide to histone deacetylases including recent progress on class II enzymes. *Experimental & Molecular Medicine* 2020; 52(2): 204–212. doi:10.1038/s12276-020-0382-4.

[21]Jones PA, Baylin SB. The fundamental role of epigenetic events in cancer. *Nat Rev Genet* 2002; 3(6): 415–428. doi:10.1038/nrg816.

[22]Muntean AG, Hess JL. Epigenetic dysregulation in cancer. *Am J Pathol* 2009; 175(4): 1353–1361. doi:10.2353/ajpath.2009.081142.

[23]Xu WS *et al.* Histone deacetylase inhibitors: molecular mechanisms of action. *Oncogene* 2007; 26(37): 5541–5552. doi:10.1038/sj.onc.1210620.

[24]New M *et al.* HDAC inhibitor-based therapies: can we interpret the code? *Mol Oncol* 2012; 6(6): 637–656. doi:10.1016/j.molonc.2012.09.003.

[25]Li Y, Seto E. HDACs and HDAC Inhibitors in Cancer Development and Therapy. *Cold Spring Harb Perspect Med* 2016; 6(10). doi:10.1101/cshperspect.a026831.

[26]Singh AK *et al.* Targeting Histone Deacetylases with Natural and Synthetic Agents: An Emerging Anticancer Strategy. *Nutrients* 2018; 10(6). doi:10.3390/nu10060731.

[27]Marks PA, Xu W-S. Histone deacetylase inhibitors: Potential in cancer therapy. *J Cell Biochem* 2009; 107(4): 600–608. doi:10.1002/jcb.22185.

[28]Glozak MA, Seto E. Histone deacetylases and cancer. *Oncogene* 2007; 26(37): 5420–5432. doi:10.1038/sj.onc.1210610.

[29]Patnaik S, Anupriya null. Drugs Targeting Epigenetic Modifications and Plausible Therapeutic Strategies Against Colorectal Cancer. *Front Pharmacol* 2019; 10: 588. doi:10.3389/fphar.2019.00588.

[30]Lutz L *et al.* Histone modifiers and marks define heterogeneous groups of colorectal carcinomas and affect responses to HDAC inhibitors in vitro. *Am J Cancer Res* 2016; 6(3): 664–676.

[31]Cancer Genome Atlas Network. Comprehensive molecular characterization of human colon and rectal cancer. *Nature* 2012; 487(7407): 330–337. doi:10.1038/nature11252.

[32]Aran D *et al.* Systematic pan-cancer analysis of tumour purity. *Nature Communications* 2015; 6(1): 8971. doi:10.1038/ncomms9971.

[33]Carter SL *et al.* Absolute quantification of somatic DNA alterations in human cancer. *Nature Biotechnology* 2012; 30(5): 413–421. doi:10.1038/nbt.2203.

[34]Yoshihara K *et al.* Inferring tumour purity and stromal and immune cell admixture from expression data. *Nature Communications* 2013; 4(1): 2612. doi:10.1038/ncomms3612.

[35]Goldman M *et al.* The UCSC Xena platform for public and private cancer genomics data visualization and interpretation. *bioRxiv* 2019: 326470. doi:10.1101/326470.

[36]Love MI *et al.* Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 2014; 15(12): 550. doi:10.1186/s13059-014-0550-8.

[37]Core Team R. A language and environment for statistical computing. *Computing* 2019; 1. doi:10.1890/0012-9658(2002)083[3097:CFHIWS]2.0.CO;2.

[38]Kassambara A *et al.* survminer: Drawing Survival Curves using “ggplot2”. Available at: https://rpkgs.datanovia.com/survminer/index.html. Accessed March 26, 2021.

[39]de Matos MR *et al.* A Systematic Pan-Cancer Analysis of Genetic Heterogeneity Reveals Associations with Epigenetic Modifiers. *Cancers* 2019; 11(3): 391. doi:10.3390/cancers11030391.

[40]Uchi R *et al.* Integrated Multiregional Analysis Proposing a New Model of Colorectal Cancer Evolution. *PLOS Genetics* 2016; 12(2): e1005778. doi:10.1371/journal.pgen.1005778.

[41]Ryser MD *et al.* Epigenetic Heterogeneity in Human Colorectal Tumors Reveals Preferential Conservation And Evidence of Immune Surveillance. *Scientific Reports* 2018; 8(1): 17292. doi:10.1038/s41598-018-35621-y.

[42]Milazzo G *et al.* Histone Deacetylases (HDACs): Evolution, Specificity, Role in Transcriptional Complexes, and Pharmacological Actionability. *Genes (Basel)* 2020; 11(5). doi:10.3390/genes11050556.

[43]Rhee J-K *et al.* Impact of Tumor Purity on Immune Gene Expression and Clustering Analyses across Multiple Cancer Types. *Cancer Immunol Res* 2018; 6(1): 87–97. doi:10.1158/2326-6066.CIR-17-0201.

[44]Hsu Y-L *et al.* Clinicopathological and molecular differences in colorectal cancer according to location. *Int J Biol Markers* 2019; 34(1): 47–53. doi:10.1177/1724600818807164.

[45]Stintzing S *et al.* Understanding the role of primary tumour localisation in colorectal cancer treatment and outcomes. *European Journal of Cancer* 2017; 84: 69–80. doi:10.1016/j.ejca.2017.07.016.

[46]Cunliffe VT. Eloquent silence: developmental functions of Class I histone deacetylases. *Current Opinion in Genetics & Development* 2008; 18(5): 404–410. doi:10.1016/j.gde.2008.10.001.

[47]Tang X *et al.* HDAC8 cooperates with SMAD3/4 complex to suppress SIRT7 and promote cell survival and migration. *Nucleic Acids Research* 2020; 48(6): 2912–2923. doi:10.1093/nar/gkaa039.

[48]Li J *et al.* HDAC3 deteriorates colorectal cancer progression via microRNA-296-3p/TGIF1/TGFβ axis. *J Exp Clin Cancer Res* 2020; 39. doi:10.1186/s13046-020-01720-w.

[49]Spurling CC *et al.* HDAC3 overexpression and colon cancer cell proliferation and differentiation. *Mol Carcinog* 2008; 47(2): 137–147. doi:10.1002/mc.20373.

[50]Nemati M *et al.* Deregulated expression of HDAC3 in colorectal cancer and its clinical significance. *Adv Clin Exp Med* 2018; 27(3): 305–311. doi:10.17219/acem/66207.

[51]Weichert W *et al.* Class I histone deacetylase expression has independent prognostic impact in human colorectal cancer: specific role of class I histone deacetylases in vitro and in vivo. *Clin Cancer Res* 2008; 14(6): 1669–1677. doi:10.1158/1078-0432.CCR-07-0990.

[52]Fischle W *et al.* Enzymatic Activity Associated with Class II HDACs Is Dependent on a Multiprotein Complex Containing HDAC3 and SMRT/N-CoR. *Molecular Cell* 2002; 9(1): 45–57. doi:10.1016/S1097-2765(01)00429-4.

[53]Park S-Y, Kim J-S. A short guide to histone deacetylases including recent progress on class II enzymes. *Experimental & Molecular Medicine* 2020; 52(2): 204–212. doi:10.1038/s12276-020-0382-4.

[54]Tong JJ *et al.* Identification of HDAC10, a novel class II human histone deacetylase containing a leucine-rich domain. *Nucleic Acids Research* 2002; 30(5): 1114–1123. doi:10.1093/nar/30.5.1114.

[55]Chalkiadaki A, Guarente L. The multifaceted functions of sirtuins in cancer. *Nature Reviews Cancer* 2015; 15(10): 608–624. doi:10.1038/nrc3985.

[56]Sebastián C *et al.* THE HISTONE DEACETYLASE SIRT6 IS A NOVEL TUMOR SUPPRESSOR THAT CONTROLS CANCER METABOLISM. *Cell* 2012; 151(6): 1185–1199. doi:10.1016/j.cell.2012.10.047.

[57]Lin Z *et al.* USP10 Antagonizes c-Myc Transcriptional Activation through SIRT6 Stabilization to Suppress Tumor Formation. *Cell Rep* 2013; 5(6): 1639–1649. doi:10.1016/j.celrep.2013.11.029.

[58]Shi L *et al.* Prognostic role of SIRT6 in gastrointestinal cancers: a meta-analysis. *Open Med (Wars)* 2020; 15(1): 358–365. doi:10.1515/med-2020-0403.

[59]Liu W *et al.* SIRT6 inhibits colorectal cancer stem cell proliferation by targeting CDC25A. *Oncol Lett* 2018; 15(4): 5368–5374. doi:10.3892/ol.2018.7989.

[60]Qi J *et al.* Downregulated SIRT6 and upregulated NMNAT2 are associated with the presence, depth and stage of colorectal cancer. *Oncol Lett* 2018; 16(5): 5829–5837. doi:10.3892/ol.2018.9400.

[61]Bai Y *et al.* Research Progress of Sirtuin4 in Cancer. *Front Oncol* 2021; 10. doi:10.3389/fonc.2020.562950.

[62]Wang Y-Q *et al.* Sirtuin5 contributes to colorectal carcinogenesis by enhancing glutaminolysis in a deglutarylation-dependent manner. *Nat Commun* 2018; 9. doi:10.1038/s41467-018-02951-4.

[63]Shi L *et al.* SIRT5-mediated deacetylation of LDHB promotes autophagy and tumorigenesis in colorectal cancer. *Mol Oncol* 2019; 13(2): 358–375. doi:10.1002/1878-0261.12408.

[64]Wang K *et al.* SIRT5 Contributes to Colorectal Cancer Growth by Regulating T Cell Activity. *J Immunol Res* 2020; 2020: 3792409. doi:10.1155/2020/3792409.

[65]Yang X *et al.* SHMT2 Desuccinylation by SIRT5 Drives Cancer Cell Proliferation. *Cancer Res* 2018; 78(2): 372–386. doi:10.1158/0008-5472.CAN-17-1912.

[66]Yu H *et al.* Overexpression of Sirt7 Exhibits Oncogenic Property and Serves as a Prognostic Factor in Colorectal Cancer. *Clin Cancer Res* 2014; 20(13): 3434–3445. doi:10.1158/1078-0432.CCR-13-2952.

[67]Barber MF *et al.* SIRT7 links H3K18 deacetylation to maintenance of oncogenic transformation. *Nature* 2012; 487(7405): 114–118. doi:10.1038/nature11043.

[68]Chen X *et al.* Dicer regulates non-homologous end joining and is associated with chemosensitivity in colon cancer patients. *Carcinogenesis* 2017; 38(9): 873–882. doi:10.1093/carcin/bgx059.

[69]Singh AK *et al.* Targeting Histone Deacetylases with Natural and Synthetic Agents: An Emerging Anticancer Strategy. *Nutrients* 2018; 10(6). doi:10.3390/nu10060731.

[70]Huang BH *et al.* Inhibition of histone deacetylase 2 increases apoptosis and p21Cip1/WAF1 expression, independent of histone deacetylase 1. *Cell Death Differ* 2005; 12(4): 395–404. doi:10.1038/sj.cdd.4401567.

[71]Ishihama K *et al.* Expression of HDAC1 and CBP/p300 in human colorectal carcinomas. *J Clin Pathol* 2007; 60(11): 1205–1210. doi:10.1136/jcp.2005.029165.

[72]Xiong W *et al.* MiR-761 inhibits colorectal cancer cell proliferation and invasion through targeting HDAC1. *Pharmazie* 2019; 74(2): 111–114. doi:10.1691/ph.2019.8756.

[73]Chen C *et al.* The histone deacetylase HDAC1 activates HIF1α/VEGFA signal pathway in colorectal cancer. *Gene* 2020; 754: 144851. doi:10.1016/j.gene.2020.144851.

[74]Higashijima J *et al.* Expression of histone deacetylase 1 and metastasis-associated protein 1 as prognostic factors in colon cancer. *Oncol Rep* 2011; 26(2): 343–348. doi:10.3892/or.2011.1312.

[75]Kumamoto K *et al.* ING2 is upregulated in colon cancer and increases invasion by enhanced MMP13 expression. *Int J Cancer* 2009; 125(6): 1306–1315. doi:10.1002/ijc.24437.

[76]Gerçeker E *et al.* Never in mitosis gene A-related kinase 6 and aurora kinase A: New gene biomarkers in the conversion from ulcerative colitis to colorectal cancer. *Oncol Rep* 2015; 34(4): 1905–1914. doi:10.3892/or.2015.4187.

[77]Kasap E *et al.* The potential role of the NEK6, AURKA, AURKB, and PAK1 genes in adenomatous colorectal polyps and colorectal adenocarcinoma. *Tumour Biol* 2016; 37(3): 3071–3080. doi:10.1007/s13277-015-4131-6.

[78]Yang H *et al.* Overexpression of histone deacetylases in cancer cells is controlled by interplay of transcription factors and epigenetic modulators. *FASEB J* 2014; 28(10): 4265–4279. doi:10.1096/fj.14-250654.

[79]Ishihama K *et al.* Expression of HDAC1 and CBP/p300 in human colorectal carcinomas. *J Clin Pathol* 2007; 60(11): 1205–1210. doi:10.1136/jcp.2005.029165.

[80]Cao L-L *et al.* The expression of histone deacetylase HDAC1 correlates with the progression and prognosis of gastrointestinal malignancy. *Oncotarget* 2017; 8(24): 39241–39253. doi:10.18632/oncotarget.16843.

[81]Ishikawa D *et al.* The Significance of MicroRNA-449a and Its Potential Target HDAC1 in Patients With Colorectal Cancer. *Anticancer Res* 2019; 39(6): 2855–2860. doi:10.21873/anticanres.13414.

[82]Liu J *et al.* SOX4 maintains the stemness of cancer cells via transcriptionally enhancing HDAC1 revealed by comparative proteomics study. *Cell Biosci* 2021; 11(1): 23. doi:10.1186/s13578-021-00539-y.

[83]Kang Y *et al.* HDAC8 and STAT3 repress BMF gene activity in colon cancer cells. *Cell Death Dis* 2014; 5(10): e1476. doi:10.1038/cddis.2014.422.

[84]Ropero S *et al.* A truncating mutation of HDAC2 in human cancers confers resistance to histone deacetylase inhibition. *Nat Genet* 2006; 38(5): 566–569. doi:10.1038/ng1773.

[85]Frolov MV, Dyson NJ. Molecular mechanisms of E2F-dependent activation and pRB-mediated repression. *J Cell Sci* 2004; 117(Pt 11): 2173–2181. doi:10.1242/jcs.01227.

[86]Ropero S *et al.* Transforming pathways unleashed by a HDAC2 mutation in human cancer. *Oncogene* 2008; 27(28): 4008–4012. doi:10.1038/onc.2008.31.

[87]Zhu P *et al.* Induction of HDAC2 expression upon loss of APC in colorectal tumorigenesis. *Cancer Cell* 2004; 5(5): 455–463. doi:10.1016/s1535-6108(04)00114-x.

[88]Giannini R, Cavallini A. Expression analysis of a subset of coregulators and three nuclear receptors in human colorectal carcinoma. *Anticancer Res* 2005; 25(6B): 4287–4292.

[89]Ozdağ H *et al.* Differential expression of selected histone modifier genes in human solid cancers. *BMC Genomics* 2006; 7: 90. doi:10.1186/1471-2164-7-90.

[90]Lutz L *et al.* Histone modifiers and marks define heterogeneous groups of colorectal carcinomas and affect responses to HDAC inhibitors in vitro. *Am J Cancer Res* 2016; 6(3): 664–676.

[91]Mao QD *et al.* MicroRNA-455 suppresses the oncogenic function of HDAC2 in human colorectal cancer. *Brazilian Journal of Medical and Biological Research* 2017; 50(6). doi:10.1590/1414-431x20176103.

[92]Tong J *et al.* FKBP3 mediates oxaliplatin resistance in colorectal cancer cells by regulating HDAC2 expression. *Oncol Rep* 2019. doi:10.3892/or.2019.7259.

[93]Ashktorab H *et al.* Global histone H4 acetylation and HDAC2 expression in colon adenoma and carcinoma. *Dig Dis Sci* 2009; 54(10): 2109–2117. doi:10.1007/s10620-008-0601-7.

[94]Benard A *et al.* Nuclear expression of histone deacetylases and their histone modifications predicts clinical outcome in colorectal cancer. *Histopathology* 2015; 66(2): 270–282. doi:10.1111/his.12534.

[95]Tang W *et al.* The p300/YY1/miR-500a-5p/HDAC2 signalling axis regulates cell proliferation in human colorectal cancer. *Nature Communications* 2019; 10(1): 663. doi:10.1038/s41467-018-08225-3.

[96]Gemoll T *et al.* HDAC2 and TXNL1 distinguish aneuploid from diploid colorectal cancers. *Cell Mol Life Sci* 2011; 68(19): 3261–3274. doi:10.1007/s00018-011-0628-3.

[97]Drew JE *et al.* Predictive Gene Signatures: Molecular Markers Distinguishing Colon Adenomatous Polyp and Carcinoma. *PLoS One* 2014; 9(11). doi:10.1371/journal.pone.0113071.

[98]LLeonart ME *et al.* New p53 related genes in human tumors: significant downregulation in colon and lung carcinomas. *Oncol Rep* 2006; 16(3): 603–608. doi:10.3892/or.16.3.603.

Appendix 1

Table A1.2 _ Differentially expressed genes in high purity-tumor samples (TCGA-COAD compared with GTEx normal samples)

Desacetilases		TCGA COAD – GTEx (together)		TCGA COAD – GTEx sigmoid		TCGA COAD - GTEx transverse	
		log2FoldChange	pvalue	log2FoldChange	pvalue	log2FoldChange	pvalue
Class I	HDAC1*	**1.13764591761371**	**7.74613462374943E-85**	**1.76259088650148**	**6.73889194260337E-118**	**1.06871304107156**	**1.53162126502258E-06**
	HDAC2*	**1.41905436935037**	**8.37426627645356E-245**	**1.47376186853702**	**4.93473625197672E-90**	**1.92388475018641**	**9.0796888965292E-76**
	HDAC3	0.573626036270983	2.30112610036866E-78	0.607585173761183	1.59946141561832E-47	**1.17698479625309**	**3.44145690740087E-50**
	HDAC8*	**1.11711503301764**	**6.07449133149368E-149**	**1.01119864698453**	**1.13457406426217E-85**	**1.19633372245703**	**2.37278967838799E-33**
Class IIa	HDAC4*	**-2.18230889387554**	**1.54079294468694E-137**	**-2.55577116409407**	**1.71593070850162E-114**	**-1.37827444983689**	**0.000270318196345**
	HDAC5	-0.559441979390788	1.1081749053502E-24	-0.774957494147516	1.2243883574185E-22	-0.38738926558395	0.071878087287335
	HDAC7	**-1.35502261409503**	**1.72668311813104E-112**	**-1.55505268583207**	**2.86775918072307E-78**	-0.911843928718794	0.000742170104049
	HDAC9*	**-1.79509121195508**	**2.30773132461429E-73**	**-2.05976842626404**	**4.47416011854407E-49**	**-2.03298776908719**	**3.63843833744929E-11**
Class IIb	HDAC6	-0.729237044573853	1.94574409026499E-96	-0.641401231095178	4.02433430743419E-26	-0.791626363892215	8.03956223237259E-16
	HDAC10*	**-1.73551737562911**	**6.82283402751707E-162**	**-1.14562275344113**	**1.27341328474095E-32**	**-2.03939596724671**	**1.47840967874464E-18**
ClassIV	HDAC11	0.365888232490835	2.61112676886274E-21	0.441862562588363	3.91919032308023E-12	0.347404642305426	0.002084178956273
Class III	SIRT1	-0.089669390546253	0.122520626750853	-0.167992499870514	0.030834654952248	0.296389069044686	0.118141390859389
	SIRT2	-0.418645130425024	1.09690761955814E-30	-0.470158914186713	4.10501854364899E-22	0.216893954752645	0.085901697098684
	SIRT3	-0.190066122056412	1.03924446716977E-08	0.010752673237478	0.832979798525915	0.005057046641551	0.969976810856301
	SIRT4	**-1.43109999421485**	**2.51399608482583E-64**	**-1.74711985775545**	**2.4310645283446E-41**	-0.105051472545098	0.743944688936988
	SIRT5	0.838399581269886	5.46772018745163E-71	**1.32400776362253**	**9.90596429148677E-75**	0.880476618685112	1.03455163330214E-07
	SIRT6	0.43493473082074	1.76455095746895E-07	**1.42617127355069**	**1.41301453338528E-41**	0.141207256187683	0.696030485172204
	SIRT7	-0.090390750475789	0.228609711466054	0.766783895635612	8.37655906143447E-13	-0.27680385380819	0.355107063867133

The highlighted values refer to the results considered significant with values of log2foldchange> 1 or <-1 and pvalue <0.0001. *Differential expressed genes that are founded in all groups

Table A1.2. Differentially expressed genes of high purity tumor samples TCGA-COAD, TCGA-READ and TCGA-CRC compared with TCGA adjacent normal samples.

Desacetilases		TCGA COAD		TCGA READ		TCGA CRC	
		log2FoldChange	pvalue	log2FoldChange	pvalue	log2FoldChange	pvalue
Class I	HDAC1	0.043539182824338	0.654744336789422	-0.171148741662198	0.286001351226813	-0.002724546557161	0.933391260202016
	HDAC2	0.947978743129871	2.06608482457144E-25	0.434932617494689	0.008112469613674	0.841455412216366	1.99855666440868E-26
	HDAC3	0.36911738982965	3.47993846076086E-08	0.133215227035898	0.384884194323738	0.320276417634993	1.44151859277514E-07
	HDAC8	0.896513989868875	4.28248871053496E-19	0.457498627805033	0.001545207268275	0.80416662330272	2.34055617145333E-21
Class IIa	HDAC4	-0.673309147698726	3.75548511726731E-10	**-1.02476039031644**	**2.3937383845604E-05**	-0.753016063856595	1.55786603558356E-13
	HDAC5	-0.280350080151659	0.004215103955082	-0.22065634827041	0.286001351226813	-0.263259187876391	0.002885922087227
	HDAC7	0.32927568611781	1.63254723244786E-05	0.106318156784177	0.444410447720296	0.286734355679645	1.23566494459595E-05
	HDAC9*	**-2.1183658749949**	**3.83714763118924E-21**	**-3.01673637795039**	**4.50042692529249E-20**	**-2.3462146951576**	**2.35391466989276E-35**
Class IIb	HDAC6	0.072126770445831	0.405219949984607	0.223593411365557	0.165115306765237	0.113794511591667	0.100188531128424
	HDAC10	0.847231168482316	9.81013791030276E-16	**1.28861754395842**	**8.55204768710206E-09**	0.951389810041816	4.06378697116843E-23
ClassIV	HDAC11	0.024149844253001	0.786443604329675	0.03122390935272	0.833036369721005	0.034375626807723	0.708985958483471
Class III	SIRT1	-0.150675454124791	0.198397189326113	**-1.00795064755903**	**2.84134384755487E-06**	-0.363010406788788	0.000339014058232
	SIRT2	-0.135550807944775	0.083729692827742	-0.236248535114381	0.15632753741399	-0.153341210385991	0.0228727691108
	SIRT3	-0.210428433206876	0.000340664676986	-0.046492077970471	0.726151631633241	-0.177602609489479	0.000772236379775
	SIRT4	-0.787128870676654	4.03902324538297E-07	**-1.16904685504109**	**0.000106121824157**	-0.898988977591906	5.1113843533837E-11
	SIRT5	0.332890043281289	4.02231784927403E-05	0.247120036788312	0.101287748849559	0.310663887236842	1.06117523976294E-05
	SIRT6	-0.714820918742879	1.31257108360457E-12	-0.648458744390317	0.001504635503501	-0.699687267293017	4.66211498927645E-15
	SIRT7	-0.037125205289276	0.706787797076444	0.305085023053405	0.117250652005067	0.035997774714051	0.708985958483471

The highlighted values refer to the results considered significant with values of log2foldchange> 1 or <-1 and pvalue <0.0001. *Differential expressed genes that are founded in all groups

Appendix 2

Table A2. Mutation frequencies of HDAC family genes in CRC patients retrieved from Cosmic and cBioPortal databases (last accessed March 2021).

Gene	Database											
	Cosmic				cBioPortal (TCGA PanCancer Atlas)				cBioPortal (All other Bowel samples)			
	Total mutated samples	Total tested samples	Mutation frequency	CNV* frequency	Total mutated samples	Total tested samples	Mutation frequency	CNV* frequency	Total mutated samples	Total tested samples	Mutation frequency	CNV* frequency
HDAC1	53	2548	2.08%	0.14% loss	6	526	1.14%	3.6% gain 31.7% loss	15	826	1.8%	22.9% loss
HDAC2	98	2547	3.85%	0.14% loss	10	526	1.9%	15% gain 14.3% loss	10	964	1%	0.95% gain 3.8% loss
HDAC3	39	2707	1.44%	–	5	526	0.95%	9.1% gain 20.7% loss	6	826	0.73%	0.95% amp** 3.8% gain 0.95% loss
HDAC4	169	2573	6.57%	–	17	526	3.23%	22.1% gain 4.8% loss	38	826	4.6%	6.7% gain
HDAC5	109	2547	4.28%	–	14	526	2.66%	0.4% amp** 20.9% gain 13.3% loss	26	826	3.1%	2.9% gain 4.8% loss
HDAC6	76	2547	2.98%	0.28% gain 0.7% loss	12	526	2.28%	0.2% amp** 15.4% gain 15;6% loss	17	826	2.1%	10.5% gain 0.95% loss
HDAC7	87	2573	3.38%	–	10	526	1.9%	19.6% gain 10.5% loss	25	826	3%	0.95% amp** 2.9% gain
HDAC8	39	2547	1.53%	0.14% gain 0.42% loss	6	526	1.14%	17.1% gain 13.5% loss	5	826	0.97%	10.5% gain 0.95% loss

HDAC9	274	2547	10.76%	3.06% gain	34	526	6.46%	0.4% amp** 57.4% gain 0.76% loss	33	826	4%	2.9% amp** 40% gain 0.95% loss
HDAC10	59	2547	2.32%	0.42% loss	7	526	1.33%	2.28% gain 36.7% loss	17	826	2.1%	0.95% gain 14.3% loss
HDAC11	50	2547	1.96%	0.14% gain 0.14% loss	6	526	1.14%	0.2% amp** 9.5% gain 14% loss	15	826	1.8%	2.86% gain 0.95% loss
SIRT1	54	2616	2.06%	0.28% gain	10	526	1.9%	0.4% amp** 4.8% gain 19.4% loss	14	826	1.7%	22% loss
SIRT2	65	2547	2.55%	0.42% gain	3	526	0.57%	0.8% amp** 20.7% gain 7.8% loss	10	826	1.2%	1.9% amp** 3.8% gain 0.95% loss
SIRT3	28	2547	1.1%	0.14% gain	3	526	0.57%	13% gain 13.4% loss	11	826	1.3%	0.95% amp** 6.7% gain 1.9% loss
SIRT4	47	2547	1.85%	–	8	526	1.52%	17.3% gain 12.5% loss	15	826	1.8%	0.95% amp** 0.95% gain 8.6% loss
SIRT5	54	2547	2.12%	–	5	526	0.95%	18.4% gain 11.2% loss	11	826	1.3%	0.95% amp** 3.8% gain 2.9% deletion
SIRT6	26	2547	1.02%	–	4	526	0.76%	16.5% gain 14.3% loss	8	826	0.97%	1.9% gain 9.5% loss
SIRT7	33	2536	1.3%	0.14% gain	4	526	0.76%	0.8% amp** 23% gain	7	826	0.85%	8.6% gain

	12.4% loss	2.86% loss

*CNV = Copy number variations; **amp = amplification; "loss" refers to both shallow deletions and deep deletions; cBioPortal considers gains and deletions: Log2 copy-number values higher than 0 and lower than 0, respectively. They are putative form GISTIC.

Appendix 3

Table A3 - Summary of correlation analysis of differentially expressed genes in high-purity COAD samples with prognostic parameters and overall survival.

DEGs	Prognostic parameter								Survival
	Gender	Age at diagnosis	BMI	TNM	*KRAS*	*BRAF*	MSI	CIMP	
Subgroup A									
HDAC1	0.7655	0.853	0.4664	0.4846	0.7477	**0.01471***	0.06991	0.9725	0.11
HDAC2	0.8241	0.1491	0.437	0.7777	0.2003	0.7206	**0.04468***	0.7564	0.99
HDAC8	0.6526	0.9066	0.3352	0.7722	0.1164	0.9412	0.2372	0.5116	0.39
HDAC4	0.5251	0.3305	0.4521	0.1272	0.7969	0.5294	0.5766	0.5573	0.78
HDAC9	0.2986	0.5408	0.379	0.4171	0.6522	0.1765	0.3742	0.7564	0.94
HDAC10	0.5203	0.08703	0.1508	0.5648	0.6522	**0.05882****	0.9068	**0.006192***	0.54
Subgroup B									
HDAC1	0.1257	0.3132	0.2169	0.8423	N.A	N.A	0.4451	1	0.34
HDAC2	0.4856	0.6885	0.5175	0.5022	N.A	N.A	0.4482	1	0.29
HDAC8	0.6255	0.5628	0.7875	0.5212	N.A	N.A	0.5258	0.6667	0.2
HDAC4	0.6833	**0.007521***	0.7174	0.3388	N.A	N.A	0.1465	0.6667	0.51
HDAC9	0.5581	0.9079	0.3709	0.5662	N.A	N.A	0.6257	0.6667	0.9
HDAC10	0.9809	0.6885	0.2395	0.8053	N.A	N.A	0.7131	1	0.35

*pvalue $p<0.05$; ** trend to with statistical significance; N.A. not available

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*).

Apêndices

Apêndice 1. Differential expressed genes analysis results of tumor samples TCGA COAD compared with GTEx normal samples

Desacetilases		TCGA COAD – GTEx		TCGA COAD – GTEx (Filtered with purity)		TCGA COAD - GTEx (Low purity)	
		log2FoldChange	pvalue	log2FoldChange	pvalue	log2FoldChange	pvalue
Class I	HDAC1*	**1.06955265503069**	**4.82241332971538E-137**	**1.07287149921766**	**5.00877306228925E-137**	**1.0023791956193**	**2.28757703822757E-91**
	HDAC2*	**1.32354731613541**	**5.09171639704838E-289**	**1.32112161718025**	**1.29945633006771E-288**	**1.25170330864666**	**3.4629628671517E-237**
	HDAC3	0.531894665467677	1.73138142705654E-99	0.53202967468075	2.35589636291556E-99	0.50076053456846	2.14497662708138E-88
	HDAC8	**1.06824014644077**	**1.22427203100327E-179**	**1.05398875190349**	**2.3198265090458E-179**	0.998216359862188	8.47998237837865E-169
Class IIa	HDAC4*	**-2.05444836579334**	**7.45226619072717E-245**	**-2.07840924303567**	**2.21440534491253E-252**	**-2.02265400656807**	**3.62826749168965E-164**
	HDAC5	-0.529323077315295	9.38991295519618E-35	-0.527972219080473	2.10381918112729E-34	-0.524266289694061	6.99155907781147E-27
	HDAC7*	**-1.17021233790945**	**3.06884127771691E-155**	**-1.17479527253482**	**5.8261354162052E-156**	**-1.0556437608537**	**2.6043094131245E-91**
	HDAC9	**-1.00657160109093**	**2.93469643539111E-27**	**-1.07860283977744**	**4.64293473345281E-32**	-0.756322323861946	2.05118771719587E-16
Class IIb	HDAC6	-0.735622817487445	8.71359235827188E-137	-0.747148775770748	3.28238781860478E-137	-0.752455934004552	2.24294265075271E-132
	HDAC10*	**-1.76050570893864**	**1.29182534252363E-258**	**-1.75732298459341**	**1.34114956533691E-257**	**-1.76850122745848**	**1.85633435347191E-198**
ClassIV	HDAC11	0.262730610815654	3.60822031622459E-14	0.264406682104956	1.91236642565643E-14	0.176378002068097	7.64857947747191E-07
Class III	SIRT1	-0.053700514191876	0.25375453923133	-0.062770822467462	0.182854293115739	-0.05833795141161	0.24317092981244
	SIRT2	-0.40278273600719	6.72301737452125E-45	-0.397796384921424	9.57585784010361E-45	-0.388450329391241	5.04518426920016E-37
	SIRT3	-0.193541896691765	3.07499877236815E-14	-0.193975398511718	3.48711093624814E-14	-0.205291360550871	1.04117904658765E-12
	SIRT4*	**-1.40600966465775**	**5.06430652051723E-103**	**-1.40792254573063**	**3.48711889506424E-104**	**-1.40856831114941**	**1.78851731501738E-84**
	SIRT5	0.683084812078104	2.11690879282293E-74	0.676475771016532	1.30812627909192E-72	0.557709234169211	2.42013977840784E-43
	SIRT6	0.41375590114318	6.19709464307202E-12	0.420023382920956	3.46923221883602E-12	0.407934562925095	1.37529330138833E-08
	SIRT7	-0.076753648925299	0.155507227754956	-0.078153383257343	0.151607356458003	-0.069212167263008	0.265788864333891

The highlighted values refer to the results considered significant with values of log2foldchange> 1 or <-1 and pvalue <0.0001. *Differential expressed genes that are founded in all groups

Apêndice 2. Differential expressed genes analysis results of tumor samples TCGA COAD, TCGA READ and TCGA CRC compared with TCGA adjacent normal samples

Desacetilases		TCGA COAD		TCGA READ		TCGA CRC	
		log2FoldChange	pvalue	log2FoldChange	pvalue	log2FoldChange	pvalue
Class I	HDAC1	-0.019584852306414	0.798999239045003	-0.228872440385108	0.135698820954204	-0.072670989710234	0.274867310966107
	HDAC2	0.87166497000233	6.14221766095891E-23	0.441815745561561	0.008992527575681	0.785359376346036	1.52529219270331E-23
	HDAC3	0.346611339017686	8.48269612374754E-09	0.154442119042419	0.235680709719586	0.304965960680576	1.27896311073433E-08
	HDAC8	0.874159579765968	1.86745966302491E-19	0.496285205693688	0.000813633956331	0.794348629232905	3.37376389026907E-22
Class IIa	HDAC4	-0.539010444927609	1.61244826269792E-07	**-1.14386738493651**	**1.26196667526848E-08**	-0.692414962748089	5.88308137416084E-14
	HDAC5	-0.250967534252439	0.009065501165713	-0.148595503662595	0.426826709380624	-0.222177486191111	0.008625477520647
	HDAC7	0.528264445883723	2.45264989504337E-11	0.231722710854125	0.114393495946943	0.46820675682846	8.43926665382799E-12
	HDAC9*	**-1.27206818646042**	**3.99266980165012E-08**	**-2.19451785836589**	**2.7047942779538E-08**	**-1.51870482341546**	**7.69280818161418E-14**
Class IIb	HDAC6	0.086602320229241	0.26628815896497	0.228638505620172	0.114393495946943	0.120623380417651	0.060401539469008
	HDAC10	0.846325856970778	7.36487568510789E-15	**1.20831052535977**	**6.63249350900425E-08**	0.922247678079722	4.17150278280212E-21
ClassIV	HDAC11	-0.069518903328795	0.447841826916174	0.017791193142895	0.910339736219256	-0.043754661068807	0.57943210393657
Class III	SIRT1	-0.104031257370311	0.374812259107003	-0.868849204703589	1.24957428855225E-05	-0.291615247205335	0.002691642119351
	SIRT2	-0.107809792287858	0.133055813376674	-0.194123457745363	0.167610274396407	-0.12792390529476	0.036638453557496
	SIRT3	-0.204137383181701	9.27461813793286E-05	-0.035540034981087	0.691648627516382	-0.177011682546237	0.000276384253047
	SIRT4	-0.758874238892528	9.02059058334978E-08	**-1.09691101702139**	**0.000145855554265**	-0.860230623554701	1.14510194167515E-11
	SIRT5	0.20489250673979	0.01120540626953	0.229706953060574	0.114393495946943	0.209987011510046	0.002691642119351
	SIRT6	-0.723937889741504	3.67194874104847E-14	-0.589440590440079	0.002989864100189	-0.701747840008799	4.94277924111335E-16
	SIRT7	-0.005517165999578	0.94012562811448	0.238522462602338	0.159907177558506	0.033412090862054	0.654108726414004

The highlighted values refer to the results considered significant with values of log2foldchange> 1 or <-1 and pvalue <0.0001. *Differential expressed genes that are founded in all groups

Apêndice 3. Differential expressed genes analysis results of tumor samples TCGA COAD, TCGA READ and TCGA CRC filtered with purity compared with TCGA adjacent normal samples

Desacetilases		TCGA COAD		TCGA READ		TCGA CRC	
		log2FoldChange	pvalue	log2FoldChange	pvalue	log2FoldChange	pvalue
Class I	HDAC1	-0.015369484537772	0.838196363637393	-0.223703890018392	0.140969911453547	-0.068035197501387	0.302540333054175
	HDAC2	0.869165791430126	6.19663788447768E-23	0.444752778708351	0.008671855424738	0.784670890657238	1.37018926944624E-23
	HDAC3	0.347670986620745	7.25379059875232E-09	0.153055857777498	0.23083651561039	0.306760328557645	1.20739409710045E-08
	HDAC8	0.856079738420008	2.65457479505887E-19	0.498271975766653	0.000807546550364	0.781120110302655	3.86509250186832E-22
Class IIa	HDAC4	-0.56330766100256	2.78631475324094E-08	**-1.13966963776765**	**1.04222308373572E-08**	-0.708718686341533	5.07568669359888E-15
	HDAC5	-0.250938721642451	0.008998890816966	-0.149717501181233	0.424095688144087	-0.222139939539317	0.008676489965397
	HDAC7	0.523416936326445	2.79303529727836E-11	0.232431202886545	0.11822966917804	0.463416351406791	7.85811394473388E-12
	HDAC9*	**-1.34131138954372**	**5.0419106179526E-09**	**-2.21683359768481**	**1.04222308373572E-08**	**-1.57695271444143**	**4.01511233698225E-15**
Class IIb	HDAC6	0.081602393450068	0.303443451656409	0.223872009018925	0.11822966917804	0.113704754438494	0.072097383884508
	HDAC10	0.848812673818391	3.82396094060018E-15	**1.21334225631406**	**5.68306991494576E-08**	0.926194056537568	1.76430687469378E-21
ClassIV	HDAC11	-0.067737899237906	0.458288281689452	0.01696920276531	0.913759567405475	-0.043451429746851	0.593383119929536
Class III	SIRT1	-0.112514802319448	0.330373711408948	-0.870676243653593	1.27071261040554E-05	-0.298506741371075	0.002291398835727
	SIRT2	-0.106370448855431	0.131094370849648	-0.196121036120196	0.163017021747333	-0.126645325262612	0.036070605455729
	SIRT3	-0.204810000674613	9.85183801641185E-05	-0.05386967871219	0.656162944315659	-0.177775308725926	0.000247785061955
	SIRT4	-0.761862396733564	5.15364788778202E-08	**-1.10191622838158**	**0.000138980652188**	-0.863462371624958	7.54590437481055E-12
	SIRT5	0.201033921658635	0.014234596174022	0.227035426492828	0.11822966917804	0.205143396449019	0.003210502770004
	SIRT6	-0.7183914679539	3.41410203746844E-14	-0.585327018013378	0.003185504147702	-0.696476173386486	5.74769230814263E-16
	SIRT7	-0.006408827098476	0.932689175922575	0.244063116430266	0.147770494318466	0.03429980726324	0.646617871454814

The highlighted values refer to the results considered significant with values of log2foldchange> 1 or <-1 and pvalue <0.0001. *Differential expressed genes that are founded in all groups

Apêndice 4. Differential expressed genes analysis results of low purity tumor samples TCGA COAD, TCGA READ and TCGA CRC compared with TCGA adjacent normal samples

Desacetilases		TCGA COAD		TCGA READ		TCGA CRC	
		log2FoldChange	pvalue	log2FoldChange	pvalue	log2FoldChange	pvalue
Class I	HDAC1	-0.060892196165869	0.381858682442873	-0.263042190061207	0.083454820671492	-0.108832893401935	0.080044617595742
	HDAC2	0.811223085825016	1.0219436745543E-20	0.459659336814225	0.00888439256964	0.755952614576261	6.609856663986E-22
	HDAC3	0.328969963065439	1.02857699148489E-08	0.185326165885484	0.120871958250135	0.301228681488218	2.11574762996703E-09
	HDAC8	0.815014907490581	3.02767183602262E-20	0.544297199624473	0.000459597906152	0.764825201666502	2.10708730451754E-23
Class IIa	HDAC4	-0.485025145039751	4.27094211327243E-06	**-1.22231137849421**	**3.67010846792045E-13**	-0.664636768726419	3.34956458388361E-13
	HDAC5	-0.235032207428248	0.016239702637214	-0.083010554135647	0.69149703368302	-0.196061668142998	0.028649509388597
	HDAC7	0.645248433807108	5.32522752413581E-15	0.337086749641756	0.019837891824476	0.584828964632806	1.90261635045388E-16
	HDAC9*	**-1.00590256179804**	**2.48783339200495E-06**	**-1.78418508767311**	**1.59735175273336E-06**	**-1.2159323222246**	**1.31918717148836E-10**
Class IIb	HDAC6	0.081314675658276	0.271700341889202	0.230556911743611	0.083454820671492	0.117353158144966	0.066531113062469
	HDAC10	0.8422712482413	4.49552171978913E-14	**1.15741884569291**	**8.39919574555498E-07**	0.909792083761157	2.6856235918299E-19
ClassIV	HDAC11	-0.139704977564957	0.131013238184311	0.016572834713664	0.918003244757537	-0.096714807713822	0.211028507386704
Class III	SIRT1	-0.088949263119503	0.400931439197137	-0.756721950930915	6.44409201866147E-05	-0.248554177223904	0.011022501329694
	SIRT2	-0.088532131768823	0.21168881808789	-0.155048490672729	0.230218080602069	-0.101117676479814	0.080044617595742
	SIRT3	-0.205447995515052	4.10327656170293E-05	-0.048718969149867	0.700076022223242	-0.177048019897968	0.000180331488962
	SIRT4	-0.741093184148053	1.4707708188126E-07	**-1.00714858181445**	**0.000709323660356**	-0.828085937141651	8.00066095818575E-11
	SIRT5	0.092835721709832	0.253390592056749	0.226494938710281	0.120871958250135	0.124008081219723	0.066802592323873
	SIRT6	-0.721396652475548	4.49552171978913E-14	-0.539310323559989	0.00888439256964	-0.687057502319009	5.66711291619486E-15
	SIRT7	0.009669153188361	0.926846350866818	0.194330438475445	0.225372950130769	0.036478087244604	0.634061117432756

The highlighted values refer to the results considered significant with values of log2foldchange> 1 or <-1 and pvalue <0.0001. *Differential expressed genes that are founded in all groups

Apêndice 5. Summary of the main studies on the expression of histone deacetylases in colorectal cancer.

Study	Method	Sample count		Cancer location and stage	Principal findings	Prognostic
		Tumor	Normal			
Zhu P. *et al* (2004)	IHC	57	46	CRC	Elevated HDAC expression was observed in 47 samples	None
Mimori K. *et al* (2005)	RT-PCR (Lightcycler Instrument, Roche, Mannheim, Germany)	61	61 (paired normal CRC)	CRC	32 patients whose tumour tissue showed overexpression of EZH-2 also had a significantly worse prognosis/ In addition, a significant correlation between EZH-2 and HDAC-1 expression was observed in 61 CRC cases with 20 cases of both high EZH-2 and high HDAC-1 expression showed poor prognoses	Depth of wall Invasion; HD; LM; LI; Stage of disease; Vascular vessel invasion
Huang B.H. *et al* (2005)	qRT-PCR (LightCycler System instrument, Roche, Mannheim, Germany); TMA; IHC; WB	45	45 (paired normal mucosa)	CRC	HDAC2 was upregulated by at least two-fold in 9/16 tumor samples compared to matched normal mucosa. In contrast, only 6/16 showed a more than two-fold upregulation of HDAC1 and HDAC1 and HDAC2 protein expression is increased in colorectal tumors.	None
Giannini R.	qRT-PCR	40	40 (surrounding	CRC	mRNA expression levels of	Age; HD; Sex; TMN

***et al* (2005)**	(iCycler Thermal Cycler System apparatus, BioRad)		normal)		HDAC1, HDAC2 were higher in the carcinomas than in the normal colorectal mucosa, but statistical significance was reached only for HDAC2.	stage; Tumor side
Ozdag H. *et al* (2006)	qRT-PCR (ABI 7900 Sequence Detection System, Applied Biosystems): HDAC 1 (SSCP/ HA) HDAC 2 (SSCP/ DHPLC)	20	20	CRC	HDAC1, HDAC5, HDAC7A, SIRT1, and SUV39H1 expression profiles were distinctive for colorectal cancers and normal colorectal mucosa. HDAC2 overexpression was observed in 50% of colorectal tumours compared to their normal pairs.	None
Lleonart M.E. *et al* (2006)	qRT-PCR (ABI PRISM 7700 instrument, Applied Biosystems)	20	20	Colon cancer	HDAC4 was found to be downregulated in 15 colon tumor tissue	Dukes Stage; Infiltration; LM; Survival; TNM Stage; TS
Ishihama K. *et al* (2007)	IHC; ISH; RT-PCR (One Step RT-PCR assay kit, Qiagen)	64	20 (colonic mucosae from the resection margin of CRC)	CRC Adenocarcinoma	High expression. levels of HDAC1 mRNA was greater in tumor tissue	Age; Depth of invasion; Gender; HD; LM; TNM Stage; TS
Stunkel W. *et al* (2007)	TMA; in situ data and IHC	230	230 (corresponding colonic mucosa)	Colon cancer	Immunostaining of normal colon tissue revealed that SIRT1 is also mostly cytoplasmic. Cytoplasmic expression was	None

					also observed at increased levels for many samples from colorectal carcinoma	
Weichert W. *et al* (2008)	IHC	140	Not included	CRC	Class I HDACs isoforms highly expressed in a subset of colorectal carcinomas with positivity for HDAC1in 36.4%, HDAC2 in 57.9%, and HDAC3 in 72.9% of cases. Expression was significantly enhanced in strongly proliferating (P = 0.002), dedifferentiated (P = 0.022) tumors. High HDAC expression levels implicated significantly reduced patient survival (P=0.001), with HDAC2 expression being an independent survival prognosticator (hazard ratio, 2.6; P = 0.03)	Age; Dukes stage; Tumor stage; Nodal stage; Metastasis; Grade
Ashktorab H. *et al* (2009)	TMA; IHC	250	57 (matched adjacent normal)	Colon adenoma and carcinoma (TNM II-IV)	HDAC2 nuclear expression was detected at high levels in 81.9%, 62.1%, and 53.1% of CRC, adenomas, and normal tissue, respectively (P = 0.002).	TNM stage
Kumamoto K. *et al* (2009)	IHC; RT-PCR (ABI prism 7500, Applied	39	39 (adjacent noncancerous tissue)	Colon cancer	Subsequent microarray analyses revealed that ING2 upregulates expression of	None

	Biosystems, Foster City, CA)				matrix metalloproteinase 13 (MMP13), which enhances cancer invasion and metastasis. The combination with ING2 and HDAC1 or mSin3A overexpression remarkably enhances MMP13 expression, when compared to only ING2 overexpression	
Nosho K. *et al* (2009)	TMA; IHC; RT-PCR (MethyLight)	485	Not described	CRC	180 (37%) tumors showed nuclear overexpression of SIRT1; SIRT1 expression is associated with CIMP and MSI, independent of other clinical and molecular variables	Age; BMI; Grade; LINE-1 methylation; Mucin; Sex; Signet ring cells; TL; TNM stage; CIMP/MSI status; *p53*; *β-catenin*; *FASN*; *COX2*; *BRAF*; *KRAS*; *PIK3CA*
Higashijima J. *et al* (2011)	IHC	74	Not described	Colon cancer	HDAC1 expression in 58 patients, associated with depth of tumor invasion, distant metastasis, stage and curability. The survival rate in the HDAC1-positive group was significantly worse than that of the HDAC1-negative group ($p<0.05$)	Age; Curability; Differentiation; DM; Gender; TI; TNM stage; VI
Chou C.W. *et al* (2011)	RT-PCR (ABI Prism 7900 Sequence	14	14 (adjacent normal)	Colon Cancer	The positive correlation between EGFR and HDAC3 expression was also observed	*EGFR*

	Detection System, Applied Biosystems, Foster City, CA)				in fourteen pairs of human colon tumor and adjacent normal tissues	
Gemoll T. *et al* (2011)	TMA-based IHC	78	78 (adjacent normal mucosa)	CRC	HDAC2 nuclear immunopositivity (score 1, 2, and 3) was more frequently present in aneuploid (91.3%) than in diploid (70%) carcinomas (p = 0.02). Western-blot analysis of selected target proteins confirmed HDAC2 as significantly differentially expressed protein	Age; Sex; Post-operative TNM stage and grading
Jang. C.H. *et al* (2012)	TMA; IHC	497 + 52 adenomas + 65 metastatic lymph nodes _ 61 distant metastatic lesions	24 (Normal colonic mucosa)	CRC	All normal colonic mucosa showed SIRT1 expression with no exception, and 42 (80.8%) of 52 adenomatous polyps were positive for SIRT1. However, only 208 (41.9%) of 497 colorectal adenocarcinomas were positive. Moreover, 45 (35.7%) of 126 metastatic tissues were positive. The associations between SIRT1 expression and clinicopathological	Age; Gender; TL; Growth pattern; Histological subtype; Depth of invasion; LM; DM; AJCC stage; Mismatch repair proteins (MLH1, MSH2)

					parameters revealed that loss of SIRT1 expression was associated with proximal tumour location, mucinous histology and defective mismatch repair protein expression. This suggests that loss of SIRT1 expression is associated with the microsatellite instability phenotype of colorectal adenocarcinoma.	
Pazienza V. *et al* (2012)	qRT-PCR (ABI PRISM 7700 Sequence Detection System,Applied Biosystems, Applera)	19	19 (matching normal)	CRC	The expression levels of SIRT1 were evaluated in 19 pairednormal and CRC tissues from 13 male and 6 female patients(mean age ± SD, 68.9 ± 8.8 years; age range 55–81 years). Compared to normal tissue with an expression profile normalized to 1, in tumor samples, SIRT1 was expressed at lower levels (median = 0.73, Q1–Q3 = 0.47–0.87, p = .003)	Age; Gender; Cancer location; Grading; Modified Dukes staging system; Histologic type
Kriegl L. *et al* (2012)	IHC	121	121 (normal mucosa)	CRC (Serrated lesions)	On classical route SIRT1 is overexpressed and are crucially involved in the alternative, serrated pathway to colorectal câncer	BRAF and KRAS mutations

Jung W. *et al* (2013)	IHC	349	349 (normal mucosa)	CRC	SIRT1 overexpression (235 patients) is a good prognostic fator for CRC	Age; HD; TNM stage
Yu H. *et al* (2014)	qRT-PCR; WB; IHC	131	131 (adjacent non-tumorous colorectal tissue)	CRC	The Sirt7 protein level significantly correlated with tumor stage (P ¼ 0.029), lymph node metastasis (P ¼ 0.046), and poor patient survival (P < 0.05)	Age; Gender; TL; TS; Histology; Tumor invasive depth; Lymph node status; Preoperative CEA; AJCC/TNM stage; Preoperative CA
Lv L. *et al* (2014)	TMA; IHC; WB	120	120 (paired distant normal mucosa)	CRC adenocarcinoma	SIRT1 expression was frequently increased in human colorectal cancer tissues correlated with advanced tumor stage and metastasis in colorectal câncer, with poor OS and DFS	Age; Depth of TI; Gender; HG; LM; Recurrence; TS
Chen X. *et al* (2014)	IHC; WB	82	Not described	CRC adenocarcinoma	high expression. of SIRT1 was associated with poor prognosis in CRC patients. SIRT1 had no significant correlation to clinicopathological features such as age, gender, location and T-category. On the other hand, the altered SIRT1 expression was significantly correlated with the number of cancer sites (P50.03),	Age; Gender; Location; Metastasis; Metastasis site; Number of sites; T-category

					metastases (P50.02) and metastatic sites (P50.02)	
Yang H. *et al* (2014)	IHC; qRT-PCR (Power SYBR Green PCR Master Mix, Bio-Rad, Hercules, CA, USA).	8	8 (adjacent normal)	CRC	HDAC1 and HDAC2 expression are up-regulated in colon câncer patients	None
Kewi J. *et al* (2014)	TMA; IHC	120	120 (paired normal mucosa	CRC adenocarcinoma	SIRT1 overexpression correlated with advanced stage and poor prognosis	Age; Depth of tumor metastasis; Gender; HG; Recurence; TS
Drew J. *et al* (2014)	IHC; RT-PCR (ABI-7500Fast, Applied Biosystems, UK).	30	30 (paired normal)	Colon cancer	Normal tissue is characterised by higher expression. Level of HDAC2 when compared to adenomatous polyp or carcinoma tissues	None
Liu C. *et al* (2014)	IHC; qRT-PCR (ABI 7900 Real-Time PCR System, Applied Biosystems, Foster City, CA, USA).	127	127 (noncancerous)	Colon cancer	High SIRT3 expression in the cytoplasm significantly correlates with high tumor grades, positive lymph node status, and poor prognosis. The overall survival was 80.2% among patients with low SIRT3 expressions and 55.9% among patients with high SIRT3 expressions	Age; Gender; LM; TNM stages; TS; Transvascular metastasis
Gerçeker E. *et al* (2015)	qRT-PCR (REST 2009, Relative Expression Software	20	20 (control group with normal colon mucosa)	CRC	Higher overexpression of HDAC1 and HDAC7 in the CRC group.	None

	Tool V.2.0.13)					
Kasap E. *et al* (2015)	qRT-PCR	25	25 (normal colon mucosa)	CRC	HDAC1 and HDAC7 were overexpressed (fold change>2), the overexpression of the was statistically significantly higher levels than the control group ($p<0.05$) and were statistically significantly more overexpressed in poorly differentiated tumors than well and moderatelly differentiated tumors in the CRC with genes were statistically significantly more overexpressed in stage TNM III–IV than stage I–II in the CRC	AJCC stage; TNM stage
Benard A. *et al* (2015)	IHC stained on TMAs	254	50 (normal colorectal)	CRC	SIRT1 showed lower nuclear expression in tumour samples, HDAC2 showed higher nuclear expression. in tumor samples and HDAC1 did not differ between normal and tumour samples.	Age; Adjuvant therapy; Gender; Patient survival; TNM stage; pT stage; pN stage; Histological subtype; TL; Location in the colon; TS; Number of lymph nodes retrieved; MSS status; Tumor in follow-up

Giardina C. *et al* (2015)	NCBI GEO database (GDS2947) analyzed using Excel and GraphPad Prism software	32	32 (adjacent normal)	CRC adenocarcinoma	Up-regulation of the class I HDACs (HDACs 1, 2, 3 and 8) in tumor	None
Zhang X. *et al* (2015)	IHC; WB; RT-PCR (SYBR Green, FP205, TIANGEN, China).	50	20 (adjacent normal paraffin)	CRC	The expressions of SIRT1 are higher in colorectal cancer tissues than normal tissues. SIRT1 expression in cancer tissues was associated with younger patient age, advanced TNM stage and mutant P53 loss.	Age; Gender; Grade; TNM stage; LM; DM; p53 mutation
Miyo M. *et al* (2015)	IHC (VECTASTAIN Elite ABC kit, Vector Laboratories, Burlingame, CA, USA)	142 CRC tumor + 38 CRC adenoma	142 (adjacent normal)	CRC	The proportions of those with high expression of SIRT4 were 61.1% (77 out of 126), 52.6% (20 out of 38), and 52.5% (75 out of 142) in normal colorectal tissue, adenoma, and colorectal cancer, respectively. SIRT4 expression in colorectal cancer decreased with the progression of invasion and metastasis, and a low expression level of SIRT4 was correlated with a worse prognosis.	Age; Sex; Histology; Depth of TI; LM; DM; Venous invasion; Lymphatic invasion

Huang G. *et al* (2016)	IHC; RT-PCR ((DNA Engine Opticon 2 system, Bio-Rad, Hercules, CA, USA); TMA; WB.	89	89 (corresponding normal colorectal)	CRC	SIRT4 was significantly downregulated in colorectal cancer tissues compared with that noted in the corresponding normal tissue. Lower SIRT4 levels were associated with worse pathological differentiation and poorer post-operative overall survival rate	Age; Gender; TS; Growth mode; Degree of differentiation; TI depth and scope; LM; UICC stage and post-operative overall survival time
Shen Z. *et al* (2016)	miRNA microarray (Agilent human miRNA 8 × 15 k microarray, V16.0); qRT-PCR	60	Not described	CRC	qRT-PCR assay showed that the expression of SIRT1 was much higher in CRC tissues than in noncancerous tissues	None
Yu D.F. *et al* (2016)	IHC; WB	40	40	CRC	Sirt1 protein was primarily expressed in the nuclei of the tumor cells, and positive staining was brownish-yellow in color. The relative expression quantities of Sirt1 in the peritumoral normal rectal mucosa and rectal carcinoma were 1.15 and 2.62, and the differences between the two groups were statistically significant ($P<0.05$). The expression level of Sirt1 in colorectal	Age; Gender; Morphological type; TD; Depth of invasion; LM; Duke's stage

					carcinoma was significantly associated with the depth of tumor invasion, differentiation and tumor size (P<0.05). Sirt1 expression was also found to be associated with tumor tissue type, lymph node metastasis, Duke's stage and patient age.	
Lutz L. *et al* (2016)	IHC	47	13 (adjacent non-dysplastic colon mucosa)	CRC	HDAC2 expression was significant higher in tumor cells compared to adjacent non-dysplastic colon mucosa (p=0.02). In contrast, HDAC1 and HDAC3 expression in tumor cells compared to adjacent non-dysplastic colon mucosa were indiscriminative.	Age; Sex; Location; pT; pN; Tumor grading; Histological subtype; MSI status
Osama A. *et al* (2016)	qRT-PCR (Reverse Transcriptase kit, Applied Biosystem, USA); WB	167	87	CRC adenocarcinoma	Sirt-1 protein level was significantly highly expressed in colorectal adenocarcinoma compared to normal and adenoma colonic tissue	Age; Abdominal pain; Bleeding; Constipation; Diarrhea; Familial polyposis; Family history of CRC; IBD; Location; Screening; Sex; Shape; Mass/ polyp; Mass/ Polyp size
Ren N.S.X.	IHC	155 (TCGA	19 (TCGA	CRC	the vast majority (about 99%)	Age at diagnosis;

***et al* (2017)**		tumour group) + 348 for protein expression	normal group)		of colorectal cancer patients had a normal SIRT1 gene, and no deletion of this gene was detected in any of these database. Comparison of human SIRT1 mRNA levels between colon tumors and normal samples from non-colorectal cancer individuals in the TCGA revealed that human SIRT1 levels were modestly but significantly reduced in colon tumor samples. SIRT1 proteins were detected in both tumor tissue and tumor-adjacent colorectal tissue, and no statistically significant difference was detected between expression in tumor tissue and tumor-adjacent tissue after stratifying by sex, race, age, or tumor stage.	Died within 5 years of surgery; HG; Race; Sex; Tumor stage
Mao Q.D. *et al* (2017)	qRT-PCR (Power SYBR-Green Master Mix, Bio-Rad, USA).	20	20 (adjacent non-tumor)	CRC	mRNA expression level of HDAC2 in CRC tissues was markedly enhanced compared with non-tumor tissues ($P<0.05$).	None
Zhang L.L. *et al* (2017)	IHC	31	26 (adjacent non-cancerous)	Colon cancer	SIRT2 is downregulated in CRC biopsy samples (n=31)	Age; Gender; LM; Location; TNM

					compared with the xpressio non-cancerous tissues. CRC xpression were positive for SIRT2 despite a lack of Expression. in the primary tumor.	staging
Nemati M. *et al* (2018)	qRT-PCR (RNeasy Mini Kit, Qiagen)	48	48 (matched tumor-adjacent normal)	CRC	HDAC3 was highly expressed in colorectal tumors compared to normal colorectal tissues (p < 0.05). Furthermore, we found significant correlations between HDAC3 expression levels and tumor differentiation grades (p < 0.05).	Age; AJCC stage; Depth; Gender; TS; TD; LM; Venous invasion
Qi J. *et al* (2018)	IHC; qRT-PCR (7500 FastReal-Time Two-step PCR system, Applied Biosystems; Thermo Fisher Scientific, Inc.); WB.	113	29 (adjacent normal colorectal)	CRC	SIRT6 protein and mRNA expression levels were significantly reduced in CRC tissues, were associated with the tumor depth invasion, stage, differentiation grade and the presence of lymph node metastasis (P<0.05).	Depth of invasion; DM; TS; LM
Zhu Y. *et al* (2018)	IHC; qRT-PCR (Bio-Rad CFX96TM Real-Time PCR system, Bio-Rad Laboratories, Inc.,	15 + 155 (TCGA)	15 (adjacent normal) + 19 (TCGA normal)	CRC	SIRT4 expression is significantly downregulated in CRC tissues	None

	Hercules, CA, USA); WB.					
Tian J. *et al* (2018)	IHC; qRT-PCR (DA7600 Real-time Nucleic Acid Amplification Fluorescence Detection System, Bio-Rad); WB.	90	22 (matched adjacent non-tumor)	Colon cancer	both mRNA and protein expression levels of SIRT6 were reduced in colon cancer tissues when compared with that of the adjacent normal tissues. The expression level of SIRT6 negatively correlated with high differentiation status, advanced TNM grade, and high occurrence rates of vascular invasion and lymph node invasion. Besides, patients with low expression of SIRT6 had shorter survival than those of patients with high SIRT6 expression	Age; Differentiation status; Gender; LI; TNM stage; VI
Wang Y.Q. *et al* (2018)	IHC; qRT-PCR (ABI Prism 7900HT Sequence Detection System, Applied Biosystems, USA); WB	88 +262 (GSE 68468) + 186 (GSE 41258)	55 (GSE 68468) + 54 (GSE 41258)	CRC	SIRT5 was strongly positive in CRC compared with the corresponding normal tissues, associated positively with larger tumor size (P = 0.036), increased lymph node metastasis (P = 0.016), advanced tumor stage (P = 0.038), and American Joint Committee on Cancer (AJCC)	AJCC stage; DM; TS; LM; Tumor stage; HG; Overall Survivor

					stage (P =0.005). The upregulation of the SIRT5 mRNA level in CRC was also validated independently in two published microarray data sets. SIRT5 upregulation is independently associated with higher risk of mortality in patients with CRC, with an average follow-up of 5 years	
Du Z.G. *et al* (2018)	IHC; RT-PCR	129	Not described	CRC	Sirt5 was predominantly expressed in tumors with wild-type Kras than in those with Kras mutations (76.9% versus 19.6%)	Age; DM; Primary site; Pathological stages; Local invasion; TS; Wild-type *KRAS*; *KRAS* mutations
Hong W.G. *et al* (2018)	IHC	265 + 2132 (meta-analysis for prognostic)	Not described	Colon cancer	SIRT1 was highly expressed in 24.5% of the 265 CRC specimens analyzed. High SIRT1 expression correlated with vascular invasion (P= 0.041) but there was no significant correlation between SIRT1 expression and other clinicopathological parameters	Age; Depth of Tumor; Distant Metastasis; Gender; TS; TL; TD; LM; LM ratio; Pathologic tumor node metastatic (pTNM) stages; Vascular; Lymphatic, and perineural invasion
He P. *et al* (2018)	IHC; qRT-PCR (7500 Fast Real-Time PCR system, Applied	96	96 (paired non-tumor adjacent normal)	CRC	HDAC3 protein expression was markedly upregulated in the tumor tissues ($\chi 2$=5.658, P=0.0174), and your	Age; AJCC stage; HG; LM; Location; Node stage; Sex; Tumor stage; Venous

	Biosystems, Carlsbad, CA, USA)				expression. was associated with T stage, N stage, AJCC stage, histological grade and venous and lymphatic metastasis.	metastasis
Wei Z. *et al* (2019)	IHC; qRT-PCR (CFX Connect Real-Time System, Bio-Rad); WB.	344	344 (adjacent normal)	CRC (primary tumor)	most tumor samples with high SHMT2 expression exhibited increased SIRT3 protein expression compared with adjacent normal tissues (P < 0.01).	None
Tong J. *et al* (2019)	qRT-PCR (cat. no. ABI-7300; Applied Biosystems; Thermo Fisher Scientific, Inc.); WB	58	58 (adjacent normal)	CRC	Compared to normal tissues, the expression of HDAC2 at the mRNA level in tumors was significantly increased.	None
Xiong W. *et al* (2019)	qRT-PCR (ABI 7300 Sequence Detection System, Applied Biosystems, Foster City, CA, USA); WB	48	48 (adjacent noncancerous)	CRC	HDAC1 expression in CRC tissues was higher than in the adjacent noncancerous tissues	None
Mo S. *et al* (2019)	Datasets that were created by Affymetrix HG-U133 plus 2.0 platform	test set (N = 386) + internal validation set (N =	Not described	Colon cancer (stages I-III)	SIRT1 was downregulated in early relapse group compared with long-term survival group. (Long-term survival refers to no relapse after a	None

		111).			minimum of 5 years follow-up. Data in test and internal validation sets from GSE39582 were divided into early relapse group and long-term survival group)	
Tang W. *et al* (2019)	ISH; IHC; qRT-PCR	81	81 (non-tumor)	CRC	Cancer tissues (T) exhibited higher HDAC2 protein expression levels compared with the corresponding noncancerous controls (N).	None
Qiao P.F. *et al* (2020)	qRT-PCR (ABI StepOne system); WB	60	60 (adjacent normal colon)	Colon cancer	SIRT1 mRNA expression. was upregulated in CRC tissues compared to the adjacent normal colon tissues	None
Saleh R. *et al* (2020)	qRT-PCR (QuantStudio 6/7 Flex realtime PCR system, Applied Biosystems, California, USA)	68	30	CRC	SIRT1 mRNA levels in tumor tissue were upregulated in advanced disease stages compared to early stages and normal tissues.	Age; Gender; Tumor budding; TNM stage
Du F. *et al* (2020)	IHC; qRT-PCR (aqMan Fast Advanced Master Mix, Applied Biosystems); WB.	100	100 (adjacent healthy CRC epitelial)	CRC (primary tumor)	SIRT2 was significantly down-regulated in CRC tissues relative to adjacent normal tissues. The absence of SIRT2 was correlated with the American Joint Committee on Cancer (AJCC) stage, distant metastasis and	Age; AJCC stage; DM; LM; Sex; TL; TS; TD; TI

					lymph node metastasis in CRC.	
Lee G. J. *et al* (2020)	IHC	101 (elderly patients) + 29 (Young patients)	Not described	CRC	Overexpression of SIRT1 and SIRT2 was observed more commonly in the very elderly patients than in young patients with CRC. SIRT1 overexpression was associated with poor outcome in the very elderly patients with CRC.	Age; BMI; Diabetes; Sex; TL; Staging; HD
Chen C. *et al* (2020)	WB; IHC; RT-PCR (ABI 7300 real-time PCR system, Applied Biosystems).	80	80 (matched normal)	CRC (primary tumor)	HDAC1 was overexpressed in CRC relative to matching pre-cancerous Tissues (Oncomine database) and the expression of mRNA for HDAC1 was elevated in clinical CRC tissue.	Not included

AJCC stage - American Joint Committee on Cancer, BMI - Body mass index, CRC - Colorectal cancer, DHPLC - Denaturing High Performance Liquid Chromatography, DM - Distant metastasis, HD - Histological differentiation, HG - Histological grade, IHC - Immunohistochemestry, ISH - in situ hybridisation, LI - Lymph node invasion, LM - Lymph node Metastasis, MSI - Microssatelite instability status, MSS - Microssatelite stability status, MTD - Tumour differentiation, SSCP/HA - Single Strand Conformation Polymorphism/Heteroduplex Analysis, TI - Tumour invasion, TL - Tumour Location, TS - Tumour Size, TMA - Tissue Microarray, UICC - Union for International Cancer Control, VI - Vascular Invasion, WB – Western blot.