

**IDENTIFICAÇÃO DE MARCADORES MOLECULARES
DE RESPOSTA AO TRATAMENTO NEOADJUVANTE
EM PACIENTES COM ADENOCARCINOMA DE RETO**

LUIA MATOS DO CANTO ALVIM

**Tese apresentada à Fundação Antônio
Prudente para a obtenção do título de Doutor
em Ciências**

Área de Concentração: Oncologia

Orientadora: Prof. Dra. Silvia Regina Rogatto

Co-Orientador: Dr. Samuel Aguiar Júnior

São Paulo

2018

FICHA CATALOGRÁFICA

Preparada pela Biblioteca da Fundação Antônio Prudente

Alvim, Luisa Matos do Canto

Identificação de marcadores moleculares de resposta ao tratamento neoadjuvante em pacientes com adenocarcinoma de reto / Luisa Matos do Canto Alvim - São Paulo 2018.

160p.

Tese (Doutorado)-Fundação Antônio Prudente.

Curso de Pós-Graduação em Ciências - Área de concentração:
Oncologia.

Orientadora: Silvia Regina Rogatto

Descritores: 1. Neoplasias Retais/Rectal Neoplasms. 2. Terapia Neoadjuvante/Neoadjuvant Therapy. 3. Biomarcadores Tumorais/Biomarkers, Tumor. 4. Metilação de DNA/DNA Methylation. 5. Transcriptoma/Transcriptome

“A tarefa não é tanto ver aquilo que ninguém viu, mas pensar o que ninguém
ainda pensou sobre aquilo que todo mundo vê.”

Arthur Schopenhauer

“A ciência nunca resolve um problema sem criar pelo menos outros dez.”

George Bernard Shaw

“Concedei-me Senhor, serenidade necessária para aceitar as coisas que
não posso modificar, coragem para modificar aquelas que posso e sabedoria
para distinguir uma das outras.”

Reinhold Niebuhr

DEDICATÓRIA

Aos pacientes

AGRADECIMENTOS

Aos pacientes, que generosamente aceitaram contribuir para esse trabalho, fornecendo os seus dados e amostras para que, por meio da pesquisa, possamos compreender melhor o mau que os aflige.

À minha orientadora, Dra. Silvia R. Rogatto, pela dedicação, eficiência, incentivo, organização, apoio, e suporte incondicional na realização desse trabalho. Por sempre fazer tudo o melhor possível com qualidade e zelo, prezando por aquilo que acha mais importante. Por ter ido sempre além do esperado, surpreendendo sempre. Pelas oportunidades de aprendizado que vão além do laboratório.

Ao meu co-orientador, Dr. Samuel Aguiar Jr. por ter proposto esse trabalho e pelos conselhos, sugestões e incentivos ao longo do seu desenvolvimento.

Aos órgãos de fomento, Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP - processos 2014/06323-9 e 2015/25803-4), Conselho Nacional de Pesquisa do Desenvolvimento Científico e Tecnológico (CNPq) e Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo financiamento do projeto e das bolsas recebidas.

Aos amigos do grupo NeoGene, que participaram da caminhada para a realização desse trabalho, André Villacis, André Guollo, Andréa Latorre, Caroline Beltrami, Fabio Marchi, Flávia Faldoni, Graziela Spilborgs, Hellen Kuasne, Isabella Maria Paião Ortiz, Juan Munoz, Julia Mello, Maísa Pinheiro, Marco Lopez, Mariana Bisarro, Mateus de Camargo Barros Filho, Priscila Mayrink, Renne Amorim, Sandra Drigo, Sara Martorelli, Tatiane Basso e Vanessa Simões e aos agregados Cadu, Juliano e Talita.

Em especial, à Hellen, que veio ao mundo para o fazer melhor e, com todo o seu carisma e seu coração gigante conseguiu me ensinar muito mais do que eu jamais sonharia em aprender. Ao Mateus, por toda a paciência e impaciência que sempre me inspirou a fazer o melhor e me ajudou a organizar ideias mais confusas que o Batman. Ao Fabio, por estar sempre

disposto a ajudar, pelas ideias e sugestões, pelas palavras de conforto e por todo o trabalho e tempo dedicados.

Aos dedicados funcionários do CIPE, por toda a ajuda e o tempo compartilhado mesmo quando todos já haviam ido embora: Jennice, Felipe, Danilo, Juliana, Rogério, Kauê, Gilmara, Maria Antonieta, Carlos.

Aos amigos e vizinhos de laboratório pela troca de experiências: Fernanda Fortes, Karina Miranda, Kelvin, Alexcia, Emne, Bianca, Hermano, Luana, Rafael, Rodrigo, Alexandre, e em especial ao Antuani Baptistella.

Ao Departamento de Patologia, em especial ao Severino e à Dra. Maria Dirlei Begnami, pela colaboração, dedicação e paciência para revisar todas as lâminas utilizadas nesse estudo.

Ao Biobanco do A.C.Camargo Cancer Center, Louise, Eloisa e Ana Paula.

Ao Núcleo de Tumores Colorretais, em especial à enfermeira de pesquisa Bruna Catin.

Aos professores da pós graduação, pela qualidade das aulas proferidas, pelo tempo e dedicação ao programa.

À equipe do Programa de Pós Graduação em Ciências (Oncologia) por todo o apoio e orientação.

À biblioteca.

Ao A.C.Camargo Cancer Center por investir em ensino e pesquisa, proporcionando esperança para os milhares de brasileiros que foram, são ou serão acometidos por câncer.

Ao Hospital de Câncer de Barretos, na pessoa do Dr. Cristóvam Scapulatempo-Neto pela contribuição a esse trabalho.

Ao Dr. Jan Baumbach e ao Simon Larsen, pela colaboração com as análises.

Ao departamento de Genética Clínica do hospital de Vejle, na Dinamarca, por toda a gentileza, carinho, eficiência, paciência e por todos os sorrisos energizantes em especial ao Mads e à Annabeth, mas também à Lene, Mette, Louise, Birgit, Birgitte, Charlotte, Lina, Sofie, Christian, Jesper, Else Marie, Helle E, Helle K, Helle J, Elisabeth, Cecilie.

À minha família, por acreditar no meu potencial e me incentivar a superar os obstáculos. A minha mãe, Rochele, por todo o suporte, amor, carinho e compreensão. Por me ensinar que nada é impossível e que tudo tem uma razão de ser e, por aguentar e ter a paciência de responder a todos os meus “por quês” que desde as fraldas até hoje a perseguem. Aos meus tios, Azair e Ricardo, pela inspiração, pelo carinho, pelo teto, por tudo.

Às minhas madrinhas, Rosi e Lu.

To my dearest Dominique, who gave me more than all the support I needed to finish my thesis, who gave me love and who has been teaching me the power of love.

Aos meus amigos.

Muito obrigada!

RESUMO

Canto LM. **Identificação de marcadores moleculares de resposta ao tratamento neoadjuvante em pacientes com adenocarcinoma de reto.** São Paulo; 2018. [Tese de Doutorado-Fundação Antônio Prudente].

Pacientes com câncer de reto (CaRe) em estadio II ou III são tratados com radioquimioterapia neoadjuvante (nCRT), a qual é responsável pela diminuição de recidivas loco-regionais e amputação do esfíncter. A resposta à terapia é variável, incluindo aqueles que não respondem ao tratamento e os que apresentam resposta patológica completa - pCR (15 a 30%). O objetivo deste estudo foi identificar alterações genéticas e epigenéticas que corroborem para a elucidação dos mecanismos de resistência ao tratamento radio e quimioterápico neoadjuvante em pacientes com adenocarcinoma de reto. Foram incluídos 34 CaRe estadio II ou III, submetidos a nCRT. Onze pacientes apresentaram pCR (ypT0N0M0) e 23, resposta patológica incompleta (pIR). Foram utilizadas 10 amostras de reto normal (necrópsias) como referência. As alterações genômicas e de expressão gênica foram avaliadas pelas plataformas *CytoScan HD Array* e *GeneChip Human Transcriptome Array* (Affymetrix/Thermo Fisher), respectivamente. O sequenciamento de 105 genes (SureSelect^{XT} Custom Panel, Agilent) relacionados ao câncer foi realizado no NextSeq 550 (Illumina). A plataforma *Infinium® Human MethylationEPIC BeadChip* (Illumina) foi utilizada para a análise do metiloma. Alterações genômicas identificadas em CaRe evidenciaram alta instabilidade cromossômica e um número significativo de casos com alterações na via de reparo por recombinação homóloga (HR), identificada por mutação em genes da via ou por escores de deficiência (tAI, LST e HDR-LOH). Essas alterações foram corroboradas por vias alteradas relacionadas à regulação das ciclinas e do ciclo celular, assim como do ponto de checagem de reconhecimento de quebras de dupla fita de DNA identificadas por meio do enriquecimento dos genes diferencialmente expressos. A análise integrada dos resultados do metiloma e transcriptômica

permitiu a identificação de genes desregulados por metilação os quais foram confirmados com dados externos (TCGA). As diferenças moleculares entre os casos com pCR e pIR incluíram alterações genômica estruturais e um maior index de instabilidade genômica (GII) em pCR. Um maior número de casos com pIR apresentou mutações em genes relacionados à via HR. Esses resultados sugerem que platina e inibidores de PARP1 poderiam ser um tratamento alternativo nos casos com deficiência dessa via. Os perfis de metilação do DNA de cada grupo revelaram maior proporção de sondas hipermetiladas em pCR (44% vs. 18% em pIR). Com base nos dados de três sondas diferencialmente metiladas, foi possível desenvolver um classificador capaz de prever a resposta à nCRT (100% de sensibilidade e 90% de especificidade). Usando os resultados do transcriptoma, foram identificadas diferentes vias alteradas em CaRe de acordo com a resposta à nCRT. Enquanto os casos com pCR apresentaram vias alteradas relacionadas à resposta do sistema imune e à sinalização de Wnt, o grupo de pIR revelou alterações de controle do ciclo celular e do reparo de quebras de dupla fita de DNA, além de desregulação de mecanismos epigenéticos e de níveis de transcritos. Utilizando os dados de expressão gênica, foi realizada uma análise do secretoma de acordo com a resposta à nCRT, identificando-se ASPH como potencial marcador de pCR. A mesma estratégia revelou UBE2C e CEMIP como marcadores de resistência à nCRT, os quais poderiam ser investigados em biópsias líquidas. Este estudo revelou potenciais marcadores moleculares e mutações que têm potencial para subclassificar os pacientes de acordo com a resposta à terapia atualmente utilizada. Em adição, nossos achados deram evidências de novas estratégias terapêuticas que poderiam ser aplicadas para os pacientes com CaRe.

SUMMARY

Canto LM. **[Identification of molecular markers of neoadjuvant treatment response in rectal adenocarcinoma patients]**. São Paulo; 2018. [Tese de Doutorado-Fundação Antônio Prudente].

Rectal cancer patients (ReCa) stage II or III undergo neoadjuvant chemoradiotherapy (nCRT), responsible for a decrease in loco-regional recurrence and sphincter amputation. The response to treatment is varied, including patients showing partial or incomplete response to therapy (pIR) and those who achieve pathological complete response - pCR (15 to 30%). The aim of this study was to identify genetic and epigenetic alterations that can be useful to understand the mechanisms of resistance to neoadjuvant chemoradiotherapy in ReCa patients. A total of 34 patients with ReCa stage II or III submitted to nCRT were included. Eleven patients had pCR (ypT0N0M0) and 23 pIR. Ten normal rectal tissue samples (necropsies) were used as reference. Genomic alterations and gene expression profiles were evaluated using the platforms CytoScan HD Array and GeneChip Human Transcriptome Array (Affymetrix/Thermo Fisher), respectively. Targeted next-generation sequencing of 105 cancer related genes was performed using the NextSeq 550 (Illumina). The Infinium® Human MethylationEPIC BeadChip (Illumina) platform was used for methylome analysis. High chromosomal instability and a significant number of cases with changes in the DNA damage repair by homologous recombination pathway (HR), identified by mutation in pathway genes or by deficiency scores (tAI, LST and HDR-LOH), were detected in the genomic analysis. These alterations are corroborated by altered pathways related to the regulation of cyclins and the cell cycle, as well as the double-stranded DNA recognition checkpoint identified by the enrichment of the differentially expressed genes. The integrative analysis of differential methylation and transcriptomic results allowed the identification of genes altered by methylation, which were confirmed with external data

(TCGA). Molecular differences between pCR and pIR cases included structural genomic changes and a high genomic instability index (GII) in pCR. A higher number of cases with pIR showed mutations in genes related to the HR pathway. These findings suggest that PARP-inhibitors and platinum may be alternative therapeutic approaches for these patients. The DNA methylation profiles of each group revealed a higher proportion of hypermethylated probes in pCR (44% vs. 18% in pIR). Based on data from three differentially methylated probes, a classifier was developed showing the ability of predicting the response to nCRT (100% sensitivity and 90% specificity). Using the transcriptome results, different altered pathways in ReCa were identified according to the response to nCRT. While altered pathways in pCR cases were related to immune system response and Wnt signaling, the pIR group revealed changes in cell cycle control and repair of double stranded DNA breaks, as well as dysregulation of epigenetic mechanisms and transcripts levels. Using the gene expression data, a secretome analysis was performed according to the nCRT response, identifying ASPH as a potential pCR marker. The same strategy revealed UBE2C and CEMIP as markers of nCRT resistance, which could be investigated in liquid biopsies. This study revealed molecular markers and mutations that showed the potential to subclassify the patients according to the response to therapy currently used in the clinical practice. Furthermore, our findings give evidences of novel therapeutic strategies for patients with CaRe.

LISTA DE FIGURAS

- Figura 1** Fluxograma da metodologia experimental utilizada neste estudo..... 36
- Figura 2** Visualização esquemática dos segmentos computados pelos três diferentes escores relacionados à deficiência da via de recombinação homóloga..... 51
- Figura 3** *Pipeline* utilizado para o processamento dos dados brutos provenientes do sequenciamento de 105 genes relacionados ao câncer em 31 biópsias de câncer de reto 53
- Figura 4** Representação esquemática da distribuição genômica e do contexto CpG das sondas diferencialmente metiladas identificadas nas amostras de câncer de reto e das sondas presentes nas plataformas utilizadas em cada estudo 60-1
- Figura 5** Descrição das sondas diferencialmente metiladas com correlação significativa com os dados de expressão gênica identificadas neste estudo..... 66-7
- Figura 6** Estratégia utilizada para análise integrada dos dados de metilação e expressão gênica a partir dos dados internos e do TCGA e identificação de vias alteradas por metilação em câncer de reto 68
- Figura 7** *Heatmap* demonstrativo das sondas cujos valores de β apresentaram desvio padrão $> 0,02$ entre as 32 amostras de câncer de reto 72

Figura 8	Representação esquemática da distribuição genômica do contexto CpG das sondas diferencialmente metiladas encontradas na comparação entre o tecido normal (TN) e resposta ao tratamento (pCR: resposta completa ou pIR: incompleta)	73
Figura 9	Etapas utilizadas para a construção de um classificador preditivo de resposta à terapia neoadjuvante em câncer de reto	75
Figura 10	Análise não supervisionada de carcinomas de reto baseada nos valores de expressão de sondas com desvio padrão > 0,02.....	77
Figura 11	Vias desreguladas exclusivamente em um dos dois grupos de câncer de reto de acordo com a resposta à quimiorradioterapia neoadjuvante.....	79
Figura 12	Diagrama de Venn representando o número de proteínas potencialmente secretadas por tumores de reto.....	81
Figura 13	Diagrama de Venn representativo da comparação dos potenciais marcadores secretados por tumores de reto com resposta patológica completa (pCR) e incompleta (pIR) ao tratamento neoadjuvante.....	84
Figura 14	Descrição das sondas diferencialmente metiladas com correlação significativa com os dados de expressão gênica identificadas em tumores de reto de pacientes com resposta patológica completa	86-7

LISTA DE TABELAS

- Tabela 1** Características clínico-patológicas associadas à resposta patológica completa (pCR) em diferentes estudos..... 12
- Tabela 2** Estudos de rastreamento de marcadores moleculares por análise de expressão gênica global comparando a resposta ao tratamento quimiorradioterápico neoadjuvante em pacientes com câncer de reto..... 22
- Tabela 3** Características clínicas dos 34 pacientes com adenocarcinomas de reto incluídos no estudo 34
- Tabela 4** Vias canônicas alteradas em câncer de reto, identificadas usando o programa *Ingenuity Pathway Analysis* a partir de genes diferencialmente expressos identificados nos dados internos e do TCGA 63
- Tabela 5** Vias canônicas identificadas pelo programa *Ingenuity Pathway Analysis* a partir de genes com correlação significativa entre metilação e expressão nas análises dos dados deste estudo (Interno) e do TCGA..... 69
- Tabela 6** Proteínas potencialmente secretadas por tumores de reto de acordo com a resposta ao tratamento neoadjuvante 82
- Tabela 7** Dez genes com correlação negativa entre os resultados de metilação e expressão em tumores de reto de pacientes com resposta completa ou incompleta à terapia neoadjuvante 88

LISTA DE SIGLAS E ABREVIATURAS

aCGH*	<i>Array-based Comparative Genomic Hybridization</i>
ACMG*	<i>American College of Medical Genetics and Genomics</i>
aDGV*	<i>Affymetrix Database of Variants</i>
AGCC*	<i>Affymetrix® GeneChip® Command Console®</i>
ATM*	<i>Ataxia telangiectasia mutated</i>
BA	Borda anal
BD	Banco de dados
BER*	<i>Base excision repair</i>
BPT	Biópsias pré-tratamento
CA 19-9	Antígeno Carboidrato 19-9
CADD*	<i>Combined Annotation Dependent Depletion</i>
CaRe	Câncer de reto
CCR	Câncer colorretal
Cdks	Cinases dependentes de ciclinas
cDNA*	<i>Terminal Deoxynucleotidyl Transferase e dATP</i>
CEA	Níveis do antígeno carcinoembriogênico
CEMIP*	<i>Cell migration inducing protein</i>
ChAS*	<i>Chromosome Analysis Suite v3.1</i>
CIMP*	<i>Island methylator phenotype</i>
CIN*	<i>Chromosomal instability</i>
CMSs*	<i>Consensus molecular subtypes</i>
cN+	Linfonodos positivos ao diagnóstico
CNA*	<i>Copy number alteration</i>
CNAs	Regiões cromossômicas alteradas
cnLOH*	<i>Copy Neutral Loss of Heterozygosity</i>
CNV*	<i>Copy number variants</i>
cT4	Estadio tumoral clínico 4
cTNM	Classificação clínica
DC	Células dendríticas

DDR*	<i>DNA damage repair</i>
DGV*	<i>Database of Genomic Variants</i>
DLDA*	<i>Análise Linear Discriminante Diagonal</i>
DSB*	<i>Double strand break</i>
EMT*	<i>Epithelial to mesenchymal transition</i>
ETM	Excisão total do mesorreto
ExAC*	<i>The Exome Aggregation Consortium</i>
FAP	Polipose Adenomatosa Familiar
FFEP	Carcinomas de reto fixadas em formalina e em blocos de parafina
FISH*	<i>Fluorescent in Situ Hybridization</i>
G1/S*	<i>Cell Cycle: G1/S Checkpoint Regulation</i>
G2/M*	<i>Cell Cycle: G2/M DNA Damage Checkpoint Regulation</i>
GAD*	<i>Genetic Association Database</i>
GC	Genoma completo
GEMT	Grupo Europeu de Marcadores Tumorais
GII	Index de instabilidade genômica
GII	Index de instabilidade genômica
HIF*	<i>Hypoxia-inducible fator</i>
HNPCC	Câncer Colorretal Hereditário não Polipóide
HR	Recombinação homóloga
HR*	<i>Homologous recombination</i>
HRD*	<i>Homologous recombination deficiency</i>
IIQ	Intervalo interquartil
INCA	Instituto Nacional do Câncer
IPA*	<i>Ingenuity Pathway Analysis</i>
LARC*	<i>Locally advanced rectal câncer</i>
LND	Linfonodos
LOH*	<i>Loss of Heterozygosity</i>
LST*	<i>Large-Scale Transitions</i>
MET	Microcirurgia endoscópica transanal
MGMT	O6-metilguanina-DNA metiltransferase

MMR*	<i>Mismatch repair</i>
MS	Ministério da Saúde
MSI*	<i>Microsatellite instability</i>
N	Casos efetivamente analisados em cada plataforma
NC	Tecido normal congelado
NC	Não consta
nCRT	Radioquimioterapia neoadjuvante
NFP	Tecido normal fixado em formalina e em blocos de parafina
NHEJ*	<i>Non-homologous end joining</i>
NK*	Células <i>natural killer</i>
OMS	Organização Mundial de Saúde
PARP	Inibidores da poli-ADP ribose polimerase
pCR	Resposta patológica completa
pE	Valores de p empíricos
PELPK	Sequência de cinco aminoácidos
pIR	Resposta patológica incompleta
QRT	Químio e radioterapia
RT	Radioterapia
SIRT1*	<i>Sirtuin Signaling Pathway</i>
SOLID*	<i>Sequencing Platform - Life Technologies</i>
SP	Sangue Periférico
SST-RMA*	<i>Signal Space Transformation - Robust Multi-array Analysis</i>
STE*	<i>Total exome sequencing</i>
SVM*	<i>Support Vector Machine</i>
T1	Lesões em estágio inicial
tAI*	<i>Telomeric allelic imbalance</i>
TC	Tecido tumoral congelado
TCGA*	<i>The Cancer Genome Atlas</i>
TCLE	Termo de Consentimento Livre e Esclarecido
TFP	Tecido fixado em formalina e em blocos de parafina
TN	Tecido de reto normal

TNM	Tamanho do tumor, comprometimento de linfonodos e metástases
TRG	Grau de regressão tumoral
TRG*	<i>Tumor regression grade</i>
Var	Variante
VFCPT	Adenoma folicular e variante folicular de carcinoma papilar de tireoide
ypTNM	Classificação após tratamento multimodal por exame histopatológico

*siglas ou abreviaturas derivadas do inglês.

ÍNDICE

1	INTRODUÇÃO	1
1.1	Câncer de Cólon e Reto	2
1.2	Conduta de Tratamento em Pacientes com Adenocarcinoma de Reto	7
1.3	Predição de Resposta à Quimiorradioterapia Neoadjuvante	10
1.3.1	Características clínico-patológicas	10
1.3.2	Marcadores moleculares	13
1.3.3	Alterações genômicas	16
1.3.4	Análises de transcritos diferencialmente expressos	20
1.3.5	miRNAs e alterações epigenéticas	25
2	OBJETIVOS	30
2.1	Objetivo Geral.....	30
2.2	Objetivos Específicos	30
3	MATERIAL E MÉTODOS	31
3.1	Casuística.....	31
3.2	Metodologia	36
3.2.1	Fracionamento celular do material congelado	37
3.2.2	Purificação dos ácidos nucleicos	37
3.2.3	Investigação de Alterações Genômicas	37
3.2.3.1	Alterações no número de cópias genômicas.....	37
3.2.3.2	Sequenciamento de alto desempenho	38
3.2.4	Análise do perfil de metilação do DNA	39
3.2.4.1	Modificação do DNA por bissulfito de sódio	40
3.2.4.2	Ensaio de Metilação Global.....	41
3.2.5	Análise de expressão de transcritos	43
3.2.5.1	Ensaio de expressão global de transcritos.....	43
3.3	Metodologia de Análise dos Dados	46
3.3.1	Módulo 1 - Avaliação das alterações encontradas em câncer de reto: Comparação entre amostras tumorais e normais.....	46
3.3.1.1	Análise dos resultados dos <i>arrays</i> de metilação	46
3.3.1.2	Análise de expressão de transcritos	46

3.3.1.3	Comparação com dados externos e análise integrada - TCGA	47
3.3.2	Módulo 2 - Avaliação das alterações encontradas em câncer de reto de acordo com resposta ao tratamento neoadjuvante	48
3.3.2.1	Investigação de Alterações Genômicas	48
3.3.3	Dados do TCGA - The Cancer Genome Atlas	53
3.3.4	Análise do perfil de metilação do DNA	54
3.3.5	Análise de expressão de transcritos	55
3.3.5.1	Análise dos resultados da expressão global de transcritos.....	55
3.3.6	Análise Integrada.....	56
3.3.7	Análise estatística.....	57
4	RESULTADOS	58
4.1	Módulo 1 - Avaliação das alterações encontradas em câncer de reto: comparação entre amostras tumorais e normais	58
4.1.1	Análise dos dados de metilação	58
4.1.1.1	Perfil de metilação de câncer de reto (CaRe) comparado ao tecido normal (TN) de reto.....	58
4.1.2	Análise de expressão global de transcritos	62
4.1.2.1	Perfil de expressão de câncer de reto.....	62
4.1.3	Análise integrada de Metilação e Expressão Gênica	65
4.2	Módulo 2 - Avaliação das alterações encontradas em câncer de reto de acordo com resposta ao tratamento neoadjuvante	70
4.2.1	Análise de alterações genômicas	70
4.2.2	Análise dos dados de metilação	70
4.2.2.1	Análise do perfil de metilação de câncer de reto de acordo com a resposta ao tratamento neoadjuvante	70
4.2.3	Análise de expressão global de transcritos	76
4.2.3.1	Análise do perfil de expressão de transcritos em câncer de reto de acordo com a resposta ao tratamento neoadjuvante.....	76
4.2.3.2	Análise do secretoma de câncer de reto de acordo com a resposta ao tratamento neoadjuvante	80
4.2.3.3	Análise integrada de metilação e expressão em câncer de reto de acordo com a resposta ao tratamento neoadjuvante.....	85

5	DISCUSSÃO	89
5.1	Módulo 1 - Avaliação das alterações encontradas em câncer de reto: Comparação entre amostras tumorais e normais.....	89
5.2	Módulo 2 - Avaliação das alterações encontradas em câncer de reto de acordo com resposta ao tratamento neoadjuvante	99
6	CONCLUSÕES	124
7	REFERÊNCIAS BIBLIOGRÁFICAS	127

ANEXOS

- Anexo 1** Cartas de aprovação dos Comitês de Ética em Pesquisa- CEP (A.C.Camargo Cancer Center e Hospital de Câncer de Barretos)
- Anexo 2** Cobertura média, número total de *reads*, porcentagem de *reads* alinhados e a fração das sequências com cobertura de pelo menos 20, 30 e 100 vezes obtidas com o sequenciador NextSeq para as 31 amostras de adenocarcinomas de reto
- Anexo 3** Manuscrito I - submetido ao periódico International Journal of Cancer
- Anexo 4** Proteínas potencialmente secretadas em câncer de reto com e sem resposta patológica completa (pCR e pIR, respectivamente) identificadas por análise *in silico* usando as ferramentas de predição dos CBS servers (SecretomeP 2.0, SignalP 4.1 e TargetP 1.1)
- Anexo 5** Proteínas potencialmente secretadas por tumores de reto de pacientes com resposta patológica completa (17) e incompleta (56) que já foram descritas em vesículas extracelulares (VE) provenientes de amostras de carcinomas colorretais

1 INTRODUÇÃO

A Organização Mundial de Saúde (OMS) considera o câncer a segunda principal causa de morte no mundo, responsável por 8,8 milhões de óbitos em 2015 (<http://www.who.int/en/news-room/fact-sheets/detail/cancer>, acesso 6 de junho de 2018). A estimativa para 2025 é de mais de 20 milhões de novos casos de câncer, comparado aos 14,1 milhões de diagnósticos de câncer realizados em 2012 (FERLAY et al. 2015). Nos Estados Unidos, câncer é a segunda maior causa de mortes, sendo esperado para 2018 um total de 1.735.350 novos casos e 609.6400 óbitos em decorrência da doença (SIEGEL et al. 2018). Atualmente, 70% das mortes por câncer são registradas em países de baixa e média renda, onde se espera encontrar nas próximas décadas aproximadamente dois terços de casos novos diagnosticados (FORMAN e FERLAY 2014).

O Brasil está entre os países que mais serão afetados pelo câncer, de acordo com a OMS (FERLAY et al. 2015). Segundo o Instituto Nacional de Câncer-INCA o câncer é a segunda maior causa de mortes no país sendo estimados para 2018, 600 mil novos casos. Os tipos mais incidentes são o câncer de pele não melanoma (165,6 mil), próstata (68,2 mil), mama (59,7 mil), cólon e reto (36,4 mil), pulmão (31,3 mil) e estômago (21,3 mil) (Ministério da Saúde 2018). No intuito de diminuir a média de crescimento anual dessas estimativas é fundamental investir na prevenção, diagnóstico e tratamento desses pacientes. A precisão no diagnóstico e a adoção de um

tratamento mais adequado às características individuais de cada paciente permitem a redução do custo de terapias, muitas vezes desnecessárias, e do sofrimento causado pelos efeitos colaterais de tratamentos, muitas vezes não assertivos. A investigação de alterações moleculares globais câncer-específicas pode permitir a identificação de perfis associados com a resposta ao tratamento e assim auxiliar na identificação de pacientes sensíveis ou resistentes a determinadas terapias.

1.1 CÂNCER DE CÓLON E RETO

O câncer colorretal (CCR, C18 - C20) abrange tumores malignos que acometem o intestino grosso (o cólon) e o reto. No mundo, é o terceiro mais frequente em homens e o segundo em mulheres, excluindo-se os tumores de pele não melanomas (FERLAY et al. 2015). Essa proporção também é observada no Brasil, onde a estimativa de novos casos para 2018-2019 é de 36.360 casos por ano, sendo 17.380 homens e 18.980 mulheres (Ministério da Saúde 2018). A incidência do CCR vem aumentando no país, sendo 70% dos casos diagnosticados como esporádicos. Considerando que os cânceres não hereditários possuem grande contribuição de fatores ambientais, é provável que esse crescimento seja decorrente do aumento da expectativa de vida no país e da mudança de hábitos da população, principalmente no que diz respeito ao aumento do consumo de alimentos pobres em fibras e ricos em gordura saturada (HUXLEY et al. 2009; MENDONÇA et al. 2012).

A idade superior a 50 anos é um dos fatores de risco mais importantes no desenvolvimento do CCR. O sedentarismo, a obesidade e o consumo de álcool e/ou, tabaco também são fatores etiológicos descritos no risco de desenvolvimento do CCR (SLATTERY et al. 2010). Além disso, a história familiar de pólipos no cólon ou reto e mutações em genes associados com síndromes hereditárias características (Câncer Colorretal Hereditário não Polipóide - HNPCC, ou Síndrome de Lynch; e Polipose Adenomatosa Familiar - FAP) são outros fatores de risco (AMERSI et al. 2005; PROVENZALE et al. 2015).

Modelos de progressão tumoral de pólipos para carcinomas invasivos e metastáticos têm sido explorados em câncer colorretal. A progressão de adenoma para carcinoma parece resultar do acúmulo de alterações genéticas e epigenéticas, gerando instabilidade cromossômica e desequilíbrio de funções celulares importantes que acabam por favorecer a carcinogênese. VOLGESTEIN et al. (1988) estabeleceram um modelo linear de alterações genéticas que levam ao CCR. Esse modelo descreve mutações no gene *KRAS* e perda de heterozigose (LOH - *Loss of Heterozygosity*) no *locus* do gene *APC* como mudanças genéticas presentes em eventos iniciais; e LOH em 18q e em 17p (*TP53*) como eventos tardios no processo de formação do CCR (FEARON e VOLGESTEIN 1990). Entretanto, esses eventos isoladamente não conseguem explicar a complexidade e a heterogeneidade da progressão nos CCR. Embora diversas alterações genômicas e epigenômicas já tenham sido descritas nesses tumores, ainda não está bem estabelecida a correlação com a

predisposição, progressão ou resposta ao tratamento da doença (The Cancer Genome Atlas-TCGA 2012a; SHI et al. 2012; BAE et al. 2013; LIANG et al. 2013; LOCHHEAD et al. 2013; MARTINS et al. 2013; GANTT et al. 2014).

O perfil mutacional dos CCR introduzida pelo estudo do TCGA (*The Cancer Genome Atlas*) revelou que um subconjunto de casos apresenta alta taxa de mutações (hipermutados) (TCGA 2012a). Os casos hipermutados possuem uma frequência menor de mutações nos genes *APC* e *TP53* quando comparados aos casos não-hipermutados, além disso, apresentam mutações em *BRAF* e *TGFBR2* (alterações recorrentes no primeiro grupo) (TCGA 2012a). Os CCR foram também classificados em quatro subtipos moleculares (CMSs - *consensus molecular subtypes*) baseados no perfil de expressão gênica de 3.962 amostras. Foi observado que a maioria dos tumores hipermutados encontrava-se no subtipo CMS1, os quais também possuíam baixo número de alterações no número de cópias genômicas, alta instabilidade de microssatélites (MSI - *microsatellite instability*) e perfil metilador das ilhas CpG (CIMP - *CpG island methylator phenotype*). Esse grupo também estava enriquecido com tumores em cólon proximal (77% comparado à 3% em reto). O subtipo CMS2 foi caracterizado com altos níveis de CNA, perfil epitelial com ativação de WNT e MYC; enquanto o CMS3 possuía poucas CNAs e CIMP com perfil de alteração em genes envolvidos no metabolismo. O subtipo CMS4 apresentava perfil mesenquimal, ativação de TGF- β e alterações em genes relacionados a angiogênese. Amostras com características mistas ou indeterminadas (13%)

não foram classificadas como um grupo específico, mas foram relatadas como casos com fenótipo transitório ou com heterogeneidade intratumoral (GUINNEY et al. 2015).

A subclassificação dos tumores baseada no perfil molecular já é explorada há algum tempo e tem contribuído significativamente para o melhor direcionamento de terapias específicas e conseqüentemente, melhor resposta e maior sobrevida dos pacientes com câncer. O câncer de mama é um exemplo da aplicabilidade dessa subclassificação cujos subtipos moleculares tem auxiliado a estratificação de pacientes com diferentes prognósticos e a seleção de tratamentos mais efetivos (PEROU et al. 2000; TCGA 2012b; PRAT et al. 2015). O mesmo tem sido investigado para o CCR, entretanto, apesar do consenso molecular dos subtipos de CCR (GUINNEY et al. 2015), ainda persistem controvérsias sobre como explorar essa subclassificação (DUNNE et al. 2016). Apesar de considerar diversas características dos tumores colorretais, a classificação molecular se baseia em dados de expressão gênica (GUINNEY et al. 2015). Além desse tipo de análise subestimar os efeitos da heterogeneidade intratumoral, os dados clínico-patológicos, genômicos e epigenéticos não são utilizados para gerar o classificador. Indiscutivelmente, a subclassificação do CCR permitirá o uso da medicina personalizada, direcionando o tratamento mais adequado de acordo com o perfil de alterações de cada tumor.

Clinicamente, o CCR é tratado de acordo com a localização do tumor em relação à flexura esplênica. Quando o tumor está localizado no ceco, cólon ascendente ou até 2/3 do cólon transversal é denominado proximal ou

direito, e quando presente no terço distal do cólon transversal, cólon descendente, sigmoide ou reto recebe o nome de distal ou esquerdo (IACOPETTA 2002; LEE et al. 2017). Essas regiões possuem diferente desenvolvimento embrionário, histologia, composição enzimática e fecal (VAN GORKOM et al. 1997; WEI et al. 2004; LI e LAI 2009; THOMAS et al. 2016). Os fatores de risco atribuídos ao CCR também diferem em relação ao local de desenvolvimento do tumor (SHIN et al. 2011; BRÄNDSTEDT et al. 2012; LEE et al. 2013). A atividade física é associada à prevenção do câncer de cólon, enquanto não é observada relação estatística com o câncer retal (THUNE e LUND 1996; STEINDORF et al. 2005; FRIEDENREICH et al. 2006; LI et al. 2007). Além disso, o câncer de reto é mais frequentemente diagnosticado em homens (PURIM et al. 2013), em idades mais jovens e há poucos casos de CCR na família quando comparados aos pacientes com câncer de cólon (SVENSSON et al. 2002; WEI et al. 2004).

Os tumores de cólon e de reto também diferem quanto ao tipo de alterações moleculares (LEE et al. 2017). Os cânceres de cólon possuem maior probabilidade de apresentar MSI e CIMP (SLATTERY et al. 2010). Tumores de reto apresentam menor prevalência de alterações na via de reparo de bases mal pareadas (MMR - *mismatch repair*) (CHARARA et al. 2004; AGUIAR et al. 2014). A frequência de mutações em oncogenes e genes supressores tumorais também difere em relação à localização dos CCR. Enquanto *BRAF V600E* é muito mais comum em tumores de cólon (24,2% vs. 2.1% em reto), mutações em *APC* (63,6% vs. 81.9% em reto) e

TP53 (34,8% vs. 64,6% em reto) são mais frequentes em câncer de reto (LEE et al. 2017).

Os carcinomas de cólon e reto diferem ainda quanto ao tratamento, sendo a quimioterapia pós-cirúrgica indicada nos dois tipos de câncer, enquanto a radioterapia (RT) ou quimiorradioterapia neoadjuvante (nCRT) é recomendada apenas para pacientes com tumores de reto localmente avançados (LARC - *locally advanced rectal cancer*) (SCHMOLL et al. 2012; VAN DE VELDE et al. 2013).

1.2 CONDOTA DE TRATAMENTO EM PACIENTES COM ADENOCARCINOMA DE RETO

A escolha do tratamento nos carcinomas de reto (CaRe) depende principalmente do tamanho, localização e extensão do tumor. A avaliação da agressividade tumoral é dada pelo exame anatomopatológico (exame físico, diagnóstico por imagem e/ ou exploração cirúrgica), sendo classificada de acordo com o sistema TNM (EDGE e COMPTON 2010). Lesões em estágio inicial (T1) são passíveis de excisão local por microcirurgia endoscópica transanal (MET) e excisão total do mesorreto (ETM) em casos de maior acometimento da camada submucosa, assim como para lesões T2. Em estágios mais avançados e com maior probabilidade de recidiva, pode ser administrada terapia complementar neoadjuvante ou após o procedimento cirúrgico (adjuvante). A radioterapia (RT) pré-operatória pode ser administrada por um curto período seguida por cirurgia, ou por um longo

período e uma pausa até a cirurgia (SCHMOLL et al. 2012; BENSON III et al. 2012).

A resposta clínica à nCRT é avaliada pela presença de tumor residual detectado por técnicas como proctoscopia rígida, exame retal digital, biópsia de qualquer lesão ou calcificação residual, níveis do antígeno carcinoembriogênico (CEA) e exames radiológicos (HABR-GAMA et al. 2010). A avaliação clínica isolada não consegue prever a resposta completa com acurácia (MARETTO et al. 2007; ISSA et al. 2012; INTVEN et al. 2013). Dessa forma, é realizada também uma análise histopatológica do material obtido após a ressecção cirúrgica. Considera-se resposta patológica completa (pCR) quando há ausência de células tumorais viáveis na peça cirúrgica, linfonodos e outros órgãos (ypT0N0) e resposta incompleta (pIR) quando há resposta parcial (redução da proporção de células tumorais viáveis) ou ausência de reposta (quando não há redução do número de células tumorais (PARK et al. 2012). Diversos sistemas avaliam essa resposta de acordo com a taxa de regressão tumoral (TRG - *tumor regression grade*), categorizando os pacientes em excelentes até péssimos respondedores (TRAKARNSANGA et al. 2014; SANTOS et al. 2014).

Embora a cirurgia possibilite a eliminação de tumor residual, o procedimento está associado com alta morbidade, incluindo disfunção sexual, intestinal e urinária (LOOS et al. 2013; WEISER et al. 2014; HABR-GAMA et al. 2015; APPELT et al. 2015). Pacientes com câncer de reto que apresentam resposta patológica completa após nCRT têm menor risco de recorrência local e melhor sobrevida global comparado a pacientes com

tumor residual (CHEN et al. 2011a; GERARD et al. 2012; SADA et al. 2018). Uma iniciativa liderada pelo grupo da Profa. Angelita Habr-Gama (São Paulo) sugere que um grupo de pacientes com boa resposta clínica e potencial resposta patológica completa à nCRT poderiam ser poupados da cirurgia, sendo acompanhados num protocolo chamado “*Watch and Wait*” (HABR-GAMA et al. 2011). Estudos subsequentes mostraram que essa estratégia permite a preservação do órgão com baixas taxas de recidivas (MAAS et al. 2011). Em um estudo avaliando 82 pacientes de dois grupos pareados (41 ETM e 41 sem cirurgia), HUPKENS et al. (2017) reportaram melhor qualidade de vida no grupo não operado. Essa estratégia resultou na criação de um banco de dados internacional visando coletar informações de diversos centros que corroborem para a compreensão dos fatores relacionados ao tratamento do CaRe (<http://www.iwwd.org/>). Até o momento, no entanto, não é possível detectar resposta patológica completa sem o tratamento cirúrgico, mas essa possibilidade vem sendo investigada por diversos grupos que vem buscando biomarcadores preditivos (WEISER et al. 2014; HABR-GAMA et al. 2015; APPELT et al. 2015; AL-SUKHNI et al. 2016; DAYDE et al. 2017; KUNDEL et al. 2018).

1.3 PREDIÇÃO DE RESPOSTA À QUIMIORRADIOTERAPIA NEOADJUVANTE

1.3.1 Características clínico-patológicas

Diversos estudos avaliaram características clínico-patológicas pré-tratamento (DE FELICE et al. 2016; JOYE et al. 2016; PROBST et al. 2016) a fim de prever a resposta à neoadjuvância (Tabela 1). A maioria dos estudos em literatura não identificou diferenças quanto as características como gênero, idade e etnia entre os pacientes com pCR e pIR (DAS et al. 2007; HUH et al. 2013; RESTIVO et al. 2013; WALLIN et al. 2013; GARLAND et al. 2014; BITTERMAN et al. 2015; JOYE et al. 2016). Entretanto, PROBST et al. (2016) relataram que mulheres possuem maiores chances de apresentarem pCR. Em adição, estadios clínicos T e N (cT e cN) baixos foram associados à pCR (KALADY et al. 2009; HUH et al. 2013; JOYE et al. 2016). Dois estudos demonstraram que pacientes com estadios cT4 apresentam menor probabilidade de atingir pCR, entretanto, a análise de regressão multivariada não foi significativa (KALADY et al. 2009; HUH et al. 2013).

O comprometimento de linfonodos ao diagnóstico também tem se mostrado preditivo de pIR (HUH et al. 2013; JOYE et al. 2016). GARLAND et al. (2014) em 297 casos de CaRe submetidos à nCRT relataram associação significativa entre cN0 e pCR (análise de regressão logística; OR = 4,384 IC 95% 1,011-19,017). Uma associação semelhante foi descrita por BITTERMAN et al. (2015) na análise de 138 casos, embora seis pacientes

tivessem resposta clínica completa, não tendo realizado cirurgia. PROBST et al. (2016) utilizaram o banco de dados nacional americano (*The National Cancer Data Base* descrito por BILLIMORIA et al. 2008) composto por 18.113 pacientes com CaRe diagnosticados em estadio II ou III e tratados com nCRT. Nessa população, não foi encontrada diferença na proporção de cN0 entre os pCR e os pIR (PROBST et al. 2016).

A localização do tumor também foi relacionada com a resposta à nCRT. Entretanto, enquanto RESTIVO et al. (2013) relataram pacientes com tumores proximais (> 5cm da borda anal) como os melhores respondedores; BITTERMAN et al. (2015) descreveram que pacientes com tumores distais poderiam ser preditivos de pCR (\geq 3cm da borda anal) (dados confirmados pela análise multivariada). Outros estudos não mostraram associação entre a localização do tumor e a resposta à nCRT (KALADY et al. 2009; QIU et al. 2011; HUH et al. 2013; GARLAND et al. 2014; KLEIMAN et al. 2015; DE FELICE et al. 2016; JOYE et al. 2016).

Em geral, é possível observar uma grande controvérsia nos achados descritos em relação à resposta à terapia neoadjuvante e as características clínicas e histopatológicas de pacientes com CaRe (Tabela 1). As características avaliadas não foram capazes de predizer a resposta ao tratamento, ressaltando a necessidade de estudos que identifiquem marcadores moleculares que possam ser úteis para essa finalidade.

Tabela 1 - Características clínico-patológicas associadas à resposta patológica completa (pCR) em diferentes estudos.

Referência /Ano	N° casos (%pCR)	Idade ao diagnóstico		cT4 [§] (pCR/pIR)	cN+ [§] (pCR/pIR)	Distância da BA (pCR/pIR)	CEA pré		CEA limite (%pCR/%pIR)
		pCR	pIR				pCR	pIR	
Das et al./ 2007	562 (19,2)		57,9 ^b	NC	NC	NC	NC	NC	≤ 2,5ng/ml (24% pCR vs. 11% >2,5)
Kalady et al./ 2009	242 (24,0)	55,5 ^b	58,5 ^b	1,7/ 5,4	32,8/ 35,8	5 (1,5-12)/ 6 ^b	2,1^b	3^b	≥2,5 ng/ml (41,4/49,5)
Qiu et al./ 2011	96 (14,6)	2 (≥65) 12 (< 65)	34 (≥65) 48 (< 65)	14,3/ 20,7	21,4/ 29,3	≥7 cm (28,6/ 9,8)	Elevado (28,6%) [¥]	Elevado (42,7%) [¥]	NC
Huh et al./ 2013	391 (14,6)	63 ^b	64 ^b	3,5/ 8,1	63,2/ 75,5	5 / 5 ^b	4,7^b		< 5,0ng/ml (71,9/ 47)
Restivo et al./ 2013	260 (16,5)	63 ^b	63 ^b	0,0/ 5,5	44,2/ 45,2	>5cm (23,2/ 86,2)	2,2^b	3^b	> 5ng/dl (4,7/ 32,3)
Wallin et al./ 2013	469 (20,5)	58 ^b	59 ^b	6,3/ 7,0	58,3/ 59,8	NC	3,4^b	9,6^b	> 5ng/ml (11,0/ 21,0)
Garland et al./ 2014	297 (11,5)	65 ^b	62 ^b	8,8/ 15,6	61,8/ 77,2	5,5 / 5,5 ^b	2,6^b	5,0^b	Pré>pós % 34,9 / 64,3^b
Bitterman et al./ 2015	138 (26,1)	59,3 ^a	54,7 ^a	0/7,8	52,8/ 72,5	NC	3,1^a	24,3^a	≥ 5ug/l (16,70/38,20)
Kleiman et al./ 2015	141 (13,5)	63,5 ^a	63,4 ^a	5,3/ 12,3	68,4/ 58,2	6 / 6,3 ^a	2,75 ^b	4,5 ^b	<3,0ug/l (52,6/34,4)
de Felice et al./2016	99 (23,2)		63,8 ^b	13,1/ 11,9	78,3/ 81,6	≤5 (60,9/55,3)	NC	NC	NC
Joye et al./2016	620 (19,4)	63,9 ^{b*}	64 ^{b*}	10/ 8,6	89,2/ 79,2	6,5 / 6,3 ^a	4,7 ± 7,7	6,9 ± 15,0	NC
Probst et al./ 2016	18 .113 (13,2)	60,5 ^a	59,9 ^a	NC	NC	NC	Elevado (31,2%)	Elevado (48,0%)	Avaliado limite de cada instituição
Gash et al./ 2017	13.742 (32,4)	60,4 ^a	59,8 ^a	5,32/8,6	43,9/ 51,7	NC	Elevado (34,79%)	Elevado (46,9)	NC

pCR: Resposta patológica completa; pIR: Resposta patológica incompleta; cT4: estadiotumoral clínico 4; cN+: linfonodos positivos ao diagnóstico; BA: Borda anal; CEA: Antígeno carcinoembrionário; [§](%pCR/%pIR); ^amédia; ^bmediana; [¥]CEA e CA19.9 juntos; NC - não consta; valores em negrito: **p<0,05**; valores em negrito e itálico: *valores não significativos (p>0,05) na análise multivariada.*

1.3.2 Marcadores moleculares

As técnicas moleculares têm se mostrado capazes de identificar alterações genéticas específicas em tumores de reto (RAHMA et al. 2013; GHADIMI et al. 2005). Como elas apresentam maior resolução que a análise histopatológica convencional, sua utilização pode facilitar a identificação de marcadores tumorais em pequenos fragmentos de tecido, como as biópsias. Estes marcadores podem auxiliar tanto na detecção de doença residual mínima como na identificação de tumores sensíveis ao tratamento pré-operatório de quimiorradioterapia e com maiores chances de apresentarem resposta completa (CHEN et al. 2011a; GERARD et al. 2012; WEISER et al. 2014; HABR-GAMA et al. 2015).

Marcadores moleculares preditivos e prognósticos são alvos de estudos realizados em CaRe (KUREMSKY et al. 2009; YOSHIDA et al. 2011; GRADE et al. 2012; MAMMANO et al. 2012). Entretanto, ainda não são utilizados na prática clínica por apresentarem resultados contraditórios e por não terem sido validados em grandes estudos prospectivos (MOLINARI et al. 2015).

O antígeno carcinoembriônico (CEA) é, até o momento, único marcador recomendado pelo o Grupo Europeu de Marcadores Tumorais (GEMT) para acompanhamento de pacientes com CCR; entretanto, não há ainda um marcador específico para CaRe (DUFFY et al. 2014). O CEA é composto por um conjunto de glicoproteínas relacionadas à adesão celular normalmente expressas no intestino e em concentrações muito baixas no sangue de indivíduos saudáveis. Altas concentrações séricas estão

normalmente associadas à presença de diversos tipos tumorais, principalmente CCR (HAMMARSTRÖM 1999). A maioria dos estudos demonstra que pacientes com altas concentrações de CEA pré nCRT apresenta grandes chances do paciente não atingir resposta completa ao tratamento (DAS et al. 2007; HUH et al. 2013; RESTIVO et al. 2013; WALLIN et al. 2013; GARLAND et al. 2014; BITTERMAN et al. 2015; JOYE et al. 2016; PROBST et al. 2016).

Os níveis séricos de CEA podem também ser afetados por outras condições benignas, como doença mamária benigna, polipose retal, pancreatite, obstrução biliar, úlcera péptica, doença inflamatória intestinal, cirrose, enfisema pulmonar em tabagistas (SAJID et al. 2007). WALLIN et al. (2013) demonstraram que a associação entre baixos níveis de CEA e pCR não se mantém em tabagistas. Outros fatores que podem afetar a quantidade do CEA na circulação são mutações no domínio PELPK (sequência de cinco aminoácidos), pois alteram a capacidade de sua eliminação pelo fígado (ZIMMER e THOMAS 2001). Ainda que muitos estudos apontem níveis basais de CEA como potencial marcador preditivo de resposta à nCRT, nenhum dos fatores mencionados acima foi avaliado em detalhes em uma grande série de casos. Os dados descritos em literatura apontam para a necessidade de estudos mais elaborados para investigar o papel do CEA na predição à neoadjuvância em CaRe.

A expressão de proteínas envolvidas no reparo do DNA, proliferação, angiogênese e apoptose como TP53 (CHEN et al. 2012), EGFR (GIRALT et al. 2005; KIM et al. 2006; DVORAK et al. 2012), TYMS (OKONKWO et al.

2001; AGUIAR et al. 2005; JAKOB et al. 2008; NEGRI et al. 2008; HUANG et al. 2013), Ki-67 (SAIGUSA et al. 2012; JAKOB et al. 2008; KIM et al. 2006), p21 (REERINK et al. 2004; RAU et al. 2003; BERTOLINI et al. 2007), BAX/BCL2 (KUDRIMOTI et al. 2007; CHANG et al. 2005; SAIGUSA et al. 2012) tem sido extensivamente avaliada, não resultando, porém, em dados conclusivos (CHANG et al. 2005; TERRAZZINO et al. 2006; BERTOLINI et al. 2007).

A grande maioria dos estudos que avaliou marcadores moleculares em CaRe utilizou a análise de expressão de proteínas por imunohistoquímica. A expressão das proteínas PDCD4, EZH2, ASNS, CPS1, HER-2, MRP3 em biópsias pré-nCRT foram individualmente analisadas (DOU et al. 2014; MENG et al. 2014a e b; LIN et al. 2014; LEE et al. 2014; YU et al. 2014). A elevada expressão da proteína PDCD4, assim como a baixa expressão de EZH2 foram associadas à boa resposta à nCRT. No entanto, no primeiro estudo apenas quatro casos apresentaram pCR e em ambos os estudos o tratamento quimioterápico pré-operatório não foi o mesmo para todos os pacientes, o que pode representar um viés na conclusão dos dados (DOU et al. 2014; MENG et al. 2014a). A associação entre a perda da expressão de ASNS e do aumento de CPS1 com pior resposta à terapia neoadjuvante foi relatada em dois estudos feitos com o mesmo grupo de pacientes asiáticos (LIN et al. 2014; LEE et al. 2014). Entretanto, as análises realizadas pelos autores não levaram em consideração pCR, e sim níveis de regressão tumoral, questionando as implicações futuras desses resultados (LIN et al. 2014; LEE et al. 2014). YU

et al. (2014) avaliaram a expressão de MRP3 em biópsias pré-tratamento, no entanto, não foi encontrada associação significativa. Além disso, nenhuma associação entre resposta à neoadjuvância e expressão de HER-2 em CaRe foi encontrada por MENG et al. (2014b). A proteína envolvida no complexo de Golgi, GOLPH3, foi avaliada por ZHU et al. (2016) em 148 biópsias de LARC. Os autores encontraram alta expressão da proteína em 62% de pacientes resistentes à nCRT (TRG0, 1 e 2 do sistema Dworak - 1997) e em 41% dos bons respondedores (TRG3 e 4). Um estudo recente descreveu a expressão de EGFR, BCL-2, p53, Ki67 e VEGFR em 47 biópsias de LARC pré nCRT. Foi encontrada maior porcentagem de células expressando EGFR em casos com pCR comparados aqueles com resposta parcial (KUNDEL et al. 2018)

1.3.3 Alterações genômicas

Ainda que estudos associem mutações no gene *KRAS*, *BRAF* e *PIK3CA* com resposta à radioterapia ou quimiorradioterapia (LUNA-PEREZ et al. 2000; GARCIA-AGUILAR et al. 2011; LOCHHEAD et al. 2013), outros grupos mostram que essas mutações não possuem relação com a resposta do paciente ao tratamento (ZAUBER et al. 2009; ERBEN et al. 2011; KIM et al. 2013). Essas alterações afetam as vias de sinalização PI3K/AKT ou RAS-RAF-MAPK, sendo um dos mecanismos mais comuns na carcinogênese colorretal (ZEESTRATEN et al. 2012; DERBEL et al. 2013).

A análise global de alterações somáticas no número de cópias genômicas (CNA - *copy number alteration*, que diverge de variações

germinativas encontradas entre as populações - CNV, *copy number variants*) é outra estratégia adotada que tem potencial para identificar marcadores moleculares úteis na predição de resposta à nCRT. Um reduzido número de estudos utilizou análises de CNAs por aCGH (array-based *Comparative Genomic Hybridization*) para avaliar resposta em CaRe. MOLINARI et al. (2011) identificaram alterações em regiões cromossômicas específicas (2q21, 3q29, 7p22-21, 7q21, 7q36, 8q23-24, 10p14-13, 13q12, 13q31-34, 16p13, 17p13-12 e 18q23), as quais foram capazes de diferenciar tumores de pacientes respondedores ao tratamento nCRT dos não-respondedores. Em outro estudo, CHEN et al. (2011b) utilizaram a plataforma Human Genomic CGH 244A (Agilent) para avaliar CNAs em 95 CaRe tratados com nCRT. Comparando o padrão de alterações em pacientes respondedores e não respondedores ao tratamento, os autores associaram perdas em 15q11.1-q26.3, 11q24.3-q25 e 8p12 com ausência de resposta e perda em 12p13.31 com resposta completa ao tratamento (CHEN et al. 2011b). Entretanto, as regiões descritas não foram relatadas em outros estudos. GONZALEZ-GONZALEZ et al. (2014) avaliaram a relação entre o número de cópias genômicas em biópsias tumorais de reto pré e pós nCRT e resposta ao tratamento com o objetivo de avaliar a regressão tumoral. Os autores utilizaram a técnica de FISH (*Fluorescent in Situ Hybridization*) para 51 sondas distribuídas pelo genoma e detectaram a del(17p) em biópsias de pacientes pobremente respondedores e del(1p) em bons respondedores (GONZALEZ-GONZALEZ et al. 2014).

A presença de aneuploidias (ganho ou perda de um ou mais cromossomos, mas não todos do lote) é uma característica comum em tumores de origem epitelial, muitas vezes causada por defeitos de segregação que levam à instabilidade cromossômica (CIN - *Chromosomal instability*) (THOMPSON e COMPTON 2008). CIN é resultado, entre outros, de altos níveis de erros durante a segregação cromossômica, instabilidade telomérica e defeitos nos mecanismos de reparo a danos no DNA (PINO et al. 2010). Durante esse processo, a deleção de genes supressores de tumor e a amplificação de oncogenes favorecem a proliferação indiscriminada de células malignas (ZACK et al. 2013). Durante a replicação, o atraso na formação do fuso mitótico devido à aneuploidia, leva a um estresse de replicação, observado em CCR com elevada CIN. Tamanho estresse, desencadeia quebras no DNA gerando cromossomos acêntricos ou rearranjados e, conseqüente, erros durante a divisão celular (TANAKA e HIROTA 2016). A via de reparo por recombinação homóloga (HR - *homologous recombination*) é crucial na correção de quebras de dupla-fita de DNA geradas nesse processo, defeito nesse mecanismo de reparo contribui para o aumento da CIN (LIU et al. 2016) e vem sendo associado à maior sensibilidade à quimioterapia (HELLEDAY 2010).

Em câncer de ovário e mama, pacientes com mutações em genes da via de HR apresentam melhor resposta à quimioterapia baseada em platina e melhor prognóstico (VENCKEN et al. 2011; MANIÉ et al. 2016; XU et al. 2017). Além disso, esses casos apresentam melhor resposta quando tratados com inibidores da poli-ADP ribose polimerase (PARP), membro da

via de reparo à danos no DNA por excisão de bases (BER - *Base excision repair*), devido à letalidade sintética alcançada (HELLEDAY 2010). O comprometimento do reparo por HR pode ser decorrente de diversos mecanismos direta ou indiretamente associados com genes da via; como mutações, hipermetilação ou alteração da regulação por RNAs não codificadores, o que torna inviável a investigação destes diversos fatores na prática clínica. Esses tumores, entretanto, apresentam elevada CIN e os padrões de alterações resultantes podem ser utilizados como cicatrizes genômicas possibilitando a identificação de deficiência na via de HR, em vez de focar em um gene específico.

Análise de alterações no número de cópias genômicas e perda de heterozigose foram utilizadas para o desenvolvimento de três escores associados à deficiência de HR (HRD - *homologous recombination deficiency*). Utilizando um limiar de 10 segmentos com LOH maiores que 15Mb, mas menores que um cromossomo inteiro, ABKEVICH et al. (2012) conseguiram detectar pacientes com mutação em *BRCA1* e *BRCA2* com uma sensibilidade de 90%. Esse escore (HRD-LOH) também foi capaz de identificar pacientes com câncer de mama com maior resposta à quimioterapia baseada em platina (TELLI et al. 2015). O escore de desequilíbrio alélico telomérico (tAI - *telomeric allelic imbalance*) avalia o número de regiões subteloméricas com desequilíbrio alélico (CNAs ou cnLOH - *copy neutral LOH*) que tem início além do centrômero estendendo-se até o telômero, sendo positivamente correlacionado com sensibilidade à platina (BIRKBAK et al. 2012). O terceiro escore, LST - *Large scale transition*

- considera o número de pontos de quebra que ocorre entre regiões de pelo menos 10Mb, sendo o limiar 15 e 20 LSTs para tumores “quase-diplóides” e “quase-tetraplóides” respectivamente (POPOVA et al. 2012).

A associação das cicatrizes genômicas em 15 tipos tumorais usando dados do TCGA foram avaliados por MARQUARD et al. (2015). Os autores compararam a presença de altos níveis dessas assinaturas com estadio e grau tumoral, tabagismo, perda de p53, CIN, índice de integridade genômica, frequência de LOH e número total de mutações por amostra. Os autores verificaram que tumores de cólon possuem escores altos para as diferentes assinaturas e sugerem que uma subpopulação desses tumores poderia responder bem ao tratamento com drogas relacionadas à platina; entretanto, os tumores de reto não foram incluídos nesse estudo.

1.3.4 Análises de transcritos diferencialmente expressos

A expressão de transcritos utilizando biópsias pré-operatórias e de sangue periférico de pacientes com tumores de reto também vem sendo utilizada na identificação de diferenças capazes de identificar casos que apresentam resposta ao tratamento neoadjuvante com CRT (Tabela 2). A maioria dos estudos utilizou biópsias obtidas previamente ao tratamento com nCRT e diferentes plataformas de análise de expressão gênica para identificar perfis de expressão que pudessem prever resposta em grupos independentes de pacientes submetidos a nCRT (revisado por AKIYOSHI et al. 2012). PALMA et al. (2013) utilizaram microarranjos de expressão (CodeLink bioarrays - Applied Microarrays, Tempe, AZ, USA) em linfócitos

de sangue periférico de 27 pacientes com LARC, detectando 8 genes diferencialmente expressos entre os dois grupos de pacientes, entre eles, o *FALZ*, associado com imunidade antitumoral. LOPES-RAMOS et al. (2015) realizaram sequenciamento de alto desempenho (SOLiD *Sequencing Platform* - Life Technologies) em 25 biópsias de reto pré-nCRT divididas em grupo treinamento (14 casos) e validação (11 casos) com o objetivo de identificar marcadores de pCR. Foram identificados 27 genes diferencialmente expressos entre os quatro casos com pCR e os 10 casos com pIR no grupo de treinamento. A partir dessa lista de genes, os autores desenvolveram algoritmos classificadores (SVM - *Support Vector Machine*) utilizando subgrupos de 2 a 5 genes que pudessem ser utilizados como preditivos de resposta ao tratamento. Os autores identificaram 35 subgrupos de genes capazes de prever resposta com uma acurácia de 81,8%. No entanto, para cada grupo de genes, quando observada alta sensibilidade, a especificidade era muito baixa e vice-versa (LOPES-RAMOS et al. 2015).

Tabela 2 - Estudos de rastreamento de marcadores moleculares por análise de expressão gênica global comparando a resposta ao tratamento quimiorradioterápico neoadjuvante em pacientes com câncer de reto.

Referência /Ano	Nº de casos	Plataforma	Nº de genes*	Principais vias ou genes alterados	Validação
Ghadimi. et al. /2005	30	microarranjo de cDNA (9,984 sondas)	54	<i>SMC1</i>	microarranjo de oligonucleotídeos (22,231 sondas) - 7 amostras independentes e 10 repetições.
Watanabe et al. /2006	35	Affymetrix U95Av2	33	<i>TDGFG3, RAC2, LUM, LGALS1, CYPD, GPX2, TYRO3, THBS2, PSPHL</i>	Mesma plataforma em 17 amostras teste
Kim et al. /2007	31	Affymetrix U133A	261		Mesma plataforma em 15 amostras teste
Rimkus et al. /2008	43	Affymetrix U133 Plus 2.0	42	<i>ETS2, Caspase1, SLC35E1</i>	Mesma plataforma em 35 dos 43 casos
Brettingham-Moore et al. /2011	51	Affymetrix U133 Plus 2.0	NC	Genes da via do TNF	Validação de classificadores previamente descritos em 3 trabalhos** e análise funcional por IPA.
Nishioka et al. /2011	17	Microarranjo customizado	17	<i>MMP, NFκB, TGFB1, TOP1, ITGB1</i>	Imuno-histoquímica em 17 amostras
Chan et al. /2013	31	Almac Diagnostics Colorectal Cancer DSA®	3	<i>EHBP1, STAT1 e GAPDH</i>	Análise funcional pelo IPA
Supiot et al. /2013	6	54K HG-U133A Plus 2.0 Affymetrix	37	<i>ADAMDEC1, AKAP7, CAPN5, CLIC5, CPE, CREB3L1, NEDD4L, RAB27A, ATP2A3, CCL28, F2RL2, SLC6A, ISX, PABPC1L, TXNIP, F2RL2, SPON1, S100A2, PXDN, TNFRSF11A, ITM2C, PVT1, PDCD4</i>	NC
Gantt et al. /2014	23	Illumina platform Human-6 v2	812	<i>APOA2, AHSG, DBH, APOA1, APOB, APOC3, LMX1A, SOAT2, SLC7A9, TF, LOC729399, SERINC5, SCNN1B, ZC3H6, SLC4A4, DTWD2, MS4A12, BEX5, MMRN1 e CLCA4</i>	Mesma plataforma em 10 amostras

Cont/ Tabela 2

Referência /Ano	Nº de casos	Amostra	Plataforma	Nº de genes*	Principais genes alterados	Validação (nº de amostras)
Agostini et al. /2015a	131	BPT	Affymetrix U133A/ Plus 2.0	277	<i>AKR1C3, CXCL11, CXCL10, CXCL9, IDO1, MMP12, HLA-DRA</i>	NC
Agostini et al. /2015b	42	BPT	Affymetrix U133A Plus 2.0	19	<i>AGRN, HFM1, CSTF3, RAB6A, PRKRIR, C12orf32, XRCC3, CDK10, CDK5R1, IL12RB1, BCKDHA, ZNF160, ASXL2, EIF3L, PSMD6, MAGI1, RAB7A, SPRY4, CNKSR2</i>	NC
Millino et al. /2017	38	BPT	Whole Human Genome Oligo microarray 4X44K (V1);	256	<i>TMEM188, MYO1B, ITGA2, GTSE1, NRG1, KLF7, TRAM1, BCL2L13</i>	Mesma plataforma em 21 amostras

*Genes diferencialmente expressos entre pacientes respondedores e não respondedores ao tratamento. **GHADIMI. et al. (2005), KIM et al. (2007), RIMKUS et al. (2008). BPT: Biópsias pré-tratamento, SP: Sangue Periférico, IPA: *Ingenuity Pathway Analysis*. NC: não consta

Utilizando a plataforma Illumina Human-6 v2, GANTT et al. (2014) identificaram um painel de 812 genes capaz de classificar pacientes não respondedores à nCRT com uma sensibilidade e especificidade de 100% ($p < 0,05$). A análise foi feita em biópsias pré-tratamento de 23 pacientes com câncer de reto em estágio II ou III encaminhados para nCRT e validadas em um grupo de 10 casos nas mesmas condições. Ainda assim, o reduzido número de casos avaliados e a grande quantidade de genes envolvidos não permitiram elucidar os mecanismos de resistência ao tratamento neoadjuvante. AGOSTINI et al. (2015a), reanalisaram os dados de três estudos realizados com a mesma plataforma (HG-U133, Affymetrix, Santa Clara, CA, EUA) (KIM et al. 2007, RIMKUS et al. 2008; AGOSTINI et al. 2015b); surpreendentemente, os resultados encontrados não coincidiram. Os principais achados da meta-análise *de novo* foram genes relacionados com o sistema imune, especialmente *CXCL10*, *IDO1* e *AKR1C3* (AGOSTINI et al. 2015a).

Uma outra estratégia potencialmente útil para identificar marcadores preditivos é a análise do secretoma, definido como o conjunto de macromoléculas liberadas pelas células na matriz extracelular. As proteínas podem ser secretadas no meio extracelular por meio da via clássica, quando apresentam um peptídeo sinal; ou da via não-clássica com o auxílio de endossomos, exossomos ou passando diretamente através da membrana plasmática (LIN et al. 2013). As proteínas secretadas pelas células tumorais tem maior probabilidade de serem detectadas no sangue, sendo potenciais biomarcadores de fácil aplicação na prática clínica. A proteína BST2, selecionada a partir de uma análise do secretoma de linhagens celulares tumorais, foi identificada em plasma de pacientes com CCR, conseguindo

discriminar pacientes de controles saudáveis (CHIANG et al. 2015). Em adição, é possível a predição *in silico* do secretoma tumoral (CACCIA et al. 2013) a partir de dados de transcritos diferencialmente expressos.

1.3.5 Alterações epigenéticas

Além das alterações genômicas e de transcritos, modificações nos mecanismos responsáveis pela modulação da expressão gênica, tem revelado potenciais marcadores de resposta ao tratamento (MOLINARI et al. 2013). A metilação do DNA é um mecanismo essencial no controle de expressão gênica, regulando a acessibilidade à cromatina e a transcrição. Existem cerca de 28 milhões de sítios CpG no genoma humano, os quais, quando agrupados, são chamados de ilhas CpG. Esses conjuntos, encontram-se normalmente na região promotora de genes e estão, de modo geral, não metilados em células normais, permitindo a expressão de genes necessários ao funcionamento celular. A metilação de ilhas CpG em regiões promotoras está, portanto, associada à repressão da expressão gênica. Regiões com concentração intermediária de CpGs são encontradas, comumente em corpo de genes, cuja metilação foi relacionada à ativação da expressão (STIRZAKER et al. 2014). Foram descritas regiões com baixa concentração de CpGs, chamadas de *shores* ou *shelf* (2kb ou 2kb – 4kb das ilhas, respectivamente) (IRIZARRY et al. 2009) e sítios de CpG isolados espalhados pelo genoma e denominados *Open Sea* (SANDOVAL et al. 2009). Alterações de metilação em regiões intergênicas tem se mostrado importantes mecanismos envolvidos no câncer, uma vez que essas regiões podem englobar elementos regulatórios distantes como *enhancers*, *super-enhancers*, regiões promotoras alternativas, sequências

codificantes de RNAs regulatórios como miRNAs e *long non-coding* RNAs (JONES 2012, MURTHA e ESTELLER 2016).

Células tumorais, frequentemente apresentam hipometilação global, que leva a uma instabilidade genômica e, hipermetilação de ilhas CpG específicas em regiões promotoras, silenciando diversos genes, incluindo supressores tumorais (ISSA 2004). Estudos mostram que assinaturas de metiloma são capazes de subclassificar tumores (HERNANDEZ-VARGAS et al. 2010), serem utilizados como marcadores prognósticos (CARVALHO et al. 2012) e de identificar marcadores de resistência inerente ou adquirida à quimioterapia (BORLEY e BROWN 2015). A metilação diferencial do DNA em sítios específicos também contribuiu para o entendimento da resposta à radioterapia com ou sem quimioterapia, em diversos tipos tumorais, incluindo glioblastomas, câncer de mama, gástrico e CCR (LV et al. 2016).

As alterações epigenéticas mais bem descritas relacionadas à resposta clínica ao tratamento é a metilação de O6-metilguanina-DNA metiltransferase (*MGMT*) (HEGI et al. 2005). Em LARC, a hipermetilação do promotor de *MGMT* foi associada com melhor resposta ao tratamento por SUN et al. (2014). Além de *MGMT*, GAEDCKE et al. (2014) investigaram 10 regiões diferencialmente metiladas capazes de prever a sobrevida livre de doença (DFS) e resposta à terapia em 71 pacientes com LARC. A alteração epigenética mais estudada é o perfil de metilação das ilhas CpG de regiões promotoras de genes supressores tumorais e de reparo a danos do DNA, denominado fenótipo metilador de ilhas CpG (*CIMP - CpG island methylator phenotype*) (OGINO et al. 2006; GOEL e SHIN 2008; KURKJIAN et al. 2008; KIM et al. 2010; LEONG et al. 2011).

Evidências experimentais dão suporte ao papel do CIMP como um mecanismo de inativação gênica em vários grupos de CCR (SAMOWITZ et al. 2005; OGINO et al. 2006). Tal fenótipo é correlacionado significativamente com o silenciamento de genes como *CDKN2A* (*p16*), *MGMT* e *MLH1* (ISSA 2004). A partir desses estudos, painéis de genes com CIMP foram sugeridos com o intuito de identificar subtipos tumorais. Uma forte correlação encontrada entre a hipermetilação e mutações no gene *BRAF* e a hipometilação e mutações em *KRAS* sugeriram a associação entre o epigenótipo de metilação e mutação em oncogenes (HUGHES et al. 2012). Poucos grupos, no entanto, avaliaram o papel preditivo de resposta ao nCRT do perfil de metilação de pacientes com CaRe (DE MAAT et al. 2008 e 2010; LEONG et al. 2011; JO et al. 2012).

MOLINARI et al. (2013) avaliaram o padrão de metilação de tecido retal normal e tumoral, identificando diferença nos genes *APC*, *CDKN2A*, *CHFR*, *RARB* e *RASSF1bis*. Além disso, o grupo relatou a associação entre a metilação do gene *TIMP3* e a resposta à nCRT (MOLINARI et al. 2013). O padrão de metilação global do CaRe é alterado após a nCRT, como demonstrado por TSANG et al. (2014) ao compararem amostras pré e pós nCRT de 46 LARC (pIR) por imuno-histoquímica, revelando o aumento e diminuição de metilação em 28% e 72% dos casos, respectivamente. A avaliação das biópsias pré nCRT, incluindo 7 casos com pCR, revelou uma correlação significativa com TRG, com os menores níveis de metilação em pCR (TSANG et al. 2014).

HA et al. (2015) avaliaram o perfil de metilação de 45 tumores de reto localmente avançados e validaram os principais achados em 67 casos (24 do grupo anterior e 43 novos) por pirosequenciamento. Entre seis genes (*DZIP1*,

ZEB1, *DKK3*, *STL*, *KLHL34* e *ARHGAP6*) selecionados e avaliados em 67 amostras, somente a metilação de *KLHL34* cg14281612 mostrou forte associação com resposta a nCRT. Dessa forma, os autores identificaram uma associação entre a hipermetilação no gene *KLHL34* com boa resposta ao tratamento, sugerindo que esse seja um marcador de radiosensibilidade nesses tumores. Esses resultados, entretanto, necessitam de validação em casuísticas maiores (HA et al. 2015).

Estudos que consigam identificar marcadores moleculares aplicáveis à prática clínica enfrentam dificuldades logísticas que impedem a aplicabilidade dos resultados. Para avaliação de marcadores preditivos de resposta recomenda-se a utilização de biópsias pré- e pós-tratamento. Entretanto, essa estratégia frequentemente não é possível, pois biópsias iniciais comumente possuem um tamanho muito pequeno ou não estão disponíveis para pesquisa. Além disso, não há tecido residual disponível em pacientes com resposta completa e a terapia limita a utilização da amostra nos casos de resposta parcial devido aos danos causados ao tecido, impossibilitando o seu uso. Desse modo, na maioria dos estudos há limitação do número de técnicas a serem utilizadas e redução do poder estatístico de análise dos resultados (MOLINARI et al. 2013). Infelizmente, em muitos estudos não há validação dos achados em um grupo independente de amostras e que sejam selecionadas com os mesmos critérios de inclusão e exclusão adotados quando da sua elaboração.

A resposta à quimiorradioterapia pré-operatória dos cânceres de reto varia substancialmente. Apesar da boa resposta ao tratamento neoadjuvante estar associada com resultado favorável, até 10% dos pacientes apresentam

recorrência e uma média de 20,3% dos casos, metástases à distância (ZORCOLO et al. 2012). A habilidade em prever resposta ao tratamento é imprescindível para adotar estratégias de tratamento personalizadas. Marcadores moleculares tem potencial para estratificar os pacientes no reconhecimento daqueles sensíveis ou resistentes à terapia neoadjuvante. Além disso, o melhor entendimento da biologia tumoral auxilia o desenvolvimento de novas estratégias terapêuticas visando menores taxas de recidiva com melhor qualidade de vida e sobrevida global para esses pacientes.

2 OBJETIVOS

2.1 OBJETIVO GERAL

Identificar alterações genéticas e epigenéticas que corroborem para a elucidação dos mecanismos de resistência ao tratamento radio e quimioterápico neoadjuvante em pacientes com adenocarcinoma de reto.

2.2 OBJETIVOS ESPECÍFICOS

- Identificar alterações genômicas, transcritos diferencialmente expressos e perfis de metilação em tumores de reto;
- Integrar os dados de expressão de transcritos com o perfil de metilação e com o de alterações genômicas para identificar candidatos a *drivers* moleculares;
- Analisar a função dos genes identificados como alterados, avaliando sua participação nas vias biológicas alteradas em câncer de reto;
- Usar as mesmas estratégias para avaliar resposta completa e resistência ao tratamento neoadjuvante nos pacientes com câncer de reto;
- Identificar moléculas alteradas capazes de classificar os pacientes de acordo com a resposta à terapia;
- Determinar transcritos que codificam proteínas potencialmente secretadas por adenocarcinomas de reto.

3 MATERIAL E MÉTODOS

3.1 CASUÍSTICA

Foram incluídos nesse estudo, 34 amostras de adenocarcinomas de reto de pacientes submetidos a tratamento neoadjuvante com quimiorradioterapia e 10 tecidos normais de reto (necropsias). Os pacientes ou seus responsáveis legais assinaram o Termo de Consentimento Livre e Esclarecido (TCLE) para doação das amostras para a pesquisa. Este projeto foi aprovado pelo Comitê de Ética em Pesquisa do A.C.Camargo Cancer Center (n° 1884/14) e do Hospital de Câncer de Barretos (HCB, n° 1030/2015) (Anexo 1).

As biópsias foram obtidas observando-se a determinação exata da distância entre o limite distal do tumor e a linha pectínea. Para o estadiamento foi solicitado ressonância magnética da pelve, tomografia computadorizada do abdome superior, radiografia simples do tórax e dosagem sérica de antígeno carcinoembrionário (CEA).

Após o diagnóstico e estadiamento, todos os pacientes foram submetidos ao tratamento neoadjuvante. A radioterapia foi realizada até dose total de 5040 cGy, fracionamento diário de 180 cGy, em campos, cinco vezes por semana, sendo 4500cGy na pelve, incluindo o sítio primário e drenagem linfática, seguido por boost de 540cGy em três frações. Concomitante à radioterapia, foi administrado 5- fluorouracil na dose de 1g/m² ao dia durante cinco dias, em infusão contínua de 120 horas, na primeira e na quinta semanas de radioterapia. A cirurgia foi realizada obedecendo aos princípios clássicos do

tratamento oncológico. Todos os dados referentes ao estadiamento inicial, tratamento, resposta patológica após cirurgia assim como outras características clínicas e patológicas dos pacientes incluídos no estudo estão descritas na Tabela 3. Todas as amostras foram reavaliadas histopatologicamente por um único patologista (Dra. Maria Dirlei Begnami, A.C.Camargo Cancer Center).

A resposta ao tratamento é classificada com base na análise anatomopatológica da peça cirúrgica e linfonodos removidos à cirurgia. Foram considerados pacientes com resposta patológica completa (pCR) aqueles com ausência de células neoplásicas viáveis na peça operatória, linfonodos ou em órgãos distantes (ypT0N0M0). A resposta patológica incompleta (pIR) refere-se àqueles que não preencherem esses critérios.

Os pacientes também foram sub-classificados de acordo com o grau de regressão tumoral (TRG), baseado em uma versão modificada de Dworak e Mandard (SANTOS et al. 2014). Os casos foram divididos em quatro categorias de acordo com a porcentagem de células neoplásicas viáveis: Grau 0: regressão completa do tumor (0% de tumor residual); Grau I: regressão subtotal do tumor (< 25% de células tumorais viáveis); Grau II: regressão parcial do tumor (25-50% de células tumorais viáveis); Grau III: regressão mínima do tumor (> 50% de células tumorais viáveis). Foram classificados como Grau II os casos com regressão completa do tumor, mas linfonodos positivos (ypT0N+) (SANTOS et al. 2014).

Foram coletados 34 casos, dos quais, 11 (32%) atingiram pCR (que corresponde a TRG0 na classificação de regressão tumoral) e 23 casos (68%) tiveram pIR dos quais 10 (43,5%), 10 (43,5%) e três (9%) foram categorizados como Grau I, II e III, respectivamente. Dois casos apresentaram tumores ou

lesões benignas prévios ao diagnóstico de CaRe (RET5: leiomioma uterino; RET43: nódulos benignos de tireoide e mama). O caso RET2 apresentou adenocarcinoma de reto e de colon sincrônicos, sendo incluído apenas nas análises genômicas.

Vinte e um casos reportaram ter história de câncer na família, dos quais, oito preenchem os critérios clínicos para Síndrome de Lynch de acordo com a classificação de Bethesda e um caso, de Amsterdam. Entre os oito casos com critérios de Bethesda, seis apresentavam expressão das enzimas de reparo de malpareamento de bases (MMR) avaliada por imunohistoquímica, dois casos não possuíam resultados para esse teste. O caso classificado com critérios de Amsterdam também não foi avaliado para as enzimas MMR. O sequenciamento do tumor desses três casos, no entanto, não revelou mutações nos genes das enzimas de MMR.

Tabela 3 - Características clínicas dos 34 pacientes com adenocarcinomas de reto incluídos no estudo.

Variáveis	Todos	Resposta Completa (pCR)	Resposta Incompleta (pIR)	p
Número de casos	34	11	23	
Gênero Masculino	20	7	13	0,73
Gênero Feminino	14	4	10	
IMC (Kg/m²)				
Mediana (variação)	26,1 (18-36,2)	25,7 (23,9-36,2)	26,6 (18-31,9)	0,57
Idade ao diagnóstico (anos)				
Mediana (variação)	56 (26-80)	61 (46-79)	52 (26-80)	0,10
História Familiar de Câncer (%)	21 (62)	7 (63,6)	14 (60,9)	1,00
Relacionados à Síndromes Hereditárias	9	1	8	0,16
Hábito Tabagista	13	4	9	1,00
Etilista	6	3	3	0,36
Outros Tumores Primários ou Lesões Benignas Prévios *#	3	1	2	1,00
Distância média da Linha Pectínea (cm)				
Mediana (variação)	4,0 (0-8)	4,5 (0-8)	4,0 (0-8)	0,66
Presença de Pólipos	16	6	10	0,73
CEA pré-tratamento				
Mediana (variação, ng/uL)	4,0 (0,3-568,3)	3,0 (1,2-17,4)	4,6 (0,3-568,3)	0,08
CA19-9 pré-tratamento				
Mediana (variação, U/mL)	9,9 (0,6-56)	14,0 (7-30,9)	7,6 (0,6-56)	0,42
Grau de diferenciação				
Bem diferenciado	8	1	7	
Moderadamente diferenciado	25	9	16	0,38
NA	1	1	0	
TNM - cT				
T2	6	3	3	
T3	26	7	19	0,48
T4	2	1	1	

Cont/ Tabela 3

Variáveis	Todos	Resposta Completa (pCR)	Resposta Incompleta (pIR)	p
TNM - cN				
N0	11	4	7	
N+	23	7	16	0,72
TNM - ypT				
T0	12	11	1	
T1	1	0	1	
T2	12	0	12	
T3	7	0	7	
T4	2	0	2	
TNM - ypN				
N0	28	11	17	
N+	6	0	6	
Tempo entre o fim da QRT e a cirurgia (dias)				
Mediana (variação)	65,5 (39-250)	63,0 (43-124)	67,0 (39-250)	0,44
Nº de LND dissecados				
Mediana (variação)	12 (0-50)	10 (0-29)	13 (4-50)	0,47
Nº de casos com LND Positivos	6	0	6	
Seguimento (meses)				
Mediana (variação)	52 (13-93)	47 (13-93)	53 (17-88)	0,79
Metástases a Distância				
Fígado	2	0	2	
Fígado e Pulmão	1	0	1	
Pélvica e Adrenal	1	0	1	
Pulmão	1	0	1	

IIQ: Intervalo interquartil, **CEA:** Antígeno Carcino-embrionário, **CA 19-9:** Antígeno Carboidrato 19-9, **TNM:** Classificação Clínica dos Tumores Malignos, onde T= extensão do tumor primário, N= metástase em linfonodo regional e M= metástase à distância. **cTNM:** classificação clínica, **ypTNM:** classificação após tratamento inicial multimodal avaliado por exame histopatológico, **QRT:** quimio e radioterapia, **LND:** linfonodos. * Leiomioma uterino (RET5) e nódulos benignos de tireoide e mama (RET43). # Tumor de reto e colon sincrônicos (RET2).

3.2 METODOLOGIA

O fluxograma abaixo (Figura 1) esquematiza as etapas realizadas neste estudo.

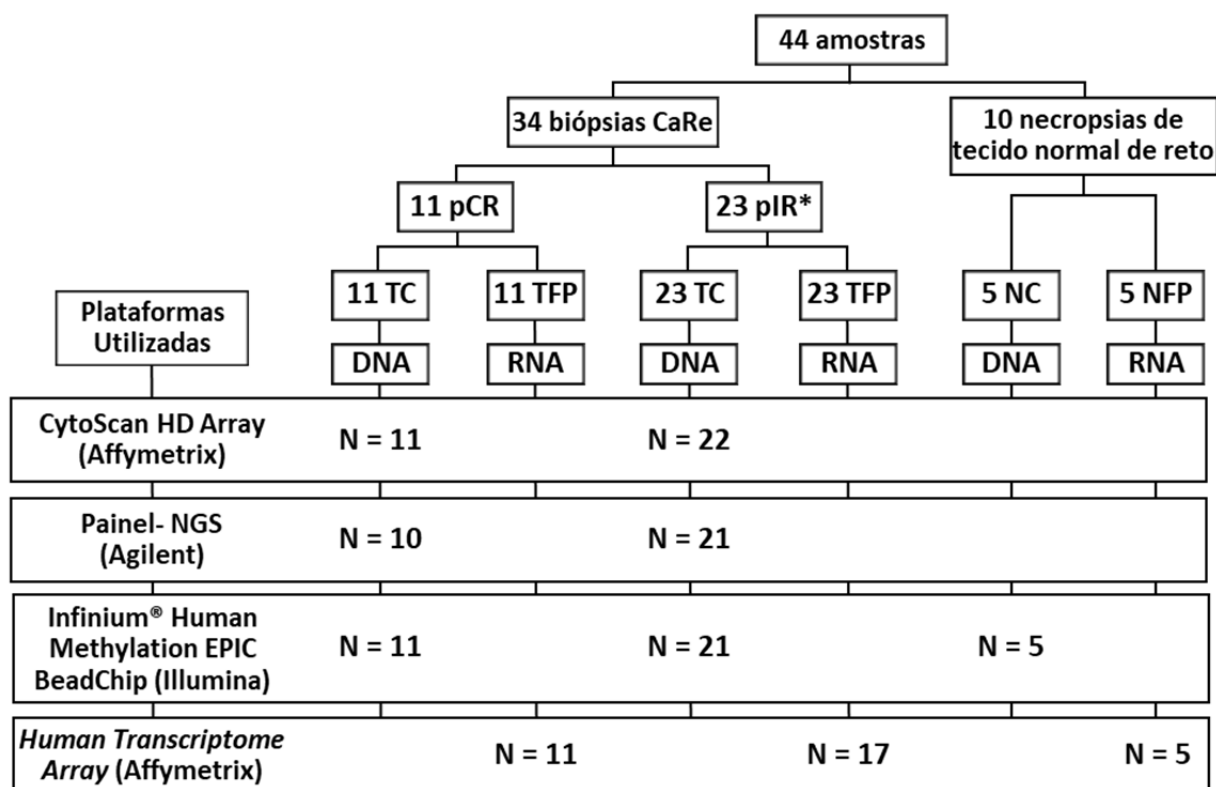


Figura 1 - Fluxograma da metodologia experimental utilizada neste estudo. CaRe: Câncer de reto, pCR: resposta patológica completa, pIR: *Resposta parcial ou ausência de resposta ao tratamento. TC: tecido tumoral congelado, TFP: tecido fixado em formalina e em blocos de parafina, NC: tecido normal congelado, NFP: tecido normal fixado em formalina e em blocos de parafina. N= representa os casos efetivamente analisados em cada plataforma nas análises de número de cópias genômicas (CytoScan HD), sequenciamento (Painel), metilação global (Infinium-EPIC) e expressão global (Human Transcriptome Array). NGS: Sequenciamento de alto desempenho.

3.2.1 Fracionamento celular do material congelado

Para a realização dos experimentos, as amostras de tecido congelado foram submetidas à separação dos componentes celulares - sendo o núcleo e o citoplasma utilizados neste estudo e as demais frações, em projetos paralelos desenvolvidos por outro grupo de pesquisadores da instituição. As biópsias foram submetidas à centrifugação com solução tampão (5 mL/g tecido, 0,25 M sacarose, 10 mM Tris-HCl, e 1 mM MgCl₂) e inibidor de protease na proporção 100:1 (Protease Inhibitor Cocktail Set V, EDTA-Free, Calbiochem®), como descrito por Han et al. (2011).

3.2.2 Purificação dos ácidos nucleicos

O DNA foi extraído das frações celulares isoladas das amostras de tecido congelado, como descrito acima, e purificado usando o *kit DNeasy Tissue and Blood* (Qiagen, Alemanha), de acordo com as recomendações do fabricante. O RNA foi extraído de blocos de parafina usando o protocolo do *kit RecoverAll™ Total Nucleic Acid Isolation Kit for FFPE* (ThermoFisher, EUA), de acordo com as recomendações do fabricante.

3.2.3 Investigação de Alterações Genômicas

3.2.3.1 Alterações no número de cópias genômicas

As alterações genômicas foram identificadas usando-se a plataforma CytoScan HD Array (ThermoFisher), que utiliza microarranjos de alta densidade com mais de 2,6 milhões de sondas polimórficas e não-polimórficas compostas por 25 pares de bases. As 750 mil sondas polimórficas ou de SNPs (para determinação alélica) e as 1,9 milhões de sondas não polimórficas (CNVs) permitem detectar

ganhos e perdas genômicas a partir de 25-50 kb, assim como regiões com mosaicismos e de cnLOH (*Copy Neutral Loss Of Heterozygosity*).

Todos os ensaios foram realizados de acordo com o protocolo padrão do fabricante. Resumidamente, 250 ng do gDNA foi digerido com a enzima de restrição *NspI*, ligado a adaptadores específicos, amplificado pela PCR com a enzima Titanium Taq DNA polimerase (Clontech, EUA) e visualizado em gel de agarose 2% (fragmentos entre 150 e 2000 pb). O material amplificado foi purificado utilizando-se *beads* magnéticas e quantificado em NanoDrop® ND1000 (NanoDrop Technologies, EUA). Posteriormente, o DNA purificado foi fragmentado com a enzima DNaseI, visualizado em gel de agarose 4% (fragmentos entre 25 e 125 pb) e marcado com biotina. O DNA marcado foi hibridado por 16-18h a 50°C no GeneChip® Hybridization Oven 645 (Affymetrix/ThermoFisher, EUA), lavado e corado no GeneChip® Fluidics Station 450 (Affymetrix/ThermoFisher, EUA) e escaneado no GeneChip® Scanner 3000 7G (Affymetrix/ThermoFisher, EUA).

3.2.3.2 Sequenciamento de alto desempenho

A integridade das amostras de DNA foi avaliada utilizando o *Agilent Genomic DNA ScreenTape* (Agilent Technologies, EUA) e o DNA de dupla fita foi quantificado usando o Qubit dsDNA BR Assay kit (Life Technologies, EUA). Uma amostra não passou nos critérios de qualidade (RET22).

Um painel customizado com 105 genes (exons e regiões flanqueadoras: 3'UTR e 5'UTR - SureSelect^{XT} Custom Panel, Agilent) relacionados ao câncer, incluindo 13 genes da via de reparo por recombinação homóloga (HR) e cinco da via de reparo a erros de pareamento (MMR) foi utilizado na construção das bibliotecas

para sequenciamento de alto desempenho de 32 amostras de câncer de reto (Anexo 3).

As bibliotecas foram construídas utilizando o kit SureSelect^{QXT} Library Prep (Agilent) de acordo com as instruções do fabricante e sequenciadas no equipamento NextSeq 550 (Illumina, EUA). Resumidamente, 50ng de DNA genômico de cada amostra foi submetido à fragmentação enzimática e ligação de adaptadores que permitem posterior amplificação pela PCR. A amplificação dos fragmentos foi realizada com a enzima *Herculase II Fusion DNA polymerase* (Agilent) e o produto foi purificado com *beads* magnéticas (Agencourt AMPure^{XP} PCR purification bead system - Beckman Coulter, EUA). As bibliotecas foram então hibridadas às sondas responsáveis por capturar as regiões de interesse. As sequências hibridadas foram capturadas por *beads* magnéticas revestidas de estreptavidina (Dynabeads MyOne Streptavidin T1 - ThermoFisher) e amplificadas com concomitante indexação dos iniciadores específicos (SureSelect^{QXT} P7 e P5 dual indexing primers, Agilent). O produto enriquecido foi purificado com *beads* magnéticas (AMPure^{XP} - Beckman Coulter, EUA). A verificação da qualidade das bibliotecas de DNA indexadas foi realizada utilizando a High Sensitivity D1000 ScreenTape (Agilent). O conjunto de bibliotecas de 31 amostras com qualidade satisfatória foi reunido em um pool de 10 nM, desnaturado em NaOH (0,2N) e diluído para 1,4 pM. As amostras foram então sequenciadas (2x75 pares de base em corrida *pair-end* no sequenciador NextSeq, Illumina) usando o kit NextSeq 500/550 mid-output (150 ciclos) (Illumina).

3.2.4 Análise do perfil de metilação do DNA

Amostras de DNA de 33 das 34 biópsias congeladas de adenocarcinoma de reto (RET21 não possuía DNA disponível para o experimento) e cinco amostras

histopatologicamente normais de reto (obtidas de necropsias) foram utilizadas na análise global de metilação. O caso RET2 foi excluído das análises subsequentes devido à presença de tumor maligno sincrônico ao diagnóstico de tumor de reto. Desta forma, foram incluídos 32 tumores e 5 amostras de tecido de reto normal (TN) para a análise final do perfil global de metilação.

3.2.4.1 Modificação do DNA por bissulfito de sódio

O tratamento do DNA com bissulfito de sódio converte citosinas não metiladas em uracilas por um processo de deaminação, mas que não afeta as 5-metilcitosinas que ficam protegidas da ação do sal pelo radical metil associado ao carbono 5 da base nitrogenada. As amostras de DNA foram submetidas ao protocolo de modificação por bissulfito utilizando o kit EZ DNA Methylation™ (Zymo Research, EUA), de acordo com as instruções do fabricante e com as modificações sugeridas pela Illumina.

Primeiramente, o DNA foi desnaturado pela adição de 100 µL da solução *CT Conversion* em 50 µL de DNA (500ng), resultando em 150µL de produto de reação. A conversão foi realizada pela incubação das amostras em um termociclador nas seguintes condições: 16 ciclos de 95°C por 30 segundos e 50°C por 1 hora, seguido de 1 ciclo de 4°C por 10 minutos. O produto da reação foi transferido para a coluna Zymo-Spin™ IC, na qual foram adicionados 400µL do tampão M-Binding. Em seguida, foram realizadas lavagens subsequentes com 100µL do tampão M-Wash e 200µL do tampão M-Desulphonation, sendo repetidas mais duas etapas de lavagem com 200µL do tampão M-Wash. Por fim, o DNA convertido foi eluído em 12µL do tampão M-Elution.

3.2.4.2 Ensaio de Metilação Global

Os experimentos de metilação global (Infinium Human MethylationEPIC BeadChip array, Illumina) foram realizados de acordo com as recomendações do fabricante. Esta plataforma permite a análise de mais de 850.000 sítios CpG por amostra, incluindo regiões promotoras de miRNAs, *enhancers* e sítios de ligação de fatores de transcrição.

As amostras convertidas foram amplificadas a partir da adição de 4µL do DNA modificado, 20µL do reagente *Multi-Sample Amplification 1 Mix* (MA1) e 4µL de NaOH 0,1N. Posteriormente, foram adicionados 68µL do reagente *Random Primer Mix* (RPM) e 75µL da solução *Multi-Sample Amplification Master Mix* (MSM) e incubadas a 37°C por 21 horas no forno de hibridação (Illumina). O DNA amplificado foi fragmentado com 50µL da enzima *Fragmentation solution* (FMS) a 37°C por 1 hora, seguido da precipitação das amostras, realizada a partir da adição de 100µL de *Precipitation solution* (PM1) e 300µL de isopropanol 100%. A solução foi mantida a 4°C por 30 minutos e precipitada por centrifugação de 3000g a 4°C, por 20 minutos. Após a precipitação, o sobrenadante foi removido por inversão da placa, a qual foi mantida em temperatura ambiente por 1 hora para secagem do *pellet*.

O DNA amplificado e fragmentado foi ressuspendido em 46µL de *Resuspension, hybridization, and wash solution* (RA1) por 1 hora a 48°C. Em seguida, as amostras foram homogeneizadas e desnaturadas a 95°C por 20 minutos. Cada BeadChip, com capacidade para oito amostras, foi colocado no HybChamberInsert (Illumina) e 26µL de cada amostra ressuspendida foi aplicada na parte apropriada do BeadChip, os quais foram então inseridos cuidadosamente na câmara de hibridação (Illumina) contendo 400mL de *Humidifying buffer* (PB2). As amostras foram hibridadas aos BeadChips a 48°C por 21 horas. As amostras de

tecido tumoral e normal foram colocadas no mesmo chip para diminuir viés de lote ou processamento.

Após hibridação, os BeadChips foram mantidos em temperatura ambiente por 30 minutos para resfriamento prévio à lavagem. O processo de lavagem foi feito em duas câmaras especiais contendo 200mL de PB1 durante 1 minuto cada. Posteriormente, cada BeadChip foi integrado a *Flow-Through Chamber* para a etapa de extensão das bases. As câmaras contendo os BeadChips foram colocadas em uma raque na posição vertical a 44°C, onde uma série de reagentes foram dispensados no reservatório apropriado na seguinte ordem: 150µL da solução RA1 por 30 segundos, 5 vezes; 450µL da solução *XStain Bead Chip solution 1* (XC1) por 10 minutos; 450µL da solução *XStain Bead Chip solution 2* (XC2) por 10 minutos; 200µL da solução *Two-Color Extension Master Mix* (TEM) por 15 minutos; 450µL de formamida 95%/1mM EDTA por 1 minuto (repetido uma vez), seguida de incubação final de 5 minutos.

Ao final da extensão, a temperatura da raque vertical foi reprogramada para 32°C e ao atingir a temperatura desejada, uma nova série de reagentes foram adicionados e todos os passos repetidos 2 vezes: 450µL de *XStain Bead Chip solution 3* (XC3) por 1 minuto (repetida mais uma vez); 250µL da solução *Superior Two-Color Master Mix* (STM) por 10 minutos; 450µL de XC3 por 1 minuto (repetida mais uma vez), seguido de incubação por mais 5 minutos; 250µL do anticorpo *Anti-StainTwo-Color Master Mix* (ATM) por 10 minutos. Após a segunda incubação com ATM, foram adicionados ainda: 450µL de XC3 por 1 minuto (repetido mais uma vez) seguido de incubação por mais 5 minutos; 250µL de STM por 10 minutos; e por fim, 450µL de XC3 por 1 minuto (repetida mais uma vez), seguido de incubação por mais 5 minutos.

A última etapa consistiu na lavagem dos BeadChips em 310mL da solução PB1 e em 310mL da solução *XStain Bead Chip solution 4 (XC4)*. Posteriormente, os BeadChips foram colocados num suporte para secagem em um dessecador a vácuo por 55 minutos e então escaneados usando HiScan Systems (Illumina, EUA).

3.2.5 Análise de expressão de transcritos

Foi extraído o RNA de 34 biópsias de carcinomas de reto fixadas em formalina e em blocos de parafina (FFEP) e de cinco amostras histopatologicamente normais de reto (FFEP). Estes casos eram dos mesmos pacientes incluídos na análise genômica, perfil mutacional e de metilação global. O caso RET2 foi excluído das análises subsequentes devido à presença de tumor maligno sincrônico ao diagnóstico de tumor de reto e três amostras (RET3, RET15 e RET26) não possuíam RNA de qualidade para a realização dos experimentos. Desta forma, foram utilizadas amostras de 30 tumores e cinco tecidos normais de reto (TN) para a investigação do perfil de expressão global.

3.2.5.1 Ensaio de expressão global de transcritos

Para identificar os perfis de expressão gênica dos tecidos normais e tumorais, foi utilizada a plataforma de alta resolução GeneChip™ Human Transcriptome Array 2.0 (HTA, ThermoFisher). A plataforma possui sondas distribuídas por toda a extensão do gene, incluindo sondas específicas cobrindo regiões de *splicing*, proporcionando assim uma imagem mais completa e exata da expressão gênica global com a capacidade adicional para análise de transcrição de todas as isoformas conhecidas produzidas por um gene. Como esta análise foi realizada em amostras FFPE, foi utilizado o kit SensationPlus™ FFPE Amplification and WT

Labeling Kit (ThermoFisher) para síntese, amplificação e marcação de cDNA. Esse kit foi especialmente desenvolvido para superar os problemas de degradação e fragmentação desse tipo de amostra, sem requerer alta concentração de RNA inicial.

As etapas de processamento das amostras de RNA foram realizadas de acordo com as recomendações do fabricante. Um conjunto de controles de RNA poli-A, fornecido pelo fabricante, projetado como controles positivos para monitorar todo o processo de marcação das amostras, foi adicionado às amostras antes do início do processamento. Resumidamente, 100ng de RNA total foram utilizados para a síntese de cDNA. Após uma incubação inicial de 10 minutos a 80°C com os iniciadores RT, as amostras foram submetidas à ação da transcriptase reversa por 1h a 42°C e 2 minutos a 25°C. O produto foi purificado usando *beads* magnéticas em lavagens seriadas com etanol 70% e ressuspendido em 14µL de água livre de nucleases a 65°C. Em seguida, foi sintetizada a região promotora no cDNA. Inicialmente, uma cauda poli-A foi incorporada à extremidade 3' do cDNA (*Terminal Deoxynucleotidyl Transferase* e dATP), onde ocorreu a hibridação de um oligo T7 dT, gerando uma região promotora de dupla fita no cDNA. Esse promotor permitiu a transcrição *in vitro* do cDNA pela T7 RNA polimerase (17h a 37°C) em RNA *sense*. A purificação do RNA *sense* foi realizada com *beads* magnéticas em lavagens seriadas com etanol 70% e ressuspendido em 23µL de água livre de nucleases a 65°C. Para realizar a marcação das amostras, 10 - 25µg de RNA *sense* foram convertidos em cDNA de dupla fita (ds-cDNA) utilizando uma primeira etapa com transcriptase reversa (42°C por 2h e 25°C por 2 minutos) e uma segunda etapa com a enzima *Klenow* (37°C por 40 minutos, 75°C por 10 minutos e 4°C por 2 minutos). Após incubar as amostras com 2µL de *WT Stop Solution* (65°C por 30 minutos e

25°C por 2 minutos) e 8µL de *WT Neutralization Mix* (temperatura ambiente), o ds-cDNA foi purificado usando *beads* magnéticas em lavagens seriadas com etanol 70% e ressuspendido em 25µL de água livre de nucleases a 65°C. Um total de 6µg de ds-cDNA foi utilizado para fragmentação e marcação das regiões terminais com biotina, feitas com incubação das amostras combinadas ao *Fragment and label Buffer and Enzyme master mix* a 37°C por 60 minutos, 93°C por 2 minutos e 4°C por 2 minutos. Às amostras marcadas foi adicionado o mix de hibridação composto de *Control oligo B2* (3nM), *20X Hybridization controls - bioB, bioC, bioD e cre*, *2X hybridization mix*, DMSO e água ultrapura estéril. Esta solução foi incubada a 99°C por 5 minutos e 47°C por pelo menos 5 minutos até colocar nos chips. A solução foi inserida nos GeneChip® Probe Array (*49-format*) e hibridada em forno por um período de 16 horas à 47°C e 60rpm. A revelação e a lavagem dos GeneChips foram realizadas de forma automatizada na GeneChip Fluidics Station 450 (Affymetrix/ ThermoFisher, EUA), segundo o protocolo Expression Wash, Stain and Scan User Manual for Cartridge Arrays (P/N 702731 Rev. 4) e o escaneamento foi realizado no Affymetrix GeneChip Scanner 7000 (Affymetrix/ ThermoFisher, EUA).

3.3 METODOLOGIA DE ANÁLISE DOS DADOS

3.3.1 Módulo 1 - Avaliação das alterações encontradas em câncer de reto: Comparação entre amostras tumorais e normais

3.3.1.1 Análise dos resultados dos arrays de metilação

Os dados brutos gerados (formato .IDAT) foram submetidos ao controle de qualidade usando o software Genome Studio v. 2.0.2 (Illumina) e normalizados usando a ferramenta de estatística R v.3.3.2 (<http://www.r-project.org>). Subsequentemente, foram removidas as sondas mapeadas nos cromossomos sexuais, sondas cross-reativas e aquelas com valor de p de detecção média baixo, como descrito por PIDSLEY et al. (2016). As sondas mapeadas em regiões de SNPs conhecidos foram removidas com o pacote *minfi* do Bioconductor (ARYEE et al. 2014). As sondas diferencialmente metiladas entre os tumores e o TN foram identificadas usando o pacote *limma* (RITCHIE et al. 2015). O valor de p ajustado (FDR) < 5% e $|\Delta\beta| > 0,15$ foram utilizados para selecionar sondas significativas, as quais foram anotadas usando o MethylationEPIC_v-1-0_B2 (Illumina).

3.3.1.2 Análise de expressão de transcritos

✓ Análise dos resultados do array de expressão de transcritos

Os arquivos CEL foram gerados e analisados pelos programas Affymetrix® GeneChip® Command Console® (AGCC) 4.0 e Expression Console v.1.4.1, respectivamente, usando o algoritmo SST-RMA (*Signal Space Transformation - Robust Multi-array Analysis*). O método RMA executa a correção do ruído de fundo,

a normalização dos dados e a sumarização das sondas para converter os níveis de intensidade das mesmas em valores de expressão (IRIZARRY et al. 2003). O método SST-RMA inclui a redução de ruído de fundo GC4 (correção que leva em conta os níveis de GC, versão 4) e normalização de intensidade SST, que ajusta os níveis de intensidade das sondas.

Dois casos apresentaram qualidade insatisfatória e foram excluídos das análises (RET18 e RET22), totalizando 33 casos avaliados (28 amostras tumorais e 5 TN). Para selecionar sondas significativas foi utilizado o valor de p ajustado (FDR) $< 0,05$ e fold change $|FC| > 1,5$.

3.3.1.3 Comparação com dados externos e análise integrada - TCGA

Os dados de metilação e de transcriptômica que estavam disponíveis pelo TCGA para câncer de reto foram obtidos pelo UCSC Xena website (<http://xena.ucsc.edu>) e utilizados para realizar análises comparativas com os dados internos desse estudo. Dessa forma, foram incluídos dados de metilação de 99 tumores (sendo 11, 32, 38 e 14 casos com estádio I, II, III e IV, respectivamente e 4 casos sem informação) e sete TN (Infinium Human Methylation 450K BeadChip, Illumina) (TCGA). Igualmente, foram incluídos os dados de expressão de transcritos provenientes de 94 tumores (sendo 12, 26, 33 e 13 casos com estádio I, II, III e IV, respectivamente e 10 casos sem informação) e 10 TN (RNAseq versão 2) (TCGA).

Os transcritos com expressão diferencial em CaRe (FDR $< 5\%$) identificados nos dois grupos (interno e TCGA) foram avaliados pelo programa IPA (Ingenuity Pathway Analysis v.2.3) para o enriquecimento de vias canônicas desreguladas nesses tumores.

A análise de integração entre os dados de metilação e expressão de transcritos (internos e externos) foi realizada usando o teste de correlação de Pearson. Os genes com correlação significativa ($p < 0,05$) negativa (expressão aumentada e hipometilação ou expressão diminuída e hipermetilação) ou positiva (expressão e metilação ambas aumentadas ou diminuídas) utilizando os dois conjuntos de amostras foram utilizados para avaliação *in silico* de vias alteradas por metilação em CaRe pelo programa IPA.

3.3.2 Módulo 2 - Avaliação das alterações encontradas em câncer de reto de acordo com resposta ao tratamento neoadjuvante

3.3.2.1 Investigação de Alterações Genômicas

✓ Alterações no número de cópias genômicas

Os arquivos CEL foram gerados e analisados pelos programas Affymetrix® GeneChip® Command Console® (AGCC) 4.0 e Chromosome Analysis Suite v3.1 (ChAS) (Affymetrix/ThermoFisher, EUA), respectivamente. Considerou-se pelo menos 25 sondas alteradas para perdas, 50 para ganhos e uma região de pelo menos 5Mb para cnLOH, como previamente descrito (VILLACIS et al. 2016). A partir dessa análise, obteve-se uma lista de segmentos ou regiões alteradas para cada amostra, mapeadas no genoma de referência hg19. Regiões com variantes comuns foram definidas como aquelas presentes em mais de 1% da população sendo pesquisadas no *Database of Genomic Variants* (DGV, <http://dgv.tcag.ca/dgv/app/home>, acesso maio de 2016) e Affymetrix Database of Variants (aDGV composto por 2.421 amostras de indivíduos fenotipicamente saudáveis avaliados pela plataforma CytoScan HD array).

✓ **Identificação de regiões preferencialmente alteradas em câncer de reto de acordo com a resposta ao tratamento neoadjuvante**

Para a identificação de regiões diferencialmente alteradas entre pCR e pIR foi utilizada a ferramenta online CoNVaQ (LARSEN et al. 2018). A ferramenta compara as regiões comumente alteradas em dois grupos de amostras usando o teste exato de Fisher. Além disso, CoNVaQ calcula valores de p empíricos (pE) para regiões correspondentes executando repetidamente o mesmo teste enquanto alterna aleatoriamente as amostras de cada grupo de estudo (teste de permutação). O pE determina a probabilidade de encontrar uma região correspondente de mesmo tamanho ou maior, quando o tamanho de cada grupo é preservado, mas as amostras são distribuídas aleatoriamente entre os grupos. A posição e o tipo dos segmentos CNV/CNA individuais não são alterados por essa estratégia.

✓ **Cálculo dos escores de instabilidade genômica e de deficiência da via de recombinação homóloga**

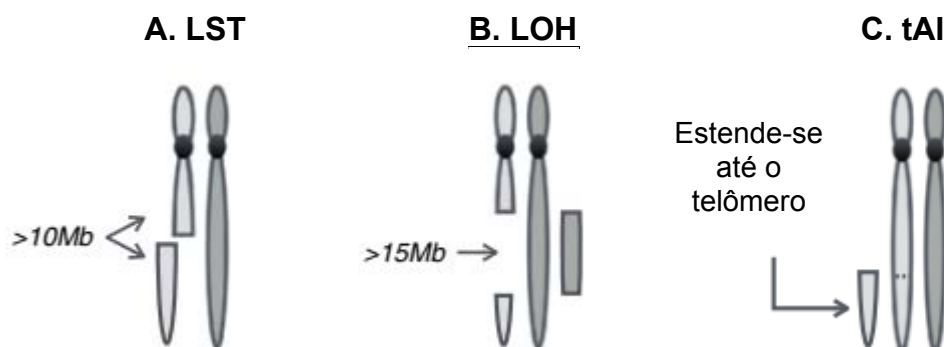
O escore ou index de instabilidade genômica (GII) foi definido como a fração de cada cromossomo com alterações no número de cópias (número total de pares de bases cobertas por segmento dividido pelo comprimento total do cromossomo) (LIU et al. 2016). O escore foi calculado com base nos segmentos obtidos da análise realizada usando o ChAS 3.1 e genoma de referência hg19.

Os seguintes escores relacionados à deficiência da via de recombinação homóloga (HRD) foram calculados de acordo com os algoritmos descritos por MARQUARD et al. (2015):

- *Large-Scale Transitions* (LST): Para o cálculo de LST, regiões com alteração no número de cópias menores que 3Mb foram filtradas e submetidas a

smoothing. O termo *smoothing* se refere à incorporação dos segmentos de extensão menor aos segmentos adjacentes com diferente número de cópias e extensão maior. Posteriormente, o número de pontos de quebra que ocorre entre regiões de pelo menos 10Mb foi contabilizado para cada braço cromossômico de uma amostra; a soma desses eventos resulta no valor de LST para determinada amostra. O limiar para diferenciar LST alto ou baixo (POPOVA et al. 2012) foi estabelecido de acordo com a ploidia de cada amostra, (calculada como descrito por MARQUARD et al. 2015), sendo o LST de 15 ou 20 para tumores “quase-diploides” ou “quase-tetraploides” respectivamente (Figura 2A)

- *Loss of Heterozygosity (LOH)*: Representa o número de regiões com $cnLOH$ maior que 15Mb e menor que um cromossomo inteiro. Escores maiores que 10 indicam a presença de HRD (ABKEVICH et al. 2012) (Figura 2B).
- *Telomeric Allelic Imbalance (tAI)*: Refere-se ao número de regiões teloméricas com desequilíbrio alélico que se estendem até a região subtelomérica sem ultrapassar o centrômero (BIRKBAK et al. 2012). Os valores de tAI foram calculados considerando-se as modificações propostas por MARQUARD et al. (2015). Em suma, analisa-se o número de cópias global para cada cromossomo de forma independente e usando esse valor como base, contabiliza-se o número de eventos com tAI para o genoma de determinado tumor (Figura 2C). A mediana foi utilizada para categorizar os casos com escore alto e baixo.



Fonte: Adaptado de MARQUARD et al. (2015).

Figura 2 - Visualização esquemática dos segmentos computados pelos três diferentes escores relacionados à deficiência da via de recombinação homóloga. Cinza claro e escuro representam cromossomos de origem materna e paterna. **A.** *Large-scale transition* (LST): número de pontos de quebra entre regiões adjacentes de pelo menos 10 Mb. **B.** *Loss of heterozygosity* (LOH): regiões com cnLOH maiores que 15 Mb, mas menores que o cromossomo inteiro. **C.** *Telomeric allelic imbalance* (tAI): regiões subteloômicas com desequilíbrio alélico que tem início além do centrômero estendendo-se até o telômero.

✓ Sequenciamento de alto desempenho

Os dados gerados a partir do sequenciamento foram analisados de acordo com o *pipeline* estabelecido pelo grupo de bioinformática e seguindo recomendações da literatura (desenvolvido por Mads Malik Aagaard, Department of Clinical Genetics, Vejle Hospital, DK). Em resumo, os dados brutos foram extraídos após o sequenciamento e as amostras identificadas e separadas do *pool* (processo denominado demultiplexação). Os *reads* resultantes foram mapeados no genoma de referência humano UCSC hg19 usando o algoritmo de alinhamento BWA (LI e DURBIN 2009). A identificação das variantes e a filtragem de qualidade das variantes foram realizadas com *Genome-Analysis-Toolkit* - GATK (DEPRISTO et al. 2011). As variantes resultantes foram anotadas principalmente utilizando as

ferramentas ANNOVAR (WANG et al. 2010), GATK e SnpEff (CINGOLANI et al. 2012).

O programa *Ingenuity Variant Analysis* (Qiagen) foi usado para classificar as variantes de acordo com o *American College of Medical Genetics and Genomics* (ACMG) (RICHARDS et al. 2015). Foram excluídas as variantes consideradas benignas ou provavelmente benignas. Em seguida, foram excluídas as variantes com frequência $> 0,01$ nos bancos de dados: *Genetic Association Database* (GAD) (BECKER et al. 2004) e/ou no *The Exome Aggregation Consortium* (ExAC) (LEK et al. 2016).

As demais variantes foram curadas manualmente usando o programa Genome Browse (Golden Helix Inc., EUA). As variantes mapeadas em regiões de homopolímeros também foram excluídas. Foi aplicada uma filtragem adicional excluindo-se variantes de significância incerta (VUS, classificadas pela ACMG) com valor do escore *Combined Annotation Dependent Depletion* (CADD) $< 3,0$. O resumo do *pipeline* básico de análise, anotação e filtragem de variantes está esquematizado na Figura 3.

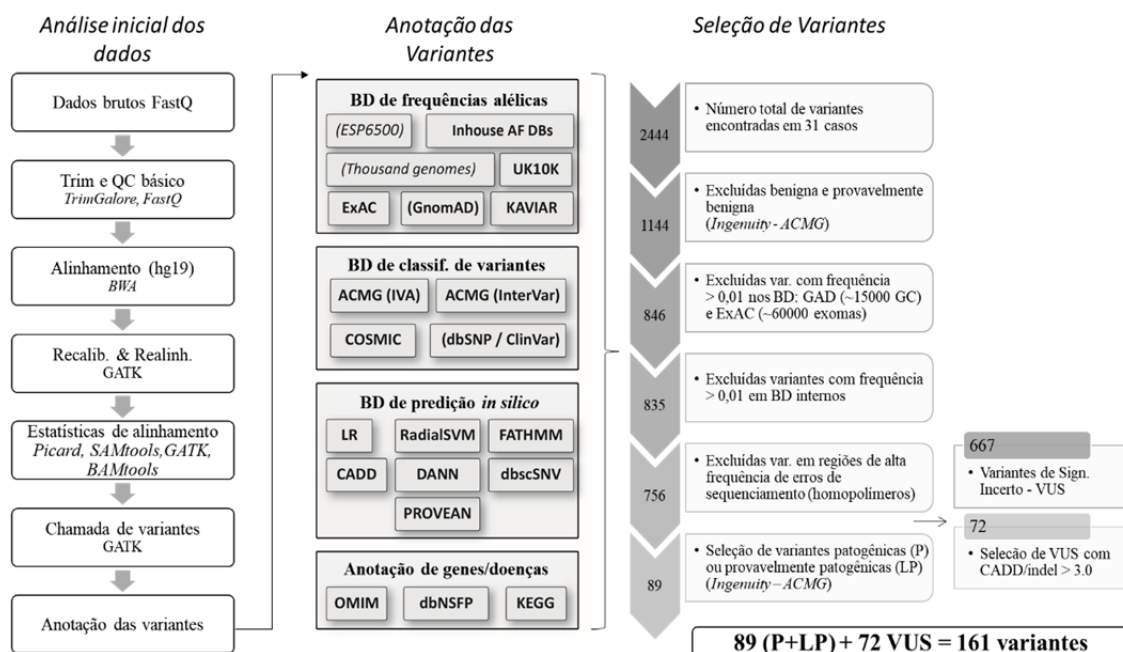


Figura 3 - Pipeline utilizado para o processamento dos dados brutos provenientes do sequenciamento de 105 genes relacionados ao câncer em 31 biópsias de câncer de reto. Estão apresentados os pacotes, programas ou banco de dados utilizados para comparações de frequências alélicas e de classificação das variantes. Também estão representados os passos para filtragem e seleção de variantes. BD: banco de dados; var.: variantes; GC: genoma completo.

3.3.3 Dados do TCGA - The Cancer Genome Atlas

Para o cálculo do GII, descrito na sessão anterior, foram obtidos os dados de alterações genômicas de 165 tumores de reto (sendo 30, 51, 50 e 25 casos com estágio I, II, III e IV, respectivamente e 9 casos sem informação) disponíveis pelo TCGA

(<https://tcga.xenahubs.net/download/>

TCGA.READ.sampleMap/SNP6_nocnv_genomicSegment.gz). O perfil mutacional de 145 desses casos (sendo 28, 48, 42 e 120 casos com estágio I, II, III e IV, respectivamente e 7 casos sem informação) estavam disponíveis e foram avaliados utilizando cBioPortal (acesso março, 2018 <http://www.cbioportal.org/>). Até o

presente, não há dados de resposta à terapia neoadjuvante nos casos de CaRe do TCGA, impedindo tais comparações com os resultados obtidos no presente estudo.

3.3.4 Análise do perfil de metilação do DNA

O controle de qualidade e o pré-processamento dos dados brutos foi realizado como descrito no item 3.3.1. As sondas com desvio padrão maior que 0,2 foram utilizadas como input para o BRB array tools (<http://linus.nci.nih.gov/BRB-ArrayTools.html>). A análise de agrupamento hierárquico não supervisionado foi realizada utilizando a distância Euclidiana. Foram identificadas sondas diferencialmente metiladas nas comparações entre cada grupo (pCR e pIR) com o TN usando o pacote limma (Ritchie ME, et al. 2015). O valor de p ajustado (FDR) $< 5\%$ e $|\Delta\beta| > 0,15$ foram utilizados para selecionar sondas significativas, as quais foram anotadas usando o MethylationEPIC_v-1-0_B2 (Illumina).

As sondas diferencialmente metiladas exclusivas da comparação de cada grupo (pCR e pIR versus TN) foram utilizadas para o desenho de um modelo preditivo de resposta à nCRT. O método de eliminação de características recursivas (Recursive Feature Elimination) foi utilizado para selecionar três sondas, as quais foram usadas na Análise Diagonal Linear Discriminante (DLDA) para o treinamento de um classificador (DUDOIT S et al. 2002). A performance do mesmo foi calculada usando o teste *leave-one-out cross-validation* (LOOCV, BRB array tools v. 4.4.0).

3.3.5 Análise de expressão de transcritos

3.3.5.1 Análise dos resultados da expressão global de transcritos

O pré-processamento e análise dos dados de expressão foram avaliados como descrito na sessão 3.3.2.1. Os valores de expressão das sondas foram comparados entre cada grupo de casos (pCR e pIR) e o grupo de TN usando o pacote limma (RITCHIE et al. 2015) disponível para o programa R (<https://www.r-project.org/>). Para selecionar sondas significativas foi utilizado $FDR < 0,05$ e $|FC| > 1,5$. O enriquecimento de vias foi realizado utilizando o programa Ingenuity Pathway Analysis (IPA, Qiagen).

Foi realizada uma análise de secretoma com o intuito de identificar proteínas secretadas pelos tumores de reto com potencial para estarem presentes na circulação, demonstrando assim, serem candidatas a biomarcadores em biópsia líquida. Para cada grupo, genes diferencialmente expressos com $FC > 1,5$ foram selecionados para a identificação de produtos possivelmente secretados pelo tumor. Inicialmente, a identificação das proteínas traduzidas pelos genes diferencialmente expressos e aquisição da sequência de aminoácidos de cada uma delas foi realizada usando UniProtKB (<http://www.uniprot.org>). Foram avaliadas somente sequências revisadas e com no máximo 4.000 aminoácidos. Os arquivos FASTA gerados por UniProtKB foram utilizados como ferramentas de predição de proteínas secretadas dos servidores CBS: SecretomeP 2.0, SignalP 4.1 e TargetP 1.1 (disponíveis em <http://www.cbs.dtu.dk/services/>).

A ferramenta TargetP é utilizada para a predição de localização de proteínas baseado na presença de sequências na região N-terminal que indiquem direcionamento mitocondrial (mTP) ou peptídeo-sinal da via secretora (SP). Em adição, SignalP prevê a presença e a localização dos sítios de clivagem do

peptídeo-sinal nas sequências de aminoácidos de diferentes organismos, identificando proteínas secretadas pela via clássica. O método incorpora uma previsão de locais de clivagem e uma predição de peptídeo-sinal/peptídeo não sinal com base em uma combinação de várias redes neurais artificiais. A ferramenta SecretomeP permite prever as proteínas secretadas pela via não clássica, ou seja, sem a presença do peptídeo-sinal. O método consulta um grande número de outros servidores para obter informações sobre vários aspectos pós-translacionais e de localização da proteína, que são integrados na previsão final de secreção. A comparação entre os resultados dos três bancos de dados foi realizada para selecionar os candidatos com maior potencial para serem investigados em biópsias líquidas.

Proteínas identificadas como secretadas pelas vias clássica (SignalP) e não clássica (SecretomeP) que não possuíam como alvo a mitocôndria (TargetP) foram selecionadas para uma investigação detalhada em bancos de dados de proteínas secretadas em vesículas e no plasma: ExoCarta (<http://www.exocarta.org/>), Vesiclepedia (<http://www.microvesicles.org/>) The Plasma/ Proteome database (<http://plasmaproteomedatabase.org>), Cancer Secretome (<http://www.cancersecretome.org/>), EVpedia (KIM et al. 2015).

3.3.6 Análise Integrada

Foi empregada uma abordagem baseada na análise integrada dos dados de expressão de mRNA (HTA 2.0, Affymetrix) e de metilação (Infinium Human MethylationEPIC BeadChip array, Illumina) para identificar sondas/genes com correlação positiva e negativa e possíveis marcadores implicados na resposta ao tratamento neoadjuvante.

Foram identificados genes diferencialmente expressos utilizando o modelo linear do pacote limma (RITCHIE et al. 2015) para a comparação de cada grupo de casos (pCR ou pIR) com o tecido normal (TN). O valor de p ajustado (FDR) $<0,05$ foi utilizado para selecionar genes/sondas significativas. Utilizando os dados de transcriptômica e metiloma, foram selecionados genes/sondas com correlação negativa (expressão aumentada e hipometilação ou expressão diminuída e hipermetilação) ou positiva (expressão e metilação ambas aumentadas ou diminuídas). Foram considerados significativos os dados de correlação de Pearson com $p < 0,05$.

3.3.7 Análise Estatística

Análises estatísticas foram realizadas usando os programas SPSS (SPSS 24 Inc, EUA), MedCalc v. 18 (<https://www.medcalc.org>) e CoNVaQ (LARSEN et al. 2018). Variáveis categóricas e contínuas relacionadas aos dados clínicos foram comparadas por teste exato de Fisher e teste U de Mann-Whitney, respectivamente. Múltiplos grupos foram comparados usando o teste não paramétrico de Kruskal-Wallis. Para a realização das análises não supervisionadas foi utilizado o programa BRB-array v.4.5.1 (SIMON et al. 2007).

4 RESULTADOS

Os resultados da análise do número de cópias genômicas (ganhos, perdas, cnLOH) e do perfil mutacional de 33 carcinomas de reto estão apresentados em detalhes no Anexo 3.

4.1 MÓDULO 1 - AVALIAÇÃO DAS ALTERAÇÕES ENCONTRADAS EM CÂNCER DE RETO: COMPARAÇÃO ENTRE AMOSTRAS TUMORAIS E NORMAIS.

4.1.1 Análise dos Dados de Metilação

As mais de 850 mil sondas incluídas na plataforma Infinium MethylationEPIC (Illumina) foram estratificadas de acordo com a distribuição genômica onde estão localizadas: região intergênica, região promotora e corpo de genes ou 3'UTR. Em relação ao contexto de CpGs, as sondas foram classificadas como: dentro de ilhas CpG, em regiões *shore/ shelf* (1 a 2.000 pares de bases/ 2.001 a 4.000 pb distante da ilha CpG) e CpGs isolados em regiões *open sea* (mais de 4.000 pb distante da ilha CpG).

4.1.1.1 Perfil de metilação de câncer de reto (CaRe) comparado ao tecido normal (TN) de reto

Inicialmente, foi realizada uma análise do perfil de metilação dos carcinomas de reto como um todo e os resultados foram comparados a dados externos para validação cruzada. Como não há dados de metilação publicados em CaRe

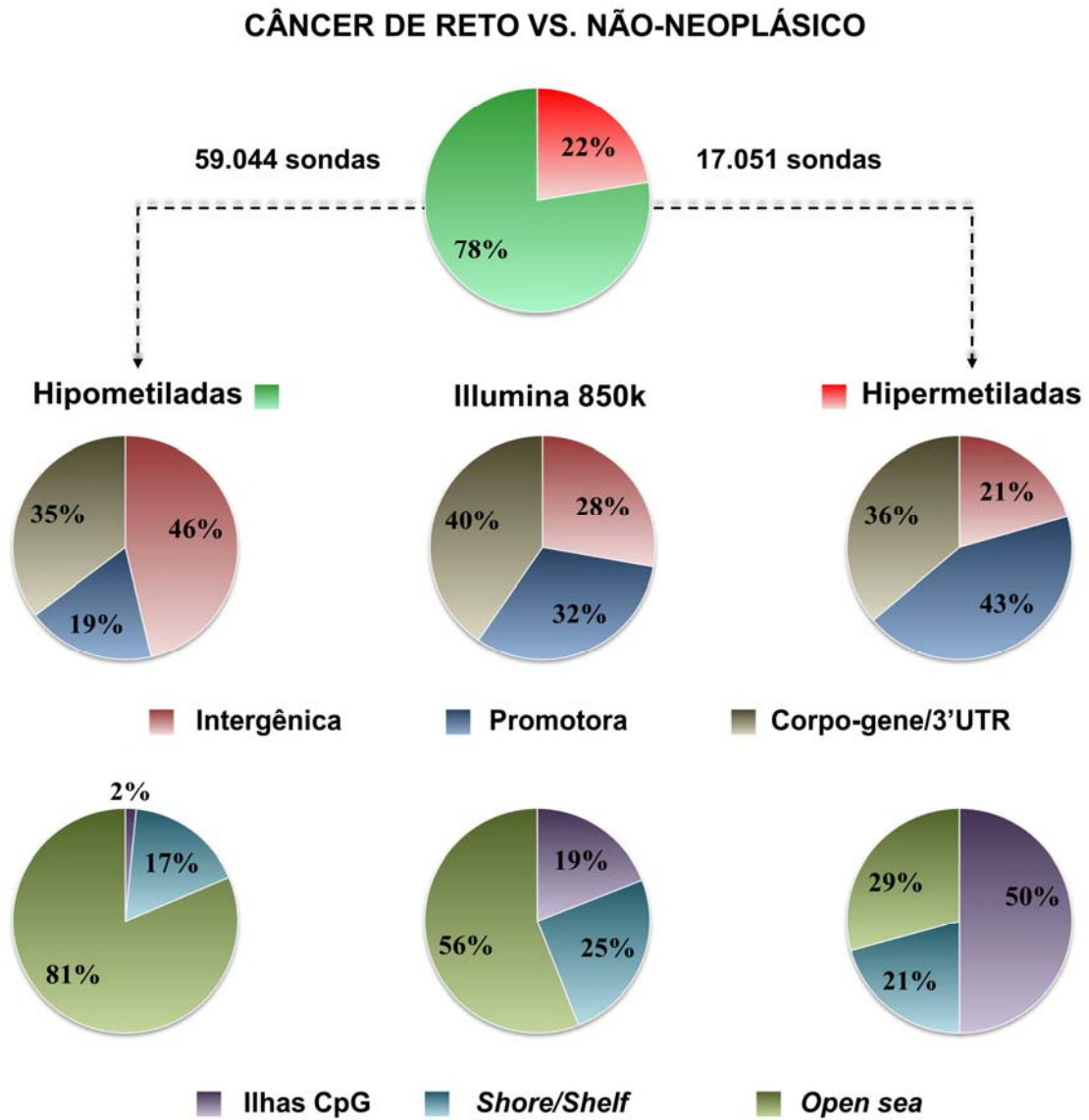
empregando a mesma plataforma que o presente estudo, foram usados dados do TCGA, com plataforma menos robusta (Infinium Methylation450, Illumina). A plataforma Infinium MethylationEPIC (Illumina), utilizada neste estudo, apresenta mais de 850 mil sondas com maioria mapeada em corpo de gene ou em região 3'UTR (40%) e nas regiões de *open sea* (56%), enquanto a plataforma Infinium Methylation450 (Illumina), com 485.512 sondas, possui a maioria das sondas localizadas em região promotora (41%) e uma distribuição mais homogênea em relação ao contexto CpG (36%: ilhas, 33%: *open sea*, 31%: *shore/shelf*). A Figura 4 representa a comparação entre os dados obtidos no presente estudo e os obtidos da análise de dados do TCGA.

A comparação entre as 32 amostras de CaRe e 5 de TN, usando a plataforma Infinium MethylationEPIC (Illumina), revelou 76.095 sondas diferencialmente metiladas (DM) (*limma*, FDR<5%; $|\Delta\beta| > 0,15$), sendo 59.044 CpGs hipometiladas (78%) e 17.051 CpGs hipermetiladas (22%). A maioria das sondas hipometiladas estava localizada em regiões intergênicas (46%) e nas regiões *open sea* (81%), enquanto a maior parte das sondas hipermetiladas estava mapeada em regiões promotoras (43%) e em ilhas CpG (50%) (Figura 4A).

Utilizando os resultados de metilação do TCGA (99 CaRe e 7 TN), foi verificado 54.328 sondas DM, das quais, 33.297 (61%) e 21.031 (39%) estavam hipometiladas e hipermetiladas, respectivamente. A maioria das sondas hipometiladas encontrava-se distribuída nas regiões intergênica (38%) e *open sea* (63%) e, hipermetiladas, nas regiões promotora (43%) e de ilhas CpG (64%) (Figura 4B). Essas proporções foram similares às descritas para o grupo de casos do presente estudo. Entre as 76.095 sondas DM identificadas neste estudo, 34,7%

estavam também na plataforma 450K e 81,2% delas foram validadas usando os dados do TCGA.

A



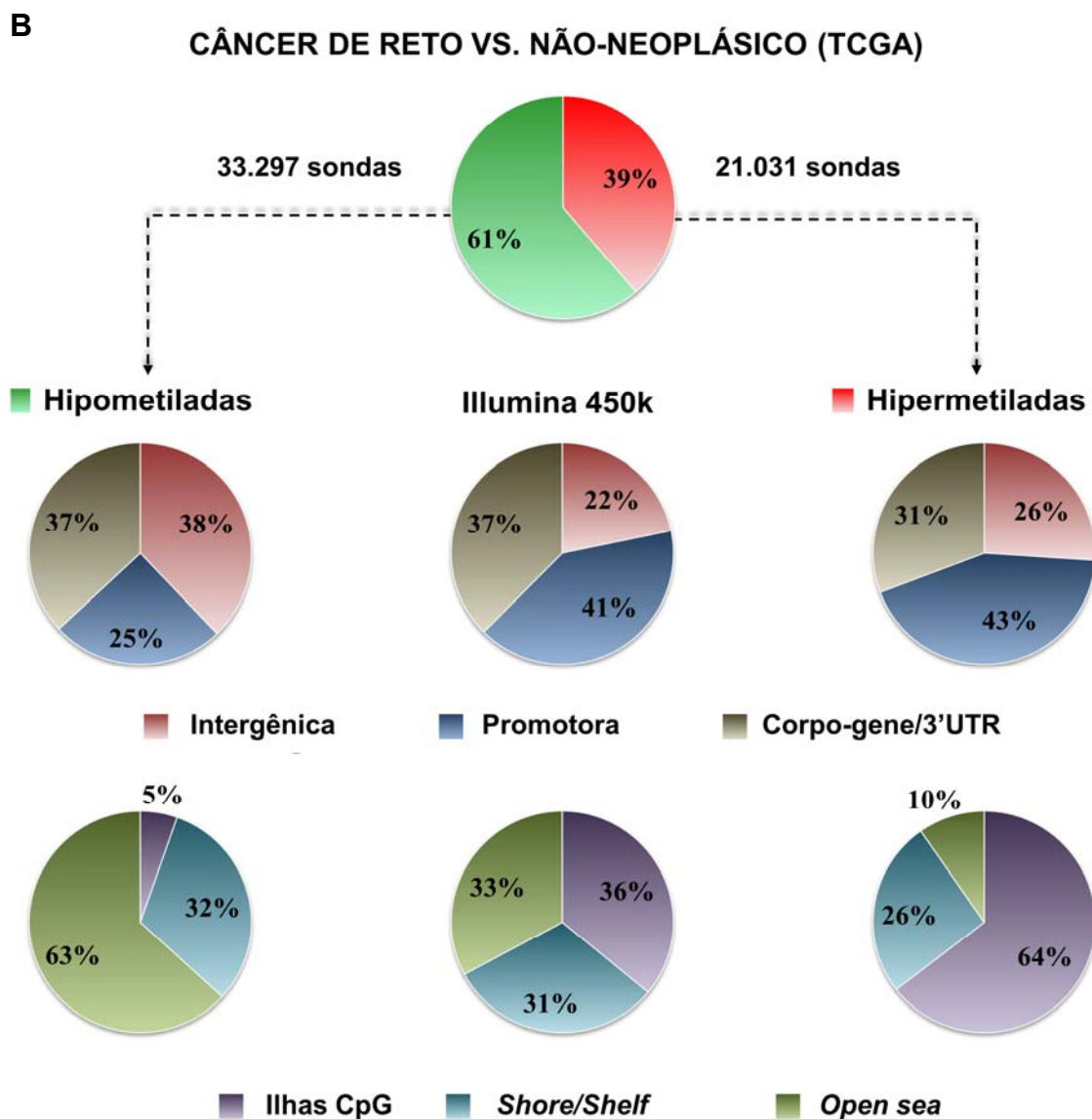


Figura 4 - Representação esquemática da distribuição genômica e do contexto CpG das sondas diferencialmente metiladas identificadas nas amostras de câncer de reto e das sondas presentes nas plataformas utilizadas em cada estudo. **A**: Sondagens com metilação diferencial detectadas neste estudo (32 CaRe e 5TN) foram obtidas usando a plataforma Infinium Human Methylation EPIC (850K, Illumina) **B**: Sondagens diferencialmente metiladas identificadas nos casos de câncer de reto disponíveis no TCGA (99 CaRe e 7 TN) e investigadas pela plataforma Infinium Human Methylation 450K (Illumina)

4.1.2 Análise de Expressão Global de Transcritos

4.2.2.1 Perfil de expressão de câncer de reto

Foram identificados 737 transcritos diferencialmente expressos (DE) em 28 CaRe comparados com cinco TN (FDR < 5%), sendo 567 sondas de transcritos codificadores (262 FC> 1,5 e 305 FC< -1,5) e 170 de não-codificadores (110 FC> 1,5 e 60 FC< -1,5). Os dados de transcritos DE disponíveis pelo TCGA (94 CaRe e 10 TN) foram utilizados para a comparação com os achados desse estudo. Entre os 9.280 transcritos DE identificados no TCGA, 454 também estavam DE no nosso estudo, sendo 437 na mesma direção.

Os genes DE nas duas análises (dados internos e do TCGA) foram utilizados para identificar vias desreguladas em CaRe usando o programa IPA, sendo detectadas 18 vias (Tabela 4). As cinco vias mais significativas foram *Cell Cycle: G2/M DNA Damage Checkpoint Regulation*, *Phospholipase C Signaling*, *Protein Kinase A Signaling*, *ILK Signaling* e *Cyclins and Cell Cycle Regulation*.

Tabela 4 - Vias canônicas alteradas em câncer de reto, identificadas usando o programa *Ingenuity Pathway Analysis* a partir de genes diferencialmente expressos identificados nos dados internos e do TCGA.

Vias Canônicas IPA	Genes	p-valor	p-valor corrigido	z-score
<i>Cell Cycle: G2/M DNA Damage Checkpoint Regulation</i>	<i>PRKDC, CKS2, YWHAG, CDK7, ABL1, CCNB2, AURKA, CDK1, CHEK1, CCNB1, SKP2, KAT2B, CKS1B, TOP2A, BORA, ATR, BRCA1, SFN</i>	7E-07	4E-04	-2,3
<i>Phospholipase C Signaling</i>	<i>RALA, MYL6, PLA2G10, ARHGEF7, FCGR2B, NFATC1, PLCD1, PLCD3, PPP1R12B, PLCE1, AHNAK, RHOB, ADCY5, MAPK3, PLA2G5, RHO, PRKCE, PLCD4, MYL3, ITK, SRC, HDAC9, ADCY2, HDAC4, HDAC2, RRAS, GNG2, ITGA2, ADCY6, PPP1R14A, CD79A, ITPR1, NFATC4, PLA2G2A, MYL9, BTK, CALM1 (includes others), PLA2G2D, GRAP2, PLCG2, MEF2D, LYN, MEF2C, FNBP1, PRKCB</i>	5E-05	2E-03	-3,5
<i>Protein Kinase A Signaling</i>	<i>MYL6, PDE3A, ACP1, PTK2, PLCD3, PLCE1, PDE7B, ADCY5, MAPK3, RYR3, CDC16, EYA2, MYL3, CDC25A, PDE2A, YWHAG, ADD2, GNG2, PPP1R14A, ITPR1, ANAPC7, CNGA3, MYL9, PTPRH, PYGM, DUSP1, PLCG2, PDE1B, PTPRS, DUSP4, PTPN21, CAMK2G, AKAP12, PTK2B, PDE7A, PPP1R3C, MYLK, PDE1A, AKAP7, PTPN12, NFATC1, AKAP11, MPPE1, CDKN3, PLCD1, PTPN4, GLI3, FLNA, DCC, PTPRZ1, PRKCE, PDE4D, CTNNB1, PLCD4, MTMR3, PLN, ADCY2, PDE9A, ADCY6, AKAP6, NFATC4, TCF7L1, CALM1 (includes others), FLNC, ADD1, PDE5A, SFN, ELK1, PRKCB</i>	1E-05	2E-03	-2,6
<i>ILK Signaling</i>	<i>MYL6, ILK, PPP2R3B, ITGB8, MYH11, CCND1, ITGB7, ITGB3, PTK2, VEGFA, TGFB111, RHOB, CFL2, FLNA, MAPK3, PPM1L, RHO, VEGFD, VCL, ACTG2, CTNNB1, MYL3, PARVA, FBLIM1, LIMS2, TNFRSF1A, FGFR1, ACTB, FERMT2, FGFR2, MYL9, FOS, ACTA2, PPP2R3A, FLNC, PPP2R2B, MAPK10, MMP9, FNBP1</i>	4E-05	2E-03	-2,3
<i>Cyclins and Cell Cycle Regulation</i>	<i>HDAC9, HDAC4, TFDP1, HDAC2, CDK7, ABL1, PPP2R3B, CCNB2, E2F3, CCND1, CDK1, CCNB1, SKP2, CCNA2, PPP2R3A, PPP2R2B, PPM1L, E2F1, E2F5, ATR, CDC25A</i>	4E-05	2E-03	3,0
<i>Estrogen-mediated S-phase Entry</i>	<i>CCNA2, TFDP1, E2F1, E2F5, E2F3, RBL1, CCND1, CDK1, CDC25A, SKP2</i>	1E-04	4E-03	2,3
<i>Aldosterone Signaling in Epithelial Cells</i>	<i>CRYAB, SGK1, HSPB8, DNAJC10, DNAJA1, SLC9A1, PLCD1, PLCD3, PLCE1, MAPK3, HSPE1, PRKCE, SCNN1B, PLCD4, HSPB6, AHCY, DNAJC9, HSPH1, SLC12A2, FGFR1, FGFR2, DNAJC2, HSPD1, ITPR1, HSPA8, SCNN1G, DUSP1, PIP5K1C, PLCG2, NR3C2, HSPB7, DNAJB5, PRKCB</i>	2E-04	5E-03	-3,4
<i>Integrin Signaling</i>	<i>RAPGEF1, RALA, ARPC1B, TSPAN7, ARHGEF7, ITGA8, ABL1, ILK, MYLK, TLN1, ITGB8, ITGB7, ITGB3, PTK2, PPP1R12B, RHOB, MAPK3, RHO, CAV1, VCL, ACTG2, CAPN5, SRC, PARVA, RRAS, FGFR1, ACTB, ITGA2, ITGA6, TSPAN2, FGFR2, GSN, ITGAL, MYL9, TLN2, ACTA2, TSPAN1, PLCG2, ITGA7, FNBP1</i>	2E-04	5E-03	-3,2

Cont/ Tabela 4

Vias Canônicas IPA	Genes	p-valor	p-valor corrigido	z-score
Paxillin Signaling	SRC,PARVA,PTK2B,RRAS,ACTB,ARHGEF7,FGFR1,ITGA8,ITGA2,ITGA6,FGFR2,TLN1,ITGB8,PTPN12,ITGAL,ITGB7,ITGB3,PTK2,TLN2,ACTA2,MAPK10,ACTG2,VCL,ITGA7,AKAP12,GPR17,HTR4,PDE7A,PDE3A,ADORA3,PDE1A,AKAP7,MPPE1,AKAP11,CHRM2,ADRB1,FFAR3,PDE7B,ADCY5,CNR2,MAPK3,PDE4D,SRC,PDE2A,RGS2,ADCY2,PDE9A,NPY1R,CNR1,ADCY6,AKAP6,CNGA3,CALM1 (includes others),P2RY14,NPR3,VIPR1,DUSP1,PKIB,PDE1B,GNAO1,PDE5A,DUSP4,AGTR1,CAMK2G,PLA2G10,GNA11,CACNA1H,PPP2R3B,PRKG2,PLCD1,PLCD3,PLCE1,IGF1,MAPK3,RYR3,PLA2G5,PPM1L,PRKCE,CACNB2,CACNA2D3,PLCD4,RRAS,CACNA1C,ITPR1,PLA2G2A,PLA2G2D,NPR3,PPP2R3A,PLCG2,PPP2R2B,CACNA2D1,GNAO1,LYN,NPR2,PRKCB	4E-04	9E-03	-2,2
cAMP-mediated signaling	ADCY2,GNG2,GNA11,CACNA1H,ADCY6,CACNA1C,ITPR1,PYY,GCG,PLCD1,PLCD3,PLCE1,FFAR3,ADCY5,PLCG2,CACNA2D1,LPAR5,PRKCE,CACNB2,CACNA2D3,PLCD4,PRKCB	4E-04	9E-03	-4,0
Synaptic Long Term Depression	PDE2A,PTK2B,ZP3,SLC12A2,PLA2G10,CACNA1H,ITPR1,PRKG2,PDE1A,PLA2G2A,CNGA3,PLCD1,PTK2,PLCD3,CALM1 (includes others),PLCE1,PLA2G2D,PLCG2,PDE1B,PLA2G5,PRKCE,PDE4D,PLCD4,PRKCB	2E-03	2E-02	-4,1
GPCR-Mediated Nutrient Sensing in Enteroendocrine Cells	KIF23,PRC1,CDC7,PPP2R3B,CCNB2,ANAPC7,CDK1,CCNB1,PLK4,PPP2R3A,PPM1L,PPP2R2B,FBXO5,CDC16,CDC25A	2E-03	3E-02	-4,3
Sperm Motility	LIF,CACNA1H,HAND1,PLCD1,PLCD3,PLCE1,IGF1,ADCY5,MAPK3,PRKCE,CACNB2,CACNA2D3,PLCD4,HDAC9,SRC,ADCY2,HDAC4,HDAC2,RRAS,FGFR1,GNG2,ADCY6,CACNA1C,FGFR2,ITPR1,NFATC4,CALM1 (includes others),PLCG2,MEF2D,CACNA2D1,MAPK10,MEF2C,SLC8A1,RCAN2,CAMK2G,PRKCB	2E-03	3E-02	-3,4
Mitotic Roles of Polo-Like Kinase	MRC1,SRC,RRAS,FGFR1,ANKFY1,USP6NL,FGFR2,EGF,ITGB8,CSF1R,ITGB7,ITGB3,MET,CSF1,PLCG2,PRKCE,PRKCB	2E-03	3E-02	2,3
Role of NFAT in Cardiac Hypertrophy	MYL6,GRIN2D,CACNA1H,TPM1,MYH11,TPM2,NFATC1,MAPK3,RYR3,CHRNA7,CACNB2,CACNA2D3,MYL3,CASQ2,HDAC9,LETM1,HDAC4,HDAC2,ATP2C1,CACNA1C,ATP2A3,ITPR1,NFATC4,GRIN3A,MYL9,CALM1 (includes others),ACTA2,MEF2D,CACNA2D1,MEF2C,SLC8A1,RCAN2,ATP2B4,CAMK2G	3E-03	3E-02	-4,2
Macropinocytosis Signaling	MYL6,CD4,GNA11,CXCL12,PTK2,RHOB,ADCY5,MAPK3,RHOV,PRKCE,MYL3,SRC,ADCY2,RRAS,FGFR1,GNG2,ADCY6,FGFR2,ITPR1,MYL9,FOS,GNAO1,LYN,MAPK10,ELK1,ELMO1,FNBP1,PRKCB	3E-03	3E-02	-2,1
Calcium Signaling		3E-03	4E-02	-3,4
CXCR4 Signaling		5E-03	5E-02	-2,0

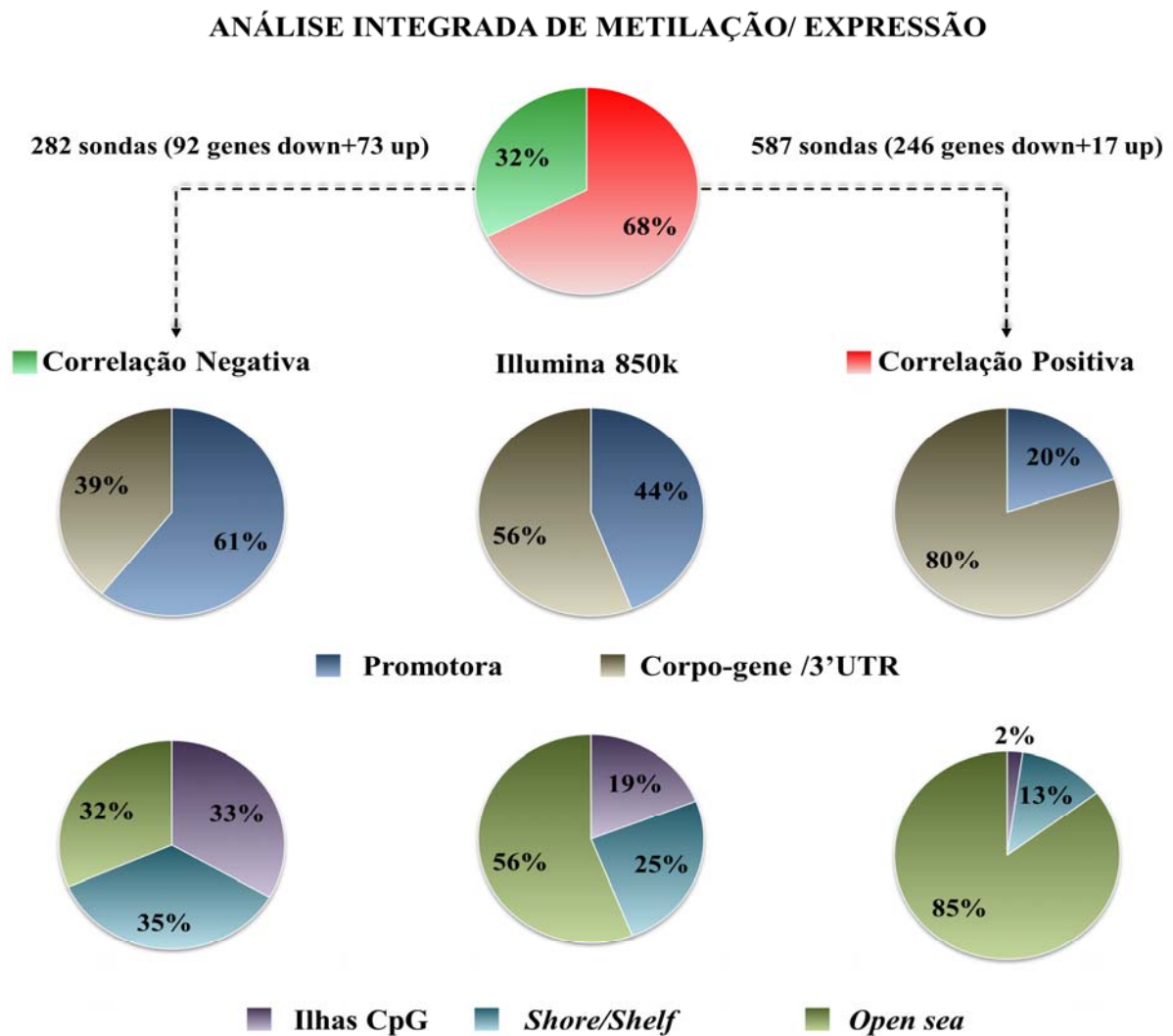
4.1.3 Análise Integrada de Metilação e Expressão Gênica

Foi realizada uma análise integrada utilizando as 76.095 sondas com metilação diferencial e 737 transcritos DE. A figura 5A detalha as características das 869 sondas DM mapeadas em 393 genes diferencialmente expressos em CaRe, os quais apresentaram correlação significativa (Pearson, $p < 0,05$). Entre as 282 sondas com correlação negativa (r^- , sondas com hiper ou hipometilação cobrindo 92 genes com diminuição de expressão e 73 genes com aumento de expressão, respectivamente), 61% estavam localizadas em região promotora e 68% em regiões regulatórias (ilhas CpG e *Shore/Shelf*). Em adição, as 587 sondas com correlação positiva (r^+ , sondas com hiper ou hipometilação cobrindo 17 genes com aumento de expressão e 246 genes com diminuição de expressão, respectivamente) estavam predominantemente mapeadas em regiões de corpo de genes ou 3'UTR (80%) e somente 15% estavam mapeadas em regiões de ilhas CpG ou *Shore/Shelf*.

Usando os dados de expressão (9.280 transcritos) e metilação (54.328 sondas) diferencial dos casos do TCGA foram identificadas 6.373 sondas DM com correlação significativa em relação à expressão de 2.425 genes (Pearson, $p < 0,05$) e uma distribuição mais homogênea entre aquelas com r^- (53%) e r^+ (47%). Ainda assim, a distribuição tipológica das sondas foi similar àquela encontrada neste estudo, considerando a diferente representatividade das sondas em cada plataforma (Figura 5B). Em suma, 67% das 3.556 sondas com r^- (sondas com hiper ou hipometilação cobrindo 805 genes com expressão diminuída e 494 genes com aumento de expressão nos tumores, respectivamente) estavam localizadas em região promotora e 85% em regiões regulatórias (ilhas CpG e *Shore/Shelf*). Já as 3.181 sondas com r^+ (sondas com hiper ou hipometilação cobrindo 127 genes com aumento de expressão e 999 genes com diminuição de expressão nos tumores,

respectivamente) localizavam-se em regiões de corpo de genes ou 3'UTR (77%). Uma fração maior de sondas com $r+$ (45%) estava mapeada em regiões de ilhas CpG ou *Shore/Shelf* nos casos do TCGA quando comparado aos deste estudo (15%). No entanto, o total de sondas distribuídas nessas regiões em cada plataforma mostrou representatividade aumentada de sondas em *Open Sea* na plataforma EPIC 850K usada nas amostras do nosso estudo (Figura 5).

A



B

ANÁLISE INTEGRADA DE METILAÇÃO/ EXPRESSÃO (TCGA)

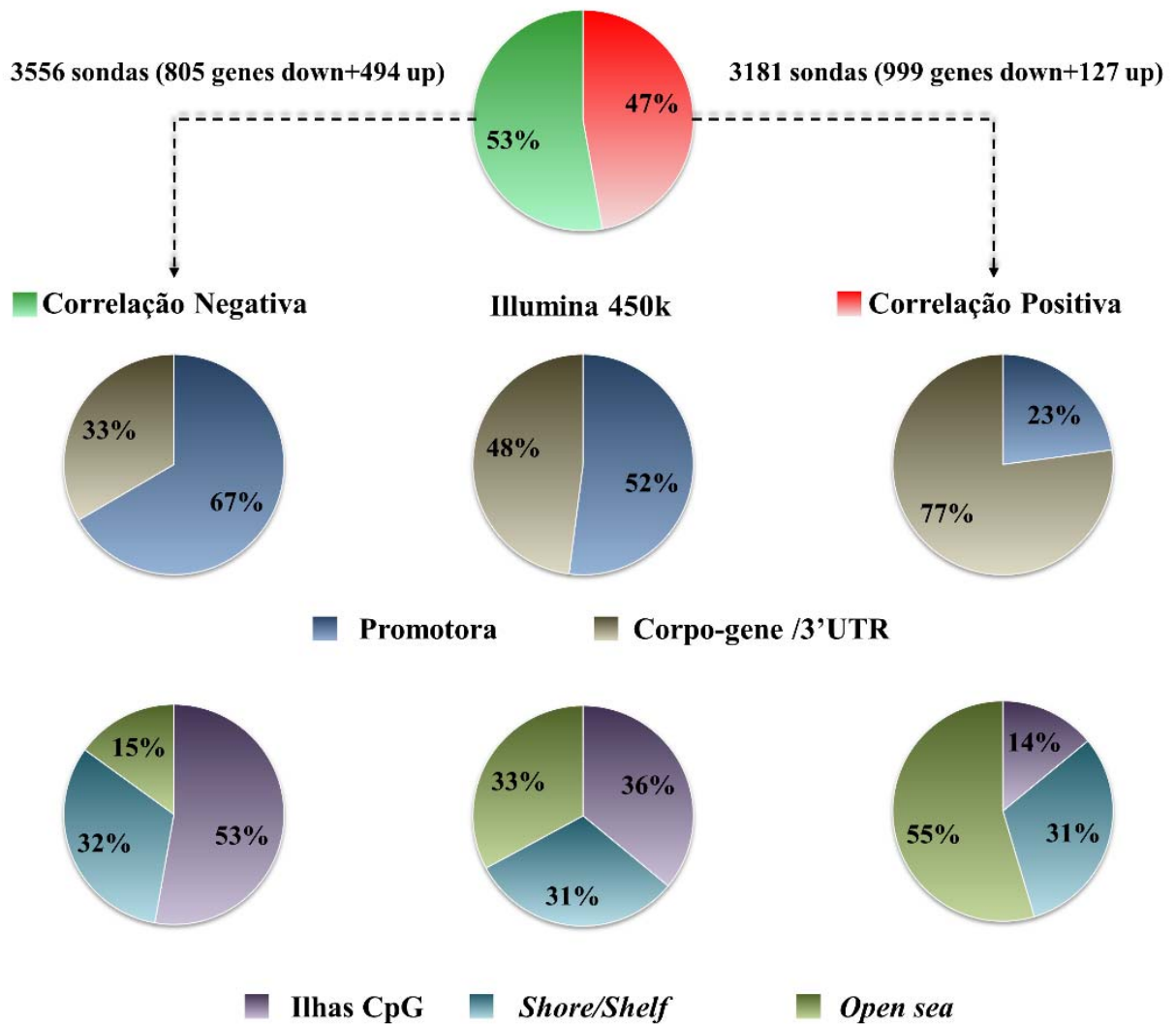


Figura 5 - Descrição das sondas diferencialmente metiladas com correlação significativa com os dados de expressão gênica identificadas neste estudo. (A) e usando dados do TCGA (B). **Down** e **up** refere-se a genes com expressão diminuída e aumentada em relação ao tecido normal, respectivamente.

Os genes diferencialmente expressos entre CaRe e TN nos dois grupos de casos (dados internos e do TCGA) que são potencialmente regulados por metilação (com base na análise integrada) foram utilizados para análise *in silico* de vias canônicas (IPA). A Figura 6 representa em detalhes a estratégia utilizada nessa análise. Foram identificadas 11 vias canônicas potencialmente desreguladas por metilação em CaRe (Tabela 5).

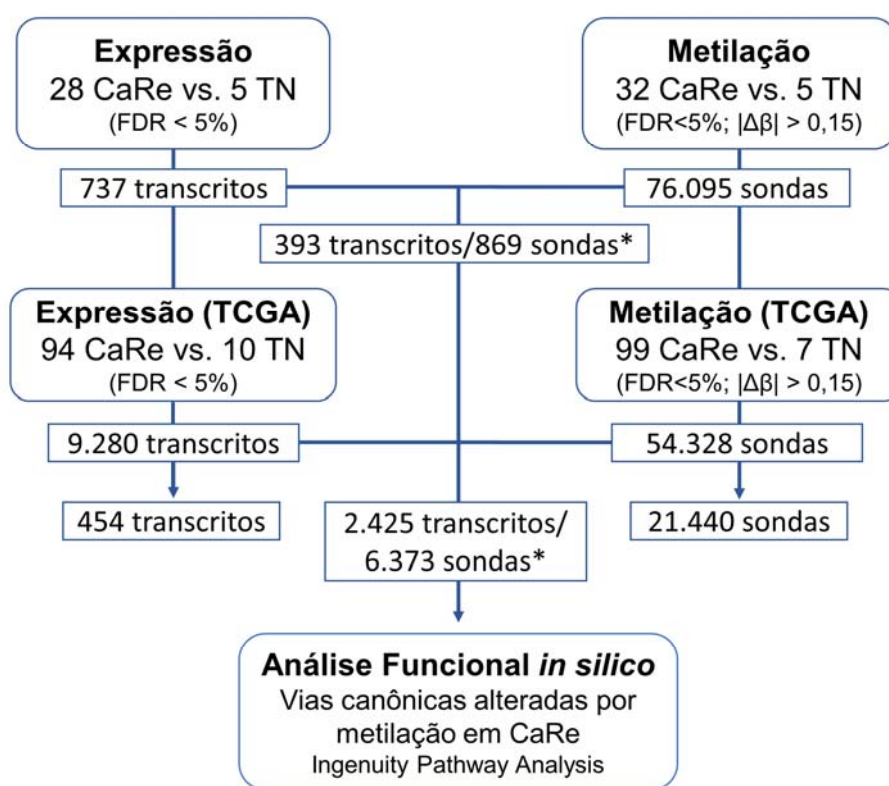


Figura 6 - Estratégia utilizada para análise integrada dos dados de metilação e expressão gênica a partir dos dados internos e do TCGA e identificação de vias alteradas por metilação em câncer de reto. *Correlação de Pearson < 0,05.

Tabela 5 - Vias canônicas identificadas pelo programa *Ingenuity Pathway Analysis* a partir de genes com correlação significativa entre metilação e expressão nas análises dos dados deste estudo (Interno) e do TCGA.

Vias Canônicas IPA	Genes	p-valor	p-valor corrigido	z-score
<i>Role of NFAT in Cardiac Hypertrophy</i>	<i>HDAC9,HDAC4,IGF1,ADCY5,FGFR1,GNG2,CACNA2D1,CACNA1H,CACNA1C,MEF2C,ITPR1,SLC8A1,PRKCB</i>	1E-07	2E-05	-2,5
<i>GPCR-Mediated Nutrient Sensing in Enteroendocrine Cells</i>	<i>ADCY5,GNG2,CACNA2D1,CACNA1H,CACNA1C,ITPR1,PRKCB</i>	7E-05	3E-03	-2,6
<i>α-Adrenergic Signaling</i>	<i>ADCY5,GNG2,ITPR1,SLC8A1,ADRA1A,PRKCB</i>	1E-04	4E-03	-2,2
<i>Cardiac Hypertrophy Signaling</i>	<i>HAND2,IGF1,ADCY5,FGFR1,GNG2,GNAO1,CACNA1C,MEF2C,ADRA1A</i>	3E-04	6E-03	-2,6
<i>Gβγ Signaling</i>	<i>GNG2,CACNA2D1,GNAO1,CACNA1H,CACNA1C,PRKCB</i>	6E-04	1E-02	-2,4
<i>GNRH Signaling</i>	<i>ADCY5,GNG2,CACNA2D1,CACNA1H,CACNA1C,ITPR1,PRKCB</i>	8E-04	1E-02	-2,0
<i>Synaptic Long Term Depression</i>	<i>IGF1,CACNA2D1,GNAO1,CACNA1H,CACNA1C,ITPR1,PRKCB</i>	1E-03	1E-02	-2,6
<i>Androgen Signaling</i>	<i>GNG2,CACNA2D1,GNAO1,CACNA1H,CACNA1C,PRKCB</i>	2E-03	2E-02	-2,0
<i>Cholecystokinin/Gastrin-mediated Signaling</i>	<i>SST,MEF2C,EPHA4,ITPR1,PRKCB</i>	2E-03	2E-02	-2,2
<i>Gaq Signaling</i>	<i>FGFR1,GNG2,ITPR1,ARHGEF25,ADRA1A,PRKCB</i>	4E-03	3E-02	-2,4
<i>Phospholipase C Signaling</i>	<i>HDAC9,HDAC4,ADCY5,GNG2,MEF2C,ITPR1,PRKCB</i>	7E-03	4E-02	-2,2

4.2 MÓDULO 2 - AVALIAÇÃO DAS ALTERAÇÕES ENCONTRADAS EM CÂNCER DE RETO DE ACORDO COM RESPOSTA AO TRATAMENTO NEOADJUVANTE

4.2.1 Análise de Alterações Genômicas

A cobertura do sequenciamento das amostras está representada no Anexo 2. O resultado das análises de alterações no número de cópias genômicas, dos escores de instabilidade genômica e de deficiência da via de reparo por recombinação homóloga, assim como os resultados de sequenciamento de alto desempenho encontram-se detalhados no Anexo 3 (manuscrito submetido para publicação no periódico *International Journal of Cancer*).

4.2.2 Análise dos dados de metilação

4.2.2.1 Análise do perfil de metilação de câncer de reto de acordo com a resposta ao tratamento neoadjuvante

A análise de agrupamento não supervisionado dos 32 tumores usando as sondas com maior variação entre os casos (desvio padrão > 0,02) permitiu a observação de três grupos distintos (Figura 7). Interessantemente, os agrupamentos 1 e 2 estavam enriquecidos com casos de pacientes não respondedores (cluster 1 e 2: 15/18 casos e cluster 3: 6/14).

A comparação entre o perfil de metilação dos tumores *versus* tecido normal de acordo com a resposta ao tratamento neoadjuvante foi realizada para identificar as particularidades de cada grupo. Os casos de CaRe disponíveis pelo TCGA não apresentam dados de resposta à terapia neoadjuvante e o único estudo em larga escala de metilação em CaRe relacionados à nCRT não possui dados de tecido

normal, por isso, não foi possível realizar uma comparação direta desses achados com dados externos.

Foram identificadas 24.428 e 75.398 sondas DM nos grupos de pacientes com pCR e pIR, respectivamente, em relação ao TN (*limma* FDR<5%; $|\Delta\beta| > 0,15$). Dessas, 20.958 sondas estavam DM nos dois grupos, enquanto 3.470 sondas estavam DM somente nos casos com pCR quando comparadas ao TN e, 54.440 somente em pIR (Figura 8). As sondas DM exclusivas de cada comparação foram selecionadas para explorar as peculiaridades de cada grupo. Os tumores de pacientes com pCR apresentaram uma maior porcentagem de sondas hipermetiladas (44% vs. 18% em pIR), que estavam principalmente localizadas em regiões promotoras (49% vs. 37% em pIR) e de ilhas CpG (66% vs. 36% em pIR), enquanto a maioria das sondas DM em pIR (82%) encontrava-se hipometilada.

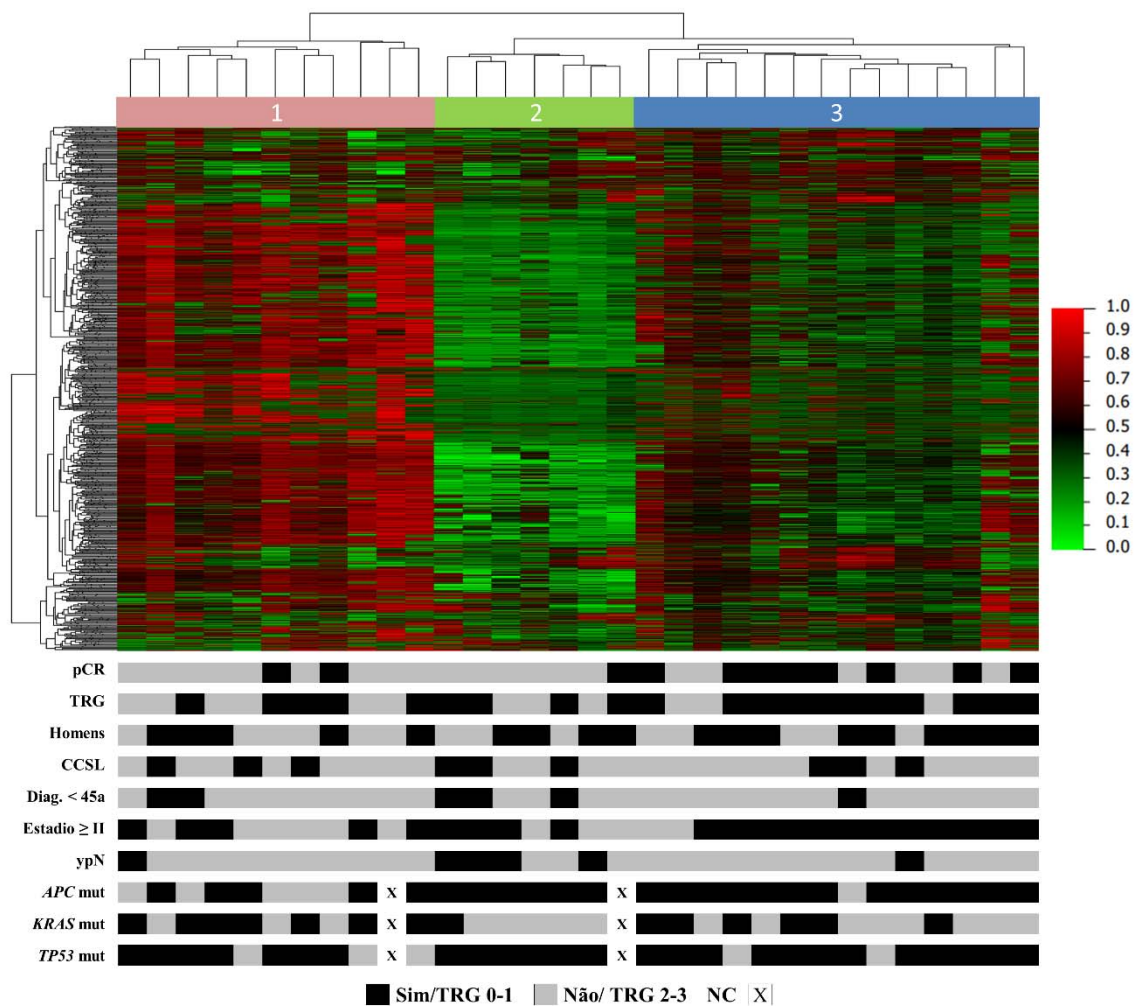


Figura 7 - *Heatmap* demonstrativo das sondas cujos valores de β apresentaram desvio padrão $> 0,02$ entre as 32 amostras de câncer de reto. Os quadros abaixo do heatmap representam as características clínicas e moleculares utilizadas para comparar os agrupamentos 1, 2 e 3 (barra superior rosa, verde e azul, respectivamente). Observa-se um enriquecimento de casos com resposta patológica incompleta nos clusters 1 e 2. **pCR**: resposta patológica completa; **CCSL**: critérios clínicos para síndrome de Lynch; **Diag. < 45a**: diagnóstico antes dos 45 anos de idade; **ypN**: status linfonodal patológico após tratamento; **Mut**: mutado; **NC**: não consta.

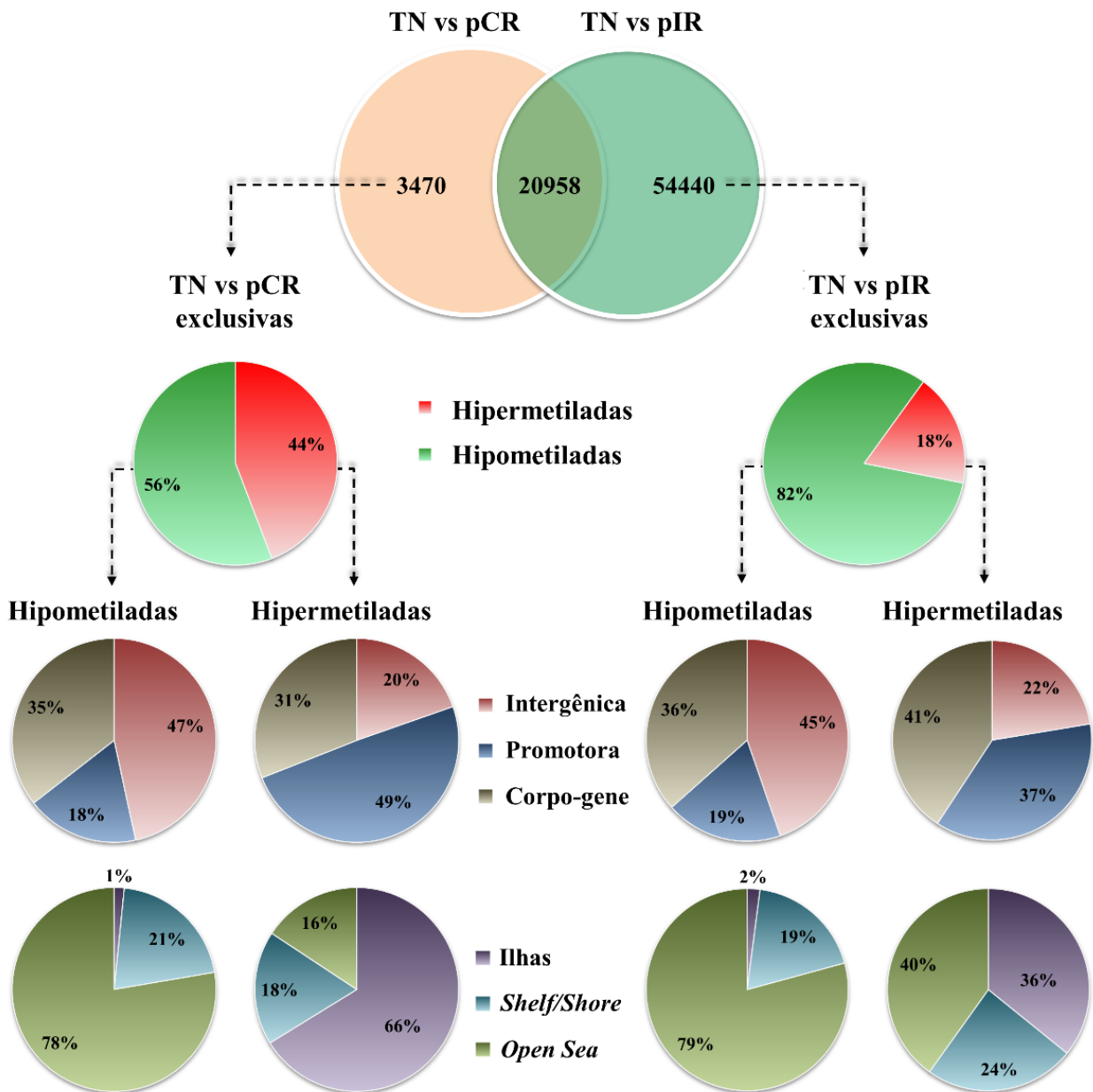


Figura 8 - Representação esquemática da distribuição genômica do contexto CpG das sondas diferencialmente metiladas encontradas na comparação entre o tecido normal (TN) e resposta ao tratamento (pCR: resposta completa ou pIR: incompleta).

As sondas diferencialmente metiladas em cada comparação (3.470 sondas em pCR vs. TN e 54.440 sondas em pIR vs. TN) foram avaliadas em relação às sondas DM na comparação entre pCR vs. pIR (*limma* $P < 0,05$; $|\Delta\beta| > 0,15$). As sondas alteradas exclusivas em cada comparação com TN, mas também DM entre os dois grupos tumorais (37 sondas exclusivas na comparação pCR vs. NT; e 157 pIR vs. NT) foram consideradas para desenhar um modelo de predição de resposta à nCRT. As 194 sondas resultantes foram posteriormente filtradas (foram selecionadas as sondas com $AUC > 0,75$ e regressão linear $< 0,05$) para identificar candidatos mais informativos e não redundantes, resultando em três sondas (Figura 9A).

As três sondas resultantes dessa análise (cg13770628, cg01072658, cg03085846) foram submetidas à *Diagonal Linear Discriminating Analysis* (DLDA) para treinar um classificador preditivo de resposta à nCRT. Foi possível discriminar corretamente os tumores de pacientes com pCR daqueles com pIR com 100% de sensibilidade e 90% de especificidade após validação cruzada usando o modelo *leave-one-out cross validation* (LOOCV) (Figura 9B). As cinco amostras normais (TN) foram classificadas como pCR. Infelizmente, somente uma das sondas usadas no classificador está presente na plataforma 450K (Illumina) impedindo a validação do mesmo com os dados do TCGA.

CLASSIFICADOR DE RESPOSTA BASEADO EM METILAÇÃO

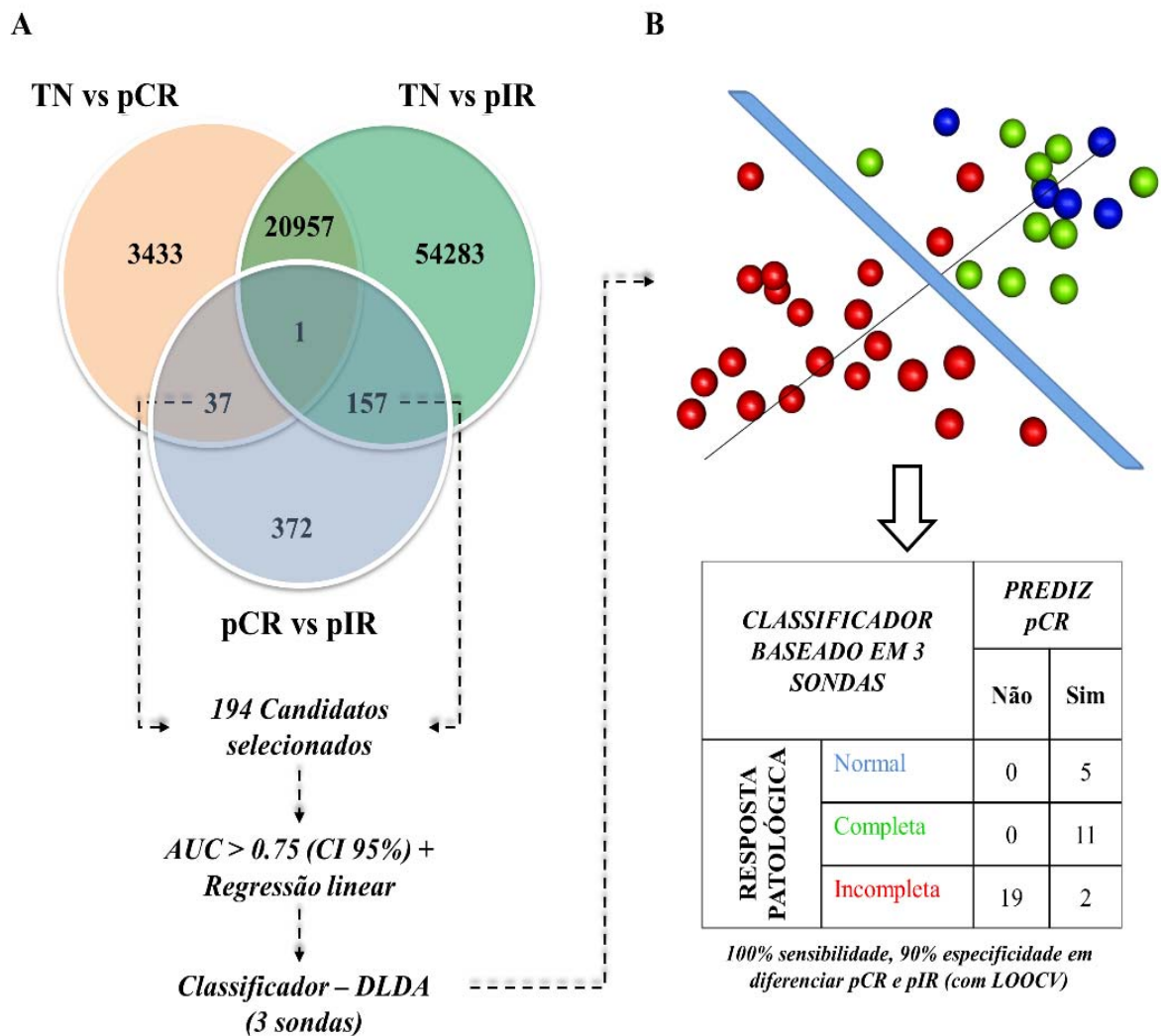


Figura 9 - Etapas utilizadas para a construção de um classificador preditivo de resposta à terapia neoadjuvante em câncer de reto. **A:** Diagrama de Venn mostrando a sobreposição das sondas identificadas nas comparações dos grupos de tumores (pCR vs. TN e pIR vs. TN) (*limma* FDR<5%; $|\Delta\beta| > 0,15$) e entre si (pCR vs. pIR, *limma* P<0,05; $|\Delta\beta| > 0,15$) e os parâmetros utilizados para construir o classificador. **B:** Distribuição tridimensional das amostras de acordo com o valor das sondas contidas no classificador e sua performance.

4.2.3 Análise de expressão global de transcritos

4.2.3.1 Análise do perfil de expressão de transcritos em câncer de reto de acordo com a resposta ao tratamento neoadjuvante

A análise não supervisionada de agrupamento hierárquico dos transcritos com variação ($DP > 0,02$) entre as 28 amostras de CaRe analisadas resultou em dois grupos não muito distintos. Não foi possível observar diferenças entre o grupo de pacientes de acordo com as características clínicas e patológicas incluindo resposta ao tratamento (Figura 10). Dessa forma, foram feitas comparações de cada grupo de CaRe (pCR e pIR) com o TN para verificar se as alterações em cada grupo estavam relacionadas a vias biológicas distintas.

Foram identificados 446 transcritos DE nos tumores dos pacientes com pCR e 913 naqueles com pIR (FDR < 5%). Dentre os transcritos DE em pCR, 378 eram codificadores (106 $FC > 1,5$ e 272 $FC < -1,5$) e 68 não-codificadores (30 $FC > 1,5$ e 38 $FC < -1,5$). Enquanto dos transcritos DE em pIR comparados ao TN, 678 eram codificadores (358 $FC > 1,5$ e 320 $FC < -1,5$) e 235 não-codificadores (149 $FC > 1,5$ e 86 $FC < -1,5$).

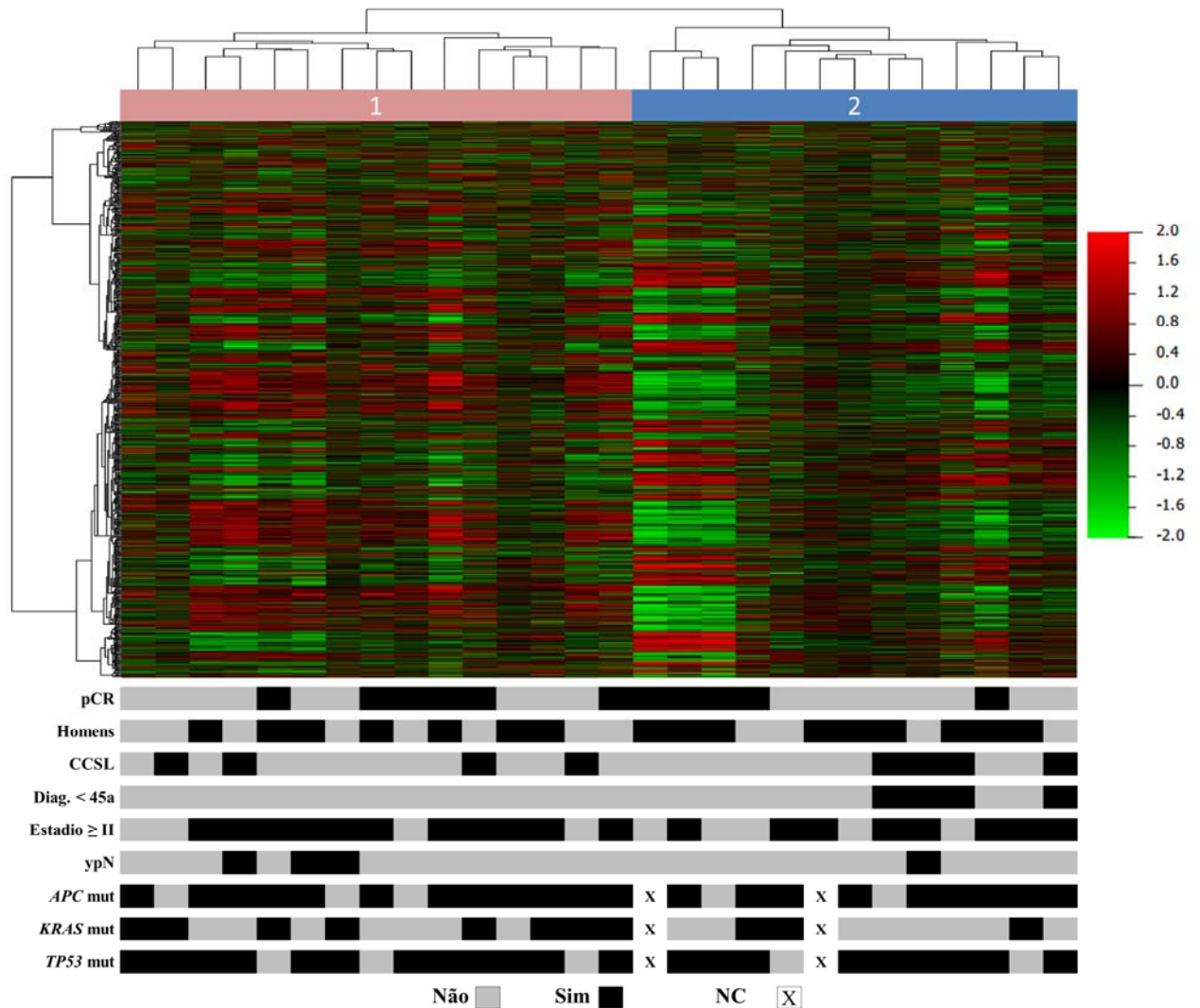


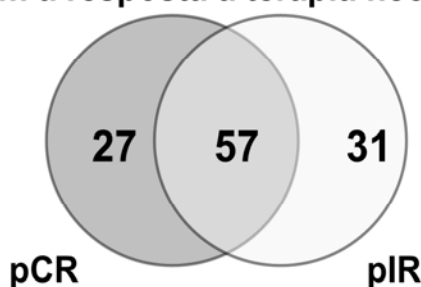
Figura 10 - Análise não supervisionada de carcinomas de reto baseada nos valores de expressão de sondas com desvio padrão $> 0,02$. Abaixo do heatmap estão representadas as características clínicas e moleculares de cada paciente. Observa-se a formação de dois grupos (1 e 2: barra superior rosa e azul, respectivamente), sem diferenças na comparação com as características avaliadas. **pCR**: resposta patológica completa; **CCSL**: critérios clínicos para síndrome de Lynch; **Diag < 45a**: diagnóstico antes dos 45 anos de idade; **ypN**: status linfonodal patológico após tratamento; **Mut**: mutado; **NC**: não consta.

Os dados de expressão dos transcritos codificadores obtidos em cada comparação entre os grupos de tumores (pCR e pIR) com o tecido de reto normal (NT) foram utilizados para avaliar as vias desreguladas em cada grupo e, dessa forma, melhor compreender as diferenças entre os tumores de pacientes com resposta completa à nCRT e os que não atingiram pCR (Figura 11).

Os 378 e 678 genes diferencialmente expressos em pCR e pIR em relação ao TN, respectivamente (FDR < 5%, |FC| > 1,5), foram analisados usando o programa IPA identificando-se 84 vias canônicas desreguladas no primeiro grupo e 88 no segundo (p<0,05). Vinte e sete vias estavam alteradas exclusivamente em pCR, enquanto 31 vias estavam alteradas somente em pIR (Figura 11).

As principais vias desreguladas somente em pCR incluem as relacionadas ao sistema imune (*Crosstalk between Dendritic Cells and Natural Killer Cells*, *Mechanisms of Viral Exit from Host Cells*, *Altered T Cell and B Cell Signaling in Rheumatoid Arthritis*), à via de sinalização de *Wnt* (*Wnt/β-catenin Signaling*, *PCP pathway*), ao metabolismo celular (*GDP-glucose Biosynthesis*), à adesão celular (*Epithelial Adherens Junction Signaling*, *Cdc42 Signaling*). Já as vias desreguladas somente em pIR abrangem funções como a regulação do ciclo celular (*Cyclins and Cell Cycle Regulation*, *Cell Cycle: G2/M DNA Damage Checkpoint Regulation*, *Cell Cycle: G1/S Checkpoint Regulation*), do reparo a danos no DNA (*DNA Double-Strand Break Repair by Homologous Recombination*, *DNA Double-Strand Break Repair by Non-Homologous End Joining*), da transcrição gênica (*Sirtuin Signaling Pathway*, *Transcriptional Regulatory Network in Embryonic Stem Cells*, *DNA Methylation and Transcriptional Repression Signaling*) e do processo de tradução de proteínas (*EIF2 Signaling*, *Regulation of eIF4 and p70S6K Signaling*) (Figura 11).

Vias desreguladas em Câncer de Reto de acordo com a resposta à terapia neoadjuvante



Crosstalk between Dendritic Cells and Natural Killer Cells

(HLA-DRB1, ACTA2, LTB, TLN1, ACTG2, TNF, CAMK2G)

Epithelial Adherens Junction Signaling

(MYL9, ACTA2, MYL6, SORBS1, TGFB3, ACTG2, MYH11, VCL, ACVR1B)

FAK Signaling

(ACTA2, ITGA5, TLN1, ACTG2, VCL, TNS1)

Wnt/ β -catenin Signaling

(SOX4, SFRP2, TGFB3, GNAO1, ACVR1B, LRP1, FZD7, WNT5A)

GDP-glucose Biosynthesis

(HK2, PGM5)

cAMP-mediated signaling

(AKAP12, RGS2, P2RY14, PDE9A, DUSP1, PDE3A, GNAO1, PRKAR1A, CAMK2G)

Glucose and Glucose-1-phosphate Degradation

(HK2, PGM5)

Altered T Cell and B Cell Signaling in Rheumatoid Arthritis

(HLA-DRB1, CCL21, LTB, IL23A, TNF)

PCP pathway

(CTHRC1, JUN, FZD7, WNT5A)

Circadian Rhythm Signaling

(PER1, CLOCK, VIP)

Cdc42 Signaling

(MYL9, B2M, FOS, HLA-DRB1, MYL6, ITGA5, MYLK)

Adipogenesis pathway

(RB1, TXNIP, CLOCK, TNF, FZD7, WNT5A)

nNOS Signaling in Skeletal Muscle Cells

(RYS3, DMD)

Eicosanoid Signaling

(PTGIS, PTGS1, HPGDS, PLA2G2A)

Mechanisms of Viral Exit from Host Cells

(ACTA2, XPO1, ACTG2)

VEGF Signaling

(ROCK2, VEGFA, ACTA2, FOXO3, EIF2S3, ACTG2, HIF1A, EIF2S2, EIF1AX, ATM)

Sirtuin Signaling Pathway

(PRKDC, CXCL8, GADD45B, SLC2A1, GLS, TOMM20, PCK1, HIF1A, HMGS2, SOD2, FOXO3, TOMM34, CLOCK, HIST1H4J, NAMPT, SLC25A5, NOS2, TNF, NDUFB2)

Regulation of eIF4 and p70S6K Signaling

(RPS15, RPS13, ITGA2, ITGA5, EIF2S3, RPS2, EIF3E, RPS4X, EIF2S2, EIF1AX, ATM, EIF3M)

Clathrin-mediated Endocytosis Signaling

(MYO6, LPA, USP9X, ITGA5, VEGFA, UBD, HSPA8, MET, ACTA2, TFRC, ACTG2, CLU, APOD, ATM)

Protein Ubiquitination Pathway

(CRYAB, DNAJC9, HSPH1, USP9X, HSPB8, DNAJC3, DNAJC10, DNAJC2, HSPD1, DNAJA1, UCHL3, UBD, HSPA8, USP49, HSPB6, UBE2C, HSPB1)

EIF2 Signaling

(EIF2S3, EIF3E, RPS4X, EIF2S2, EIF3M, VEGFA, RPS15, ACTA2, RPS13, RPL21, RPL37, ACTG2, RPS2, EIF1AX, ATM)

Cell Cycle: G2/M DNA Damage Checkpoint Regulation

(PRKDC, CKS2, TOP2A, ABL1, AURKA, ATM)

Cell Cycle: G1/S Checkpoint Regulation

(RB1, TFDP1, HDAC2, CDK6, ABL1, GNL3, ATM)

Aryl Hydrocarbon Receptor Signaling

(RB1, CCNA2, FOS, TFDP1, GSTM3, CDK6, IL1B, TNF, PTGES3, ATM, HSPB1)

Transcriptional Regulatory Network in Embryonic Stem Cells

(SET, HIST1H4K, HIST1H4D, HAND1, HIST1H4J, SKIL)

DNA Double-Strand Break Repair by Non-Homologous End Joining

(PRKDC, DCLRE1C, ATM)

DNA Double-Strand Break Repair by Homologous Recombination

(ABL1, BRCA2, ATM)

Cyclins and Cell Cycle Regulation

(RB1, CCNA2, TFDP1, HDAC2, CDK6, ABL1, ATM)

Renal Cell Carcinoma Signaling

(MET, UBD, VEGFA, FOS, SLC2A1, HIF1A, ATM)

DNA Methylation and Transcriptional Repression Signaling

(HDAC2, HIST1H4K, HIST1H4D, HIST1H4J)

Figura 11 - Vias desreguladas exclusivamente em um dos dois grupos de câncer de reto de acordo com a resposta à quimiorradioterapia neoadjuvante. As 15 vias alteradas com maior significância estão descritas para cada grupo: resposta patológica completa (pCR) e incompleta (pIR).

4.2.3.2 Análise do secretoma de câncer de reto de acordo com a resposta ao tratamento neoadjuvante

Para identificar proteínas potencialmente secretadas pelos tumores de reto dos pacientes com pCR e pIR com potencial para serem utilizadas como marcadores detectáveis em biópsia líquida, foram utilizados os 106 genes que apresentaram aumento de expressão (FDR < 5%, FC > 1,5) em pCR e os 358 em pIR. Todos os genes do grupo de pCR e 353 genes de pIR foram associados à proteínas identificadas usando UniProtKB (<http://www.uniprot.org>) cujas sequências foram utilizadas para predição *in silico* - cinco genes foram associados a dois UniProt IDs (*HIST1H2BI* e *HIST1H2BG* ao ID: P62805; *HIST1H4D*, *HIST1H4J* e *HIST1H4K* ao ID: P62805) e dois genes codificam proteínas com mais de 4,000 aminoácidos e não puderam ser avaliados.

A análise usando as três ferramentas de predição SecretomeP 2.0, SignalP 4.1 e TargetP 1.1 (CBS servers: <http://www.cbs.dtu.dk/services/>) resultou em 40 proteínas potencialmente secretadas em pCR e 104 proteínas em pIR (Anexo 4). A busca em bancos de dados por essas proteínas ou respectivos mRNAs descritos em vesículas extracelulares (VE) provenientes de amostras de carcinomas colorretais (linhagens celulares, urina, saliva, plasma, tumores ou lavagem abdominal) (Vesiclepedia: <http://www.microvesicles.org/> ou Exocarta: <http://www.exocarta.org/>) e em plasma (Plasma Proteome: <http://plasmaproteomedatabase.org>) resultou em 17 proteínas em pCR e 56 em pIR (Figura 12 e Anexo 5).

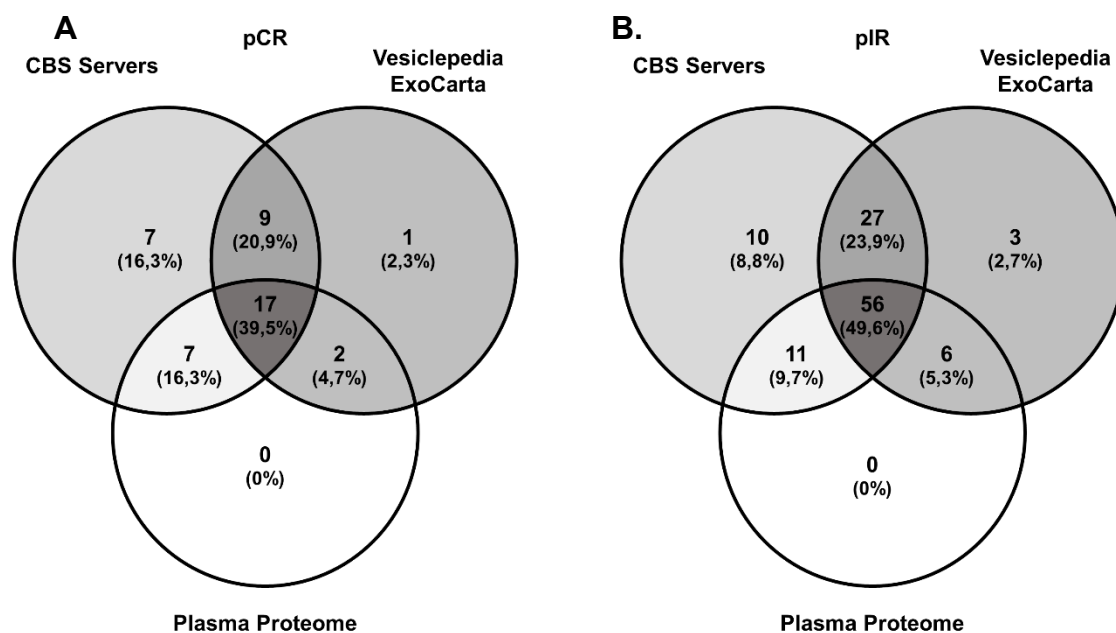


Figura 12 - Diagrama de Venn representando o número de proteínas potencialmente secretadas por tumores de reto. **A.** Pacientes que apresentaram resposta patológica completa (pCR) ou **B.** incompleta (pIR) identificadas por análise *in silico* (CBS servers). Também está representado o conjunto dessas proteínas descritas em bancos de dados de vesículas extracelulares (proteínas ou mRNAs de amostras de câncer colorretal em Vesiclepedia ou ExoCarta) e de plasma (Plasma Proteome).

Doze dessas proteínas são comuns aos dois grupos de amostras (pCR e pIR) e não foram consideradas potenciais marcadores de resposta ao tratamento (Anexo 5). Também foi realizada uma busca nos dados do *Cancer Secretome* (<http://www.cancersecretome.org/>) sendo identificadas uma proteína da lista de pCR e 20 da lista de pIR. Dessa forma, esse banco de dados foi usado como filtro apenas para os pIR. Ao final, foram identificados cinco potenciais marcadores secretados em pCR e 14 em pIR (Tabela 6).

Tabela 6 - Proteínas potencialmente secretadas por tumores de reto de acordo com a resposta ao tratamento neoadjuvante.

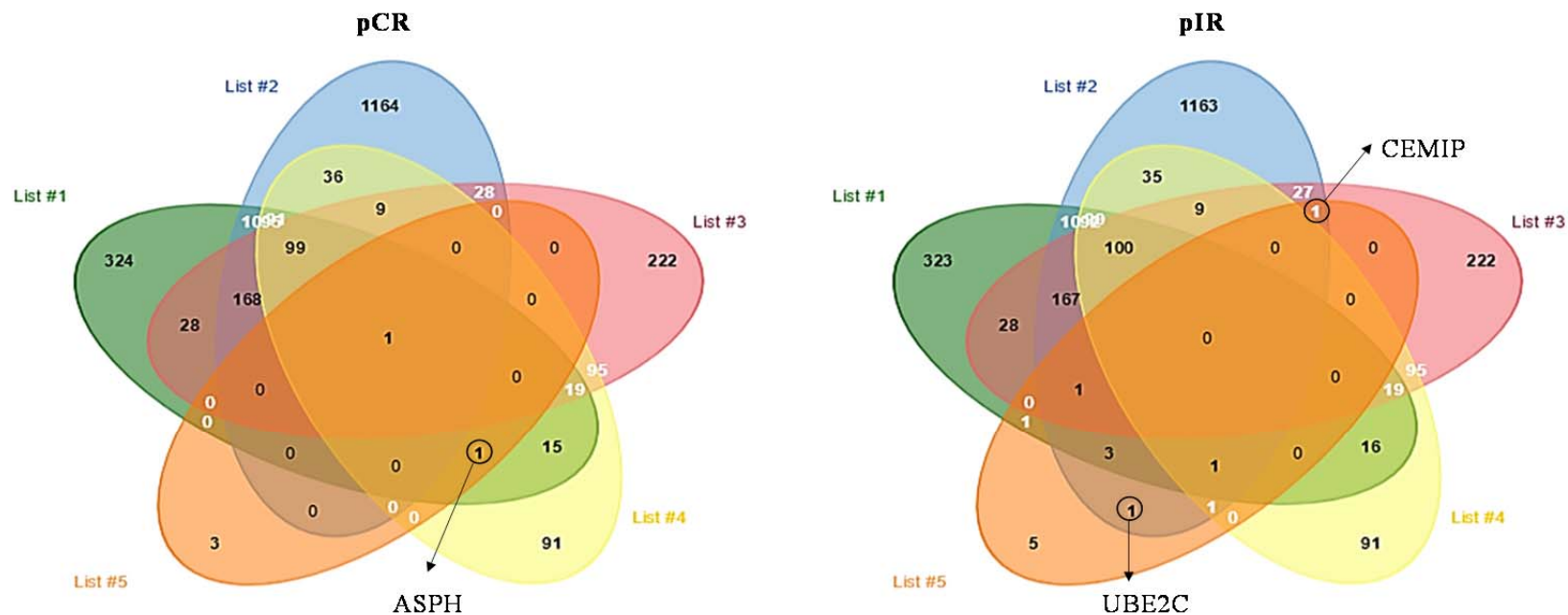
Gene	Nome da Proteína	UniProt	FC	FDR
Resposta Patológica Completa				
ASPH	<i>Aspartyl/asparaginyl beta-hydroxylase</i>	Q12797	1,70	0,039
B2MG	<i>Beta-2-microglobulin</i>	P61769	1,94	0,031
ERG1	<i>Squalene monooxygenase</i>	Q14534	1,54	0,046
UGGG2	<i>UDP-glucose:glycoprotein glucosyltransferase 2 (UGT2)</i>	Q9NYU1	1,56	0,027
WNT5A	<i>Protein Wnt-5a</i>	P41221	1,58	0,023
Resposta Patológica Incompleta				
AREG	<i>Amphiregulin (AR) (Colorectum cell-derived growth factor) (CRDGF)</i>	P15514	2,44	0,004
BACE2	<i>Beta-secretase 2</i>	Q9Y5Z0	1,54	0,015
CD44	<i>CD44 antigen (Extracellular matrix receptor III) (ECMR-III)</i>	P16070	2,00	0,005
CD47	<i>Leukocyte surface antigen CD47</i>	Q08722	1,75	1,09E-04
CEMIP	<i>Cell migration-inducing and hyaluronan-binding protein</i>	Q8WUJ3	2,26	0,002
CXCL3	<i>C-X-C motif chemokine 3</i>	P19876	1,83	0,002
CXCL5	<i>C-X-C motif chemokine 5</i>	P42830	2,25	0,026
DPEP1	<i>Dipeptidase 1</i>	P16444	1,52	0,039
GDF15	<i>Growth/differentiation factor 15</i>	Q99988	2,17	0,001
LIF	<i>Leukemia inhibitory factor</i>	P15018	1,60	0,013
MET	<i>Hepatocyte growth factor receptor (Proto-oncogene c-Met) (Tyrosine-protein kinase Met)</i>	P08581	1,66	8,46E-05
PDCD5	<i>Programmed cell death protein 5</i>	O14737	1,59	1,36E-04
PHF6	<i>PHD finger protein 6</i>	Q8IWS0	1,51	0,002
UBE2C	<i>Ubiquitin-conjugating enzyme E2 C</i>	O00762	1,58	0,007

UniProt: Número de identificação das proteínas dados pelo UniProtKB (<http://www.uniprot.org>); **FC.:** *Fold Change*; **FDR.:** *False discovery rate*

A expressão dessas 19 proteínas foi avaliada utilizando o banco de dados EVpedia (KIM et al. 2015) em VE provenientes de sangue total de indivíduos saudáveis, de soro de pacientes com câncer colorretal e de duas linhagens celulares de câncer colorretal; uma sensível à 5-FU (HCC2998) e outra resistente

(SW620) (<http://colonatlas.org/>). Esses grupos continham 457, 669, 1841 e 2691 moléculas descritas, respectivamente. Não haviam dados provenientes de amostras de câncer de reto.

Entre as cinco proteínas avaliadas em pCR, uma estava presente nas VE de todos os grupos (B2MG), três não foram encontradas em nenhum outro grupo (EREG1, UGGG2 e WNT5A) e a proteína ASPH estava presente em pCR, no sangue de indivíduos saudáveis e nas VE da linhagem sensível à 5-FU. Dentre as 14 proteínas utilizadas na comparação com pIR, cinco não estavam presentes em nenhum outro grupo (AREG, CXCL3, CXCL5, LIF e PHF6), enquanto UBE2C foi identificada em pIR e na linhagem resistente (SW620). CEMIP foi encontrada em pIR, na linhagem resistente à 5-FU e no soro de pacientes com CCR (Figura 13).



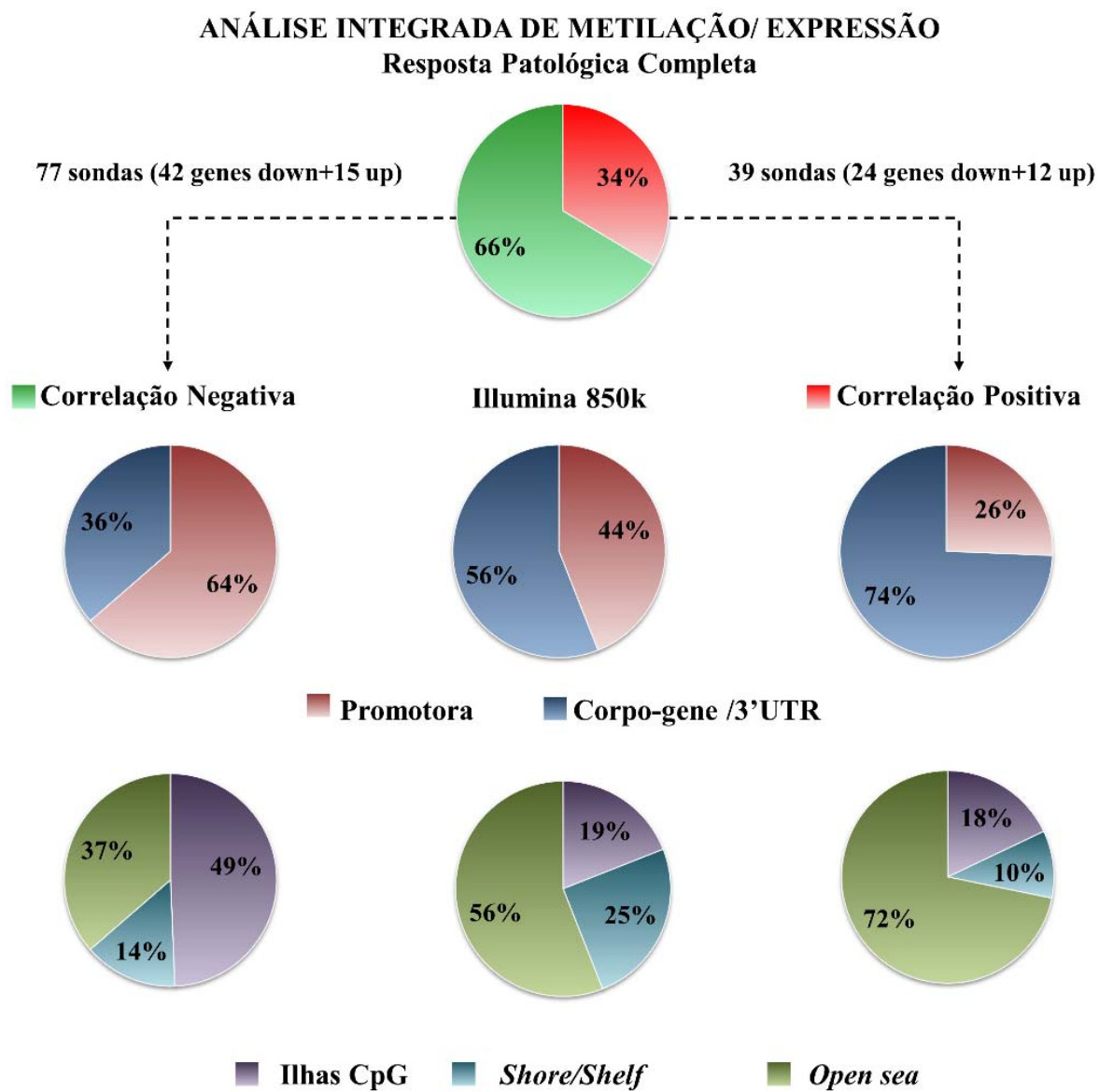
Listas	Amostra, grupo, número EVpedia	# de moléculas
List#1	Linhagem de câncer de colon (HCC 2998); Sensível à 5-FU; 278941040114	1841
List#2	Linhagem de câncer de colon (SW620); Resistente à 5-FU; 278941040119	2691
List#3	Soro; Câncer Colorretal; 289865850101	669
List#4	Sangue Total; Indivíduos saudáveis; 267750130105	457
List#5	Tumor, pCR ou pIR	5 ou 14

Figura 13 - Diagrama de Venn representativo da comparação dos potenciais marcadores secretados por tumores de reto com resposta patológica completa (pCR) e incompleta (pIR) ao tratamento neoadjuvante (List#5). Estão listadas as proteínas descritas em vesículas extracelulares provenientes de sangue total de indivíduos saudáveis (List#4), soro de pacientes com câncer colorretal (List#3) e duas linhagens celulares de câncer colorretal sendo uma sensível (List#1) e outra resistente à 5-fluorouracil (List#2).

4.2.3.3 Análise integrada de metilação e expressão em câncer de reto de acordo com a resposta ao tratamento neoadjuvante

As sondas DM em cada grupo pacientes de acordo com a resposta ao tratamento (pCR e pIR) em relação ao tecido normal e que apresentaram correlação significativa ($p < 0,05$) com os dados de expressão gênica estão descritas na Figura 14. Foram identificadas 116 sondas mapeadas em 93 genes nos casos com pCR, a maioria com r- (66%). O oposto ao identificado em pIR onde 72% das 2824 sondas DM apresentaram r+. Em geral, as sondas com r- encontravam-se, principalmente, em região promotora nos dois grupos (64% em pCR e 53% em pIR) e aquelas com r+, estavam em regiões de corpo de gene ou 3'UTR (74% em pCR e 76% em pIR). A distribuição das sondas em relação às ilhas CpG, no entanto, foi diferente nos dois grupos. Enquanto 49% das sondas com r- e 18% com r+ estavam localizadas em ilhas CpG nos casos com pCR, somente 24% daquelas com r- e 6% com r+ localizavam-se nessa região em não respondedores. Os dez genes cujas sondas mapeadas em região promotora apresentaram maior correlação negativa em relação à expressão estão representados na Tabela 6.

A



B

ANÁLISE INTEGRADA DE METILAÇÃO/ EXPRESSÃO
Resposta Patológica Incompleta

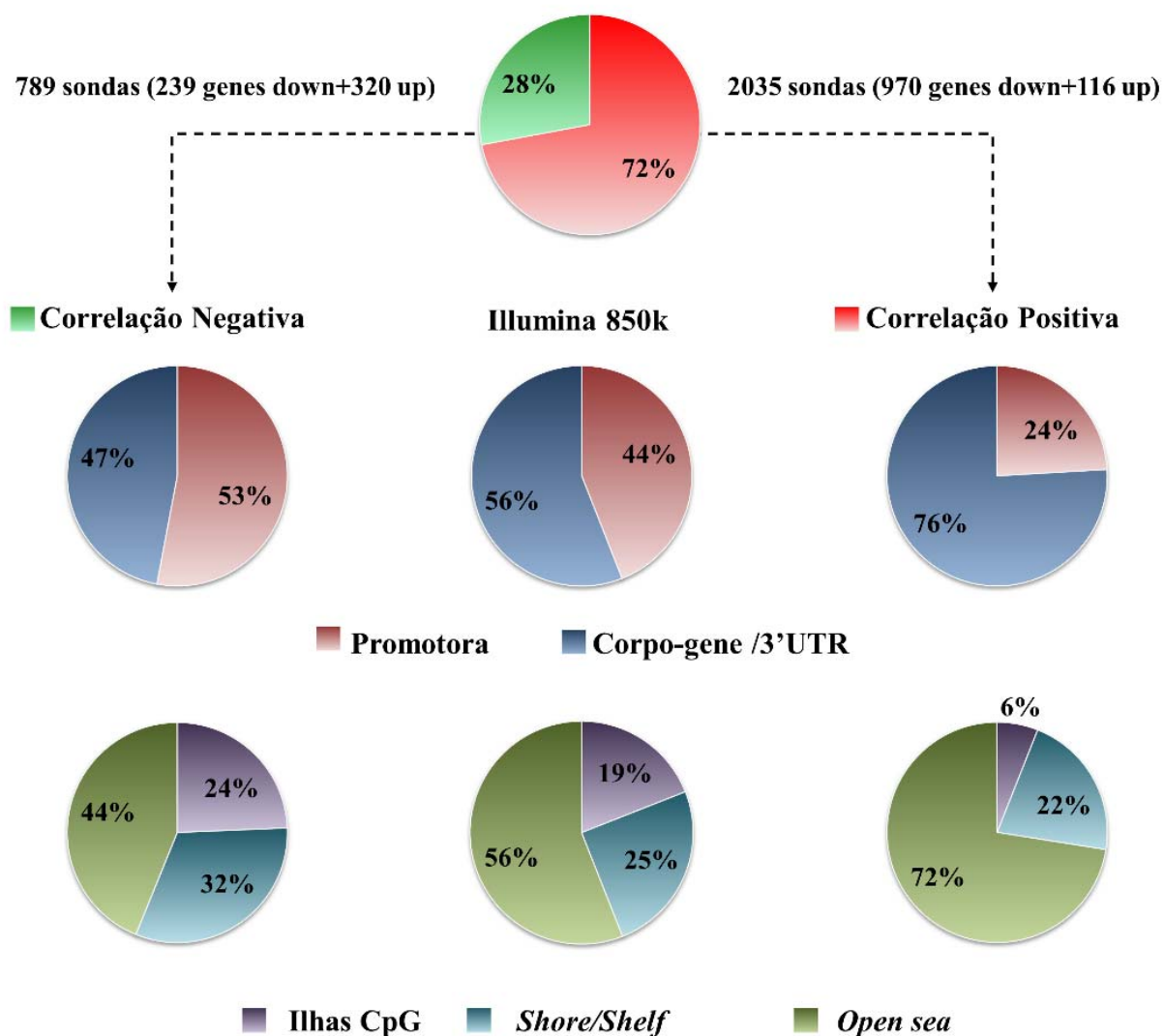


Figura 14 - Descrição das sondas diferencialmente metiladas com correlação significativa com os dados de expressão gênica identificadas em tumores de reto de pacientes com resposta patológica completa. (A) ou incompleta (B) à terapia neoadjuvante. **Down** e **up** refere-se a genes com expressão diminuída e aumentada em relação ao tecido normal, respectivamente.

Tabela 7 - Dez genes com correlação negativa entre os resultados de metilação e expressão em tumores de reto de pacientes com resposta completa ou incompleta à terapia neoadjuvante.

Gene	N° sondas	Metilação	Expressão	r	p-valor
Resposta Patológica Completa					
<i>TBX20</i>	3	Hiper	Down	-0.66	1.77E-04
<i>ZNF85</i>	1	Hiper	Down	-0.65	2.46E-04
<i>ZNF75A</i>	10	Hiper	Down	-0.63	4.22E-04
<i>ZNF518B</i>	6	Hiper	Down	-0.63	4.30E-04
<i>ST3GAL4</i>	2	Hiper	Down	-0.56	2.19E-03
<i>PTPRF</i>	1	Hiper	Down	-0.56	2.26E-03
<i>ZNF730</i>	1	Hiper	Down	-0.56	2.30E-03
<i>MIR1197</i>	1	Hipo	Up	-0.54	3.46E-03
<i>ZNF608</i>	1	Hiper	Down	-0.54	3.55E-03
<i>ITIH5</i>	2	Hiper	Down	-0.46	1.55E-02
Resposta Patológica Incompleta					
<i>TRIM29</i>	4	Hipo	Up	-0.74	9.50E-06
<i>UNC5CL</i>	1	Hipo	Up	-0.71	2.78E-05
<i>KIAA0226L</i>	1	Hipo	Up	-0.69	5.86E-05
<i>CLIC6</i>	5	Hiper	Down	-0.68	9.28E-05
<i>INPP5D</i>	9	Hipo	Up	-0.68	9.99E-05
<i>FAM155A</i>	4	Hiper	Down	-0.66	1.69E-04
<i>TUSC3</i>	3	Hiper	Down	-0.65	2.11E-04
<i>LTBP2</i>	1	Hiper	Down	-0.65	2.68E-04
<i>GSTA4</i>	2	Hiper	Down	-0.64	3.15E-04
<i>SMAD9</i>	1	Hiper	Down	-0.61	8.02E-04

N° sondas: número de sondas diferencialmente metiladas mapeadas no gene. **Metil.** status de metilação das sondas: **Hiper**metiladas ou **Hipo**metiladas. Expressão: **Down** e **up** refere-se a genes com expressão diminuída e aumentada em relação ao tecido normal, respectivamente. **r.** correlação de Pearson.

5 DISCUSSÃO

5.1 MÓDULO 1 - AVALIAÇÃO DAS ALTERAÇÕES ENCONTRADAS EM CÂNCER DE RETO: COMPARAÇÃO ENTRE AMOSTRAS TUMORAIS E NORMAIS

A metilação do DNA é um mecanismo essencial no controle de expressão gênica, regulando a acessibilidade à cromatina e a transcrição. Células tumorais, frequentemente apresentam hipometilação global, a qual leva a instabilidade genômica e, hipermetilação de ilhas CpG específicas em regiões promotoras, a qual promove o silenciamento de diversos genes, incluindo supressores tumorais (ISSA 2004).

No intuito de compreender melhor a biologia dos tumores de reto foi realizada, primeiramente, uma análise das sondas diferencialmente metiladas (DM) em relação ao tecido normal. Com o objetivo de confirmar se os achados encontrados nesse estudo eram reprodutíveis, foi feita comparação com os dados de metilação depositados pelo TCGA. Todos os casos disponíveis para a plataforma 450K (Illumina) foram utilizados, (sendo 11, 32, 38 e 14 casos com estágio I, II, III e IV, respectivamente e 4 casos sem informação). Ainda que utilizando plataformas diferentes e que nossos casos englobassem somente tumores em estádios II e III, 81,2% das sondas DM presentes em ambas as plataformas foram validadas com os dados externos. De forma geral, os tumores de reto apresentaram hipometilação

acentuada, principalmente em regiões intergênicas e de *Open Sea*, enquanto as sondas hipermetiladas encontraram-se principalmente em regiões promotoras e de ilhas CpG.

WEI et al. (2016) realizaram análise semelhante em doze amostras pareadas de seis casos de CaRe (estádios II e III) e TN usando a plataforma Illumina Infinium Methylation 450K. Os autores encontraram a maioria das sondas com hipometilação (56,6% das sondas DM) em regiões não cobertas por ilhas CpG (67,3%) e hipermetilação em regiões próximas (*Shore/Shelf*) ou de ilhas CpG (94,9%). Os autores selecionaram as CpGs localizadas em TSS200, TSS1500, 5'UTR ou no 1° exon para avaliar o padrão de metilação de 36 genes com mais de duas sondas DM em CaRe. No presente estudo, utilizamos a mesma estratégia de seleção descrita pelos autores para avaliar esses mesmos genes. Foi encontrado o mesmo padrão de metilação em relação ao TN em 31 genes, enquanto não foram identificadas sondas DM em 3 genes (*FRZB*, *KCNQ1*, *ZNF655*) e dois genes tinham padrões inversos de metilação (*KCNIP4* e *TNFRSF8* estavam hipo e hipermetilados em WEI et al. (2016), respectivamente; enquanto apresentaram $\Delta\beta^+$ e $\Delta\beta^-$ no presente estudo).

A mesma estratégia e plataforma empregadas por WEI et al. (2016) foram utilizadas por VYMETALKOVA et al. (2016) em 25 amostras de CaRe (estádios I, II, III e IV) pareadas com 25 tecidos normais adjacente. Os autores descreveram 5.929 CpGs DM, com predominância de hipometilação em regiões intergênicas e de *Open Sea* e hipermetilação em regiões de ilhas CpG. Somente as CpGs com maior hiper e hipometilação, assim como 33

genes identificados na análise de agrupamento hierárquico, foram comparadas com os dados do TCGA, observando-se 77% de concordância. A comparação entre os dados desse e do presente estudo revelaram que 94% dos 48 genes apresentaram o mesmo padrão de metilação, reforçando que o padrão alterado de metilação aqui descrito, representa de fato modificações encontradas nos tumores de reto. Os autores não abordaram o papel da metilação e resposta à nCRT em CaRe, uma vez que nem todos os pacientes incluídos no estudo foram submetidos a esse tratamento (VYMETALKOVA et al. 2016).

A análise global de expressão gênica em CaRe revelou 737 transcritos DE, dos quais 437 (59%) foram validados com dados externos (TCGA), ainda que usando diferentes técnicas (*microarray* de expressão no presente estudo e sequenciamento de RNA no TCGA) e que os dados não estivessem pareados para estadiamento. Todos os casos disponíveis pelo TCGA foram utilizados para que se obtivesse um maior poder estatístico para comparação. Uma meta-análise de três diferentes estudos (GEO datasets: GSE75548, GSE35982, GSE12225) comparando os resultados de expressão gênica de 65 CaRe e 42 tecidos normais adjacentes (TNA) identificou 4.091 mRNAs DE, que comparados aos dados do TCGA (READ: 95 CaRe e 10 TN) revelaram 2.118 genes (52%) em comum (SUN et al. 2017). Entre esses 2.118 genes, 204 também foram identificados no presente estudo, sendo que 97 não estavam disponíveis para comparação.

Os 437 genes DE identificados no presente estudo e validados com os dados do TCGA foram usados para identificar vias desreguladas em

CaRe. Destacaram-se entre os achados as alterações nas vias relacionadas ao controle do ciclo celular, as quais podem estar associadas com a promoção de instabilidade genômica por bloqueio do ponto de verificação G2/M de danos ao DNA (*G2/M DNA Damage Checkpoint Regulation*) e indução de proliferação por ativação de ciclinas e do ciclo celular (*Cyclins and Cell Cycle Regulation, Estrogen-mediated S-phase Entry, Mitotic Roles of Polo-Like Kinase*) (Tabela 4). Esses resultados corroboram a alta instabilidade genômica observada nestes tumores assim como o número significativo de casos com escores indicativos de falha nos mecanismos de reparo a danos no DNA por recombinação homóloga (Anexo 3).

A baixa sensibilidade às lesões no DNA pode ser também decorrente da desregulação da via das integrinas (*Integrin Signaling*), identificada como alterada nos carcinomas de reto (Tabela 4). As integrinas tem papel fundamental na comunicação com a matriz extracelular e participam da inibição do crescimento celular por meio da inibição por contato. A perda de contato célula-célula causa modificações da arquitetura celular que podem levar à diminuição da expressão da proteína p53, responsável pela indução de apoptose ou bloqueio da proliferação celular na presença de danos no DNA (BIANCONI et al. 2016). Além da via de sinalização das integrinas, outras duas vias foram identificadas como alteradas em CaRe (*ILK e Paxillin Signaling*) (Tabela 4). Essas vias estão relacionadas ao papel das integrinas na manutenção do citoesqueleto e contato com células vizinhas (ATTWELL et al. 2003).

Uma série de vias de sinalização interconectadas também estava alterada em CaRe (Tabela 4). Foi identificada a desregulação da via de sinalização de resposta à estímulos externos relacionada a receptores associados à proteína G (GPCR) em células enteroendócrinas (*GPCR-Mediated Nutrient Sensing in Enteroendocrine Cells*). Essas células agem como sensores aos componentes liberados no intestino, reconhecidos pelos GPCRs, resultando em sinais transmitidos a outros órgãos do organismo, como por exemplo a liberação de insulina e controle da glicose livre no sangue (MORAN et al. 2008). Os GPCRs são receptores transmembrânicos (CXCR4 é um dos GPCRs: *CXCR4 Signaling*) que controlam as vias de transdução de sinais intracelulares ativando a proteína G. Esta proteína regula uma série de eventos de sinalização à jusante, como a formação de AMP-cíclica (*cAMP-mediated signaling*), a qual ativa proteínas quinase A (*Protein Kinase A Signaling*) que ativam, por exemplo, fatores de transcrição (REIMANN et al. 2012; LAW et al. 2016; HUSTED et al. 2017). Outro evento é a ativação das fosfolipases C (PLC: *Phospholipase C Signaling*), que realizam a hidrólise de fosfatidilinositol-4,5-bifosfato (PIP₂), produzindo inositol-1, 4, 5-trifosfato (IP₃) e diacilglicerol (DAG). IP₃ liga-se a receptores no retículo endoplasmático ativando canais de cálcio (Ca^{+2} : *calcium signaling*) (YANG et al. 2013). A regulação dos níveis de cálcio citosólico foi correlacionada à morte celular por apoptose e autofagia (HARR e DISTELHORST 2010). Recentemente, RAYNAL et al (2016) relataram que a modulação da via de sinalização de cálcio pode reverter o silenciamento

epigenético de genes supressores de tumor em múltiplas linhagens celulares de câncer, incluindo o câncer de cólon.

Ao avaliar os dados brutos adquiridos pelo TC-GA-Assembler de 163 CaRe e 9 TN disponíveis pelo TCGA (READ), ZUO et al. (2016) identificaram 620 genes DE (389 com aumento e 231 com diminuição de expressão). Os processos biológicos associados a esses genes, identificados usando o *Gene Ontology* (GO), incluíam biogênese e organização de componentes celulares, processos metabólicos, celulares e do sistema imune. A partir dos genes DE, foi construída uma rede de genes com alta correlação de expressão ($r > 0,6$) composta de 71 genes e 253 interações. Esses genes estavam relacionados com contração endotelial, ribossomos e adesão focal. Três módulos foram extraídos da rede (contração muscular e organização do citoesqueleto; regulação da proliferação de células da glia e organização da matriz extracelular) e foi realizado o enriquecimento dos genes contidos em cada módulo usando o *Kyoto Encyclopedia of Genes and Genomes* (KEGG) (ZUO et al. 2016). Entre os 27 genes presentes nos três módulos, 13 também estavam DE no presente estudo. No presente estudo também foram identificadas vias desreguladas relacionadas à organização do citoesqueleto e adesão focal (*Integrin, Paxilin e ILK Signaling*; Tabela 4).

SUN et al. (2017) usaram somente dados de transcritos DE para realizar uma meta-análise usando três estudos prévios como grupo de treinamento e os dados do TCGA como teste (descrito na p. 95). Os autores utilizaram os 2.118 genes DE identificados para avaliar 66 dos 95 CaRe com

dados clínico-patológicos completos disponíveis pelo TCGA, resultando em 16 módulos cujos genes apresentaram expressão semelhante. Cada módulo foi comparado (correlação de Pearson) às características clínicas e patológicas (estadiamento patológico, pTNM, gênero, tipo histológico, idade ao diagnóstico, invasão linfática e sobrevida global) dos CaRe. O módulo com maior correlação significativa com estadiamento ($r = 0,38$; $p = 0,002$), apresentou 54 genes envolvidos em adesão celular (GO e KEGG) que foram posteriormente filtrados, resultando em cinco genes (*SCG3*, *SYP*, *CDK5R2*, *AP3B2* e *RUNDC3A*) capazes de diferenciar estádios I/II de III/IV (AUC = 0,744). Dentre esses cinco genes, o aumento de expressão de *SCG3* foi validado em um outro grupo de casos (GSE29621) como associado a estadiamento III/IV e pior prognóstico em CaRe. O módulo com associação à sobrevida global (*Hazard Ratio* = 13,534; $p = 0,014$) apresentou 41 genes relacionados ao câncer ($p = 0,046$) e à sinalização mediada por cálcio ($p = 0,039$; GO e KEGG). Dentre os genes filtrados nesse grupo, *HOMER2* foi o único validado no grupo independente (GSE29621) como associado a menor sobrevida global e pior prognóstico em CaRe. No presente estudo, foi identificado somente um (*IL23A*) dos 95 genes DE presentes nos dois módulos. Curiosamente, os dois genes (*SCG3* e *HOMER2*), cuja expressão aumentada foi associada ao pior prognóstico no relato de SUN et al. (2017), apresentaram expressão diminuída em CaRe comparado ao TN nos casos do TCGA.

A análise integrada dos dados de transcritos DE e sondas DM em CaRe comparado ao TN revelou uma maior proporção de sondas com

correlação positiva ($r+$) quando comparado aos dados do TCGA (68% vs. 47%, respectivamente). Esse achado é provavelmente decorrente da composição das diferentes plataformas usadas nas duas análises, uma vez que a plataforma EPIC 850K (Illumina) apresenta um acréscimo de mais de 400 mil sondas em relação a 450K (Illumina), das quais, 40% localizam-se em região de corpo de gene e 78,2% em *open sea* (MORAN et al. 2016). A metilação em corpo de gene tem sido descrita como um mecanismo epigenético que apresenta correlação positiva em relação à expressão gênica (YANG et al. 2014). Estes achados corroboram os resultados da análise integrada do presente estudo e dos dados do TCGA, dos quais, respectivamente, 80% e 77% das sondas com $r+$ estavam em corpo de gene (Figura 5). Como esperado, a maioria das sondas com correlação negativa estavam localizadas em região promotora de genes, nos dois grupos.

HUA et al. (2017) avaliaram os genes com metilação diferencial (93 CaRe e 7 TN) em ilhas CpG que também apresentaram alteração nos níveis de expressão (89 CaRe e 10 TN) disponíveis pelo TCGA (READ). Os autores reportaram 50 genes alterados, dos quais 11 exibiram aumento e 39, diminuição de expressão, sendo que todos apresentaram CpGs hipermetiladas. Dentre esses 50 genes, 12 foram identificados em nossa análise integrada, mas somente 7 apresentaram metilação diferencial em ilhas CpG: seis genes hipermetilados e com diminuição de expressão (*DCLK1*, *GNAO1*, *HAND2*, *LONRF2*, *NBEA*, *RELN*, *SLC8A1*, *SLIT3*, *TLL1*, *TMTC1* e *UNC5C*) e um gene com hipermetilação e aumento de expressão (*GRIN2D*). Um dos motivos pelos quais não foi encontrada uma elevada

concordância entre os dois estudos, pode ser a diferença nas etapas de pré-processamento dos dados de metilação. Os autores, por exemplo, não excluíram as sondas mapeadas nos cromossomos sexuais, de modo que 170 das 561 sondas DM descritas estavam mapeadas no cromossomo X e duas no cromossomo Y. Além disso, os autores obtiveram os dados de outro repositório ([https:// tcga-data.nci.nih.gov/tcga/](https://tcga-data.nci.nih.gov/tcga/), Oct 20, 2015) e utilizaram um número diferente de casos em relação ao nosso estudo. Os autores ainda descreveram as vias alteradas identificadas a partir dos 50 genes (KEGG), incluindo: *cAMP signaling pathway*, *circadian entrainment*, *glutamatergic synapse*, *nicotine addiction*, *amphetamine addiction*, *morphine addiction*, *melanogenesis*, *retrograde endocannabinoid signaling* e *cholinergic synapse*. A via de sinalização de cAMP também foi identificada em nosso estudo.

A análise funcional *in silico* dos genes identificados a partir da integração dos dados de metilação e expressão gênica dos dois grupos de casos aqui analisados (dados internos e TCGA) revelou 11 vias canônicas desreguladas. Quatro dessas vias (*GPCR-Mediated Nutrient Sensing in Enteroendocrine Cells*, *Phospholipase C Signaling*, *Role of NFAT in Cardiac Hypertrophy* e *Synaptic Long Term Depression*) já haviam sido identificadas na análise de expressão global de transcritos (CaRe vs. TN), confirmando sua relevância em CaRe assim como o mecanismo de regulação desses genes por meio da metilação. Além disso, também foram identificadas vias específicas ligadas diretamente (*Gβγ Signaling*, *Gαq Signaling*, *α-Adrenergic Signaling*) ou indiretamente (*GNRH Signaling* e *Androgen Signaling*) à sinalização dos GPCRs e das PLCs.

Receptores adrenérgicos também são GPCRs e estão envolvidos na resposta ao estresse (LOHSE 2015). Estudos reportaram a relação da ativação dessa via com o desenvolvimento do câncer de mama (OBEID e CONZEN 2013), enquanto o bloqueio por metilação da região promotora de *alpha-1B-adrenergic receptor* foi identificado em câncer gástrico (NODA et al. 2007). Os receptores GPCRs atuam na regulação de enzimas e canais de cálcio pela ativação das proteínas-G. Essas proteínas são compostas das subunidades α , β e γ , sendo G β e γ altamente ligadas e consideradas uma unidade. Quando ativada por GPCRs, há separação das duas subunidades (o dímero G $\beta\gamma$ e G α - G α_q , representa uma família de G α), ambas funcionais, as quais participam de diversas vias de sinalização intracelular, incluindo divisão celular e reorganização do citoesqueleto promovido pela ativação de quinases (SYROVATKINA et al. 2016).

A via relacionada a liberação do hormônio gonadotrofina (GNRH) é mediada pela ativação de proteínas G com ativação de PLC e subsequente ativação dos canais de cálcio. Além disso, a via cAMP-PKA pode ser ativada pela proteína G após estímulo de GNRH, culminando em ativação da expressão de mais receptores de GNRH (PERRETT e MCARDLE 2013). Androgênios também apresentam uma cascata de sinalização à jusante dependente de cAMP-PKA (MERKLE e HOFFMANN 2011). As duas vias também estão relacionadas à mobilidade e migração celular e já foram associadas ao desenvolvimento de câncer (PERRETT e MCARDLE 2013; MERKLE e HOFFMANN 2011).

5.2 MÓDULO 2 - AVALIAÇÃO DAS ALTERAÇÕES ENCONTRADAS EM CÂNCER DE RETO DE ACORDO COM RESPOSTA AO TRATAMENTO NEOADJUVANTE

A radioquimioterapia neoadjuvante à cirurgia é o tratamento padrão para pacientes com câncer de reto localmente avançado, sendo a principal responsável pela diminuição de recidivas loco-regionais e de cirurgias com amputação do esfíncter (ABDUL-JALIL et al. 2014). A resposta desses pacientes ao tratamento varia substancialmente entre aqueles resistentes à terapia e os que apresentam resposta completa, os quais poderiam ser poupados da ressecção cirúrgica. No entanto, os mecanismos de resistência desses tumores ainda não estão bem estabelecidos e a predição de resposta por exames clínicos não são suficientemente precisos para predizer resposta patológica completa (HABR-GAMA et al. 2015). Dessa forma, a busca por marcadores preditivos é fundamental para proporcionar um tratamento personalizado, diminuindo o número de pacientes que não necessitariam ser submetidos a este tratamento e revelando outras potenciais estratégias terapêuticas.

Estudos mostram que assinaturas de metiloma são capazes de subclassificar tumores (HERNANDEZ-VARGAS et al. 2010), serem utilizados como marcadores prognósticos (CARVALHO et al. 2012) e de identificar marcadores de resistência inerente ou adquirida à quimioterapia (BORLEY e BROWN 2015). A metilação diferencial do DNA em sítios específicos também contribuiu para o entendimento da resposta à

radioterapia com ou sem quimioterapia, em diversos tipos tumorais, incluindo glioblastoma, câncer de mama, gástrico e CCR (LV et al. 2016).

A comparação das sondas DM exclusivas de cada grupo de acordo com a resposta à nCRT em CaRe revelou maior proporção de sondas hipermetiladas em pCR (44%) que em pIR (18%). Enquanto as sondas hipometiladas estavam distribuídas proporcionalmente nas mesmas regiões nos dois grupos, as sondas hipermetiladas estavam, em sua maioria, em ilhas CpG (66%) em pCR e em *Open Sea* (40%) em pIR. Para o nosso conhecimento, há apenas um estudo de análise de metilação global em CaRe com foco na resposta à nCRT, o qual não fez comparação entre as amostras tumorais com o TN (HA et al. 2015). Portanto, não foi possível comparar os nossos achados com os descritos pelos autores para avaliar se essas proporções se equivalem.

A análise global de metilação por imunistoquímica foi realizada por TSANG et al. (2014) com o intuito de verificar diferenças entre respondedores e não respondedores. Essa técnica é baseada na ligação de um anticorpo monoclonal contra 5-metilcitosina a regiões ricas em metil nos núcleos interfásicos. A captura e quantificação do sinal foi realizada com auxílio do programa Aperio (Vista, CA, USA). Utilizando o index de metilação baseado na intensidade da marcação das células tumorais, os autores reportaram correlação negativa entre TRG e os valores de metilação. Tumores com pCR apresentaram uma média de metilação menor quando comparado aos pacientes com resposta parcial (TSANG et al. 2014). Como não há comparação com o tecido normal no estudo, a interpretação dos

resultados e comparação com o presente trabalho fica prejudicada. Ainda que os autores tenham observado menor intensidade de metilação em pCR, não foi descrito como esse perfil compara-se ao TN. Além disso, não é possível observar a localização das regiões metiladas e se essas se encontram em regiões regulatórias ou se são consequência de alterações pré-existentes no tumor.

A alteração epigenética mais estudada em CCR é o perfil metilador de ilhas CpG (CIMP), que é baseado no status de metilação de regiões promotoras de genes supressores tumorais e de reparo a danos no DNA. Entretanto, não há um consenso sobre quais genes especificamente constituem o painel de CIMP (OGINO et al. 2006; GOEL e SHIN 2008; KURKJIAN et al. 2008; KIM et al. 2010; LEONG et al. 2011). Uma avaliação do status de metilação da região promotora dos genes mais comumente descritos como indicadores de CIMP (*CACNA1G*, *IGF2*, *NEUROG1*, *RUNX3* e *SOCS1*) em cada grupo (pCR e pIR) comparado ao TN revelou duas sondas alteradas em *IGF2* (uma na região 5'UTR com $\Delta\beta = 0,19$ e outra no primeiro éxon: $\Delta\beta = -0,27$) e uma sonda em *RUNX3* (TSS1500, $\Delta\beta = 0,16$) apenas em pIR. Não foram encontradas alterações em pCR e nos genes *CACNA1G*, *NEUROG1* e *SOCS1* nas comparações dos dois grupos. Com base apenas nesses genes, concluímos que nosso grupo de casos não possui CIMP. Poucos grupos avaliaram o papel do CIMP em carcinomas de reto (DE MAAT et al. 2010; LEONG et al. 2011; JO et al. 2012; MOLINARI et al. 2013). Nesses tumores, *CIMP-high* (presença de metilação em ilhas CpG mapeadas na região promotora de seis ou mais genes, incluindo *CACNA1G*,

CDKN2A, *CRABP1*, *IGF2*, *MLH1*, *NEUROG1*, *RUNX3* e *SOCS1*) foi descrito em 2 a 10% dos casos (YAMAUCHI et al. 2012).

Em uma análise global do perfil de metilação (450K, Illumina) de 45 tumores de reto localmente avançados, foram selecionadas 24 CpGs localizadas em ilhas e regiões promotoras entre 331 CpGs relacionadas à resposta à nCRT ($P < 0,005$ e $|\Delta\beta| > 0,15$) (HA et al. 2015). Dentre essas, sete CpGs localizadas em seis genes diferencialmente expressos em CCR (avaliados no banco de dados: <http://medicalgenome.kribb.re.kr/GENT>) foram avaliadas por pirosequenciamento em 67 casos (24 no grupo teste e 43 novos casos). A hipermetilação em *KLHL34* cg14281612 foi relacionada com melhor resposta à nCRT, com um valor de corte de 39,7% (sensibilidade e especificidade 0,625 e 0,727 respectivamente) (HA et al. 2015). A associação entre a hipermetilação e expressão diminuída foi demonstrada em linhagens celulares, mas não em amostras clínicas. Em nosso estudo, não foram encontradas sondas DM nesse gene nos dois grupos de pacientes, pCR e pIR comparados ao TN, pois esse gene localiza-se no cromossomo X, cujas sondas não foram avaliadas. Além de não compararem os achados com TN, HA et al. (2015) agruparam pacientes com TRG1-3 (de acordo com MANDARD et al. 1994) como bons respondedores, tornando questionável o valor preditivo desses achados. Outro fato curioso nesse estudo é a descrição de três diferentes sondas no gene *KLHL34*. A informação quanto a sonda diverge na tabela, figura e discussão, não sendo possível identificar quais das três sondas (cg14281612, cg14232291 e cg01828474) tem relevância.

A avaliação do status de metilação de genes específicos tem sido mais explorada em literatura na busca de marcadores de resposta à nCRT em CaRe. A hipermetilação dos genes *ATM*, *CRBP1*, *MGMT*, *TFAP2E* e *TIMP3* foi associada à resposta a radioterapia e/ ou quimioterapia baseada em 5-FU em CaRe (KIM et al. 2002; EBERT et al. 2012; MOLINARI et al. 2013; SUN et al. 2014; YOKOI et al. 2017). No presente estudo, não foram encontradas sondas DM nos genes *ATM* e *CRBP1* nas comparações realizadas. Hipermetilação da região promotora de *ATM* foi relacionada à perda da expressão desse gene e aumento da sensibilidade à radioterapia em linhagens celulares de CCR, mas esse achado não foi reportado em amostras de pacientes (KIM et al. 2002). No nosso estudo, foi identificada expressão aumentada em *ATM* nos casos com pIR, essa alteração pode ser decorrente de outros mecanismos, incluindo a mutação no gene, a qual foi observada em dois casos com pIR (RET7 e RET30). Hipermetilação do gene *CRBP1* foi relatada em linhagens celulares de CCR resistentes à radiação (YOKOI et al. 2017). Os autores também avaliaram 33 biópsias de CaRe, das quais 14 apresentaram hipermetilação; 13 deles eram de pacientes não respondedores (YOKOI et al. 2017). SUN et al. (2014) descreveram a hipermetilação na região promotora de *MGMT* no DNA tumoral circulante de pacientes com CaRe coletado antes da neoadjuvância com rádio e quimioterapia, como associada a melhor resposta ao tratamento. Embora não tenham sido identificadas em nosso estudo sondas DM na região promotora de *MGMT*, pacientes com pIR apresentaram cinco sondas hipometiladas e três hipermetiladas na região do corpo do gene. Alterações

nos níveis de expressão gênica devido à metilação em corpo de genes foram descritas previamente (YANG et al. 2014; BELTRAMI et al. 2017). No entanto, a não concordância do perfil encontrado e o fato desse gene não ter sido identificado como diferencialmente expresso no presente estudo sugerem que essas sondas não possuem papel regulatório nesse gene. Cinco sondas com $\Delta\beta^+$ e uma sonda com $\Delta\beta^-$ foram identificadas no corpo dos genes *TFAP2E* e *TIMP3*, respectivamente, nos dois grupos comparados ao TN. EBERT et al. (2012) descreveram associação entre a hipermetilação de *TFAP2E* e pior resposta à nCRT. Os autores relataram que apenas 10% dos casos com hipermetilação nesse gene apresentaram resposta à nCRT, em contraste com 82% dos casos com hipometilação. As sondas localizadas na região promotora do gene não estavam DM em nenhum dos grupos do presente estudo.

Um classificador baseado em *Diagonal Linear Discriminating Analysis* (DLDA) foi treinado usando as sondas diferencialmente metiladas nos dois grupos em relação ao TN e na comparação entre pCR e pIR ($p < 0,05$) após a aplicação dos filtros ($AUC > 0,75$ e regressão linear $< 0,05$). As três sondas resultantes permitiram classificar corretamente os tumores com 100% de sensibilidade e 90% de especificidade após validação cruzada usando o modelo *leave-one-out cross validation* (LOOCV). Os dois casos classificados erroneamente (RET19 e RET 31) apresentaram, respectivamente, TRG 1 e 2, diagnosticados aos 26 e 67 anos de idade, submetidos à cirurgia após 9,7 e 35,7 semanas, respectivamente. Ambos do sexo masculino, com estadiamento clínico III, sem comorbidades, vivos e livres de doença até a

data do último seguimento em setembro e junho de 2016, respectivamente. Somente uma das sondas estava presente na plataforma 450K (Illumina), de forma que o classificador não pode ser aplicado aos dados do TCGA ou aos achados de HA et al (2015). A análise em um grupo independente de amostras tem potencial para esclarecer o papel destes genes/sondas como marcadores de resposta a terapia neoadjuvante.

GAEDCKE et al. (2014) propuseram um painel de 10 regiões diferencialmente metiladas (DMR) para discriminar pacientes com CaRe com melhor e pior prognóstico. Os autores utilizaram espectrometria de massa baseada em MALDI-TOF (MassArray® - Sequenom) e identificaram 20 DMRs associadas ao prognóstico. Dez (*ADAP1*, *BARHL2*, *CABLES2*, *DOT1L*, *ERAS*, *ESRRG*, *RNF220*, *ST6GALNAC5*, *TAF4* e *SLC20A2*) das 20 DMR foram correlacionadas com sobrevida livre de doença em um grupo de validação. O teste Log rank revelou associação de melhor prognóstico nos casos que apresentaram altos níveis de metilação. Embora em nosso estudo não tenham sido analisadas DMR, foram identificadas sondas diferencialmente metiladas em oito destes 10 genes (com exceção de *CABLES* e *ERAS*) descritos pelos autores. No grupo com pCR, *BARHL2*, *ESRRG*, *DOT1L* e *TAF4* estavam hipermetilados na região promotora e em corpo de gene, respectivamente. *ADAP1* apresentou sondas hipometiladas em região promotora e de corpo de gene, enquanto *ST6GALNAC5* teve sondas hipo e hipermetiladas tanto em região promotora quanto de corpo de gene. *SLC20A2* possuía sondas hipermetiladas na região 5'UTR, assim como *RNF220*, que também apresentou sondas hipermetiladas em corpo de

gene. Nos casos com pIR, foram identificadas sondas com hiper e hipometilação em região promotora e de corpo dos genes *ESRRG*, *ST6GALNAC5* e *SLC20A2*. *TAF4* exibiu sondas hipo e hipermetiladas em corpo de gene. *ADAP1* apresentou hipometilação em região promotora e em corpo de gene; *DOT1L*, na região promotora, e hipermetilação em região de corpo de gene. *RNF220* possuía sondas hipermetiladas na região 5'UTR e em corpo de gene e *BARHL2* continha somente uma CpG hipermetilada. Aparentemente, não foi detectada associação entre a hipermetilação nesses genes e resposta à nCRT, como observada no nosso estudo.

Marcadores de resposta à nCRT baseados em dados de expressão gênica tem sido extensivamente avaliados, com diversas listas de genes publicadas nos últimos anos (Tabela 2). Vários autores que tentaram reproduzir os achados publicados, não tiveram sucesso, seja porque as listas não coincidiam (AGOSTINI et al. 2015a) ou porque a sensibilidade e especificidade dos marcadores descritos não eram satisfatórias quando aplicadas em outro grupo de casos (LOPES-RAMOS et al. 2015). Alguns fatores podem influenciar a reprodutibilidade dos resultados, por exemplo, alguns estudos compararam pCR e pIR, enquanto outros agrupam bons e maus respondedores (de acordo com o TRG). Em adição, as diferentes plataformas e métodos de análise utilizados em cada estudo, não permitiram até o momento que os achados fossem traduzidos para prática clínica (PEREZ e HABR-GAMA 2016).

Ainda assim, o perfil de expressão gênica desses tumores pode ser utilizado como uma valiosa ferramenta para identificar mecanismos

regulatórios específicos alterados em cada grupo de pacientes. As diferentes vias desreguladas em tumores de pacientes com pCR ou pIR podem elucidar se mecanismos alternativos podem ser explorados como potenciais novos alvos terapêuticos e auxiliar na compreensão da biologia dos carcinomas de reto. A partir do enriquecimento dos 378 e 678 genes diferencialmente expressos em pCR e pIR em relação ao TN foram identificadas 84 e 88 vias canônicas (IPA), respectivamente. Para avaliar as diferenças entre os dois grupos, as vias desreguladas identificadas somente em um grupo ou no outro foram avaliadas mais detalhadamente.

Os tumores de pacientes com pCR apresentaram 27 vias exclusivas (Figura 11). Verificou-se alteração em vias relacionadas à resposta imune inata (*Crosstalk between Dendritic Cells and Natural Killer Cells*) e adaptativa (*Altered T Cell and B Cell Signaling in Rheumatoid Arthritis*). As células dendríticas (DC) e as células *natural killer* (NK) são componentes do sistema imune inato que regulam respostas adaptativas na vigilância imunológica de reconhecimento do câncer e infecções. As DC são apresentadoras de antígenos que também podem modular a atividade de outras células do sistema imune inato e adaptativo por meio da liberação de citocinas e quimiocinas, podendo ativar ou bloquear a ação dessas células. Algumas dessas moléculas secretadas podem, por exemplo, ativar as células NK, que são células efetoras citotóxicas. As NKs também podem ser ativadas ao encontrar antígenos, pois expressam uma grande quantidade de receptores, inibidores e ativadores, cujo balanço desencadeia a resposta da célula. A interação entre DC e NKs é bidirecional, NKs podem induzir a

maturação das DC, que por sua vez, podem estimular a proliferação, sobrevivência e citotoxicidade das NKs, preparando uma resposta imune anti-tumoral eficiente (MAHMOOD et al. 2015). Estudos mostram que a interação de DC CXCL16^{+/+} e células NK CXCR6⁺, resultam em ativação das NKs com alta produção de INF γ , o qual possui atividade pró- apoptótica, anti-proliferativa e anti-angiogênica em câncer (VEINOTTE et al. 2016). O papel das NKs no auxílio à resposta ao tratamento com vacinas de DC também foi descrito por diversos autores e revisado em detalhes por PAMPENA e LEVY (2015). As células NKs também apresentam direta associação com a regulação das células T. A infiltração das células NK e T foram associadas com bom prognóstico em tumores colorretais (SCONOCCHIA et al. 2014).

A via de sinalização alterada de células T e B em artrite reumatóide (*Altered T Cell and B Cell Signaling in Rheumatoid Arthritis*) inclui alteração de moléculas envolvidas na apresentação de antígenos (HLAs) e na regulação de citocinas pró-inflamatórias, como TNF e LTB (RAO et al. 2017). Células tumorais expressam antígenos que as diferem das células normais tornando-as identificáveis pelo sistema imunológico, o qual normalmente leva a morte das células tumorais. Entretanto, as células tumorais conseguem evadir a resposta imune por meio de diversos mecanismos, sendo esse um dos novos *Hallmarks* do câncer (HANAHAN e WEINBERG 2011). Nossos resultados mostraram diminuição de expressão de moléculas classe II (*HLA-DRB1*) e aumento de expressão de TNF nos tumores de pacientes com pCR. Esses resultados evidenciam potenciais mecanismos

de evasão do sistema imune e promoção da tumorigênese pela ativação de vias relacionadas ao TNF, como, por exemplo, a Wnt/ β -catenina. A relação entre essas vias foi reportada previamente em CCR (BALKWILL 2009; LIU et al. 2012).

Alterações na via Wnt promovendo o acúmulo de β -catenina (*Wnt/ β -catenin Signaling*) são considerados eventos iniciais na transição adenoma-carcinoma na maioria dos tumores colorretais por meio da inativação de *APC*, gene supressor tumoral responsável pela degradação da β -catenina citoplasmática (FEARON 2011). O excesso da proteína no citoplasma, leva a migração da mesma para o núcleo, onde há a formação de um complexo de transcrição, resultando na expressão dos genes-alvo de Wnt. Essa via de sinalização canônica de Wnt, modula diversos processos celulares, como proliferação, sobrevivência, apoptose, diferenciação e adesão celular (NOVELLASDEMUNT et al. 2015). No presente estudo, 25 casos apresentaram mutação no gene *APC* (8 pCR, 32%) o que pode contribuir para a desregulação da via Wnt/ β -catenina. No entanto, não foi identificada expressão diferencial desse gene nos casos analisados, possivelmente pelo tamanho amostral ou estratégia de análise dos dados usada na comparação entre os grupos. Em adição, foi identificada alterações na via não-canônica de Wnt (*PCP pathway*), a qual está relacionada a reorganização do citoesqueleto durante a mobilidade celular. A ativação dessa via por meio da expressão aumentada da proteína *Collagen triple helix repeat containing 1* (CTHRC1) foi associada ao aumento da proliferação, migração e invasão de linhagens de câncer colorretal em ensaios *in vitro* (YANG et al. 2015). Esse

gene estava com expressão aumentada nos nossos casos com pCR e não estava diferencialmente expresso nos casos com pIR. A via de sinalização de Wnt também está associada à adipogênese (*Adipogenesis pathway*) em adultos, reprimindo-a pelo bloqueio da indução de *Peroxisome proliferator - activated receptor γ* (PPARG) e *CCAAT/enhancer-binding protein- α* (CEBPA) (CHRISTODOULIDES et al. 2009). A desregulação de adipócitos associada à inflamação crônica pode levar à tumorigênese induzida por adipócitos. PPARG é um receptor nuclear que participa da inibição da sinalização de Wnt, assim como diversas outras vias de sinalização anti-tumorais, por meio de ações não-genômicas (como interação e ativação de proteínas por meio de contato) e genômicas (quando está associado a outras proteínas no núcleo da célula formando um complexo de modulação de transcrição gênica). Agonistas de PPARG foram descritos como nova estratégia terapêutica, sendo demonstrada sua eficácia antineoplásica em diversos estudos *in vitro* e *in vivo* (VELLA et al. 2017). Os genes *PPARG* e *CEBPA* não estavam diferencialmente expressos nos casos analisados nesse estudo, o que poderia ser explicado pelo tamanho da casuística ou estratégias de análise dos dados. Outros genes da via de adipogênese, como *RB1* foram identificados com expressão aumentada em pCR. A inibição de PPARG por *RB1* foi demonstrada em linhagens celulares de fibroblastos embrionários, fibroblastos-*like* de camundongos e de osteosarcoma (FAJAS et al. 2002) podendo ser um dos mecanismos regulatórios de PPARG em ReCa.

As vias exclusivas desreguladas em tumores de pacientes com p1R (31 vias, Figura 11) incluíram a desregulação dos pontos de checagem do ciclo celular, tanto em G1/S (*Cell Cycle: G1/S Checkpoint Regulation*) como em G2/M (*Cell Cycle: G2/M DNA Damage Checkpoint Regulation*). Além disso, foi identificada alterações nas vias de reparo de quebras de dupla fita de DNA que seriam ativadas durante esses pontos de checagem na presença de lesões no material genético (*DNA Double-Strand Break Repair by Non-Homologous End Joining* e *DNA Double-Strand Break Repair by Homologous Recombination*, respectivamente). Lesões no DNA são eventos frequentes decorrentes da exposição a agentes genotóxicos, como radiações e espécies reativas de oxigênio e, são usualmente reparadas sem gerar consequências drásticas para as células. As quebras de dupla fita (DSB - *double strand break*) são as lesões mais nocivas às células, pois podem resultar em grandes rearranjos cromossômicos e erros de segregação cromossômicas quando incorretamente reparadas. Os dois principais mecanismos de reparo de DSB são o reparo por recombinação homóloga (HR - *homologous recombination*) e a junção de extremidades não-homólogas (NHEJ - *non-homologous end joining*) (APARICIO et al. 2014).

O processo de reparo ao dano no DNA (DDR - *DNA damage repair*) abrange o reconhecimento das lesões, recrutamento das proteínas de reparo e, nas células em processo de divisão, bloqueio do ciclo celular. Lesões muito grandes ou contínuas podem levar à morte celular ou senescência. Durante o ciclo celular, diversos pontos de checagem avaliam

as condições necessárias para continuar o processo de divisão e a maquinaria de reparo está diretamente ligada a esse mecanismo de controle (SHALTIEL et al. 2015). Primeiramente, o reconhecimento das DSBs é feito pelo complexo MRN (Mre11, Rad50 e Nbs1), que recruta ATM (*ataxia telangiectasia mutated*). Durante a fase G1, não há participação da via de HR, visto que os cromossomos ainda não foram duplicados e não existe a cromátide irmã como referência para o pareamento homólogo. Dessa forma, durante essa fase do ciclo celular, o principal mecanismo de reparo as DSBs é NHEJ.

Após recrutamento de ATM, há ativação de CHK2, estabilizando p53 e desencadeando a transcrição de diversos genes, incluindo o gene da proteína p21, inibidor de proteínas quinases dependentes de ciclinas (Cdks) e p38 que estabiliza o mRNA de p21. A ação de ATM, CHK2 e p38 impede a continuação do ciclo celular para a fase S, bloqueando a ação de CDK2, iniciadora da replicação do DNA durante a fase S. Durante G2, Wee1, CHK1 e ATR são os principais moduladores do ponto de checagem, o qual não depende de p53, p21 e ATM. A manutenção na fase G2, durante reparo, por exemplo, são dependentes de BRCA2 e PALB2 (SHALTIEL et al. 2015). Alterações na via de reparo a danos no DNA por HR também foram identificadas por análise de mutações em um número significativo de casos no presente estudo (23%) e nos dados do TCGA (25%). Escores indicativos de defeitos em HR por meio de análises de alterações genômicas foram identificados em 61% dos casos, dos quais, 70% eram pIR. Esses resultados estão discutidos em detalhes no Anexo 5.

A via do processo de ubiquitinação (*Protein Ubiquitination Pathway*) também desregulada em casos pIR, pode contribuir para as alterações no ciclo celular, uma vez que agem diretamente na coordenação rígida dos níveis das proteínas envolvidas nos pontos de checagem do ciclo. As ubiquitinas atuam na degradação proteica pela formação do proteassomo, regulando os níveis de ciclinas e de p21 e p53, por exemplo, sendo que perda ou degradação de p21 ou p53 leva à inativação do ponto de checagem G1/S (BASSERMANN et al. 2014). Além disso, histonas deacetilases, como a sirtuína SIRT1 (*Sirtuin Signaling Pathway*) também podem sofrer ubiquitinação estimulada por sinais decorrentes da via de reparo a danos no DNA (PENG et al. 2015).

Mecanismos epigenéticos parecem exercer um papel crucial nos tumores de pacientes com pIR, pois tanto a via de sinalização das sirtuínas quanto da metilação de DNA (*DNA Methylation and Transcriptional Repression Signaling*) encontravam-se alteradas. As sirtuínas representam um conjunto de enzimas deacetilases envolvidas na resposta a estresses metabólicos e genotóxicos, atuando na modificação da cromatina pelo controle da acetilação de histonas e consequente regulação da expressão gênica (BOSCH-PRESEGUE e VAQUERO 2013; OLIVEIRA et al. 2013). Assim como as sirtuínas, a metilação do DNA é um dos mecanismos de regulação da expressão gênica, o qual ocorre pela ação de DNA metiltransferases, que adicionam um grupo metil (CH₃) à quinta posição do carbono de uma CpG, formando a 5-metilcitosina (5mC) (EGGER et al. 2004; HOQUE et al. 2005). No presente estudo, foi encontrado um maior

número de genes diferencialmente expressos em pIR (678) que em pCR (378), podendo ser consequência de alterações nas vias de regulação da expressão gênica nesse grupo.

A rede regulatória transcricional em células-tronco embrionárias (*Transcriptional Regulatory Network in Embryonic Stem Cells*) estava alterada nos casos com pIR. Nessa rede está incluído o proto-oncogene nuclear *SET* que se apresentou com expressão aumentada. Este gene inibe a ação de histonas acetilases, ligando-se a regiões específicas de histonas, regulando também a transcrição gênica (ALMEIDA et al. 2017). A expressão aumentada de *SET* já foi descrita em CCR e associada a pior prognóstico (CRISTOBAL et al. 2015). Assim, o gene *SET* poderia ter um papel relevante também em CaRe nos pacientes que apresentaram pIR.

Os genes diferencialmente expressos em cada grupo comparado com o TN também foram utilizados para identificar proteínas potencialmente secretada por carcinomas de reto, de acordo com a resposta à nCRT. As moléculas secretadas/excretadas pelas células tumorais tem sido descritas como potenciais novos biomarcadores de avaliação em plasma/soro ou até mesmo nas fezes (SCHAAIJ-VISSER et al. 2013). A análise de predição *in silico* resultou em 40 proteínas de 106 genes com expressão aumentada em pCR e 104 proteínas de 353 genes em pIR potencialmente secretadas por esses tumores. Dessas, 17 do primeiro grupo e 56 do segundo já haviam sido descritas, ou seus mRNAs, em bancos de dados de vesículas extracelulares provenientes de amostras de CCR. Como o intuito do presente estudo foi avaliar as diferenças entre os tumores de pacientes com

pCR e pIR, as 12 proteínas identificadas em ambos os grupos não foram consideradas nas análises subsequentes, restando 5 proteínas de interesse em pCR e 24 em pIR. O banco de dados *Cancer Secretome* (<http://www.cancersecretome.org/>) foi utilizado para filtrar esta lista nos casos pIR, sendo encontradas 14 proteínas em CaRe (Tabela 6). Essa estratégia foi aplicada no intuito de pré-selecionar candidatos para posterior validação.

A investigação do secretoma em CCR foi descrita em estudos utilizando meios condicionados (MCs) provenientes de linhagens celulares, líquido peritoneal, urina, fezes e soro/plasma (IMPERLINI et al. 2013; SCHAAIJ-VISSER et al. 2013; ZENG et al. 2013; SHIN et al. 2014; TATO-COSTA et al. 2016; BHARDWAJ et al. 2017). A maioria desses estudos teve como objetivo identificar marcadores diagnósticos de CCR. TATO-COSTA et al. (2016) avaliaram o papel do secretoma de linhagens celulares de cólon associado à indução de senescência (SAS - *senescence-associated secretome*) devido ao tratamento com 5-FU ou doxorubicina. Os autores identificaram diferentes citocinas no MC das células senescentes (IL-8, TGF- α , VEGF, cistatina C, lipocalina 2 (LCN2 ou NGAL), fator de inibição da migração de macrófagos (MIF), indutor de metaloproteinases de matriz (EMMPRIN) e receptor do ativador de plasminogênio tipo uroquinase), em comparação com o MC da mesma linhagem não induzida. O cultivo de linhagens não tratadas de cólon e de reto com o MC das células senescentes levou ao aumento de proliferação nessas células e ativação da transição epitélio-mesênquima (EMT - *epithelial to mesenchymal transition*)

sendo observado um aumento da capacidade de invasão. Os tumores de pacientes com CaRe (avaliados após intervenção terapêutica) submetidos ao tratamento com 5-FU neoadjuvante em comparação com os tumores de pacientes não tratados, apresentaram maior expressão de marcadores de células senescentes (IL-18, p21 e p16) e de proteínas envolvidas na EMT (Snail, Slug e vimentina). Com base nesses achados, os autores sugeriram que a terapia com 5-FU induz senescência em algumas células tumorais e que o secretoma dessas células modifica o microambiente por meio da liberação de citocinas ativadoras da EMT, podendo estar relacionado com recorrência e pior prognóstico (os quais não foram avaliados) (TATO-COSTA et al. 2016).

A citocina IL-8 também foi identificada em nossa análise, no entanto, tanto tumores de paciente com pCR como os daqueles com pIR apresentaram aumento de expressão dessa citocina. Ainda que os níveis de expressão (baseados em FC) nos casos com pIR tenha sido maior que nos casos com pCR (FC = 4,5 e 2,4 respectivamente), essa diferença não foi considerada significativa em nossa casuística ($p > 0,05$). A IL-8 é uma citocina pró-inflamatória que foi descrita como associada a pior prognóstico em CCR devido ao aumento de proliferação, migração, invasão e resistência ao tratamento com oxaliplatina associados ao seu aumento de expressão (NING e LENZ 2012).

Uma das vias de sinalização intracelular ativada pela IL-8 se dá por meio das proteínas G e PLCs, identificadas como alteradas em nossos casos de CaRe (na análise individual de expressão e na análise integrada

com os dados de metilação). À jusante das proteínas G, IL-8 pode ativar a via de VEGF e FAK (alteradas em pIR e pCR, respectivamente, Figura 11), às quais estão relacionadas a angiogênese e remodelamento do citoesqueleto, respectivamente (LIU et al. 2016). A associação da expressão aumentada de IL-8 com diferentes vias de sinalização, que também podem estar alteradas, pode explicar os diferentes desfechos observados (pCR e pIR) em pacientes com elevada expressão dessa citocina. Desta forma, esses achados indicam cautela ao considerar a proteína IL-8 como biomarcador em CaRe.

As interleucinas 3 (*CXCL3*) e 5 (*CXCL5*) foram identificadas em nosso estudo como potenciais biomarcadores em pacientes com pIR. O aumento de expressão de *CXCL5* foi descrito em soro de pacientes com CCR e associado a pior prognóstico e metástases no fígado (KAWAMURA et al. 2012). A expressão do gene *CXCL3* também foi identificada em tumores de pacientes com CCR, os quais apresentaram sobrevida global menor que aqueles com níveis normais de expressão (XIONG et al. 2017). No entanto, não foram encontrados estudos descrevendo o papel dessas citocinas em CaRe especificamente, ou em relação à resposta à rádio ou quimioterapia.

Para identificar quais das proteínas secretadas poderiam estar relacionadas à resposta ao tratamento neoadjuvante baseado em 5-FU, foi realizada uma busca no banco de dados EVpedia (KIM et al. 2015). Este banco de dados possui informações de VE provenientes de diversas linhagens de células expostas ou não a condições variáveis. Para uma análise mais detalhada, selecionamos duas linhagens celulares (HCC2998 e

SW620) de CCR descritas como sensível e resistente à 5-FU respectivamente, e realizada a comparação com as proteínas descritas nas VEs dessas células. Além disso, as proteínas identificadas em VE no sangue de um grupo de indivíduos saudáveis e no soro de pacientes com CCR também foram utilizadas na comparação. Para identificar potencial marcadores de sensibilidade à terapia, foram selecionadas proteínas que estivessem presentes tanto em pCR quanto na linhagem sensível à 5-FU. A única proteína identificada em ambos foi a ASPH (*Aspartate beta-hydroxylase*), que também foi identificada nas VEs do sangue de indivíduos saudáveis. A ASPH foi descrita como potencial alvo terapêutico em gliomas (STURLA et al. 2016) e como ativador de resposta imune anti-tumoral em carcinomas hepatocelulares (TOMIMARU et al. 2015), mas não foram encontrados estudos sobre seu papel em CCR e na resposta a 5-FU.

Para a identificação dos potenciais marcadores de resistência à nCRT, foram observadas as proteínas identificadas em pIR que também estivessem nas VEs da linhagem resistente à 5-FU, e não na linhagem sensível ou em pCR. As proteínas UBE2C (*Ubiquitin-conjugating enzyme E2 C*) e CEMIP (*Cell migration-inducing and hyaluronan-binding protein*) foram as únicas que preenchem esses critérios, sendo que CEMIP também foi identificada no soro dos pacientes com CCR. A enzima UBE2C faz parte do complexo proteassomo envolvido na via de ubiquitinação de proteínas (via também identificada como alterada em pIR), cuja elevada expressão tem sido descrita em diversos tipos de câncer e associada a características clínicas de pior prognóstico (XIE et al. 2014). BAJAJ et al. (2016) utilizando

linhagens celulares relataram que a expressão de *UBE2C* estava intimamente associada à p53 durante tratamento com 5-FU, sendo reprimida por p53 selvagem (wt) e induzida por p53 mutada (BAJAJ et al. 2016). A frequência de CaRe com mutação em *TP53* foi elevada no presente estudo (84%, Anexo 3), a alta expressão de *UBE2C* em pacientes com mutação em *TP53* pode representar um dos mecanismos de resistência à nCRT e seu uso como biomarcador deve ser explorado nesses pacientes.

A proteína CEMIP (*Cell migration inducing protein*), está associada ao desenvolvimento e prognóstico desfavorável de diversos tipos tumorais, incluindo CCR. O aumento de expressão desse gene e da proteína foram identificados em câncer de cólon estando associado a menores taxas de sobrevida em 5 anos (FINK et al. 2015). O silenciamento de *CEMIP* diminuiu o potencial de invasão tumoral e metástase em CCR (FINK et al. 2015). Segundo os autores, como esta proteína é secretada, foi sugerido seu papel como potencial marcador em CCR. Interessantemente, GOMBODORJ et al. (2017) avaliando resistência à radiação em condições de hipóxia em células de CCR, identificaram CEMIP como modulador da via de HIF (*Hypoxia-inducible factor*). Usando o inibidor de UBE2 (NSC697923) em combinação com radioterapia, os autores demonstraram maior citotoxicidade na linhagem SW480 (resistentes à radioterapia) e inibição do crescimento tumoral em modelo animal em comparação com tratamento único com radioterapia (GOMBODORJ et al. 2017). Esses resultados corroboram os achados do potencial uso de *UBE2C* e *CEMIP* como marcadores de resposta em biópsia líquida de pacientes com CaRe.

Por fim, foi realizada uma análise integrada dos genes diferencialmente expressos com os dados de metilação diferencial nos dois grupos de CaRe. Os tumores de pacientes com pIR apresentaram um número muito maior de genes desregulados que os pCR. Essa diferença pode ser decorrente do maior número de casos presentes neste grupo de tumores (17 pIR vs. 5 TN e 11 pCR vs. 5 TN) ou pode ser resultado das alterações envolvidas no controle de expressão gênica nesse grupo (não identificada no grupo com pCR), evidenciado pela análise funcional *in silico* realizada no presente estudo (Figura 11). Enquanto a maioria das sondas DM, resultantes da análise integrada, apresentaram correlação negativa em pCR (66%), somente 28% tinham r- no grupo de casos com pIR. Além disso, foi evidenciada uma maior proporção de sondas em regiões regulatórias (região promotora e em ilhas CpG) nos casos com pCR e uma distribuição mais homogênea das sondas com r- nos casos com pIR (Figura 14).

Entre os genes regulados por metilação em pCR, diversos estão envolvidos em modulação de transcrição (cinco genes membros da família *Zinc Finger* e o miR1197) e de modificações pós-traducionais, como *ST3GAL4* (glicosiltransferase). Os *zinc fingers* podem atuar como oncogenes ou supressores de tumor. Ainda que não tenha sido encontrado nenhum estudo descrevendo a relação da metilação dos genes identificados no presente estudo, a metilação da região promotora de outros genes da família *zinc finger* foi descrita como associada à diminuição da expressão dos mesmos e consequente inibição de apoptose (*ZNF545*) e angiogênese (*ZNF24*), entre outros mecanismos pro-tumorais (JEN e WANG 2016). O

aumento da expressão de *ST3Gal4* foi associado aos tumores gástricos e renais, induzindo invasão e metástase pela ligação à selectinas, que são expressas por células vasculares endoteliais inflamadas (RODRIGUES e MACAULEY 2018).

A expressão do gene *ITIH5* foi identificada como inibida por metilação nos casos com pCR. A expressão diminuída dos genes da família *ITIH* foi descrita em diversos tumores, incluindo cólon, mama e pulmão (HAMM et al. 2008). *ITIH5* possui um papel na manutenção da estabilidade da matriz extracelular. A associação da metilação da região promotora desse gene com a diminuição da sua expressão em câncer de cólon foi reportada, especialmente em tumores positivos para CIMP (KLOTEN et al. 2014). O miR1197 foi identificado com hipometilação na região promotora nos tumores com pIR. Esse miRNA tem 317 alvos preditos (<http://mirdb.org>). O enriquecimento de processos biológicos dos genes-alvo usando o Gene Ontology (<http://geneontology.org>) indicou a participação dos genes em processos biossintéticos de glicosaminoglicanos (*Fold Enrichment* = + 5,29; FDR = 0,04), regulação de processos metabólicos (FE = + 1,47; FDR = 0,0003), via de sinalização do receptor intracelular (FE = + 4,59; FDR = 0,01), entre outros (dados não mostrados). Não foi encontrado nenhum estudo descrevendo o papel desse microRNA em câncer, sendo, portanto, um bom candidato para futuras investigações.

Os genes regulados por metilação em pIR incluíram o *SMAD9*, um membro de uma família de fatores de transcrição, cuja inibição por metilação foi descrita em carcinoma gástrico associada à infecção por *Helicobacter*

pylori (KANG et al. 2008). Em uma avaliação de expressão gênica global em amostras de adenoma folicular e variante folicular de carcinoma papilar de tireoide (VFCPT), foi identificado, entre outros, a diminuição da expressão de *SMAD9* somente em VFCPT, enquanto o mesmo estava com expressão aumentada em adenomas com relação ao TN (SCHULTEN et al. 2015). Mutações em genes da família SMAD foram associados com a síndrome de polipose juvenil e câncer colorretal esporádico, sendo também associadas com a desregulação da via TGF- β , frequentemente alterada em CCR (FLEMING et al. 2013).

O gene *TRIM29* foi detectado com aumento de expressão e hipometilação nos casos com pIR. Ao contrário, em câncer de mama a inibição da expressão de *TRIM29* pela metilação de sua região promotora foi associada à EMT mediada por TWIST (AI et al. 2014). Os autores relataram uma associação entre a hipermetilação/expressão diminuída de *TRIM29* com menor sobrevida livre de doença, maior tamanho tumoral, alto grau e características metastáticas. A participação de *TRIM29* na indução da transição epitélio-mesênquima mediada por TWIST foi confirmada em ensaios funcionais (CHOI et al. 2017). SNAI1 (Snail), SNAI2 (Slug) e TWIST são importantes reguladores da transcrição de genes que codificam proteínas envolvidas na adesão celular e na manutenção das características das células epiteliais e, quando desregulados, induzem as células epiteliais a transformarem-se em células com características mesenquimais (EMT). Esse processo foi associado a formação de metástases e características mais agressivas em tumores (TIWARI et al. 2012). Foram reportadas a

associação de marcadores da EMT e resistência à nCRT em CaRe, assim como a associação desses marcadores e aumento da capacidade de invasão em linhagens celulares de CCR (BHANGU et al. 2014; TATO-COSTA et al. 2016). No entanto, são necessários estudos adicionais para melhor compreender o papel de *TRIM29* em CaRe.

Os resultados apresentados nesse estudo confirmaram achados importantes descritos previamente em literatura e revelaram novos genes e vias associados ao câncer de reto. Em especial, foram identificadas alterações na via de reparo a danos no DNA por recombinação homóloga (por meio de mutações, assinaturas genômicas e expressão diferencial de transcritos) e vias de sinalização intracelulares relacionadas à divisão, migração e reorganização do citoesqueleto. Além disso, foram descritas diferenças moleculares entre tumores provenientes de pacientes com resposta patológica completa ou incompleta. Interessantemente, foi desenvolvido um classificador baseado em sondas diferencialmente metiladas que foi capaz de discriminar os pacientes quanto a resposta a terapia. A análise do transcriptoma também permitiu a identificação de proteínas potencialmente secretadas pelos carcinomas de reto, os quais podem servir de marcadores de resposta à terapia neoadjuvante.

6 CONCLUSÕES

- Foram identificadas alterações genômicas, transcriptômicas e epigenéticas em câncer de reto ainda não descritas em literatura e que contribuem para o melhor entendimento da biologia do tumor;
- Alterações genômicas já descritas em CaRe foram identificadas, com prevalência de ganhos em 7p, 8q, 13q e 20q e perdas em 8p, 14, 17p e 18. Foi demonstrada uma alta instabilidade cromossômica quantificada pelo *Genomic Instability Index* (GII). Pela primeira vez, escores de deficiência da via de reparo de DNA por recombinação homóloga (HRD) foram calculados para CaRe, resultando em um número significativo de casos com escores altos.
- Foram identificadas variantes em oito genes relacionados a vias de reparo de lesões no DNA. A análise dos dados do TCGA confirmou uma associação entre mutação em genes da via de HR e GII em casos não-hipermutados;
- Foi identificada hipometilação acentuada nos CaRe, com a maioria de sondas localizadas em regiões intergênicas e *open sea*, enquanto as sondas hipermetiladas encontraram-se principalmente em regiões promotoras e de ilhas CpG;
- Os genes diferencialmente expressos em CaRe comparados ao TN e validados com dados externos revelaram alterações em vias biológicas. Entre essas vias estão a de a inibição do ponto de checagem G2/M de lesões no DNA e ativação de ciclinas e do ciclo

celular , as quais podem estar relacionadas a alta instabilidade genômica observada nesses tumores. Também estavam alteradas as vias de sinalização relacionadas aos receptores associados à proteína G, PLCs e cascatas de sinalização intracelulares com alterações no equilíbrio de Ca^{+2} ;

- A análise integrada dos dados de metilação e expressão diferencial dos CaRe comparados ao TN (validada com dados externos, TCGA), revelou que uma parcela das vias alteradas já identificadas com os dados de expressão gênica global estavam reguladas por metilação. As demais vias alteradas por metilação estavam associadas à proteína G e PLCs, previamente relacionadas ao desenvolvimento do câncer;
- Observou-se maior GII em casos respondedores, não havendo diferença de escores de HRD entre os grupos pIR e pCR. Entretanto, seis de sete casos com mutação em genes dessa via tinham pIR, indicando que diferentes mecanismos levam à instabilidade genômica nestes tumores.
- Foi observada maior proporção de sondas hipermetiladas localizadas em ilhas em pCR que em pIR. A falta de estudos de metilação global em CaRe e resposta à nCRT não permitiu uma comparação mais detalhada com os achados deste estudo.
- Com base nos dados de três sondas diferencialmente metiladas, foi possível desenvolver um classificador capaz de prever resposta à nCRT com 100% de sensibilidade e 90% de especificidade após ensaios de validação cruzada. A confirmação desses resultados em

uma casuística independente e maior pode contribuir para o uso destes marcadores na prática clínica;

- Os resultados da expressão gênica comparada ao TN permitiram identificar 84 e 88 vias canônicas desreguladas em pCR e pIR, respectivamente. Os tumores de pacientes com pCR apresentaram alterações em vias relacionadas à modulação do sistema imune com ativação de vias de sinalização, incluindo as canônicas (Wnt/ β -catenina) e não-canônicas (PCP) de Wnt;
- Os tumores de pacientes com pIR apresentaram vias alteradas relacionadas ao descontrole do ciclo celular e dos mecanismos de reparo das quebras de dupla fita de DNA. Foram também identificados alterações de vias envolvidas nos mecanismos epigenéticos incluindo a metilação do DNA e a acetilação de histonas, sendo características descritas em tumores de pior prognóstico;
- Foram identificadas proteínas alteradas potencialmente secretadas por CaRe, que podem servir como marcadores avaliados por biópsia líquida. ASPH foi revelada como potencial marcador de sensibilidade à nCRT, apresentando expressão elevada em pCR; enquanto UBE2C e CEMIP foram selecionados como potenciais marcadores de resistência ao tratamento.

7 REFERÊNCIAS BIBLIOGRÁFICAS

Abdul-Jalil KI, Sheehan KM, Kehoe J, et al. The prognostic value of tumour regression grade following neoadjuvant chemoradiation therapy for rectal cancer. **Colorectal Dis** 2014; 16:O16-25.

Abkevich V, Timms KM, Hennessy BT, et al. Patterns of genomic loss of heterozygosity predict homologous recombination repair defects in epithelial ovarian cancer. **Br J Cancer** 2012; 107:1776-82.

Agostini M, Janssen KP, Kim IJ, et al. An integrative approach for the identification of prognostic and predictive biomarkers in rectal cancer. **Oncotarget** 2015a; 6:32561-74.

Agostini M, Zangrando A, Pastrello C, et al. A functional biological network centered on XRCC3: a new possible marker of chemoradiotherapy resistance in rectal cancer patients. **Cancer Biol Ther** 2015b; 16:1160-71.

Aguiar S Jr, Lopes A, Soares FA, et al. Prognostic and predictive value of the thymidylate synthase expression in patients with non-metastatic colorectal cancer. **Eur J Surg Oncol** 2005; 31:863-8.

Aguiar S Jr, Taglietti PM, Begnami MD, et al. Value of mismatch repair deficiency for predicting complete pathologic response to preoperative chemoradiation in rectal carcinoma [abstract]. **J Clin Oncol** 2014, 32(3 suppl):639. [Presented at 2014 Annual Meeting of the American Society of Clinical Oncology; 2014 May 30 - June 03; Chicago, Illinois (USA)]

Ai L, Kim W, Alpay M, et al. TRIM29 suppresses TWIST1 and invasive breast cancer behavior. **Cancer Res** 2014; 74:4875-87.

Akiyoshi T, Kobunai T, Watanabe T. Predicting the response to preoperative radiation or chemoradiation by a microarray analysis of the gene expression profiles in rectal cancer. **Surg Today** 2012; 42:713-9.

Almeida LO, Neto MPC, Sousa LO, et al. SET oncoprotein accumulation regulates transcription through DNA demethylation and histone hypoacetylation. **Oncotarget** 2017; 8:26802-18.

Al-Sukhni E, Attwood K, Mattson DM, Gabriel E, Nurkin SJ. Predictors of pathologic complete response following neoadjuvant chemoradiotherapy for rectal cancer. **Ann Surg Oncol** 2016; 23:1177-86.

Amersi F, Agustin M, Ko CY. Colorectal cancer: epidemiology, risk factors, and health services. **Clin Colon Rectal Surg** 2005; 18:133-40.

Aparicio T, Baer R, Gautier J. DNA double-strand break repair pathway choice and cancer. **DNA Repair** 2014; 19:169-75.

Appelt AL, Pløen J, Harling H, et al. High-dose chemoradiotherapy and watchful waiting for distal rectal cancer: a prospective observational study. **Lancet Oncol** 2015; 16:919-27.

Attwell S, Mills J, Troussard A, Wu C, Dedhar S. Integration of cell attachment, cytoskeletal localization, and signaling by integrin-linked kinase (ILK), CH-ILKBP, and the Tumor Suppressor PTEN. **Molecular Biol Cell** 2003; 14:4813-25.

Bae JM, Kim JH, Cho NY, Kim TY, Kang GH. Prognostic implication of the CpG island methylator phenotype in colorectal cancers depends on tumour location. **Br J Cancer** 2013; 109:1004-12.

Bajaj S, Alam SK, Roy KS, et al. E2 ubiquitin-conjugating enzyme, UBE2C gene, is reciprocally regulated by wild-type and gain-of-function mutant p53. **J Biol Chem** 2016; 291:14231-47.

Balkwill F. Tumour necrosis factor and cancer. **Nat Rev Cancer** 2009; 9:361-71.

Bassermann F, Eichner R, Pagano M. The ubiquitin proteasome system — Implications for cell cycle control and the targeted treatment of cancer. **Biochim Biophys Acta** 2014; 1843:150-62.

Becker KG, Barnes KC, Bright TJ, Wang SA. The Genetic Association Database. **Nat Genet** 2004; 36:431-32.

Beltrami CM, Dos Reis MB, Barros-Filho MC, et al. Integrated data analysis reveals potential drivers and pathways disrupted by DNA methylation in papillary thyroid carcinomas. **Clin Epigenetics** 2017; 9:45.

Benson III ALB, Bekaii-Saab T, Chan E, et al. Rectal cancer. **J Natl Compr Canc Netw** 2012; 10:1527-64.

Bertolini F, Bengala C, Losi L, et al. Prognostic and predictive value of baseline and posttreatment molecular marker expression in locally advanced rectal cancer treated with neoadjuvant chemoradiotherapy. **Int J Radiat Oncol Biol Phys** 2007; 68:1455-61.

Bhangu A, Wood G, Brown G, et al. The role of epithelial mesenchymal transition and resistance to neoadjuvant therapy in locally advanced rectal cancer. **Colorectal Dis** 2014; 16:133-43.

Bhardwaj M, Erben V, Schrotz-King P, Brenner H. Cell line secretome and tumor tissue proteome markers for early detection of colorectal cancer: a systematic review. **Cancers Cancers (Basel)** 2017; 9:156.

Bianconi D, Unseld M, Prager GW. Integrins in the spotlight of cancer. **Int J Mol Sci** 2016; 17:2037-64.

Billimoria KY, Stewart AK, Winchester DP, Ko CY. The National Cancer Data Base: a powerful initiative to improve cancer care in the United States. **Ann Surg Oncol** 2008; 15:683-90.

Birkbak NJ, Wang ZC, Kim JY, et al. Telomeric allelic imbalance indicates defective DNA repair and sensitivity to DNA-damaging agents. **Cancer Discov** 2012; 2:366-75.

Bitterman DS, Salgado LR, Moore HG, et al. Predictors of complete response and disease recurrence following chemoradiation for rectal cancer. **Front Oncol** 2015; 5:286.

Borley J, Brown R. Epigenetic mechanisms and therapeutic targets of chemotherapy resistance in epithelial ovarian cancer. **Ann Med** 2015; 47:359-69.

Bosch-Presegue L, Vaquero A. Sirtuins in stress response: guardians of the genome. **Oncogene** 2014; 33:3764-75.

Brändstedt J, Wangefjord S, Nodin B, et al. Gender, anthropometric factors and risk of colorectal cancer with particular reference to tumour location and TNM stage: a cohort study. **Biol Sex Differ** 2012; 3:23.

Brettingham-Moore KH, Duong CP, Greenawalt DM, et al. Pretreatment transcriptional profiling for predicting response to neoadjuvant chemoradiotherapy in rectal adenocarcinoma. **Clin Cancer Res** 2011; 17:3039-47.

Caccia D, Dugo M, Callari M, Bongarzone I. Bioinformatics tools for secretome analysis. **Biochim Biophys Acta** 2013; 1834:2442-53.

Carvalho RH, Haberle V, Hou J, et al. Genome-wide DNA methylation profiling of non-small cell lung carcinomas. **Epigenetics Chromatin** 2012; 5:9.

Chan J, Kinsella MT, Willis JE, et al. A predictive genetic signature for response to fluoropyrimidine-based neoadjuvant chemoradiation in clinical stage II and III rectal cancer. **Front Oncol** 2013; 3:288.

Chang HJ, Jung KH, Kim DY, et al. Bax, a predictive marker for therapeutic response to preoperative chemoradiotherapy in patients with rectal carcinoma. **Hum Pathol** 2005; 36:364-371.

Charara M, Edmonston TB, Burkholder S, et al. Microsatellite status and cell cycle associated markers in rectal cancer patients undergoing a combined regimen of 5-FU and CPT-11 chemotherapy and radiotherapy. **Anticancer Res** 2004; 24:3161-7.

Chen MB, Wu XY, Yu R, et al. P53 status as a predictive biomarker for patients receiving neoadjuvant radiation-based treatment: a meta-analysis in rectal cancer. **PloS One** 2012; 7:e45388.

Chen Z, Duldulao MP, Li W, Lee W, Kim J, Garcia-Aguilar J. Molecular diagnosis of response to neoadjuvant chemoradiation therapy in patients with locally advanced rectal cancer. **J Am Coll Surg** 2011a; 212:1008-17.

Chen Z, Liu Z, Li W, et al. Chromosomal copy number alterations are associated with tumor response to chemoradiation in locally advanced rectal cancer. **Genes Chrom Cancer** 2011b; 50:689-99.

Chiang SF, Kan CY, Hsiao YC, et al. Bone marrow stromal antigen 2 is a novel plasma biomarker and prognosticator for colorectal carcinoma: a secretome-based verification study. **Disease Markers** 2015; 2015:874054.

Choi SK, Pandiyan K, Eun JW, et al. Epigenetic landscape change analysis during human EMT sheds light on a key EMT mediator TRIM29. **Oncotarget** 2017; 8:98322-35.

Christodoulides C, Lagathu C, Sethi JK, Vidal-Puig A. Adipogenesis and WNT signalling. **Trends Endocrinol Metabol** 2009; 20:16-24.

Cingolani P, Platts A, Wang LL, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. **Fly** 2012; 1:46-57.

Cristobal I, Rincon R, Manso R, et al. Deregulation of the PP2A inhibitor SET shows promising therapeutic implications and determines poor clinical outcome in patients with metastatic colorectal cancer. **Clin Cancer Res** 2015; 21:347-56.

Das P, Skibber JM, Rodriguez-Bigas MA, et al. Predictors of tumor response and downstaging in patients who receive preoperative chemoradiation for rectal cancer. **Cancer** 2007; 109:1750-5.

Dayde D, Tanaka I, Jain R, Tai MC, Taguchi A. Predictive and prognostic molecular biomarkers for response to neoadjuvant chemoradiation in rectal cancer. **Int J Mol Sci** 2017; 18:573.

de Felice F, Izzo L, Musio D, et al. Clinical predictive factors of pathologic complete response in locally advanced rectal cancer. **Oncotarget** 2016; 7:33374-80.

de Maat MF, van de Velde CJH, van der Werff MPJ, et al. Quantitative analysis of methylation of genomic loci in early-stage rectal cancer predicts distant recurrence. **J Clin Oncol** 2008; 26:2327-35.

de Maat MF, van de Velde CJ, Benard A, et al. Identification of a quantitative MINT locus methylation profile predicting local regional recurrence of rectal cancer. **Clin Cancer Res** 2010; 16:2811-8.

Depristo MA, Banks E, Poplin R, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. **Nat Genet** 2011; 43:491-501.

Derbel O, Wang Q, Desseigne F, et al. Impact of KRAS, BRAF and PI3KCA mutations in rectal carcinomas treated with neoadjuvant radiochemotherapy and surgery. **BMC Cancer** 2013; 13:200.

Dou X, Wang RB, Meng XJ, et al. PDCD4 as a predictor of sensitivity to neoadjuvant chemoradiotherapy in locally advanced rectal cancer patients. **Asian Pac J Cancer Prev** 2014; 15:825-30.

Duffy MJ, Lamerz R, Haglund C, et al. Tumor markers in colorectal cancer, gastric cancer and gastrointestinal stromal cancers: European group on tumor markers 2014 guidelines update. **Int J Cancer** 2014; 134:2513-22.

Dunne PD, McArt DG, Bradley CA, et al. Challenging the cancer molecular stratification dogma: intratumoral heterogeneity undermines consensus molecular subtypes and potential diagnostic value in colorectal cancer. **Clin Cancer Res** 2016; 22:4095-104.

Dvorak J, Sitorova V, Ryska A, et al. Prognostic significance of changes of tumor epidermal growth factor receptor expression after neoadjuvant chemoradiation in patients with rectal adenocarcinoma. **Strahlenther Onkol** 2012; 188:833-8.

Ebert MPA, Tänzer M, Balluff B, et al. TFAP2E-DKK4 and Chemoresistance in colorectal cancer. **N Engl J Med** 2012; 366:44-53.

Edge SB, Compton CC. The American Joint Committee on Cancer: the 7th edition of the AJCC cancer staging manual and the future of TNM. **Ann Surg Oncol** 2010; 17:1471-4.

Egger G, Liang G, Aparicio A, Jones PA. Epigenetics in human disease and prospects for epigenetic therapy. **Nature** 2004; 429:457-63.

Erben P, Ströbel P, Horisberger K, et al. KRAS and BRAF mutations and PTEN expression do not predict efficacy of cetuximab-based chemoradiotherapy in locally advanced rectal cancer. **Int J Radiat Oncol Biol Phys** 2011; 81:1032-8.

Fajas L, Egler V, Reiter R, et al. The retinoblastoma-histone deacetylase 3 complex inhibits PPAR γ and adipocyte differentiation. **Developmental Cell** 2002; 3:903-10.

Fearon ER, Vogelstein B. A genetic model for colorectal tumorigenesis. **Cell** 1990; 61:759-67.

Fearon ER. Molecular genetics of colorectal cancer. **Ann Rev Pathol** 2011; 6:479-507.

Ferlay J. The global and regional burden of cancer. In: Stewart BW, Wild CP, editors. **W cancer report 2014**. Lyon: IARC; 2014. p.16-63.

Ferlay J, Soerjomataram I, Dikshit R, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. **Int J Cancer** 2015; 136:E359-86.

Fink SP, Myeroff LL, Kariv R, et al. Induction of KIAA1199/CEMIP is associated with colon cancer phenotype and poor patient survival. **Oncotarget** 2015; 6:30500-15.

Fleming NI, Jorissen RN, Mouradov D, et al. SMAD2, SMAD3 and SMAD4 mutations in colorectal cancer. **Cancer Res** 2013; 73:725-35.

Friedenreich C, Norat T, Steindorf K, et al. Physical activity and risk of colon and rectal cancers: the European prospective investigation into cancer and nutrition. **Cancer Epidemiol Biomarkers Prev** 2006; 15:2398-407.

Gaedcke J, Leha A, Claus R, et al. Identification of a DNA methylation signature to predict disease-free survival in locally advanced rectal cancer. **Oncotarget** 2014; 5:8123-35.

Gantt GA, Chen Y, DeJulius K, Mace AG, Barnholtz-Sloan J, Kalady MF. Gene expression profile is associated with chemoradiation resistance in rectal cancer. **Colorectal Dis** 2014; 16:57-66.

Garcia-Aguilar J, Chen Z, Smith D, et al. Identification of a biomarker profile associated with resistance to neoadjuvant chemoradiation therapy in rectal cancer. **Ann Surg** 2011; 254:486-93.

Garland ML, Vather R, Bunkley N, Pearse M, Bissett IP. Clinical tumour size and nodal status predict pathologic complete response following neoadjuvant chemoradiotherapy for rectal cancer. **Int J Colorectal Dis** 2014; 29:301-7.

Gash KJ, Baser O, Kiran RP. Factors associated with degree of tumour response to neo-adjuvant radiotherapy in rectal cancer and subsequent corresponding outcomes. **Eur J Surg Oncol** 2017; 43:2052-9.

Gerard JP, Rostom Y, Gal J, Benchimol D, Ortholan C, Aschele C, Levi JM. Can we increase the chance of sphincter saving surgery in rectal cancer with neoadjuvant treatments: lessons from a systematic review of recent randomized trials. **Crit Rev Oncol Hematol** 2012; 81:21-8.

Ghadimi BM, Grade M, Difilippantonio MJ, et al. Effectiveness of gene expression profiling for response prediction of rectal adenocarcinomas to preoperative chemoradiotherapy. **J Clin Oncol** 2005; 23:1826-38.

Giralt J, de las Heras M, Cerezo L, et al. The expression of epidermal growth factor receptor results in a worse prognosis for patients with rectal cancer treated with preoperative radiotherapy: a multicenter, retrospective analysis. **Radiother Oncol** 2005; 74:101-8.

Goel A, Shin SK. CpG Island methylator phenotype in colorectal cancer: a current perspective. **Curr Colorectal Cancer Rep** 2008; 4:77-83.

Gombodorj N, Yokobori T, Yoshiyama S, et al. Inhibition of ubiquitin-conjugating enzyme E2 may activate the degradation of hypoxia-inducible factors and, thus, overcome cellular resistance to radiation in colorectal cancer. **Anticancer Res** 2017; 37:2425-36.

Gonzalez-Gonzalez M, Garcia J, Alcazar JAJA, et al. Association between the cytogenetic profile of tumor cells and response to preoperative radiochemotherapy in locally advanced rectal cancer. **Medicine** 2014; 93:e153.

Grade M, Wolff HA, Gaedcke J, Ghadimi BM. The molecular basis of chemoradiosensitivity in rectal cancer: implications for personalized therapies. **Langenbecks Arch Surg** 2012; 397:543-55.

Guinney J, Dienstmann R, Wang X, et al. The consensus molecular subtypes of colorectal cancer. **Nat Med** 2015; 21:1350-6.

Ha YJ, Kim CW, Roh SA, et al. Epigenetic regulation of KLHL34 predictive of pathologic response to preoperative chemoradiation therapy in rectal cancer patients. **Int J Radiat Oncol Biol Phys** 2015; 91:650-58.

Habr-Gama A, Perez RO, São Julião GP, Proscurshim I, Gama-Rodrigues J. Nonoperative approaches to rectal cancer: a critical evaluation.. **Semin Radiat Oncol** 2011; 21:234-9.

Habr-Gama A, São Julião GP, Perez RO. Nonoperative management of rectal cancer: identifying the ideal patients. **Hematol Oncol Clin North Am** 2015; 29:135-51.

Hamm A, Veeck J, Bektas N, et al. Frequent expression loss of Inter-alpha-trypsin inhibitor heavy chain (ITI-H) genes in multiple human solid tumors: a systematic expression analysis. **BMC Cancer** 2008; 8:25.

Hammarström S. The carcinoembryonic antigen (CEA) family: structures, suggested functions and expression in normal and malignant tissues. **Semin Cancer Biol** 1999; 9:67-81.

Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. **Cell** 2011; 144:646-74.

Harr MW, Distelhorst CW. Apoptosis and autophagy: decoding calcium signals that mediate life or death. **Cold Spring Harb Perspect Biol** 2010; 2:a005579.

Hegi ME, Diserens AC, Gorlia T, et al. MGMT gene silencing and benefit from temozolomide in glioblastoma. **N Engl J Med** 2005; 352:997-1003.

Helleday T. Homologous recombination in cancer development, treatment and development of drug resistance. **Carcinogenesis** 2010; 31:955-60.

Hernandez-Vargas H, Lambert MP, Calvez-Kelm FL, et al. Hepatocellular carcinoma displays distinct DNA methylation signatures with potential as clinical predictors. **PLoS One** 2010; 5:e9749.

Hoque MO, Rosenbaum E, Westra WH, et al. Quantitative assessment of promoter methylation profiles in thyroid neoplasms. **J Clin Endocrinol Metab** 2005; 90:4011-8.

Hua Y, Ma X, Liu X, et al. Abnormal expression of mRNA, microRNA alteration and aberrant DNA methylation patterns in rectal adenocarcinoma. **PLoS One** 2017; 12:e0174461.

Huang MY, Wu CH, Huang CM, et al. DPYD, TY, TK1, and TK2 genetic expressions as response markers in locally advanced rectal cancer patients treated with fluoropyrimidine-based chemoradiotherapy. **BioMed Res Internat** 2013; 2013:931028.

Hughes LA, Khalid-de Bakker CA, Smits KM, et al. The CpG island methylator phenotype in colorectal cancer: progress and problems. **Biochim Biophys Acta** 2012; 1825:77-85.

Huh JW, Kim HR, Kim YJ. Clinical prediction of pathological complete response after preoperative chemoradiotherapy for rectal cancer. **Dis Colon Rectum** 2013, 56:698-703.

Hupkens BJP, Martens MH, Stoot JH, et al. Quality of life in rectal cancer patients after chemoradiation: watch-and-wait policy versus standard resection - A matched-controlled study. **Dis Colon Rectum** 2017; 60:1032-40.

Husted AS, Trauelsen M, Rudenko O, Hjorth SA, Schwartz TW. GPCR-mediated signaling of metabolites. **Cell Metabol** 2017; 25:777-96.

Huxley RR, Ansary-Moghaddam A, Clifton P, et al. The impact of dietary and lifestyle risk factors on risk of colorectal cancer: a quantitative overview of the epidemiological evidence. **Int J Cancer** 2009; 125:171-80.

Iacopetta B. Are there two sides to colorectal cancer? **Int J Cancer** 2002; 101:403-8.

Imperlini E, Colavita I, Caterino M, et al. The secretome signature of colon cancer cell lines. **J Cell Biochem** 2013; 114:2577-87.

Intven M, Reerink O, Philippens ME. Diffusion-weighted MRI in locally advanced rectal cancer: pathological response prediction after neo-adjuvant radiochemotherapy. **Strahlenther Onkol** 2013; 189:117-22.

Irizarry RA, Ladd-Acosta C, Wen B, et al. The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissuespecific CpG island shores. **Nature Genet** 2009; 41:178-86.

Issa JP. CpG island methylator phenotype in cancer. **Nat Rev Cancer** 2004; 4:988-93.

Issa N, Murninkas A, Powsner E, Dreznick Z, et al. Long-term outcome of local excision after complete pathological response to neoadjuvant chemoradiation therapy for rectal cancer. **World J Surg** 2012, 36:2481-7.

Jakob C, Liersch T, Meyer W, Becker H, Baretton GB, Aust DE. Predictive value of Ki67 and p53 in locally advanced rectal cancer: correlation with thymidylate synthase and histopathological tumor regression after neoadjuvant 5-FU-based chemoradiotherapy. **World J Gastroenterol** 2008; 14:1060-6.

Jen J, Wang Y. Zinc finger proteins in cancer progression. **J Biomed Science** 2016; 23:53-62.

Jo P, Jung K, Grade M, et al. CpG island methylator phenotype infers a poor disease-free survival in locally advanced rectal cancer. **Surgery** 2012; 151:564-70.

Jones PA. Functions of DNA methylation: Islands, start sites, gene bodies and beyond. **Nature Reviews Genetics** 2012; 13:484–92.

Joye I, Debucquoy A, Fieuws S, et al. Can clinical factors be used as a selection tool for an organ-preserving strategy in rectal cancer? **Acta Oncol** 2016, 55:1047-52.

Kalady MF, de Campos-Lobato LF, Stocchi L, et al. Predictive factors of pathologic complete response after neoadjuvant chemoradiation for rectal cancer. **Ann Surg** 2009; 250:582-9.

Kang G. H, Lee S, Cho N, et al. DNA methylation profiles of gastric carcinoma characterized by quantitative DNA methylation analysis. **Lab Invest** 2008; 88:161-70.

Kawamura M, Toiyama Y, Tanaka K, et al. CXCL5, a promoter of cell proliferation, migration and invasion, is a novel serum prognostic marker in patients with colorectal cancer. **Eur J Cancer** 2012; 48:2244-51.

Kim DK, Lee J, Kim SR, et al. EVpedia: a community web portal for extracellular vesicles research. **Bioinformatics** 2015; 31:933-9.

Kim IJ, Lim SB, Kang HC, et al. Microarray gene expression profiling for predicting complete response to preoperative chemoradiotherapy in patients with advanced rectal cancer. **Dis Colon Rectum** 2007; 50:1342-53.

Kim JS, Kim JM, Li S, et al. Epidermal growth factor receptor as a predictor of tumor downstaging in locally advanced rectal cancer patients treated with preoperative chemoradiotherapy. **Int J Radiat Oncol Biol Phys** 2006; 66:195-200.

Kim MS, Lee J, Sidransky D. DNA methylation markers in colorectal cancer. **Cancer Metastasis Rev** 2010; 29:181-206.

Kim SY, Shim EK, Yeo HY, et al. KRAS mutation status and clinical outcome of preoperative chemoradiation with cetuximab in locally advanced rectal cancer: a pooled analysis of 2 phase II trials. **Int J Radiat Oncol Biol Phys** 2013; 85:201-7.

Kim W, Vo QN, Shrivastav M, Lataxes TA, Brown KD. Aberrant methylation of the ATM promoter correlates with increased radiosensitivity in a human colorectal tumor cell line. **Oncogene** 2002; 21:3864-71.

Kleiman A, Al-Khamis A, Farsi A, et al. Normalization of CEA levels post-neoadjuvant therapy is a strong predictor of pathologic complete response in rectal cancer. **J Gastrointest Surg** 2015; 19:1106-12.

Kloten V, Rose M, Kaspar S, et al. Epigenetic inactivation of the novel candidate tumor suppressor gene ITIH5 in colon cancer predicts unfavorable overall survival in the CpG island methylator phenotype. **Epigenetics** 2014, 9:1290-301.

Kudrimoti M, Lee EY, Kang Y, Ahmed M, Mohiuddin M. Genetic markers predictive of response to induction chemoradiotherapy for locally advanced rectal cancers. **J Ky Med Assoc** 2007; 105:18-22.

Kundel Y, Nasser NJ, Rath-Wolfson L, et al. Molecular predictors of response to neoadjuvant chemoradiation for rectal cancer. **Am J Clin Oncol** 2018; 41:613-8.

Kuremsky JG, Tepper JE, Mcleod HL. Biomarkers for response to neoadjuvant chemoradiation for rectal cancer. **Int J Radiat Oncol Biol Phys** 2009; 74:673-88.

Kurkjian C, Kummar S, Murgo A. DNA methylation: its role in cancer development and therapy. **Curr Probl Cancer** 2008; 32:187-235.

Larsen SJ, Canto LM, Rogatto SR, Baumbach J. CoNVaQ: a web tool for copy number variation-based association studies. **BMC Genomics** 2018; 19:369.

Law NC, White MF, Hunzicker-Dunn ME. G protein-coupled receptors (GPCRs) that signal via protein kinase A (PKA) cross-talk at insulin receptor substrate 1 (IRS1) to activate the phosphatidylinositol 3-kinase (PI3K)/AKT pathway. **J Biol Chem** 2016; 291:27160-69.

Lee MS, Menter DG, Kopetz S. Right versus left colon cancer biology: integrating the consensus molecular subtypes. **J Natl Compr Canc Netw** 2017; 15:411-9.

Lee YC, Lee YL, Chuang JP, Lee JC. Differences in survival between colon and rectal cancer from SEER data. **PLoS One** 2013; 8:e78709.

Lee YY, Li CF, Lin CY, et al. Overexpression of CPS1 is an independent negative prognosticator in rectal cancers receiving concurrent chemoradiotherapy. **Tumor Biol** 2014; 35:11097-105.

Lek M, Karczewski KJ, Minikel, EV, et al. Analysis of protein-coding genetic variation in 60,706 humans. **Nature** 2016; 536:285-91.

Leong KJ, Wei W, Tannahill LA, et al. Methylation profiling of rectal cancer identifies novel markers of early-stage disease. **Br J Surg** 2011; 98:724-34.

Li F, Lai M. Colorectal cancer, one entity or three. **J Zhejiang Univ Sci B** 2009; 10:219-29.

Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. **Bioinformatics** 2009; 25:754-60.

Li M, Li JY, Zhao AL, GU J. Colorectal cancer or colon and rectal cancer? **Oncol** 2007; 73:52-57.

Liang JW, Shi ZZ, Zhang TT, et al. Analysis of genomic aberrations associated with the clinicopathological parameters of rectal cancer by array-based comparative genomic hybridization. **Oncol Rep** 2013; 29:1827-34.

Lin CY, et al. Deficiency in asparagine synthetase expression in rectal cancers receiving concurrent chemoradiotherapy: negative prognostic impact and therapeutic relevance, **Tumor Biol** 2014; 35:6823-30.

Lin Q, Tan HT, Lim HSR, Chung MCM. Sieving through the cancer secretome. **Biochim Biophys Acta** 2013; 1834:2360-71.

Liu C, Srihari S, Lal S, et al. Personalised pathway analysis reveals association between DNA repair pathway dysregulation and chromosomal instability in sporadic breast cancer. **Mol Oncol** 2016; 10:179-93.

Liu Q, Li A, Tian Y, et al. The CXCL8-CXCR1/2 pathways in cancer. **Cytokine e Growth Factor Rev** 2016; 31:61-71.

Liu Z, Brooks RS, Ciappio ED, et al. Diet-induced obesity elevates colonic TNF- α in mice and is accompanied by an activation of Wnt signaling: a mechanism for obesity-associated colorectal cancer. **J Nutritional Biochem** 2012; 23:1207-13.

Lochhead P, Kuchiba A, Imamura Y, et al. Microsatellite instability and BRAF mutation testing in colorectal cancer prognostication. **J Natl Cancer Inst** 2013; 105:1151-6.

Lohse MJ. The ins and outs of adrenergic signaling. **J Molec Med** 2015; 93:955-62.

Loos M, Quentmeier P, Schuster T, et al. Effect of preoperative radio(chemo)therapy on long-term functional outcome in rectal cancer patients: a systematic review and meta-analysis. **Ann Surg Oncol** 2013; 20:1816-28.

Lopes-Ramos C, Koyama FC, Habr-Gama A, et al. Comprehensive evaluation of the effectiveness of gene expression signatures to predict complete response to neoadjuvant chemoradiotherapy and guide surgical intervention in rectal cancer. **Cancer Genet** 2015; 208:319-26.

Luna-Perez P, Segura J, Alvarado I, Labastida S, Santiago-Payán H, Quintero A. Specific c-K-ras gene mutations as a tumor- response marker in locally advanced rectal cancer treated with preoperative chemoradiotherapy. **Ann Surg Oncol** 2000; 7:727-731.

Lv JF, Hu L, Zhuo W, Zhang CM, Zhou HH, Fan L. Epigenetic alternations and cancer chemotherapy response. **Cancer Chemother Pharmacol** 2016; 77:673-84.

Maas M, Beets-Tan RGH, Lambregts DMJ, et al. Wait-and-see policy for clinical complete responders after chemoradiation for rectal cancer. **J Clin Oncol** 2011; 29:4633-40.

Mahmood S, Upreti D, Sow I, et al. Bidirectional interactions of NK cells and dendritic cells in immunotherapy: current and future perspective. **Immunotherapy** 2015; 7:301-8.

Mammano E, Galdi F, Pierobon M, et al. Multiplexed protein signal pathway mapping identifies patients with rectal cancer that responds to neoadjuvant treatment. **Clin Colorectal Cancer** 2012; 11:268-74.

Mandard AM, Dalibard F, Mandard JC, et al. Pathologic assessment of tumor regression after preoperative chemoradiotherapy of esophageal carcinoma. clinicopathologic correlations. **Cancer** 1994; 73:2680-6.

Manié E, Popova T, Battistella A, et al. Genomic hallmarks of homologous recombination deficiency in invasive breast carcinomas. **Int J Cancer** 2016; 138:891-900.

Maretto I, Pomerri F, Pucciarelli S, et al. The potential of restaging in the prediction of pathologic response after preoperative chemoradiotherapy for rectal cancer. **Ann Surg Oncol** 2007; 14:455-61.

Marquard AM, Eklund AC, Joshi T, et al. Pan-cancer analysis of genomic scar signatures associated with homologous recombination deficiency suggests novel indications for existing cancer drugs. **Biomarker Res** 2015; 3:9.

Martins SF, Garcia EA, Luz MA, Pardal F, Rodrigues M, Filho AL. Clinicopathological correlation and prognostic significance of VEGF-A, VEGF-C, VEGFR-2 and VEGFR-3 expression in colorectal cancer. **Cancer Genomics Proteomics** 2013; 10:55-67.

Mendonça R, Valadão M, Couto AC, Koifman S. Tendência de mortalidade por câncer colorretal em cinco capitais brasileiras de 1980 a 2009. **Cad Saúde Colet** 2012; 20:329-35.

Meng X, Huang Z, Wang R, et al. The prognostic role of EZH2 expression in rectal cancer patients treated with neoadjuvant chemoradiotherapy. **Rad Oncol** 2014a; 9:188.

Meng X, Wang R, Huang Z, et al. Human epidermal growth factor receptor-2 expression in locally advanced rectal cancer: association with response to neoadjuvant therapy and prognosis. **Cancer Sci** 2014b, 105:818-24.

Merkle D, Hoffmann R. Roles of cAMP and cAMP-dependent protein kinase in the progression of prostate cancer: cross-talk with the androgen receptor. **Cellular Signalling** 2011; 23:507-15.

Millino C, Maretto I, Pacchioni B, et al. Gene and microRNA expression are predictive of tumor response in rectal adenocarcinoma patients treated with preoperative chemoradiotherapy. **J Cell Physiol** 2017; 232:426-35.

Ministério da Saúde. Instituto Nacional do Câncer José de Alencar Gomes da Silva. **Estimativa/2018 Incidência de Câncer no Brasil**. Rio de Janeiro: INCA; 2018.

Molinari C, Ballardini M, Teodorani N, et al. Genomic alterations in rectal tumors and response to neoadjuvant chemoradiotherapy: an exploratory study. **Radiat Oncol** 2011; 6:161.

Molinari C, Casadio V, Foca F, et al. Gene methylation in rectal cancer: predictive marker of response to chemoradiotherapy? **J Cell Physiol** 2013; 228:2343-9.

Molinari C, Matteucci F, Caroli P, Passardi A. Biomarkers and molecular imaging as predictors of response to neoadjuvant chemoradiotherapy in patients with locally advanced rectal cancer. **Clin Colorectal Cancer** 2015; 14:227-38.

Moran GW, Leslie FC, Levison SE, Mclaughlin JT. Review: enteroendocrine cells: neglected players in gastrointestinal disorders? **Therap Adv Gastroenterol** 2008; 1:51-60.

Moran S, Arribas C, Esteller M. Validation of a DNA methylation microarray for 850,000 CpG sites of the human genome enriched in enhancer sequences. **Epigenomics** 2016; 8:389-99.

Murtha M, Esteller M. Extraordinary Cancer Epigenomics: Thinking Outside the Classical Coding and Promoter Box. **Trends in Cancer** 2016; 2:572–84.

Negri FV, Campanini N, Camisa R, et al. Biological predictive factors in rectal cancer treated with preoperative radiotherapy or radiochemotherapy. **Br J Cancer** 2008; 98:143-7.

Ning Y, Lenz H. Targeting IL-8 in colorectal cancer. **Expert Opin Ther Targets** 2012; 16:491-7.

Nishioka M, Shimada M, Kurita N, et al. Gene expression profile can predict pathological response to preoperative chemoradiotherapy in rectal cancer. **Cancer Genomics e Proteomics** 2011; 8:87-92.

Noda H, Miyaji Y, Nakanishi A, Konishi F, Miki Y. Frequent reduced expression of alpha-1B-adrenergic receptor caused by aberrant promoter methylation in gastric cancers. **Br J Cancer** 2007; 96:383-90.

Novellademunt L, Antas P, Li VSW. Targeting Wnt signaling in colorectal cancer: a review in the theme: cell signaling: proteins, pathways and mechanisms. **Am J Physiol Cell Physiol** 2015; 309:C511-C521.

Obeid Ei, Conzen Sd. The role of adrenergic signaling in breast cancer biology. **Cancer Biomarkers** 2013; 13:161-9.

Ogino S, Kawasaki T, Kirkner GJ, Loda M, Fuchs CS. CpG island methylator phenotype-low (CIMP-low) in colorectal cancer: possible associations with male sex and KRAS mutations. **J Mol Diagn** 2006; 8:582-8.

Okonkwo A, Musunuri S, Talamonti M, et al. Molecular markers and prediction of response to chemoradiation in rectal cancer. **Oncol Rep** 2001; 8:497-500.

Oliveira MVM, Andrade JMO, Paraíso AF, Santos SHS. Sirtuins and cancer: new insights and cell signaling. **Cancer Invest** 2013; 31:645-53.

Palma P, Cuadros M, Conde-Muñoz R, et al. Microarray profiling of mononuclear peripheral blood cells identifies novel candidate genes related to chemoradiation response in rectal cancer. **PLoS One** 2013; 8:e74034.

Pampena MB, Levy EM. Natural killer cells as helper cells in dendritic cell cancer vaccines. **Front Immunol** 2015; 6:13.

Park IJ, You YN, Agarwal A, et al. Neoadjuvant treatment response as an early response indicator for patients with rectal cancer. **J Clin Oncol** 2012; 30:1770-6.

Peng L, Yuan Z, Li Y, et al. Ubiquitinated Sirtuin 1 (SIRT1) Function is modulated during DNA damage-induced cell death and survival. **J Biol Chem** 2015; 290:8904-12.

Perez RO, Habr-Gama A. Putting down the scalpel in rectal cancer management - a historical perspective. **Colorectal Dis** 2018; 20:12-5.

Perou CM, Sørlie T, Eisen MB, et al. Molecular portraits of human breast tumours. **Nature** 2000; 406:747-52.

Perrett RM, Mcardle CA. Molecular mechanisms of gonadotropin-releasing hormone signaling: integrating cyclic nucleotides into the network. **Front Endocrinol** 2013; 4:180.

Pino MS, Chung DC. The chromosomal instability pathway in colon cancer. **Gastroenterology** 2010; 138:2059-72.

Popova T, Manié E, Rieunier G, et al. Ploidy and large-scale genomic instability consistently identify basal-like breast carcinomas with BRCA1/2 inactivation. **Cancer Res** 2012; 72:5454-62.

Prat A, Pineda E, Adamo B, et al. Clinical implications of the intrinsic molecular subtypes of breast cancer. **Breast** 2015; 24 Suppl 2:S26-35.

Probst CP, Becerra AZ, Aquina CT, et al. Watch and wait? Elevated pretreatment CEA is associated with decreased pathological complete response in rectal cancer. **J Gastrointest Surg** 2016; 20:43-52.

Provenzale D, Jasperson K, Ahnen DJ, et al. Colorectal Cancer Screening, Version 1.2015. **J Natl Compr Canc Netw** 2015; 13:959-68.

Purim O, Gordon N, Brenner B. Cancer of the colon and rectum: potential effects of sex-age interactions on incidence and outcome. **Med Sci Monitor** 2013; 19:203-9.

Qiu HZ, Wu B, Xiao Y, Lin GL. Combination of differentiation and T stage can predict unresponsiveness to neoadjuvant therapy for rectal cancer. **Colorectal Dis** 2011; 13:1353-60.

Rahma OE, Burotto M, Do Canto LM, Germanos AA, Haddad BR, Marshall JL. Striking similarities in genetic aberrations between a rectal tumor and its lung recurrence. **World J Gastrointestinal Oncol** 2013; 5:198-203.

Rao DA, Gurish MF, Marshall JL, et al. Pathologically expanded peripheral T helper cell subset drives B cells in rheumatoid arthritis. **Nature** 2017; 542:110-4.

Rau B, Sturm I, Lage H, et al. Dynamic expression profile of p21WAF1/CIP1 and Ki-67 predicts survival in rectal carcinoma treated with preoperative radiochemotherapy. **J Clin Oncol** 2003; 21:3391-401.

Raynal NJ-M, Lee JT, Wang Y, et al. Targeting calcium signaling induces epigenetic reactivation of tumor suppressor genes in cancer. **Cancer Res** 2016; 76:1494-505.

Reerink O, Karrenbeld A, Plukker JTM, et al. Molecular prognostic factors in locally irresectable rectal cancer treated preoperatively by chemoradiotherapy. **Anticancer Res** 2004; 24:1217-22.

Reimann F, Tolhurst G, Gribble FM. G-protein-coupled receptors in intestinal chemosensation. **Cell Metabolism** 2012; 15:421-31.

Restivo A, Zorcolo L, Cocco IMF, et al. Elevated CEA levels and low distance of the tumor from the anal verge are predictors of incomplete response to chemoradiation in patients with rectal cancer. **Ann Surg Oncol** 2013; 20:864-71.

Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. **Genet Med** 2015; 17:405-24.

Rimkus C, Friederichs J, Boulesteix AL, et al. Microarray-based prediction of tumor response to neoadjuvant radiochemotherapy of patients with locally advanced rectal cancer. **Clin Gastroenterol Hepatol** 2008; 6:53-61.

Rodrigues E, Macauley M. Hypersialylation in cancer: modulation of inflammation and therapeutic opportunities. **Cancer** 2018; 10:207-26.

Sada YH, Tran Cao HS, Chang GJ, et al. Prognostic value of neoadjuvant treatment response in locally advanced rectal cancer. **J Surg Res** 2018; 226:15-23.

Saigusa S, Tanaka K, Toiyama Y, et al. Gene expression profiles of tumor regression grade in locally advanced rectal cancer after neoadjuvant chemoradiotherapy. **Oncol Rep** 2012; 28:855-61.

Sajid KM, Parveen R, Durr-e-Sabih, et al. Carcinoembryonic antigen (CEA) levels in hookah smokers, cigarette smokers and non-smokers. **J Pak Med Assoc** 2007; 57:595-99.

Samowitz WS, Albertsen H, Herrick J, et al. Evaluation of a large, population-based sample supports a CpG island methylator phenotype in colon cancer. **Gastroenterology** 2005; 129:837-45.

Santos MD, Silva C, Rocha A, Matos E, Nogueira C, Lopes C. Prognostic value of mandard and dworak tumor regression grading in rectal cancer: study of a single tertiary center. **ISRN Surg** 2014; 2014:310542.

Schaaij-Visser TBM, De Wit M, Lam SW, Jiménez CR. The cancer secretome, current status and opportunities in the lung, breast and colorectal cancer context. **Biochim Biophys Acta** 2013; 1834:2242-58.

Schmoll HJ, Van Cutsem E, Stein A, et al. ESMO Consensus Guidelines for management of patients with colon and rectal cancer. a personalized approach to clinical decision making. **Ann Oncol** 2012, 23:2479-516.

Schulten H, Al-Mansouri Z, Baghallab I, et al. Comparison of microarray expression profiles between follicular variant of papillary thyroid carcinomas and follicular adenomas of the thyroid. **BMC Genomics** 2015; 16:Suppl 1-S7.

Sconocchia G, Eppenberger S, Spagnoli GC, et al. NK cells and T cells cooperate during the clinical course of colorectal cancer. **Onc Immunology** 2014; 3:e952197.

Shaltiel IA, Krenning L, Bruinsma, W. The same, only different - DNA damage checkpoints and their reversal throughout the cell cycle. **J Cell Sci** 2015; 128:607-20.

Shi ZZ, Zhang YM, Shang L, et al. Genomic profiling of rectal adenoma and carcinoma by array-based comparative genomic hybridization. **BMC Med Genomics** 2012; 5:52.

Shin A, Joo J, Bak, et al. Site-specific risk factors for colorectal cancer in a Korean population. **PloS One** 2011; 6:e23196.

Shin J, Kim H.-J, Kim G, et al. Discovery of melanotransferrin as a serological marker of colorectal cancer by secretome analysis and quantitative proteomics. **J Proteome Res** 2014; 13:4919-31.

Siegel RL, Miller KD, Jemal A. Cancer statistics, 2018. **CA Cancer J Clin** 2018; 68:7-30.

Simon R, Lam A, Li MC, Ngan M, Menezes S, Zhao Y. Analysis of gene expression data using BRB-ArrayTools. **Cancer Inform** 2007; 3:11-7.

Slattery ML, Curtin K, Wolff RK, Herrick JS, Caan BJ, Samowitz W. Diet, physical activity, and body size associations with rectal tumor mutations and epigenetic changes. **Cancer Causes Control** 2010; 21:1237-45.

Steindorf K, Jedrychowski W, Schmidt M, et al. Case-control study of lifetime occupational and recreational physical activity and risks of colon and rectal cancer. **Eur J Cancer Prev** 2005; 14:363-71.

Stirzaker C, Taberlay PC, Statham AL, Clark SJ. Mining cancer methylomes: prospects and challenges. **Trends Genet** 2014; 30:75-84.

Sturla L, Tong M, Hebda N, et al. Aspartate- β -hydroxylase (ASPH): a potential therapeutic target in human malignant gliomas. **Heliyon** 2016; 2:e00203.

Sun M, Sun T, He Z, Xiong B. Identification of two novel biomarkers of rectal carcinoma progression and prognosis via co-expression network analysis. **Oncotarget** 2017; 8:69594-609.

Sun W, Sun Y, Zhu M, et al. The role of plasma cell-free DNA detection in predicting preoperative chemoradiotherapy response in rectal cancer patients. **Oncol Rep** 2014; 31:1466-72.

Supiot S, Gouraud W, Campion L, et al. Early dynamic transcriptomic changes during preoperative radiotherapy in patients with rectal cancer: a feasibility study. **World J Gastroenterol** 2013; 19:3249-54.

Svensson E, Grotmol T, Hoff G, Langmark F, Norstein J, Tretli S. Trends in colorectal cancer incidence in Norway by gender and anatomic site: an age-period-cohort analysis. **Eur J Cancer Prev** 2002; 11:489-95.

Syrovatkina V, Alegre KO, Dey R, Huang X. Regulation, signaling, and physiological functions of G-proteins. **J Mol Biol** 2016; 428:3850-68.

Tanaka K, Hirota T. Chromosomal instability: a common feature and a therapeutic target of cancer. **Biochim Biophys Acta** 2016; 1866:64-75.

Tato-Costa J, Casimiro S, Pacheco T, et al. Therapy-induced cellular senescence induces epithelial-to-mesenchymal transition and increases invasiveness in rectal cancer. **Clin Colorectal Cancer** 2016; 15:170-8.e3.

Telli ML, Jensen KC, Vinayak S, et al. Phase II study of gemcitabine, carboplatin, and iniparib as neoadjuvant therapy for triple-negative and BRCA1/2 mutation-associated breast cancer with assessment of a tumor-based measure of genomic instability: PrECOG 0105. **J Clin Oncol** 2015; 33:1895-901.

Terrazzino S, Agostini M, Pucciarelli S, et al. A haplotype of the methylenetetrahydrofolate reductase gene predicts poor tumor response in rectal cancer patients receiving preoperative chemoradiation. **Pharmacogenet Genomics** 2006; 16:817-24.

The Cancer Genome Atlas Network. Comprehensive molecular characterization of human colon and rectal cancer. **Nature** 2012a; 487:330-7.

The Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. **Nature** 2012b; 490:61-70.

Thomas AM, Jesus EC, Lopes A, et al. Tissue-associated bacterial alterations in rectal carcinoma patients revealed by 16S rRNA community profiling. **Front Cell Infec Microbiol** 2016; 6:179.

Thompson SL, Compton DA. Examining the link between chromosomal instability and aneuploidy in human cells. **J Cell Biol** 2008; 180:665-72.

Thune I, Lund E. Physical activity and risk of colorectal cancer in men and women. **Br J Cancer** 1996, 73:1134-40.

Tiwari N, Gheldof A, Tatarski M, Christofori G. EMT as the ultimate survival mechanism of cancer cells. **Sem Cancer Biol** 2012; 22:194-207.

Tomimaru Y, Mishra S, Safran H, et al. Aspartate- β -hydroxylase induces epitope-specific T cell responses in hepatocellular carcinoma. **Vaccine** 2015; 33:1256-66.

Trakarnsanga A, Gönen M, Shia J, et al. Comparison of tumor regression grade systems for locally advanced rectal cancer after multimodality treatment. **J Natl Cancer Inst** 2014; 106:dju248.

Tsang JS, Vencken S, Shara O, et al. Global DNA methylation is altered by neoadjuvant chemoradiotherapy in rectal cancer and may predict response to treatment - A pilot study. **Eur J Surg Oncol** 2014; 40:1459-66.

van de Velde CJ, Aristei C, Boelens PG, et al. EURECCA colorectal: multidisciplinary mission statement on better care for patients with colon and rectal cancer in Europe. **Eur J Cancer** 2013, 49:2784-90.

van Gorkom BA, Cats A, van der Meer R, et al. Effects of hemicolectomy on bile acid metabolism in relation to colon carcinogenesis in man. **Eur J Clin Invest** 1997; 27:589-94.

Veinotte L, Gebremeskel S, Johnston B. CXCL16-positive dendritic cells enhance invariant natural killer T cell-dependent IFN γ production and tumor control. **Oncol Immunology** 2016; 5:e1160979.

Vella V, Nicolosi ML, Giuliano S, et al. PPAR- γ agonists as antineoplastic agents in cancers with dysregulated IGF axis. **Front Endocrinol** 2017; 8:31.

Vencken PMLH, Kriege M, Hoogwerf D, et al. Chemosensitivity and outcome of BRCA1- and BRCA2-associated ovarian cancer patients after first-line chemotherapy compared with sporadic ovarian cancer patients. **Ann Oncol** 2011; 22:1346-52.

Villacis RAR, Miranda PM, Gomy I, et al. Contribution of rare germline copy number variations and common susceptibility loci in Lynch syndrome patients negative for mutations in the mismatch repair genes. **Int J Cancer** 2016; 138:1928-35.

Volgestein B, Fearon ER, Hamilton SR, et al. Genetic alterations during colorectal-tumor development. **N Engl J Med** 1988; 319:525-32.

Vymetalkova V, Vodicka P, Pardini B, et al. Epigenome-wide analysis of DNA methylation reveals a rectal cancer-specific epigenomic signature. **Epigenomics** 2016; 8:1193-207.

Wallin U, Rothenberger D, Lowry A, Luepker R, Mellgren A. CEA - a predictor for pathologic complete response after neoadjuvant therapy for rectal cancer. **Dis Colon Rectum** 2013; 56:859-68.

Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. **Nucleic Acids Res** 2010; 38:e164.

Watanabe T, Komuro Y, Kiyomatsu T, et al. Prediction of sensitivity of rectal cancer cells in response to preoperative radiotherapy by DNA microarray analysis of gene expression profiles. **Cancer Res** 2006; 66:3370-4.

Wei EK, Giovannucci E, Wu K, et al. Comparison of risk factors for colon and rectal cancer. **Int J Cancer** 2004; 108:433-42.

Wei J, Li G, Zhang J, et al. Integrated analysis of genome-wide DNA methylation and gene expression profiles identifies potential novel biomarkers of rectal cancer. **Oncotarget** 2016; 7:62547-58.

Weiser MR, Beets-Tan R, Beets G. Management of complete response after chemoradiation in rectal cancer. **Surg Oncol Clin North Am** 2014; 23:113-25.

Xie C, Powell C, Yao M, Wu J, Dong Q. Ubiquitin-conjugating enzyme E2C: a potential cancer biomarker. **Int J Biochem e Cell Biol** 2014; 47:113-17.

Xiong Y, You W, Wang R, Peng L, Fu Z. Prediction and validation of hub genes associated with colorectal cancer by integrating PPI network and gene expression data. **BioMed Res Int** 2017; 2017:2421459.

Xu K, Yang S, Zhao Y. Prognostic significance of BRCA mutations in ovarian cancer: an updated systematic review with meta-analysis. **Oncotarget** 2017; 8:19-23.

Yamauchi M, Morikawa T, Kuchiba A, et al. Assessment of colorectal cancer molecular features along bowel subsites challenges the conception of distinct dichotomy of proximal versus distal colorectum. **Gut** 2012; 61:847-854.

Yang X, Han H, De Carvalho Dd, et al. Gene body methylation can alter gene expression and is a therapeutic target in cancer. **Cancer Cell** 2014; 26:577-90.

Yang XM, You HY, Li Q, et al. CTHRC1 promotes human colorectal cancer cell proliferation and invasiveness by activating Wnt/PCP signaling. **Int J Clin Exp Pathol** 2015; 8:12793-801.

Yang YR, Follo MY, Cocco L, Suh P. The physiological roles of primary phospholipase C. **Adv Biol Regulation** 2013; 53:232-241.

Yokoi K, Yamashita K, Ishii S, et al. Comprehensive molecular exploration identified promoter DNA methylation of the CRBP1 gene as a determinant of radiation sensitivity in rectal cancer. **Br J Cancer** 2017; 116:1046-56.

Yoshida T, Kinugasa T, Akagi Y, et al. Decreased expression of claudin-1 in rectal cancer: a factor for recurrence and poor prognosis. **Anticancer Res** 2011; 31:2517-25.

Yu Z, Zhang C, Wang H, et al. Multidrug resistance-associated protein 3 confers resistance to chemoradiotherapy for rectal cancer by regulating reactive oxygen species and caspase-3-dependent apoptotic pathway. **Cancer Lett** 2014; 353:182-93.

Zack TI, Schumacher SE, Carter SL, et al. Pan-cancer patterns of somatic copy number alteration. **Nat Genet** 2013; 45:1134-40.

Zauber NP, Marotta SP, Berman E, et al. Molecular genetic changes associated with colorectal carcinogenesis are not prognostic for tumor regression following preoperative chemoradiation of rectal carcinoma. **Int J Radiat Oncol Biol Phys** 2009; 74:472-6.

Zeestraten ECM, Kuppen PJK, Van De Velde CJH, Marijnen CAM. Prediction in rectal cancer. **Semin Radiat Oncol** 2012; 22:175-83.

Zeng X, Yang P, Chen B, Jin X, Liu Y. Science direct quantitative secretome analysis reveals the interactions between epithelia and tumor cells by in vitro modulating colon cancer microenvironment. **J Proteomics** 2013; 89:51-70.

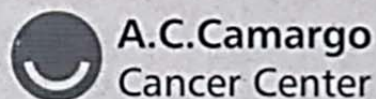
Zhu K, Zhao Q, Yue J, et al. GOLPH3 overexpression correlates with poor response to neoadjuvant therapy and prognosis in locally advanced rectal cancer. **Oncotarget** 2016; 7:68328-38.

Zimmer R, Thomas P. Mutations in the carcinoembryonic antigen gene in colorectal cancer patients: implications on liver metastasis. **Cancer Res** 2001; 61:2822-6.

Zorcolo L, Rosman AS, Restivo A, et al. Complete pathologic response after combined modality treatment for rectal cancer and long-term survival: a meta-analysis. **Ann Surg Oncol** 2012; 19:2822-32.

Zuo ZG, Zhang XF, Ye XZ, et al. Bioinformatic analysis of RNA-seq data unveiled critical genes in rectal adenocarcinoma. **Eur Rev Med Pharmacol Sci** 2016; 20:3017-25.

**Anexo 1 - Cartas de aprovação dos Comitês de Ética em Pesquisa – CEP
(A.C.Camargo Cancer Center e Hospital de Câncer de Barretos)**



**A.C. Camargo
Cancer Center**

Comitê de Ética em
Pesquisa - CEP

São Paulo, 16 de maio de 2014.

A
Dra. Silvia Regina Rogatto

Aluna: Luisa Matos do Canto Alvim (Doutorado).

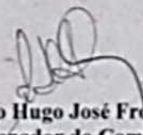
Ref.: Projeto de Pesquisa nº. 1884/14
"Identificação de marcadores moleculares de resposta ao tratamento neoadjuvante em
pacientes com adenocarcinoma de reto".

Os membros do Comitê de Ética em Pesquisa em Seres Humanos da Fundação Antonio Prudente – Hospital do Câncer - A.C. Camargo/SP, em sua última reunião de 29/04/2014, aprovaram a realização do projeto (datado de 11 de Março de 2014) e tomaram conhecimento dos seguintes documentos:

- Folha de Rosto para Pesquisa Envolvendo Seres Humanos;
- Termo de Compromisso do Pesquisador com Resoluções do Conselho Nacional de Saúde;
- Termo de Dispensa do Consentimento Livre e Esclarecido;
- Declaração Sobre os Dados Coletados, Publicação dos Dados e Propriedade das Informações Geradas;
- Declaração sobre Uso e Destino do Material Biológico, Publicação dos Dados e Propriedade das Informações Geradas;
- Cronograma do estudo;
- Orçamento Financeiro Detalhado;
- Declaração sobre a publicação dos dados de pesquisa utilizando amostras fornecidas e processadas pelo Biobanco do A.C. Camargo Cancer Center;
- Declaração de Infraestrutura e Instalações do CIPE;
- Declaração de Ciência e Comprometimento do Núcleo de Tumores Colorretais;
- Declaração de Ciência e Comprometimento do Biobanco;
- Declaração de Ciência e Comprometimento do Departamento de Endoscopia.

Informações a respeito do andamento do referido projeto deverão ser encaminhadas ao CEP dentro de 06 meses em relatório (modelo CEP).

Atenciosamente,


**Dr. Antônio Hugo José Froes Marques Campos
2º Vice-Coordenador do Comitê de Ética em Pesquisa**

1/1



FUNDAÇÃO PIO XII -
HOSPITAL DE CÂNCER DE
BARRETOS



PARECER CONSUBSTANCIADO DO CEP

DADOS DO PROJETO DE PESQUISA

Título da Pesquisa: Identificação de marcadores moleculares de resposta ao tratamento neoadjuvante em pacientes com adenocarcinoma de reto

Pesquisador: Cristovam Scapulatempo Neto

Área Temática: Genética Humana:
(Trata-se de pesquisa envolvendo Genética Humana que não necessita de análise ética por parte da CONEP.);

Versão: 2

CAAE: 30387414.4.2001.5437

Instituição Proponente: Fundação Pio XII

Patrocinador Principal: Conselho Nacional de Desenvolvimento Científico e Tecnológico
FUNDAÇÃO DE AMPARO A PESQUISA DO ESTADO DE SÃO PAULO

DADOS DO PARECER

Número do Parecer: 1.368.303

Apresentação do Projeto:

O projeto apresenta-se bem escrito. E trata-se de um estudo de tumores de cólon e de reto que diferem não só quanto à localização, mas também quanto ao tipo e a frequência de alterações genômicas e epigenômicas, assim como genes e vias moleculares envolvidos no seu desenvolvimento. Para pacientes com tumores em estágio inicial, nos dois tipos de câncer, é recomendada a exérese tumoral por cirurgia, podendo ser indicada quimioterapia pós-cirúrgica. Entretanto, em estágios mais avançados a conduta de tratamento dessas neoplasias diverge, sendo a radioterapia ou radioquimioterapia neoadjuvante recomendada apenas para pacientes com tumores de reto em estágio II ou III. A resposta clínica ao tratamento é avaliada pela presença de tumor residual. Porém, apenas 25% a 50% dos pacientes que apresentam resposta clínica completa também apresentarão resposta patológica completa, sendo essa detectada apenas por cirurgia. Estudos moleculares tem sido realizados em carcinomas de reto para a busca de marcadores com o intuito de identificar pacientes respondedores ao tratamento neoadjuvante. Até o presente momento, o antígeno carcinoembrionário (CEA) é o único marcador utilizado na prática clínica como fator prognóstico de acompanhamento pós-cirúrgico nesses pacientes. Este projeto tem

Endereço: Rua Antenor Duarte Vilela, 1331

Bairro: Dr. Paulo Prata

CEP: 14.784-400

UF: SP

Município: BARRETOS

Telefone: (17)3321-0347

Fax: (17)3321-6600

E-mail: cep@hcancerbarretos.com.br



FUNDAÇÃO PIO XII - HOSPITAL DE CÂNCER DE BARRETOS



Continuação do Parecer: 1.368.303

como objetivo avaliar os padrões de metilação e de alterações genômicas, assim como a expressão de transcritos em biópsias de pacientes diagnosticados com adenocarcinoma de reto submetidos à terapia neoadjuvante. Todas as biópsias foram coletadas previamente ao tratamento quimio- e ou radioterápico. A análise integrada dos dados genômicos, transcriptômicos e de metilação tem potencial para evidenciar drivers candidatos a marcadores de resposta à terapia e para revelar alvos terapêuticos mais eficazes para o tratamento de pacientes com câncer de reto.

Objetivo da Pesquisa:

Objetivo Primário do projeto informado nas Informações Básicas da Plataforma Brasil:

Identificar alterações genéticas e epigenéticas que corroborem para a elucidação dos mecanismos de resistência ao tratamento de rádio e quimioterapia em pacientes com adenocarcinoma de reto esporádicos, atendidos no AC Camargo Cancer Center.

Objetivos informados no corpo do projeto anexada a Plataforma

Objetivo Primário

Identificar alterações genéticas e epigenéticas que corroborem para a elucidação dos mecanismos de resistência ao tratamento de rádio e quimioterapia em pacientes com adenocarcinoma de reto esporádicos, atendidos no AC Camargo Cancer Center, SP e no Hospital de Câncer de Barretos, Barretos, SP.

Objetivos específicos

- Identificar alterações no número de cópias (ganhos e perdas) genômicas em tumores de reto de pacientes com resposta completa e com resistência ao tratamento rádio e quimioterápico;
- Identificar transcritos diferencialmente expressos entre tumores de reto de pacientes segundo a resposta à rádio e quimioterapia neoadjuvante;
- Identificar perfis de metilação em tumores de reto de pacientes segundo a resposta ao tratamento rádio e quimioterápico;
- Integrar os dados de expressão de transcritos com o perfil de metilação e com o de alterações no número de cópias de DNA para identificar candidatos a drivers moleculares;
- Analisar a função dos genes identificados como alterados, avaliando sua participação nas vias biológicas e possível contribuição para resistência ao tratamento rádio e quimioterápico.
- Selecionar candidatos e confirmar os achados em um grupo independente de casos.

Endereço: Rua Antenor Duarte Vilela, 1331

Bairro: Dr. Paulo Prata

CEP: 14.784-400

UF: SP

Município: BARRETOS

Telefone: (17)3321-0347

Fax: (17)3321-6600

E-mail: cep@hcancerbarretos.com.br



FUNDAÇÃO PIO XII -
HOSPITAL DE CÂNCER DE
BARRETOS



Continuação do Parecer: 1.368.303

Avaliação dos Riscos e Benefícios:

Riscos e Benefícios especificados pelo pesquisador na Plataforma Brasil

Riscos:

As amostras que já foram coletadas fazem parte do banco de amostras do Hospital AC Camargo, assim como aquelas que serão coletadas no serviço de endoscopia, não apresentando riscos adicionais aos pacientes.

Benefícios:

As amostras que já foram coletadas fazem parte do banco de amostras do Hospital AC Camargo, assim como aquelas que serão coletadas no serviço de endoscopia, não apresentando benefícios adicionais aos pacientes.

Comentários e Considerações sobre a Pesquisa:

Sem comentários e considerações.

Considerações sobre os Termos de apresentação obrigatória:

Todos os documentos foram apresentados de forma adequada.

Recomendações:

Conclusões ou Pendências e Lista de Inadequações:

1. O estudo prevê a utilização de 22 amostras de tecido congelado e 22 blocos de parafina de pacientes do HCB. Para as extrações de RNA serão utilizados os blocos de parafina mas não está claro na metodologia do projeto quantos cortes serão utilizados, espessura e se sobrar material remanescente do paciente em Biobanco, bem como se seguirá as normas do pré-estabelecidas pelo CRAMMB do HCB. Solicita-se especificar o mencionado acima.

RESPOSTA: SERÃO REALIZADOS 11 CORTES DE 10 MICRA DE ESPESSURA DE CADA BLOCO DE PARAFINA PARA EXTRAÇÃO. DURANTE A REALIZAÇÃO DOS CORTES SERÁ TOMADO O CUIDADO DE NÃO SE ESGOTAR O BLOCO DE PARAFINA. O POP DO CRAMMB AUTORIZA A REALIZAÇÃO DE 10 CORTES.

ANÁLISE CEP: PENDÊNCIA ATENDIDA.

2. Solicita-se esclarecer se as amostras do HCB serão utilizadas a partir de Julho de 2014 contendo

Endereço: Rua Antenor Duarte Vilela, 1331
Bairro: Dr. Paulo Prata **CEP:** 14.784-400
UF: SP **Município:** BARRETOS
Telefone: (17)3321-0347 **Fax:** (17)3321-6600 **E-mail:** cep@hcancerbarretos.com.br



FUNDAÇÃO PIO XII -
HOSPITAL DE CÂNCER DE
BARRETOS



Continuação do Parecer: 1.368.303

o novo TCLE do Biobanco.

RESPOSTA: Não serão utilizadas amostras colhidas à partir de Julho de 2014.

ANÁLISE CEP: PENDÊNCIA ATENDIDA.

Considerações Finais a critério do CEP:

O Comitê de Ética em Pesquisa da Fundação Pio XII – Hospital do Câncer de Barretos de acordo com as atribuições definidas na Resolução CNS 466/2012 e na Norma Operacional Nº 001/2013 do CNS, e após a análise das respostas as pendências emitidas, manifesta-se pela APROVAÇÃO do projeto de pesquisa proposto.

Solicitamos que sejam encaminhados ao CEP:

1. Relatórios semestrais, sendo o primeiro previsto para 14/06/2016.
2. Comunicar toda e qualquer alteração do Projeto e Termo de Consentimento Livre e Esclarecido. Nestas circunstâncias a inclusão de participantes deve ser temporariamente interrompida até a aprovação do Comitê de Ética em Pesquisa.
3. Comunicar imediatamente ao Comitê qualquer Evento Adverso Grave ocorrido durante o desenvolvimento do estudo.
4. Para projetos que utilizam amostras criopreservadas, procurar o BIOBANCO para início do processamento.
5. Os dados individuais de todas as etapas da pesquisa devem ser mantidos em local seguro por 5 anos, após conclusão da pesquisa, para possível auditoria dos órgãos competentes.
6. Este projeto está cadastrado no CEP-HCB sob o número 1030/2015.

Este parecer foi elaborado baseado nos documentos abaixo relacionados:

Tipo Documento	Arquivo	Postagem	Autor	Situação
Informações Básicas do Projeto	PB_INFORMAÇÕES_BASICAS_DO_PROJETO_612037.pdf	30/11/2015 15:43:43		Aceito
Outros	Carta_de_resposta.pdf	30/11/2015 15:25:25	Cristovam Scapulatempo Neto	Aceito
Outros	Carta_de_resposta.docx	30/11/2015	Cristovam	Aceito

Endereço: Rua Antenor Duarte Vilela, 1331

Bairro: Dr. Paulo Prata

CEP: 14.784-400

UF: SP

Município: BARRETOS

Telefone: (17)3321-0347

Fax: (17)3321-6600

E-mail: cep@hcancerbarretos.com.br



FUNDAÇÃO PIO XII -
HOSPITAL DE CÂNCER DE
BARRETOS



Continuação do Parecer: 1.368.303

Outros	Carta_de_resposta.docx	15:24:48	Scapulatempo Neto	Aceito
Folha de Rosto	Folha_de_rosto.pdf	09/11/2015 08:18:11	CRISTOVAM SCAPULATEMPO NETO	Aceito
Projeto Detalhado / Brochura Investigador	PROJETO_CA_DE_RETO.docx	06/11/2015 22:47:18	CRISTOVAM SCAPULATEMPO NETO	Aceito
Outros	Ciencia_e_autorizacao_departamentos. pdf	23/10/2015 15:25:27	CRISTOVAM SCAPULATEMPO NETO	Aceito
Outros	Levantamento_de_amostras_criopreserv adas_BIOBANCO.pdf	23/10/2015 09:36:40	CRISTOVAM SCAPULATEMPO NETO	Aceito
Outros	Fonte_financiamento.pdf	23/10/2015 09:35:52	CRISTOVAM SCAPULATEMPO NETO	Aceito
Outros	Corresponsabilidade_pesquisador_exter no.pdf	23/10/2015 09:35:09	CRISTOVAM SCAPULATEMPO NETO	Aceito
Outros	Cadastro_projeto.pdf	23/10/2015 09:33:59	CRISTOVAM SCAPULATEMPO NETO	Aceito
Outros	Responsabilidade_pesquisador.pdf	23/10/2015 09:30:59	CRISTOVAM SCAPULATEMPO NETO	Aceito
Outros	MABIN.pdf	23/10/2015 09:30:07	CRISTOVAM SCAPULATEMPO NETO	Aceito
Outros	Formulário para submissão do projeto.pdf	31/03/2014 18:18:40		Aceito
Outros	Declaração sobre uso e destino do material biológico.pdf	31/03/2014 18:12:30		Aceito
Outros	Declaração sobre os dados coletados.pdf	31/03/2014 18:12:01		Aceito
Outros	Orçamento financeiro detalhado.pdf	28/03/2014 15:01:40		Aceito
Outros	Cronograma de estudo.pdf	28/03/2014 15:01:23		Aceito
Outros	Identificação dos currículos na Plataforma Lattes.pdf	28/03/2014 15:00:26		Aceito
Outros	Declaração de Ciência e Comprometimento Samuel.pdf	28/03/2014 14:59:58		Aceito
Outros	Declaração de Ciência e comprometimento Endoscopia.pdf	28/03/2014 14:59:48		Aceito
Outros	Declaração de Ciência e Comprometimento Biobanco.pdf	28/03/2014 14:59:39		Aceito
Outros	Declaração de Infraestrutura e	28/03/2014		Aceito

Endereço: Rua Antenor Duarte Vilela, 1331

Bairro: Dr. Paulo Prata

CEP: 14.784-400

UF: SP

Município: BARRETOS

Telefone: (17)3321-0347

Fax: (17)3321-6600

E-mail: cep@hcancerbarretos.com.br



FUNDAÇÃO PIO XII -
HOSPITAL DE CÂNCER DE
BARRETOS



Continuação do Parecer: 1.368.303

Outros	Instalações.pdf	14:59:09		Aceito
Outros	Termo de compromisso do pesquisador com resoluções do conselho nacional de saúde.pdf	28/03/2014 14:58:26		Aceito
TCLE / Termos de Assentimento / Justificativa de Ausência	Termo de dispensa do consentimento livre e esclarecido.pdf	28/03/2014 14:57:19		Aceito
Outros	Declaração sobre a publicação dos dados de pesquisa utilizando amostras fornecidas e processadas pelo biobanco do ac camargo cancer center.pdf	19/03/2014 10:01:37		Aceito

Situação do Parecer:

Aprovado

Necessita Apreciação da CONEP:

Não

BARRETOS, 15 de Dezembro de 2015

Assinado por:
Thiago Buosi Silva
(Coordenador)

Endereço: Rua Antenor Duarte Vilela, 1331
Bairro: Dr. Paulo Prata **CEP:** 14.784-400
UF: SP **Município:** BARRETOS
Telefone: (17)3321-0347 **Fax:** (17)3321-6600 **E-mail:** cep@hcancerbarretos.com.br

Anexo 2 - Cobertura média, número total de *reads*, porcentagem de *reads* alinhados e a fração das sequências com cobertura de pelo menos 20, 30 e 100 vezes obtidas com o sequenciador NextSeq para as 31 amostras de adenocarcinomas de reto.

Caso	Cobertura Média	# Total de Reads	% Reads alinhados	% de regiões cobertas		
				x20	x30	x100
RET2	139,18	1580974	99,75351	93,9	92,0	70,1
RET3	187,43	2275089	99,71979	95,3	94,1	81,7
RET4	165,01	2015777	99,72373	94,6	93,0	75,5
RET5	171,29	1969516	99,73821	95,1	93,8	78,8
RET6	156,61	2000552	99,72783	94,8	93,5	77,1
RET7	167,68	2159331	99,74057	94,8	93,3	76,6
RET8	203,50	3092943	99,72227	95,1	94,0	83,0
RET9	169,58	2216756	99,73660	95,0	93,6	77,8
RET10	149,17	1851027	99,74414	94,6	92,8	73,3
RET11	178,31	2244858	99,73914	95,1	93,7	79,2
RET12	196,08	2429632	99,71675	95,8	94,8	83,7
RET13	163,42	2554653	99,77797	94,5	93,2	77,1
RET15	192,04	2874946	99,78274	94,6	93,3	80,4
RET17	152,54	2299094	99,78109	94,3	92,9	74,1
RET18	187,31	2749994	99,74185	95,0	93,9	80,2
RET19	172,71	3126395	99,76529	94,5	93,2	78,2
RET20	171,64	2623036	99,78231	94,7	93,5	78,8
RET24	109,15	1689621	99,68182	92,5	89,6	56,6
RET25	284,42	7327494	99,70383	96,1	95,3	88,8
RET26	163,35	1967094	99,73077	94,9	93,6	77,5
RET27	163,66	2065246	99,74236	94,7	93,3	77,9
RET28	162,27	1908455	99,74812	94,3	92,5	75,4
RET29	164,86	1852606	99,73011	94,7	93,3	76,0
RET30	168,09	1892555	99,68429	95,0	93,8	78,1
RET31	165,79	1979443	99,75230	94,9	93,3	77,2
RET32	191,96	2262387	99,75261	95,1	93,9	81,0
RET33	178,83	2112018	99,75123	94,8	93,4	79,0
RET34	159,84	1812336	99,74955	94,4	92,9	74,6
RET41	153,07	1911588	99,74728	95,0	93,4	74,8
RET43	184,80	2400005	99,74158	95,4	94,2	79,3
RET44	171,82	1968201	99,71111	94,5	93,1	77,3

Anexo 3 - Manuscrito I - submetido ao periódico International Journal of Cancer

Research article - Tumor Markers and Signatures

Influence of the genomic instability and mutational profile in the response to therapy in locally advanced rectal carcinomas

Luisa Matos do Canto¹, Simon J Larsen², Bruna Elisa Catin Kupper³, Maria Dirlei Ferreira de Souza Begnami⁴, Cristovam Scapulatempo Neto⁵, Annabeth Høgh Petersen⁶, Mads M. Aagaard⁶, Jan Baumbach⁷, Samuel Aguiar Jr³, Silvia R. Rogatto^{6,8§}

¹International Research Center - CIPE, A.C.Camargo Cancer Center, São Paulo, Brazil;

²Department of Mathematics and Computer Science, University of Southern Denmark, Odense, DK

³Department of Pelvic Surgery, A.C.Camargo Cancer Center, Sao Paulo, Brazil

⁴Department of Pathology, A.C.Camargo Cancer Center, São Paulo, Brazil;

⁵Molecular Oncology Research Center, Barretos, and Diagnósticos da América (DASA), São Paulo, Brazil.

⁶Department of Clinical Genetics, Vejle Hospital, Institute of Regional Health Research, University of Southern Denmark, DK

⁷TUM School of Life Sciences Weihenstephan, Technical University of Munich, Germany

⁸Danish Colorectal Cancer Center South, Denmark

Short title: Genomic profile of rectal carcinomas

§ Corresponding author:

Professor Silvia Regina Rogatto,

Department of Clinical Genetics, Vejle Hospital

Institute of Regional Health Research, University of Southern Denmark

Beriderbakken 4, 7100 Vejle - DK

Telephone number: +45 -7940 6669

silvia.regina.rogatto@rsyd.dk

Anexo 4 - Proteínas potencialmente secretadas em câncer de reto com e sem resposta patológica completa (pCR e pIR, respectivamente) identificadas por análise *in silico* usando as ferramentas de predição dos CBS servers (SecretomeP 2.0, SignalP 4.1 e TargetP 1.1: <http://www.cbs.dtu.dk/services/>).

Resposta Patológica Completa (IDs.)				Resposta Patológica Incompleta (IDs.)											
UniProt	Gene	UniProt	Gene	UniProt	Gene	UniProt	Gene	UniProt	Gene	UniProt	Gene	UniProt	Gene	UniProt	Gene
P36896	ACV1B	P03956	MMP1	P19876	CXCL3	P03956	MMP1	Q06643	TNFC	Q86XA0	MET23	Q8IWS0	PHF6		
Q9Y4X0	AMMR1	P39900	MMP12	Q68DV7	RNF43	P39900	MMP12	Q9Y4X0	AMMR1	Q9P1F3	ABRAL	P24666	PPAC		
Q9NXR5	ANR10	Q8TAD7	OCC1	P15692	VEGFA	P78334	GBRE	Q9NXR5	ANR10	Q53H80	AKIR2	P46778	RL21		
Q12797	ASPH	Q8WY50	PLAC4	P15018	LIF	Q15582	BGH3	Q5RI15	COX20	P98196	AT11A	P61927	RL37		
P61769	B2MG	Q8N131	PORIM	P15514	AREG	Q8N131	PORIM	P43268	ETV4	Q9H7F0	AT133	Q9Y508	RN114		
P16442	BGAT	P10114	RAP2A	P08476	INHBA	P16070	CD44	Q9C009	FOXQ1	Q8IU99	CAHM1	P62277	RS13		
P56880	CLD20	P15880	RS2	P78556	CCL20	Q8WUJ3	CEMIP	Q8TAD7	OCC1	Q9BWT1	CDCA7	P62841	RS15		
Q5RI15	COX20	Q8NFV5	SPDE1	P56880	CLD20	P15529	MCP	Q8WY50	PLAC4	P13569	CFTR	P62701	RS4X		
Q96CG8	CTHR1	A6NIY4	SPDE5	Q9Y5Z0	BACE2	P08581	MET	P10114	RAP2A	Q5VZT2	CJ113	P08579	RU2B		
P19875	CXCL2	O15269	SPTC1	P25063	CD24	Q96HE7	ERO1A	P15880	RS2	P78545	ELF3	P55011	S12A2		
P80162	CXCL6	P01375	TNFA	P80162	CXCL6	P08174	DAF	P01375	TNFA	P32189	GLPK	Q5TGI4	SAMD5		
P21860	ERBB3	Q06643	TNFC	Q08722	CD47	Q99988	GDF15	O14562	UBFD1	P68431	H31	Q8TCJ2	STT3B		
Q14534	ERG1	O14562	UBFD1	P17301	ITA2	Q13217	DNJC3	Q9UPY5	XCT	Q71DI3	H32	Q15185	TEBP		
P43268	ETV4	Q9NYU1	UGGG2	P19875	CXCL2	P16444	DPEP1	O14980	XPO1	P09067	HXB5	Q92664	TF3A		
Q9C009	FOXQ1	P00749	UROK	P21860	ERBB3	P00749	UROK	Q14592	ZN460	P17482	HXB9	P02786	TFR1		
P78334	GBRE	P41221	WNT5A	P40199	CEAM6	P20382	MCH	P08254	MMP3	P41134	ID1	Q15785	TOM34		
Q96P67	GPR82	Q9UPY5	XCT	O14944	EREG	Q96DZ1	ERLEC	Q9HD45	TM9S3	P01584	IL1B	O00762	UBE2C		
P09341	GROA	O14980	XPO1	P09341	GROA	P23229	ITA6	P34910	EVI2B	Q9BYG3	MK67I	O76080	ZFAN5		
Q9NPF7	IL23A	Q14592	ZN460	Q9NPF7	IL23A	A6NM11	L37A2	Q9Y5L2	HLPDA	Q06416	P5F1B	Q9NS37	ZHANG		
P10145	IL8			P10145	IL8	Q96P67	GPR82	Q8TBQ9	KISHA	O14737	PDCD5	P17029	ZKSC1		
O95868	LY66D			P42830	CXCL5	O15269	SPTC1	O00453	LST1	P53611	PGTB2				

IDs.: Número de identificação das proteínas e nome dos genes dados pelo UniProtKB (<http://www.uniprot.org>)

Anexo 5 - Proteínas potencialmente secretadas por tumores de reto de pacientes com resposta patológica completa (17) e incompleta (56) que já foram descritas em vesículas extracelulares (VE) provenientes de amostras de carcinomas colorretais.

Símbolo	Nome da Proteína
Resposta Patológica Completa	
<i>AMMR1*</i>	AMME syndrome candidate gene 1 protein
<i>ANR10*</i>	Ankyrin repeat domain-containing protein 10
<i>ASPH</i>	Aspartyl/asparaginyl beta-hydroxylase (Aspartate beta-hydroxylase)
<i>B2MG</i>	Beta-2-microglobulin
<i>CXCL6*</i>	C-X-C motif chemokine 6 (Chemokine alpha 3) (CKA-3) (Small-inducible cytokine B6)
<i>ERBB3*</i>	Receptor tyrosine-protein kinase erbB-3 (Proto-oncogene-like protein c-ErbB-3)
<i>ERG1</i>	Squalene monooxygenase
<i>ETV4*</i>	ETS translocation variant 4 (Adenovirus E1A enhancer-binding protein) (E1A-F)
<i>FOXQ1*</i>	Forkhead box protein Q1 (HNF-3/forkhead-like protein 1)
<i>IL8*</i>	Interleukin-8 (C-X-C motif chemokine 8)
<i>MMP1*</i>	Interstitial collagenase (Fibroblast collagenase) (Matrix metalloproteinase-1) (MMP-1)
<i>RS2*</i>	40S ribosomal protein S2
<i>SPTC1*</i>	Serine palmitoyltransferase 1
<i>UGGG2</i>	UDP-glucose:glycoprotein glucosyltransferase 2
<i>WNT5A</i>	Protein Wnt-5a
<i>XCT*</i>	Cystine/glutamate transporter
<i>XPO1*</i>	Exportin-1
Resposta Patológica Incompleta	
<i>ABRAL</i>	Costars family protein ABRACL (ABRA C-terminal-like protein)
<i>AMMR1*</i>	AMME syndrome candidate gene 1 protein
<i>ANR10*</i>	Ankyrin repeat domain-containing protein 10
<i>AREG</i>	Amphiregulin (AR) (Colorectum cell-derived growth factor) (CRDGF)
<i>AT11A</i>	Probable phospholipid-transporting ATPase IH
<i>AT133</i>	Probable cation-transporting ATPase 13A3 (ATPase family homolog up-regulated in senescence cells 1)
<i>BACE2</i>	Beta-secretase 2 (Membrane-associated aspartic protease 1)
<i>BGH3</i>	Transforming growth factor-beta-induced protein ig-h3
<i>CCL20</i>	C-C motif chemokine 20 (Beta-chemokine exodus-1) (Macrophage inflammatory protein 3 alpha)
<i>CD44</i>	CD44 antigen (Extracellular matrix receptor III) (ECMR-III)
<i>CD47</i>	Leukocyte surface antigen CD47 (Antigenic surface determinant protein OA3) (Integrin-associated protein)
<i>CEAM6</i>	Carcinoembryonic antigen-related cell adhesion molecule 6
<i>CEMIP</i>	Cell migration-inducing and hyaluronan-binding protein
<i>CXCL3</i>	C-X-C motif chemokine 3 (Growth-regulated protein gamma) (GRO-gamma)
<i>CXCL5</i>	C-X-C motif chemokine 5 (Small-inducible cytokine B5)
<i>CXCL6*</i>	C-X-C motif chemokine 6 (Chemokine alpha 3) (CKA-3) (Small-inducible cytokine B6)
<i>DAF</i>	Complement decay-accelerating factor (CD antigen CD55)

DPEP1	Dipeptidase 1 (Dehydropeptidase-I) (Microsomal dipeptidase) (Renal dipeptidase) (hRDP)
ERBB3*	Receptor tyrosine-protein kinase erbB-3 (Proto-oncogene-like protein c-ErbB-3)
ERO1A	ERO1-like protein alpha (ERO1-L) (Endoplasmic reticulum oxidoreductase alpha)
ETV4*	ETS translocation variant 4 (Adenovirus E1A enhancer-binding protein) (E1A-F)
EVI2B	Protein EVI2B (Ecotropic viral integration site 2B protein homolog)
FOXQ1*	Forkhead box protein Q1 (HNF-3/forkhead-like protein 1)
GDF15	Growth/differentiation factor 15 (GDF-15) (Macrophage inhibitory cytokine 1)
GLPK	Glycerol kinase (GK) (Glycerokinase) (ATP:glycerol 3-phosphotransferase)
HXB9	Homeobox protein Hox-B9
ID1	DNA-binding protein inhibitor ID-1
IL8*	Interleukin-8 (C-X-C motif chemokine 8)
INHBA	Inhibin beta A chain (Activin beta-A chain)
ITA2	Integrin alpha-2 (CD49 antigen-like family member B)
ITA6	Integrin alpha-6 (CD49 antigen-like family member F)
LIF	Leukemia inhibitory factor
MCP	Membrane cofactor protein (TLX) (Trophoblast leukocyte common antigen)
MET	Hepatocyte growth factor receptor (Proto-oncogene c-Met) (Tyrosine-protein kinase Met)
MMP1*	Interstitial collagenase (Fibroblast collagenase) (Matrix metalloproteinase-1) (MMP-1)
PDCD5	Programmed cell death protein 5
PGTB2	Geranylgeranyl transferase type-2 subunit beta
PHF6	PHD finger protein 6
PPAC	Low molecular weight phosphotyrosine protein phosphatase
RN114	E3 ubiquitin-protein ligase RNF114
RS13	40S ribosomal protein S13
RS2*	40S ribosomal protein S2
RS4X	40S ribosomal protein S4, X isoform
RU2B	U2 small nuclear ribonucleoprotein B
S12A2	Solute carrier family 12 member 2 (Basolateral Na-K-Cl symporter)
SPTC1*	Serine palmitoyltransferase 1
TEBP	Prostaglandin E synthase 3 (Telomerase-binding protein p23)
TF3A	Transcription factor IIIA
TFR1	Transferrin receptor protein 1
TM9S3	Transmembrane 9 superfamily member 3
TOM34	Translocase of outer membrane 34 kDa subunit
UBE2C	Ubiquitin-conjugating enzyme E2 C
XCT*	Cystine/glutamate transporter
XPO1*	Exportin-1
ZFAN5	AN1-type zinc finger protein 5
ZKSC1	Zinc finger protein with KRAB and SCAN domains 1

* Proteínas potencialmente secretadas encontradas nos dois grupos (pCR e pIR). Em **negrito**, proteínas descritas no banco de dados Cancer Secretome (<http://www.cancersecretome.org/>), usado como filtro adicional para pacientes com pIR.