

Fundação Pio XII-Hospital de Câncer de Barretos

**Wellington dos Santos**

**Perfil mutacional de um painel de oncogenes e genes supressores tumorais no câncer colorretal e suas lesões precursoras.**

Tese de doutorado apresentada ao Programa de Pós-Graduação da Fundação PIO XII – Hospital de Câncer de Barretos para obtenção do título em Ciências da Saúde.

Área de Concentração: Oncologia

**Orientadora:** Denise Peixoto Guimarães

**Co-orientador:** Rui Manuel Vieira Reis

**Barretos – São Paulo**

**2020**

## **SUPORTE À PESQUISA POR AGÊNCIA DE FOMENTO**

Este trabalho foi desenvolvido com bolsa de doutorado da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) por meio do Programa de Apoio à Pós-Graduação (PROAP)

As opiniões, hipóteses e conclusões ou recomendações expressas neste material são de responsabilidade dos autores e não necessariamente refletem a visão da CAPES

Esta tese foi elaborada e está apresentada de acordo com as normas da Pós-Graduação do Hospital de Câncer de Barretos – Fundação Pio XII, baseando-se no Regimento do Programa de Pós-Graduação em Oncologia e no Manual de Apresentação de Dissertações e Testes do Hospital de Câncer de Barretos. Os pesquisadores declaram ainda que este trabalho foi realizado em concordância com o Código de Boas Práticas Científicas (FAPESP), não havendo nada em seu conteúdo que possa ser considerado como plágio, fabricação ou falsificação de dados. As opiniões, hipóteses e conclusões ou recomendações expressas neste material são de responsabilidade dos autores e não necessariamente refletem a visão da Fundação Pio XII – Hospital de Câncer de Barretos.

Embora o Núcleo de Apoio ao Pesquisador do Hospital de Câncer de Barretos tenha realizado as análises estatísticas e orientado sua interpretação, a descrição da metodologia estatística, a apresentação dos resultados e suas conclusões são de inteira responsabilidade dos pesquisadores envolvidos estudo.

Os pesquisadores declaram não ter qualquer conflito de interesse relacionado ao presente estudo.

Dedico este trabalho à minha companheira e amada Cristiane...

## AGRADECIMENTOS

É difícil elencar e agradecer todos que contribuíram para que este trabalho fosse concluído e, sem dúvida, sem eles nada teria sido possível. Deixo aqui meu singelo agradecimento a todos que, direta ou indiretamente, contribuíram para conclusão desse trabalho.

À minha eterna amiga e companheira e amada **Cris**, que esteve ao meu lado em todos os momentos, nunca permitindo que eu deixasse de enxergar o futuro de maneira positiva e, mais do que isso, nunca permitindo que eu me deixasse sucumbir ao pessimismo interno. Obrigado por me fazer lutar pelo melhor de mim, e por me permitir enxergar um futuro ao seu lado. Obrigado por se manter firme nos momentos de ausência. Obrigado por ser minha mulher, minha amiga, minha conselheira (quando eu permito) e, acima de tudo, eterna paixão. Obrigado por me fazer olhar para mim mesmo da mesma forma que me vê.

À minha **família**, que sempre esteve disposta a me ajudar em todos os momentos, sempre preocupados com meu bem-estar e saúde. À minha querida e amada **avó**, pelo amor, preocupação e dedicação que distribui a todos, principalmente a este neto. Um agradecimento especial à minha **mãe**, que sempre vou levar em meu coração e que sempre guiará meus passos, mesmo estando ausente; seu sorriso e olhar doces irão me acompanhar por onde quer que eu vá.

À minha orientadora, **Dra. Denise Peixoto Guimarães**, por todo o conhecimento e incentivo durante todo este período. Obrigado por me permitir continuar trabalhando, mesmo nos momentos delicados nos quais eu não acreditava no meu trabalho. Obrigado pelos ensinamentos e discussões que me proporcionaram grande crescimento profissional e pessoal.

Ao meu co-orientador, **Dr. Rui Reis**, que, muito além de me auxiliar durante todo o trabalho, acreditou no meu potencial e sempre me fez extrair o melhor de mim. Obrigado por me emprestar o entusiasmo com que vê a ciência e a busca por novos conhecimentos.

Aos membros da Banca de acompanhamento, **Dra Patricia Pintor dos Reis** e **Dr. René Aloísio da Costa Vieira**, pela disponibilidade e por todas considerações que ajudaram a enriquecer o trabalho ao longo de todo o processo.

À **Banca de Defesa**, pelo tempo e disponibilidade cedidos para avaliação desta tese.

Ao **Hospital de Amor**, pela construção de um corpo de excelência que permite geração de conhecimento de ponta ao mesmo tempo em que preza pelo cuidado aos pacientes.

À todos os **amigos e funcionários do Centro de Pesquisa em Oncologia Molecular (CPOM)**, pela convivência e possibilidade de troca de experiência. Um agradecimento especial aos **biólogos do CPOM**, pela disponibilidade em ajudar sempre que necessário.

Às pesquisadoras e amigas **Ana Carolina de Carvalho** e **Letícia Ferro Leal**, as quais aprendi a admirar e nas quais busco inspiração como pesquisador.

Ao **Diagnóstico Molecular**, especialmente ao **Gustavo**, pela ajuda e troca de experiência e discussão, e pelo fornecimento de dados moleculares utilizados no presente trabalho.

Ao **Núcleo de Epidemiologia e Bioestatística (NEB)**, pela ajuda na condução das análises estatísticas. Em especial ao bioestatístico **Marco Antônio de Oliveira**, que me auxiliaram nas análises e interpretação dos resultados do presente trabalho.

Ao **Departamento de Patologia** e ao **Biobanco** do Hospital de Amor, pela ajuda com as amostras utilizadas. Especialmente ao **Dr. Marcus Matsushita**, pelo auxílio na classificação das amostras e à **Bia**, pela dedicação e competência na busca e corte dos blocos de parafina.

Ao **Programa de Pós-graduação**, por todo auxílio e suporte durante todo o período.

Ao **SAME**, pela prontidão em disponibilizar os prontuários médicos.

A todos os amigos que sempre estiveram ao meu lado e sempre me auxiliaram a seguir em frente. À família **Batista** que, mesmo distante, sempre me guiaram e apoiaram. Ao meu amigo **Marcos Alves** por todo o companheirismo e filosofia de boteco. Ao meu grande amigo **André**, cuja amizade se iniciou no laboratório e transbordou para todos os cantos da vida; sua amizade me inspira a sempre buscar o melhor.

Ao **Coral Acordes Vocais**, que me ensinou encontrar o prazer na música. Em especial às amizades conquistadas durante as cantorias: **Dani, Rafa, Renata, Ângela** e **José Renato**, vocês conquistaram um lugar especial no meu coração.

Finalmente, deixo meus agradecimentos a todos que, direta ou indiretamente, fizeram parte desta jornada e me auxiliaram a trilhar o caminho até aqui. Muito obrigado.

## ÍNDICE

INTRODUÇÃO	1
História natural das lesões precursoras ao CCR	2
Características clínico-patológicas do câncer colorretal	3
Alterações moleculares do CCR	4
Impacto clínico da identificação de alterações moleculares do CCR	7
JUSTIFICATIVA	9
OBJETIVOS	10
1. Geral	10
2. Específicos	10
PARTE 1: CARACTERIZAÇÃO DO PERFIL MUTACIONAL DO CÂNCER COLORRETAL	
COLORRETAL	12
MATERIAL E MÉTODOS	12
1. População com câncer colorretal (CCR)	12
2. Extração de DNA	12
3. Análise do perfil mutacional	12
RESULTADOS	15
1. Características clínicas e anátomo-patológicas dos pacientes com CCR	15
2. Caracterização mutacional das amostras de câncer colorretal	17
DISCUSSÃO	35
CONCLUSÕES	40
PARTE 2: CARACTERIZAÇÃO DO PERFIL MUTACIONAL DAS LESÕES PRECURSORAS DO CÂNCER COLORRETAL	
PRECURSORAS DO CÂNCER COLORRETAL	41
MATERIAL E MÉTODOS	41
1. População com lesões precursoras do câncer colorretal (CCR)	41
2. Extração de DNA	41
3. Análise do perfil mutacional	42
RESULTADOS	44
1. Características dos pacientes com lesões colorretais precursoras do CCR	44
2. Caracterização mutacional das lesões precursoras do câncer colorretal	46
2.1 Cobertura e qualidade do sequenciamento das lesões colorretais não câncer	46
2.2 Perfil de mutações nas amostras de lesões precursoras do câncer colorretal	47
2.3 Perfil de mutações <i>drivers</i> de acordo com o tipo histológico nas amostras de lesões precursoras do câncer colorretal	53
2.4 Diferenças entre os tipos de lesões e frequência de mutações <i>drivers</i> em cada gene	55
2.5 Alterações em vias moleculares associadas ao CCR nas lesões não câncer	59
DISCUSSÃO	61
CONCLUSÕES	65
CONCLUSÃO GERAL	66
CONSIDERAÇÕES ÉTICAS	67
REFERÊNCIAS	68
Anexos	79

## LISTA DE FIGURAS

- Figura 1.** Estimativa de incidência dos tipos tumorais mais frequentes no Brasil para o biênio ano de 2018-2019. **1**
- Figura 2.** Progressão lesões-carcinoma. Atualmente duas sequencias principais são reconhecidas: a tradicional progressão adenoma-carcinoma e a progressão pólipó serrilhado-carcinoma, ambas com características moleculares distintas. Adaptado de Kuipers et al, 2015<sup>7</sup>. **3**
- Figura 3.** Frequência de mutações encontradas no câncer colorretal <sup>23</sup>. O estudo classifica os tumores colorretais em hipermutados e não hipermutados. Podemos observar a taxa de mutação em cada uma das 224 pacientes (a) e os genes alterados com maior frequência (b). **8**
- Figura 4.** Genes frequentemente alterados segundo dados do TCGA (*The Cancer Genome Atlas*). Os genes incluídos no painel analisado no presente estudo (azul claro) englobam os genes alterados com maior frequência (acima de 6% de casos) em CCR **13**
- Figura 5.** Cobertura média dos genes mais frequentemente alterados nas amostras sequenciadas. Indicado em cada caixa está a mediana (linha) e a média (+) de cobertura por gene das 96 amostras sequenciadas. Podemos observar grande variância da cobertura e uma amostra ficou abaixo de 30x (linha vermelha tracejada) e foi retirada das análises posteriores. **17**
- Figura 6.** Variantes encontradas nas amostras sequenciadas, incluindo todas as variantes alteradas em relação ao tecido germinativo pareado (a) e mutações mantidas após o filtro de patogenicidade (b). **18**
- Figura 7.** Genes alterados com maior frequência em nossa população. **19**



**Figura 8.** Mutações *drivers* encontradas nos genes analisados em nossa população de estudo classificados por frequência de tumores alterados (apenas os genes com alterações em mais de 3% das amostras são mostrados). A parte central da figura mostra os tipos de mutações de cada gene distribuídas por cada amostra. No topo podemos observar a contagem de mutação por amostra. As laterais mostram a frequência de mutação em cada gene de acordo com a localização (cólon proximal ou distal). Na parte inferior podemos observar o estadiamento tumoral bem como status de instabilidade de microssatélite (MSI). **20**

**Figura 9.** Diagrama das mutações no gene *APC*. A maioria das mutações encontradas foram truncantes, incluindo mutações nonsense e frameshift. Em destaque, podemos observar a posição 1556, que apresentou maior frequência de alteração (6 mutações). **21**

**Figura 10.** Diagrama das mutações no gene *TP53*. A maioria das mutações encontradas no gene foram do tipo *missense* (pontos verdes) e se concentravam na região de ligação ao DNA (caixa em vermelho). Em destaque podemos observar a posição 273, que apresentou a maior frequência de alteração (6 mutações). **22**

**Figura 11.** Genes alterados na via MAPK (a) e a distribuição de mutações ao longo dos principais domínios proteicos destes genes (b). MSI: instabilidade microssatélite. **23**

**Figura 12.** Oncoplot dos genes alterados na via PI3K-AKT (a) e a distribuição das mutações ao longo dos principais domínios proteicos destes genes (b). MSI: instabilidade microssatélite. **25**

**Figura 13.** Genes alterados na via Notch. MSS: estabilidade microssatélite; MSI-L: baixa instabilidade microssatélite; MSI-H: alta instabilidade microssatélite. **26**

**Figura 14.** Genes alterados na família de tirosina quinases. MSI: instabilidade microssatélite. **28**

- Figura 15.** Comparação do número de mutação entre os casos com alta instabilidade de microssatélite (MSI-H) e os casos com baixa instabilidade (MSI-L) ou sem instabilidade (MSS). **33**
- Figura 16.** Cobertura dos 207 amplicons sequenciados pela plataforma Ion Torrent, distribuídos pelos 50 genes analisados. Foi obtida uma cobertura média de 1631,5x. **46**
- Figura 17.** Variantes encontradas nas amostras de lesões sequenciadas, incluindo todas as variantes alteradas nas 90 amostras (a) e mutações mantidas após o filtro de patogenicidade (b) em 60 amostras. **47**
- Figura 18.** Mutações *drivers* encontradas nas lesões colorretais. Os genes são exibidos de acordo com a frequência (esquerda). Amostras sem nenhuma mutação *driver* foram omitidas da figura. PH: pólipó hiperplásico; LSS: lesão serrilhada séssil **48**
- Figura 19.** Diagrama das mutações no gene *APC*. A maioria das mutações encontradas foram truncantes, incluindo mutações nonsense e frameshift. Em destaque, podemos observar a mutação p.Arg1450Ter, que apresentou maior frequência (4 mutações). **49**
- Figura 20.** Diagrama das mutações no gene *TP53*, incluindo mutações do tipo missense (pontos verdes) e mutações nonsense (pontos pretos). A maioria das mutações encontradas no gene concentrava-se na região de ligação ao DNA (caixa vermelha). **49**
- Figura 21.** Diagrama das mutações no gene *KRAS*, todas mutações do tipo missense (pontos verdes) afetando hotspots do gene no domínio Ras da proteína (caixa verde). O resíduo proteico mais afetado teve 41,2% das mutações na proteína (códon 13). **50**

**Figura 22.** Diagrama das mutações no gene *BRAF*, todas mutações do tipo missense (pontos verdes) no domínio tirosina quinase da proteína (caixa azul). A mutação mais frequente foi a p.Val600Glu, com 8 mutações. Duas amostras apresentaram duas mutações no gene: p.Val600Glu e p.Ala598Val em uma amostra e p.Asp594Asn e p. Ala598Val em outra. **50**

**Figura 23.** Diagrama das mutações no gene *EGFR*, todas mutações do tipo missense (pontos verdes) afetando majoritariamente o domínio no domínio tirosina quinase da proteína (caixa azul). A mutação mais frequente foi a p.Val600Glu, com 8 mutações. Duas amostras apresentaram duas mutações no gene: p.Val600Glu e p.Ala598Val em uma amostra e p.Asp594Asn e p. Ala598Val em outra. **50**

**Figura 24.** Diagrama das mutações no gene *FBXW7*, que apresentou mutações do tipo *missense* e *nonsense* nos domínios repetitivos de transducina beta (caixa vermelhas). O sítio 465 foi o códon mais afetado do gene, afetando 66,7% das amostras. **51**

**Figura 25.** Diagrama das mutações nos genes com frequência abaixo de 6% das amostras de lesões colorretais não câncer. **52**

**Figura 26.** Número de mutações *drivers* diferem entre os tipos de lesões (teste de Kruskal-Wallis  $p < 0,001$ ). Linhas horizontais representam a média de mutações *drivers* em adenomas precoces, adenomas avançados, pólipos hiperplásicos (PH) e lesões serrilhadas sésseis (LSS). Comparações par a par baseadas no teste Mann-Whitney: \* $p < 0,05$ ; \*\* $p < 0,01$ ; \*\*\* $p < 0,001$ . **54**

**Figura 27.** Mutações *drivers* encontradas em adenomas (à esquerda) e pólipos serrilhados (à direita). Genes diferencialmente mutados incluem *APC* e *TP53* com maior frequência em adenomas e *BRAF* com maior frequência em pólipos serrilhados. Amostras sem mutações *drivers* foram omitidas da figura. PH: pólipo hiperplásico; LSS (lesão serrilhada sésil). **56**

**Figura 28.** Mutações *drivers* encontradas em adenomas precoces (à esquerda) e adenomas avançados (à direita). Apesar de apresentarem algumas variações, como mutações exclusivas em alguns genes, apenas mutações nos genes *KRAS* e *GNAS* variaram significativamente entre os adenomas, com maior frequência em adenomas avançados. Amostras sem mutações *drivers* foram omitidas. **58**

**Figura 29.** Mutações *drivers* encontradas em pólipos hiperplásicos (PH, à esquerda) e lesões serrilhadas sésseis (LSS, à direita). Alterações em genes relacionados à progressão adenoma-carcinoma, como *APC* e *TP53*, foram encontrados com baixa frequência. Mutações no gene *BRAF* diferiram significativamente entre PH e LSS e mutações em alguns genes foram específicas para cada tipo de lesão. Amostras sem mutações *drivers* foram omitidas da figura **59**

**Figura 30.** Alterações em vias relacionadas ao câncer colorretal encontradas nas lesões precursoras. PH: pólipo hiperplásico; LSS: lesão serrilhada séssil **61**

## LISTA DE TABELAS

<b>Tabela 1.</b> Estadiamento TNM de tumores colorretais.	<b>4</b>
<b>Tabela 2 -</b> Características clínico-patológicas dos pacientes com câncer colorretal.	<b>16</b>
<b>Tabela 3.</b> Associação entre os genes com maior frequência de mutações e as características sócio-demográficas	<b>29</b>
<b>Tabela 4.</b> Associação entre os genes com maior frequência de mutações e as características clínico-patológicas dos tumores	<b>30</b>
<b>Tabela 5.</b> Associações entre a presença de mutações e localização do tumor	<b>31</b>
<b>Tabela 6.</b> Frequência de alteração nos diferentes genes em relação ao status MSI	<b>34</b>
<b>Tabela 7.</b> Comparação da frequência de mutações em diferentes bancos de dados públicos (disponíveis no cBioPortal) e outros estudos.	<b>36</b>
<b>Tabela 8 –</b> Características sócio-demográficas dos pacientes e morfológicas e histopatológicas das lesões colorretais não-câncer.	<b>45</b>
<b>Tabela 9.</b> Presença de mutações drivers nas lesões colorretais	<b>54</b>
<b>Tabela 10.</b> Comparação entre a frequência de mutações em pólipos serrilhados e adenomas	<b>55</b>
<b>Tabela 11.</b> Comparação entre a frequência de mutações em adenomas precoces e adenomas avançados e entre lesões serrilhadas sésseis e pólipos hiperplásicos	<b>57</b>

### LISTA DE ABREVIACÕES

ABL1	ABL proto-oncogene 1, non-receptor tyrosine kinase
ABL1	ABL proto-oncogene 1, non-receptor tyrosine kinase
AKT	RAC-alfa serina/treonina-proteína quinase, do inglês " <i>RAC-alpha serine/threonine-protein kinase</i> "
AKT1	AKT serine/threonine kinase 1
AKT2	AKT serine/threonine kinase 2
AKT3	AKT serine/threonine kinase 3
ALK	anaplastic lymphoma receptor tyrosine kinase
APC	APC, WNT signaling pathway regulator
AR	androgen receptor
ARAF	A-Raf proto-oncogene, serine/threonine kinase
ATM	ATM serine/threonine kinase
AURKA	aurora kinase A
AURKB	aurora kinase B
AXIN1	axin 1
AXL	AXL receptor tyrosine kinase
BAP1	BRCA1 associated protein 1
BARD1	BRCA1 associated RING domain 1
BCL2	BCL2, apoptosis regulator
BCL2L1	BCL2 like 1
BCL2L2	BCL2 like 2
BLM	Bloom syndrome RecQ like helicase
bp	pares de base, do inglês " <i>base pair</i> "
BRAF	B-Raf proto-oncogene, serine/threonine kinase
BRCA1	BRCA1, DNA repair associated
BRCA2	BRCA2, DNA repair associated
BRD2	bromodomain containing 2
BRD3	bromodomain containing 3
BRD4	bromodomain containing 4
BRDT	bromodomain testis associated

BRIP1	BRCA1 interacting protein C-terminal helicase 1
BTK	Bruton tyrosine kinase
BWA	do inglês " <i>Burrows – Wheeler Aligner</i> "
CBFB	core-binding factor beta subunit
CCND1	cyclin D1
CCND2	cyclin D2
CCND3	cyclin D3
CCNE1	cyclin E1
CCR	câncer colorretal
CD22	CD22 molecule
CD274	CD274 molecule
CD79A	CD79a molecule
CD79B	CD79b molecule
CDH1	cadherin 1
CDK12	cyclin dependent kinase 12
CDK4	cyclin dependent kinase 4
CDKN1A	cyclin dependent kinase inhibitor 1A
CDKN2A	cyclin dependent kinase inhibitor 2A
CGI	do inglês <i>Cancer Genome Interpreter</i>
CHEK1	checkpoint kinase 1
CHEK2	checkpoint kinase 2
CMS	Subtipos molecular consensual do inglês " <i>Consensus Molecular Subtypes</i> "
COSMIC	Catálogo das mutações somáticas no câncer, do inglês " <i>Catalogue of Somatic Mutations in Cancer</i> "
CpG	5'-citosina-fosfato-guanina-3', do inglês
CREBBP	CREB binding protein
CRKL	CRK like proto-oncogene, adaptor protein
CSF1R	colony stimulating factor 1 receptor
CTNNB1	catenin beta 1

dbSNP	Banco de dados de polimorfismo de nucleotídeo único, do inglês <i>"Single Nucleotide Polymorphism Database"</i>
DDR2	discoidin domain receptor tyrosine kinase 2
DDX3X	DEAD-box helicase 3, X-linked
DNA	ácido desoxirribonucléico, do inglês <i>"deoxyribonucleic acid"</i>
EGFR	epidermal growth factor receptor
EPHA7	EPH receptor A7
ERBB2	erb-b2 receptor tyrosine kinase 2
ERBB3	erb-b2 receptor tyrosine kinase 3
ERBB4	erb-b2 receptor tyrosine kinase 4
ERK	quinase regulada por sinal extracelular, inglês <i>"extracellular signal-regulated kinase"</i>
ESR1	estrogen receptor 1
EZH2	enhancer of zeste 2 polycomb repressive complex 2 subunit
FAM175A	family with sequence similarity 175 member A
FAP	polipose adenomatosa familiar (do inglês <i>"Familial Adenomatous polyposis"</i> )
FAS	Fas cell surface death receptor
FBXW7	F-box and WD repeat domain containing 7
FGFR1	fibroblast growth factor receptor 1
FGFR2	fibroblast growth factor receptor 2
FGFR3	fibroblast growth factor receptor 3
FGFR4	fibroblast growth factor receptor 4
FHIT	fragile histidine triad
FLT1	fms related tyrosine kinase 1
FLT3	fms related tyrosine kinase 3
FLT4	fms related tyrosine kinase 4
FRS2	fibroblast growth factor receptor substrate 2
GATA2	GATA binding protein 2
GNA11	G protein subunit alpha 11
GNAQ	G protein subunit alpha q



GNAS	GNAS complex locus
HBC	Hospital de câncer de Barretos
HDAC1	histone deacetylase 1
HDAC4	histone deacetylase 4
HDAC7	histone deacetylase 7
hg19	genoma humano 19, do inglês " <i>human genome 19</i> "
HGF	hepatocyte growth factor
HNF1A	HNF1 homeobox A
HNPCC	do inglês " <i>hereditary nonpolyposis colorectal cancer</i> "
HRAS	HRas proto-oncogene, GTPase
IDH1	isocitrate dehydrogenase
IDH2	isocitrate dehydrogenase
IGF1R	insulin like growth factor 1 receptor
INCA	Instituto Nacional do Câncer
InDel	inserção ou deleção, do inglês " <i>insertion or deletion</i> "
JAK1	Janus kinase 1
JAK2	Janus kinase 2
JAK3	Janus kinase 3
KDR	kinase insert domain receptor
KIT	KIT proto-oncogene receptor tyrosine kinase
KRAS	KRAS proto-oncogene, GTPase
LSS	lesão serrilhada séssil
MAP2K1	mitogen-activated protein kinase kinase 1
MAP2K2	mitogen-activated protein kinase kinase 2
MAP2K4	mitogen-activated protein kinase kinase 4
MAP3K1	mitogen-activated protein kinase kinase kinase 1
MAPK	Proteína quinase ativada por mitogênio, do inglês " <i>Mitogen-activated protein kinase</i> "
MAPK1	mitogen-activated protein kinase 1
MCL1	BCL2 family apoptosis regulator
MDM2	MDM2 proto-oncogene

MET	MET proto-oncogene, receptor tyrosine kinase
MLH1	mutL homolog 1
MPL	MPL proto-oncogene, thrombopoietin receptor
MRE11A	meiotic recombination 11 homolog A
MS4A1	membrane spanning 4-domains A1
MSH2	mutS homolog 2
MSH6	mutS homolog 6
MSI	instabilidade de microssatélite (do inglês " <i>microsatellite instability</i> ")
MSI-H	alta instabilidade de microssatélite
MSI-L	baixa instabilidade de microssatélite
MSS	microssatélite estável
MTOR	mechanistic target of rapamycin
MUTYH	mutY DNA glycosylase
MYC	v-myc avian myelocytomatosis viral oncogene homolog
MYD88	myeloid differentiation primary response 88
NBN	nibrin
NF1	neurofibromin 1
NF2	neurofibromin 2
NFE2L2	nuclear factor, erythroid 2 like 2
NFKBIA	NFKB inhibitor alpha
NOTCH1	notch 1
NOTCH2	notch 2
NOTCH3	notch 3
NPM1	nucleophosmin
NRAS	neuroblastoma RAS viral oncogene homolog
NT5C2	5'-nucleotidase, cytosolic II
NTRK1	neurotrophic receptor tyrosine kinase 1
OMS	Organização Mundial da Saúde
PALB2	partner and localizer of BRCA2
PDGFRA	platelet derived growth factor receptor alpha
PDGFRB	platelet derived growth factor receptor beta

PDK1	pyruvate dehydrogenase kinase 1
PH	pólipo hiperplásico
PI4K	Phosphatidylinositol 4-kinase
PIK3	Fosfoinositida 3-quinases, do inglês " <i>phosphoinositide 3-kinases</i> "
PIK3CA	phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit alpha
PIK3CB	phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit beta
PIK3CG	phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit gamma
PIP <sub>3</sub>	phosphatidylinositol 3,4,5 trisphosphate
PMS2	PMS1 homolog 2, mismatch repair system component
PTCH1	patched 1
PTEN	phosphatase and tensin homolog
PTPN11	protein tyrosine phosphatase, non-receptor type 11
RAC1	ras-related C3 botulinum toxin substrate 1
RAD51	RAD51 recombinase
RAF1	Raf-1 proto-oncogene, serine/threonine kinase
RANBP2	RAN binding protein 2
RARA	retinoic acid receptor alpha
RB1	RB transcriptional corepressor 1
RET	ret proto-oncogene
RICTOR	RPTOR independent companion of MTOR complex 2
RNA	ácido ribonucléico, do inglês " <i>ribonucleic acid</i> "
ROS1	ROS proto-oncogene 1, receptor tyrosine kinase
RRM1	ribonucleotide reductase catalytic subunit M1
RTK	Receptor de tirosina quinase, do inglês " <i>Receptor tyrosine kinases</i> "
RUNX1	runt related transcription factor 1
SDHB	succinate dehydrogenase complex iron sulfur subunit B
SMAD4	SMAD family member 4
SMARCB1	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily b, member 1
SMO	smoothened, frizzled class receptor

SOX2	SRY-box 2
SRC	SRC proto-oncogene, non-receptor tyrosine kinase
STAT3	signal transducer and activator of transcription 3
STAT5B	signal transducer and activator of transcription 5B
STK11	serine/threonine kinase 11
TCGA	O atlas do genoma do câncer (do inglês, " <i>The Cancer Genome Atlas</i> ")
TERC	telomerase RNA component
TERT	telomerase reverse transcriptase
TGFBR2	transforming growth factor beta receptor 2
TP53	tumor protein p53
TSC1	tuberous sclerosis 1
TSC2	tuberous sclerosis 2
UTR	região não traduzida, do inglês " <i>untranslated region</i> "
VEGFA	vascular endothelial growth factor A
VHL	von Hippel-Lindau tumor suppressor
Wnt	do inglês " <i>Wingless-related integration site</i> "
WT1	Wilms tumor 1
XPO1	exportin 1

**LISTA DE SÍMBOLOS**

%	Porcentagem
$\beta$	beta
x	vezes
$\chi^2$	qui-quadrado
=	igual
>	maior
<	menor
*	asterisco
ng	nanogramas

## RESUMO

dos Santos W. Perfil mutacional de um painel de oncogenes e genes supressores tumorais no câncer colorretal e suas lesões precursoras. **Tese (Doutorado)**. Barretos: Hospital de câncer de Barretos; 2020.

**JUSTIFICATIVA:** O câncer colorretal (CCR) é o terceiro em incidência mundial, ocupando o segundo lugar em causas de morte entre os cânceres. O desenvolvimento da doença geralmente ocorre por meio do surgimento de lesões precursoras na mucosa intestinal normal com posterior progressão ao câncer. Fatores ambientais, estilo de vida, idade, dieta, fatores genéticos e moleculares contribuem para o surgimento e progressão do CCR. As vias moleculares envolvidas na progressão e desenvolvimento da doença têm sido bem caracterizadas e incluem mutações genéticas, instabilidade de microssatélite e alterações epigenéticas. Tais características, entretanto, não são bem conhecidas na população brasileira.

**OBJETIVOS:** Investigar as mutações somáticas no CCR e nas suas lesões precursoras em uma amostra da população brasileira.

**MATERIAL E MÉTODOS:** O trabalho foi desenvolvido no Hospital de Câncer de Barretos (HCB) incluindo pacientes adultos, com CCR que foram submetidos à cirurgia para remoção do câncer, bem como pacientes com lesões colorretais não câncer diagnosticadas e removidas durante a colonoscopia. Foi realizado o perfil mutacional de toda a região codificante de 150 oncogenes e genes supressores tumorais por meio de NGS para amostras de CCR, e perfil mutacional em 50 oncogenes e genes supressores tumorais (2800 mutações *hotspot*) para amostras de lesões colorretais.

**RESULTADOS:** Foram sequenciadas 91 amostras de CCR de pacientes com média de idade de 61,2 anos, a maioria sendo do sexo masculino (53,8%). Excluindo alteração silenciosas e possivelmente não patogênicas, os genes mais frequentemente mutados foram *APC* (66 pacientes, 72,5%), *TP53* (51, 56,0%), *KRAS* (48, 52,7%), *PIK3CA* (14, 15,4%) e *FBXW7* (11, 12,1%). Tumores localizados no cólon distal tiveram maior frequência de mutações nos genes *APC* (77,5% vs 50,0%,  $p=0,025$ ), *TP53* (67,6% vs 20,0%,  $p<0,001$ ) e *KRAS* (63,4% vs 15,0%,  $p<0,001$ ). Tumores localizados no cólon proximal apresentaram maior frequência de mutações nos genes *ATM* (20,0% vs 2,8%,  $p=0,02$ ), *BLM* (25,0% vs 4,2%,  $p=0,012$ ), *BRAF* (40,0%

vs 0,0%,  $p < 0,001$ ), *NOTCH2* (15,0% vs 1,4%,  $p = 0,032$ ), *NOTCH3* (15,0% vs 1,4%,  $p = 0,032$ ), *PTEN* (15,0% vs 1,4%,  $p = 0,032$ ), *RANBP2* (15,0% vs 1,4%,  $p = 0,032$ ) e *TGFBR2* (30% vs 2,8%,  $p < 0,001$ ). Tumores com instabilidade de microssatélite foram mais frequentes no cólon proximal (83,3%,  $p < 0,001$ ) e apresentaram maior média de mutações do que tumores sem instabilidade ( $p < 0,001$ ). Os 87 pacientes com lesões precursoras colorretais apresentaram uma média de idade de 62,9 anos, cujas 90 lesões incluíram adenomas (74,4%, 67/90), lesões serrilhadas sésseis (7,8%, 7/90) e pólipos hiperplásicos (17,8%, 16/90). Adenomas apresentaram um número maior de mutações *drivers* quando comparados com pólipos serrilhados (73,1% vs 47,8%,  $p = 0,039$ ), bem como maior frequência de mutações nos genes *APC* (41,8% vs 4,3%,  $p = 0,001$ ) e *TP53* (25,4% vs 4,3%,  $p = 0,034$ ) e menor frequência de mutações no gene *BRAF* (6,0% vs 34,8%,  $p < 0,001$ ). Entre adenomas, os adenomas avançados apresentaram um número maior de mutações *drivers* quando comparados com adenomas precoces (92,9% vs 59,0%,  $p = 0,002$ ) e maior frequência de mutações nos genes *KRAS* (46,4% vs 5,1%,  $p = 0,002$ ) e *GNAS* (17,9% vs 0,0%,  $p = 0,010$ ). Entre os pólipos serrilhados, as lesões serrilhadas sésseis também apresentaram número maior de mutações *drivers* quando comparados com pólipos hiperplásicos (85,7% vs 31,3%,  $p = 0,027$ ), além de maior frequência de mutações no gene *BRAF* (71,4% vs 18,8%,  $p = 0,026$ ).

**CONCLUSÕES:** Nosso trabalho demonstrou, de forma inédita, que a frequência de mutações encontradas no câncer colorretal da população brasileira é similar àquela encontrada em outras populações, com pequenas variações e mostrou a presença de mutações *drivers* em lesões precursoras do câncer colorretal, com adenomas avançados e lesões serrilhadas sésseis os maiores portadores de tais mutações, bem como diferenças no perfil de mutações de acordo com o tipo de lesão analisada.

**PALAVRAS-CHAVE:** câncer colorretal; adenoma; pólipos serrilhados; patologia molecular; mutação; biomarcadores tumorais.

## ABSTRACT

dos Santos W. Mutational profile of a panel of oncogenes and tumor suppressor genes in colorectal cancer and its precursor lesions. **Thesis (Doctorate degree)**. Barretos: Hospital de câncer de Barretos; 2020.

**BACKGROUND:** Colorectal cancer (CRC) is the third most common cancer worldwide, ranking as second in causes of death among cancers. CRC usually develops from precursor lesions within intestinal mucosa with subsequent progression to cancer. Environmental factors, lifestyle, age, diet, genetics, and molecular factors contribute to the development and progression of CRC. The molecular pathways involved in disease progression and development have been well characterized. They include genetic mutations, microsatellite instability and epigenetic changes. These alterations, however, are not well studied in the Brazilian population.

**AIM:** To investigate somatic mutations in CRC and its precursor lesions in the Brazilian population.

**MATERIAL AND METHODS:** We carried out the study at the Barretos Cancer Hospital (HCB). We included adult patients with CRC who underwent surgery for cancer removal, as well as patients with non-cancer colorectal lesions diagnosed and removed during the colonoscopy. The mutational profile of the entire coding region of 150 oncogenes and tumor suppressor genes by using NGS was performed in CRC samples. In addition, we performed a mutational profile of 50 oncogenes and tumor suppressor (containing 2800 hotspot mutations) for colorectal lesion samples.

**RESULTS:** A total of 91 CRC samples from patients with a mean age of 61.2 years old were sequenced, the majority being male (53.8%). Excluding silent and possibly non-pathogenic changes, the most frequently mutated genes were *APC* (66 patients, 72.5%), *TP53* (51, 56.0%), *KRAS* (48, 52.7%), *PIK3CA* (14, 15.4%) and *FBXW7* (11, 12.1%). Tumors from distal colon had a higher frequency of mutations than proximal tumors in the genes *APC* (77.5% vs 50.0%,  $p=0.025$ ), *TP53* (67.7% vs 20.0%,  $p<0.001$ ) and *KRAS* (63.4% vs 15.0%,  $p<0.001$ ). Tumors from proximal colon had a higher frequency of mutations in the genes *ATM* (20.0% vs 2.8%,  $p=0.02$ ), *BLM* (25.0% vs 4.2%,  $p = 0.012$ ), *BRAF* ( 40.0% vs 0.0%,  $p <0.001$ ), *NOTCH2* (15.0% vs 1.4%,  $p = 0.032$ ), *NOTCH3* (15.0% vs 1.4%,  $p = 0.032$ ), *PTEN* (15.0% vs 1.4%,  $p = 0.032$ ), *RANBP2* (15.0% vs 1.4%,  $p = 0.032$ ) and *TGFBR2* (30% vs 2.8%,  $p <0.001$ ). Tumors with microsatellite instability were more frequent in the proximal colon (83.3%,  $p<0.001$ ) and





had a higher mutation burden ( $p < 0.001$ ). The 87 patients with colorectal precursor lesions had a mean age of 62.0 years old, whose 90 lesions included adenomas (74.4%, 67/90), sessile serrated lesions (7.8%, 7/90) and hyperplastic polyps (17.8%, 16/90). Adenomas showed a higher number of driver mutations when compared to serrated polyps (73.1% vs 47.8%,  $p = 0.039$ ), as well as a higher frequency of mutations in the genes *APC* (41.8% vs 4.3%,  $p = 0.001$ ) and *TP53* (25.4% vs 4.3%,  $p = 0.034$ ) and had lower frequency of mutation in the *BRAF* gene (6.0% vs 34.8%,  $p < 0.001$ ). Advanced adenomas showed a higher number of drivers mutations when compared to early adenomas (92.9% vs 59.0%,  $p = 0.002$ ) and a higher frequency of mutations in the genes *KRAS* (46.4% vs 5.1%,  $p = 0.002$ ) and *GNAS* (17.9% vs 0.0%,  $p = 0.010$ ). Sessile serrated lesions showed a higher number of driver mutations when compared to hyperplastic polyps (85.7% vs 31.3%,  $p = 0.027$ ), in addition to a higher frequency of mutations in the *BRAF* gene (71.4% vs 18.8%,  $p = 0.026$ ).

**CONCLUSIONS:** Our study demonstrated that the mutation profile found in CRC in the Brazilian population is similar to that found in other populations, with small variations. Furthermore, we demonstrated the presence of driver mutations in precursor lesions of colorectal cancer, with advanced adenomas and sessile serrated lesions presenting the highest frequencies of these mutations.

**KEYWORDS:** colorectal cancer; adenoma; serrated polyp; molecular pathology; mutation; tumor biomarkers.

## INTRODUÇÃO

O câncer colorretal (CCR) é o terceiro tipo de câncer mais comum no mundo, sendo o segundo tipo de câncer que mais causa mortes<sup>1-3</sup>. Os casos de CCR, juntamente com os casos de câncer de pulmão, mama, próstata estômago e fígado, representam 55% da incidência global de casos de câncer<sup>1,2</sup>. A incidência de CCR é importante tanto em países desenvolvidos, quanto em países em vias de desenvolvimento. No Brasil, segundo estimativa de incidência para 2018 do Instituto Nacional de Câncer (INCA) foram estimados 18.980 novos casos de CCR em mulheres (9,4% das neoplasias exceto pele não melanoma) e 17.380 em homens (8,1% das neoplasias exceto pele não melanoma, **Figura 1**) totalizando 36.360 novos casos<sup>4</sup>.

Localização Primária	Casos	%			Localização Primária	Casos	%		
Próstata	68.220	31,7%	<b>Homens</b>		Mama Feminina	59.700	29,5%		
Traqueia, Brônquio e Pulmão	18.740	8,7%			<b>Mulheres</b>		Cólon e Reto	18.980	9,4%
Cólon e Reto	17.380	8,1%					Colo do Útero	16.370	8,1%
Estômago	13.540	6,3%					Traqueia, Brônquio e Pulmão	12.530	6,2%
Cavidade Oral	11.200	5,2%					Glândula Tireoide	8.040	4,0%
Esôfago	8.240	3,8%					Estômago	7.750	3,8%
Bexiga	6.690	3,1%					Corpo do Útero	6.600	3,3%
Laringe	6.390	3,0%					Ovário	6.150	3,0%
Leucemias	5.940	2,8%					Sistema Nervoso Central	5.510	2,7%
Sistema Nervoso Central	5.810	2,7%					Leucemias	4.860	2,4%

**Figura 1.** Estimativa de incidência dos tipos tumorais mais frequentes no Brasil para o biênio ano de 2018-2019.

Os mecanismos patogênicos e fatores etiológicos que promovem o desenvolvimento do CCR são complexos e heterogêneos. Existem interações complexas que envolvem idade, fatores ambientais, estilo de vida, dieta e fatores genéticos que podem alterar o risco para desenvolver CCR<sup>5,6</sup>. Alta incidência de CCR tem sido observada em populações com uma dieta rica em gorduras insaturadas, combinadas com um estilo de vida sedentário. Além disso, estudos epidemiológicos indicam que obesidade, alto consumo de carne vermelha e carnes processadas, fumo e consumo excessivo de álcool são fatores importantes que podem aumentar o risco de desenvolvimento de CCR<sup>5,6</sup>.

Os fatores hereditários que influenciam na predisposição do desenvolvimento de CCR podem contabilizar cerca de 15-20% dos casos dado a presença de histórico familiar<sup>5,7</sup>. Dentre os pacientes que contam com histórico familiar positivo, um subgrupo é afetado por síndromes hereditárias que predispõem ao CCR. A principal dessas síndromes é denominada Síndrome de Lynch ou HNPCC (do inglês *“hereditary nonpolyposis colorectal cancer”*), que é

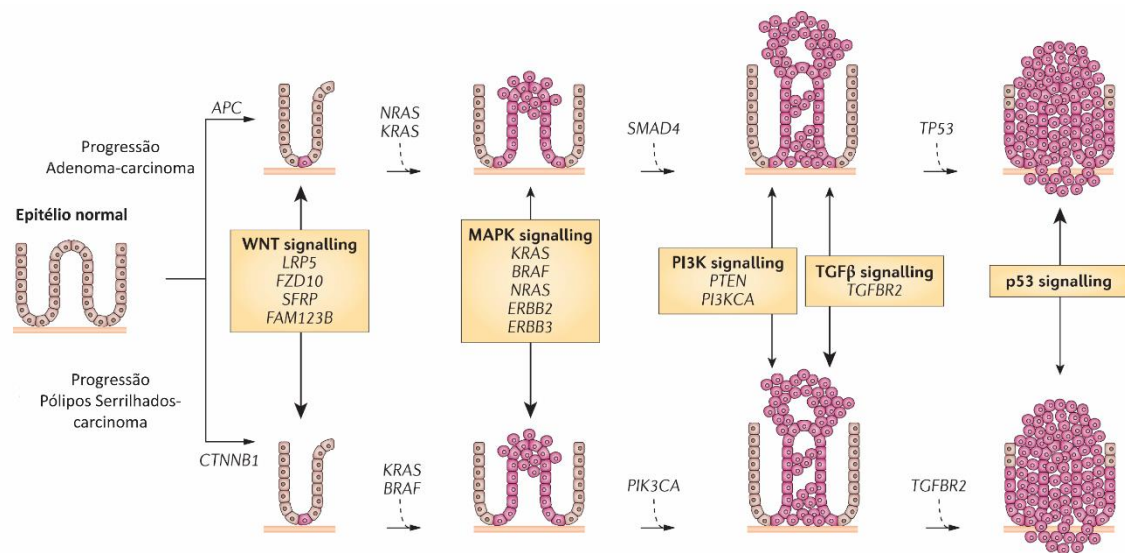
caracterizada por mutações em um dos genes de reparo do sistema de pareamento do DNA - *mismatch-repair* (*MLH1*, *MSH2*, *MSH6*, *PMS2* ou *EPCAM*), o que ocasiona no acúmulo de mutações que ocorrem, particularmente, em regiões de microssatélite do DNA, acarretando em instabilidade de microssatélite (MSI)<sup>5,7</sup>. Outra importante síndrome hereditária do CCR é a polipose adenomatosa familiar (FAP, do inglês "*familial adenomatous polyposis*"), que está relacionada às mutações no gene *APC*, e é caracterizada pela desenvolvimento de numerosos adenomas e subsequente CCR precoce<sup>7</sup>.

### **História natural das lesões precursoras ao CCR**

A maioria dos casos de CCR desenvolve-se por meio de uma sequência de eventos que levam da mucosa normal do trato intestinal às lesões precursoras e posterior desenvolvimento do tumor maligno (**Figura 2**). O adenoma é a principal lesão que contribui para o desenvolvimento do CCR<sup>5</sup>. Na sequência clássica do adenoma carcinoma, o desenvolvimento do câncer colorretal origina-se de criptas aberrantes, evoluindo para o adenoma precoce e então para o adenoma avançado ( $\geq 10\text{mm}$  ou com mais de 25% de componente viloso ou com alto grau de displasia) e deste ao câncer. Apenas 10% dos adenomas evoluirão ao câncer e aquele com maior probabilidade de evolução é o adenoma avançado<sup>7,8</sup>. Existe um alto risco de desenvolvimento de CCR em indivíduos em que os adenomas não são removidos enquanto que a polipectomia diminui esse risco<sup>5</sup>. Atualmente, o adenoma representa o principal alvo do rastreamento do CCR devido à sua alta incidência, possibilidade de remoção e ao tempo prolongado, entre 7 e 10 anos, de desenvolvimento adenoma-carcinoma<sup>9</sup>.

Os adenomas surgem a partir do epitélio glandular e são caracterizados pela presença de displasia e alteração na diferenciação das células epiteliais. O modelo clássico de progressão adenoma-carcinoma foi proposto por Fearon e Vogelstein<sup>10</sup>, no qual o desenvolvimento do tumor é dirigido pelo progressivo acúmulo de mutações que pode levar de 3 a 10 anos<sup>5</sup>. Embora a progressão adenoma-carcinoma seja a principal via de desenvolvimento do CCR, recentemente os pólipos serrilhados têm sido reconhecidos como uma importante lesão precursora podendo ser responsável por cerca de 7,5% de todos os casos de CCR<sup>11,12</sup>. Pólipos serrilhados formam um grupo heterogêneo de lesões caracterizado morfológicamente pela arquitetura serrilhada do tecido epitelial no local da lesão<sup>8,13</sup> e incluem pólipo hiperplásico, lesão serrilhada séssil e adenoma serrilhado tradicional<sup>8,13-15</sup>.

Dentre os pólipos serrilhados aqueles considerados com potencial de evoluir para o câncer é a lesão serrilhada sésil e os pólipos hiperplásicos do colón direito e maiores do que 10mm 11,13.



**Figura 2.** Progressão lesões-carcinoma. Atualmente duas sequencias principais são reconhecidas: a tradicional progressão adenoma-carcinoma e a progressão pólipos serrilhado-carcinoma, ambas com características moleculares distintas. Adaptado de Kuipers et al, 2015<sup>7</sup>.

### Características clínico-patológicas do câncer colorretal

O tipo mais comum de CCR é o adenocarcinoma, contabilizando por cerca de 85% dos casos<sup>16</sup>. Segundo a Organização Mundial da Saúde (OMS), os adenocarcinomas podem ser histologicamente bem diferenciados (>95% de arquitetura glandular), moderadamente diferenciados (50-95% de arquitetura glandular), pouco diferenciados (0-49% de arquitetura glandular) ou indiferenciados<sup>8</sup>. Baseado no sistema TMN, os tumores colorretais podem ser classificados em estádios 0-IV (**Tabela 1**)<sup>15</sup>. As taxas de sobrevivência são diretamente relacionadas ao estadiamento no qual o tumor é detectado, sendo superiores a 90% para pacientes que tem tumores no estágio I ao diagnóstico e inferiores a 10% no estágio IV<sup>17</sup>.

**Tabela 1.** Estadiamento TNM de tumores colorretais<sup>15</sup>

<b>Estadio</b>	<b>T</b>	<b>N</b>	<b>M</b>
0	Tis	N0	M0
I	T1	N0	M0
	T2	N0	M0
II	T3,T4	N0	M0
IIA	T3	N0	M0
IIB	T4a	N0	M0
IIC	T4b	N0	M0
III	Qualquer T	N1, N2	M0
IIIA	T1, T2	N1	M0
	T1	N2a	M0
	T1, T2	N2b	M0
IIIB	T2, T3	N2a	M0
	T3, T4a	N1	M0
	T3, T4a	N2b	M0
IIIC	T4a	N2a	M0
	T4b	N1, N2	M0
IV	Qualquer T	Qualquer N	M1
IVA	Qualquer T	Qualquer N	M1a
IVB	Qualquer T	Qualquer N	M1b
IVC	Qualquer T	Qualquer N	M1c

Tis: carcinoma *in situ*; T1: tumor invade a submucosa; T2: tumor invade a camada muscularis própria; T3: tumor invade a subserosa ou em tecidos pericólicos não peritonealizados; T4: tumor invade outros órgãos ou estruturas; N0: sem linfonodos regionais; N1: metástase em 1-3 linfonodos regionais; N1: metástase em 1-3 linfonodos regionais; N1a: metástase em 1 linfonodo regional; N1b: metástase em 2-3 linfonodos regionais; N1c: depósitos tumorais na subserosa ou em tecidos moles pericólicos ou periretais não peritonealizados sem metástase de linfonodos regionais; N2: 4 ou mais linfonodos regionais; N2a: 4-6 linfonodos regionais; N2b: 7 ou mais linfonodos regionais; M0: sem metástase à distância; M1: metástase à distância; M1a: metástase restrita à um órgão (fígado, pulmão, ovário, linfonodos não regionais) sem metástase peritoneal; M1b: metástase em mais de um órgão; M1c: metástase no peritônio com ou sem envolvimento em outros órgãos

O CCR é mais observado com maior frequência no cólon distal quando comparado com o cólon proximal. Além de diferenças na incidência, pacientes com tumores no cólon proximal frequentemente apresentam tumores mais avançados ao diagnóstico e sintomas característicos, tal como perda de peso e anemia microcítica, em contraste com pacientes com tumores no cólon distal, os quais apresentam tumores menos avançados ao diagnóstico e os sintomas mais frequentes são sangramento retal e alteração no hábito intestinal<sup>18,19</sup>. Além disso, os casos de CCR no cólon proximal e distal apresentam diferenças significativas no que diz respeito à sua epidemiologia, alterações moleculares e diversidade microbiota<sup>20-23</sup>.

### **Alterações moleculares do CCR**

Avanços na investigação dos aspectos moleculares do CCR possibilitou a descoberta de vários genes e vias importantes na iniciação e progressão da doença<sup>5,23</sup>. Tem sido reconhecido

que o desenvolvimento do CCR envolve diferentes vias moleculares e inclui a aquisição de mutações, a instabilidade de microssatélite e as alterações epigenéticas. Os pólipos serrilhados estão associados à instabilidade em regiões de microssatélites e metilação anormal do DNA, enquanto no desenvolvimento do adenoma ao câncer, a inativação do gene supressor de tumor *APC* e a instabilidade cromossômica são as alterações encontradas com maior frequência<sup>24</sup>.

As vias moleculares na progressão adenoma-carcinoma são bem caracterizadas e envolvem o acúmulo de mutações em genes supressores de tumores, tais como *APC*, e oncogenes como *KRAS*<sup>5,25</sup>. A inativação do gene *APC* é observada em aproximadamente 70-80% de casos de adenomas e carcinomas colorretais, e possivelmente desempenha um papel importante na iniciação e progressão adenoma-carcinoma<sup>5</sup>. A proteína *APC* interage com o citoesqueleto, está envolvida com a regulação dos níveis - de  $\beta$ -catenina e contribui para fidelidade de processos como migração, diferenciação, morte celular e mitose<sup>26</sup>. Na maioria dos casos, mutações no gene *APC* resultam em uma proteína truncada, alterando a diferenciação, apoptose e o controle do ciclo da célula, além de ocasionar em defeitos durante a segregação cromossômica<sup>5,26,27</sup>. O acúmulo de  $\beta$ -catenina resultante da inativação do *APC* altera regulações transcricionais e diminui a diferenciação celular. O gene *APC* desempenha um papel fisiológico importante na regulação de  $\beta$ -catenina e é regulado pela via de sinalização Wnt. Esta via está envolvida na diferenciação, proliferação e motilidade celular e é ativada pela acumulação de  $\beta$ -catenina no núcleo e subsequente ativação transcricional de genes alvos<sup>28</sup>. Cerca de 94% dos casos de CCR apresentam alguma alteração na via Wnt/  $\beta$ -catenina<sup>23</sup>.

Embora *APC* tenha um papel importante na tumorigênese do CCR, outras mutações somáticas colaboram para promover a progressão de lesões adenomatosas em carcinoma. Tais alterações incluem a ativação de proto-oncogenes como *KRAS*, *BRAF* e *PIK3CA*, e/ou inativação de genes supressores de tumores como *TP53*, *SMAD4*, *ARIAD1A* e *FBXW7*<sup>29</sup>. O gene *TP53* parece ser um importante alvo de perda de heterozigiosidade em casos de CCR e, na maioria dos casos em que há perda de um alelo *TP53*, o alelo remanescente está mutado<sup>5</sup>. A proteína p53 codificada é um fator de transcrição que desempenha um papel na regulação de genes que participam em pontos de checagem do ciclo celular, promoção da apoptose e restrição da angiogênese<sup>30</sup>.

Mutações somáticas no gene *KRAS* são encontradas em cerca 40% dos casos de CCR<sup>5</sup>. Os membros da família Ras estão comumente mutados em muitos tipos de câncer humano. Estas proteínas exercem efeitos em várias cascatas de sinalização e atuam regulando múltiplos receptores para fatores de crescimento, tal como o fator de crescimento epidérmico (*EGFR*). As mutações no gene *KRAS* em CCR, em sua grande maioria, resultam em ganho de função da proteína. Mutações em outros genes parecem atuar intensificando o efeito a jusante da proteína K-ras. Por exemplo, mutações somáticas no gene *PIK3CA* são encontradas em aproximadamente 15-25% dos casos de CCR e estão associadas com o aumento de produção de PIP<sub>3</sub> (do inglês, “*phosphatidylinositol 3,4,5 trisphosphate*”), um mensageiro com efeitos na proliferação e crescimento celular que também é regulado pela proteína K-ras<sup>5</sup>. O gene *PTEN* também influencia nos níveis de PIP<sub>3</sub> e está alterado em cerca de 10% dos casos de CCR<sup>5</sup>.

Diferentemente da caracterização molecular da progressão adenoma-carcinoma, os mecanismos moleculares que promovem a progressão do CCR a partir de outras lesões precursoras, tais como lesões serrilhadas sésseis, não são tão bem compreendidos<sup>25</sup>. Muitos carcinomas serrilhados apresentam instabilidade em regiões de microssatélite e frequentemente estão associados com hipermetilação em ilhas CpG<sup>5,14,23</sup>. Embora a instabilidade de microssatélite (MSI) seja um marcador de casos hereditários associados à HNPCC, cerca de 15-20% de CCR esporádicos apresentam MSI<sup>31,32</sup>. Tais tumores geralmente apresentam hipermetilação e frequente inativação epigenética de genes relacionados ao sistema de reparo de mal pareamento do DNA, principalmente o gene *MLH1*. Estes tumores apresentam, frequentemente, mutação ativante do gene *BRAF*, que parece ser uma alteração comum nos estádios iniciais do desenvolvimento de CCR a partir de lesões serrilhadas<sup>33</sup>. O gene que codifica para proteína B-Raf, que é regulada pelas proteínas Ras, está mutado em aproximadamente 5-10% dos casos de CCR<sup>34</sup>.

Alterações no perfil de expressão gênica também são observadas em diferentes subgrupos do CCR. Recentemente, foram determinados quatro diferentes subtipos consensuais de CCR (CMS, do inglês “*Consensus Molecular Subtypes*”), baseado no perfil de expressão de RNA<sup>22</sup>: CMS1, que são tumores enriquecidos para MSI, acentuada expressão de proteínas relacionados ao sistema de reparo, alta ativação imunológica e hipermetilação; CMS2, que abrigam tumores com instabilidade cromossômica e ativação das via WNT e MYC;

CMS3, que abriga tumores com desregulação epitelial e metabólica; e o CMS4, com tumores com ativação do fator de crescimento  $\beta$ , infiltração estromal e angiogênese<sup>22</sup>

Tais características moleculares diferem, ainda, em relação à localização tumoral. Tumores do cólon proximal (compreendendo ceco, cólon ascendente, flexura hepática, cólon transversal e flexura esplênica) são associados com presença de instabilidade de microssatélite e mutações ativantes no gene *BRAF*, frequentemente apresentando um alto índice (*burden*) mutacional. Diferentemente, tumores do cólon distal (cólon descendente, cólon sigmoide e reto) são associados com instabilidade cromossômica e não apresentam hipermutação e MSI, frequentemente<sup>22,23</sup>. Além disso, tumores do cólon proximal são enriquecidos para o subtipo CMS1, ao passo que tumores distal apresentam maior frequência do subtipo CMS2<sup>22</sup>.

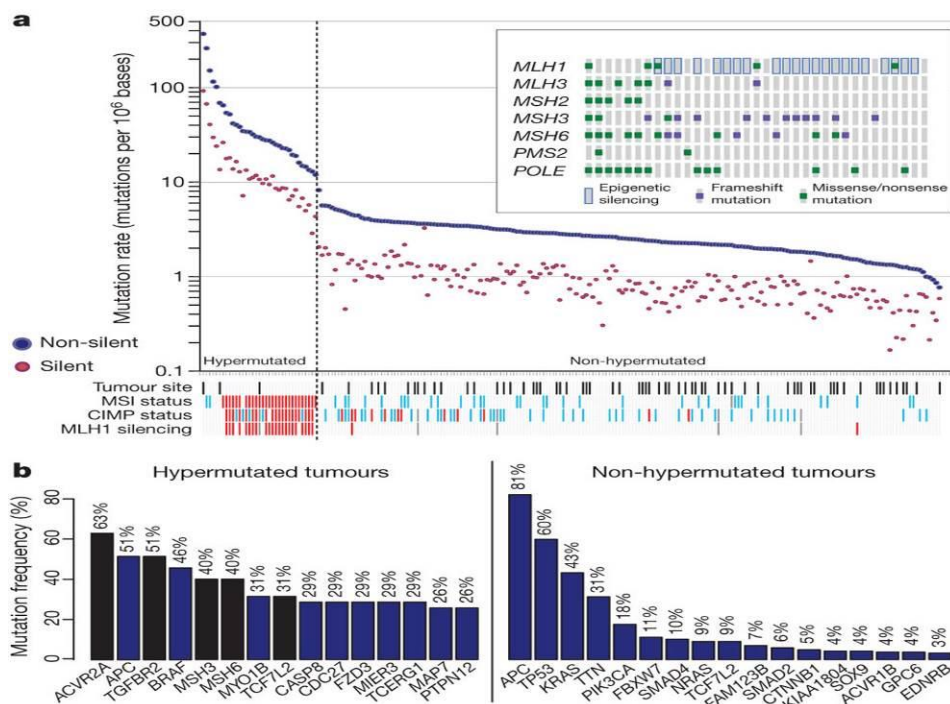
### **Impacto clínico da identificação de alterações moleculares do CCR**

A caracterização das alterações moleculares pode levar à identificação de marcadores genéticos presentes nos diferentes estágios de CCR e nas lesões precursoras, e assim ajudar na melhor compreensão da carcinogênese colorretal e na identificação de marcadores de diagnóstico/rastreamento, prognóstico e tratamento, com significativo impacto na história natural do CCR. A identificação de alterações genéticas em estágios iniciais de CCR e mesmo em lesões precursoras podem levar a possibilidade de rastreamento molecular da doença com o objetivo de aumentar a sensibilidade e especificidade de testes atualmente empregados com o objetivo final de redução na incidência e da mortalidade deste câncer. Estima-se que cerca de 95% dos casos de CCR poderiam se beneficiar de cirurgias curativas se o diagnóstico fosse feito em estágios iniciais ou ainda nos estádios pré-malignos<sup>35</sup>. De fato, muitos estudos demonstraram redução na mortalidade em regiões onde foram implantados programas de rastreamento de CCR<sup>35</sup>.

O Departamento de Prevenção do Hospital de Câncer de Barretos implantou, recentemente, o programa de rastreamento do câncer colorretal esporádico em indivíduos de médio risco (acima de 50 anos) com pesquisa de sangue oculto nas fezes seguida da colonoscopia nos casos de teste positivo. Entretanto, a sensibilidade do teste de sangue oculto nas fezes é ainda aquém do que se espera de um método de rastreamento. Tem sido cogitada a possibilidade de detecção de lesões precursoras e do CCR em estádios precoces através de testes moleculares em amostras do sangue, fezes e urina no contexto do rastreamento do CCR, com maior sensibilidade<sup>35,36</sup>.



A identificação de biomarcadores relevantes e confiáveis e a construção de um painel de alterações com elevada sensibilidade e especificidade na detecção de lesões colorretais aparecem como uma das metas mais desafiadoras para esta triagem molecular. Embora a caracterização molecular de CCR já tenha sido investigada em outros estudos<sup>23,37</sup> (**Figura 3**), o perfil mutacional do CCR na população brasileira ainda permanece em grande parte desconhecido<sup>25,38,39</sup>, principalmente no que diz respeito às lesões precursoras. Assim, a caracterização das alterações moleculares em lesões potencialmente malignas e em estádios iniciais de CCR é de extrema importância para obtenção de biomarcadores relevantes para o rastreamento e melhora da sobrevida na doença.



**Figura 3.** Frequência de mutações encontradas no câncer colorretal<sup>23</sup>. O estudo classifica os tumores colorretais em hipermutados e não hipermutados. Podemos observar a taxa de mutação em cada uma das 224 pacientes (a) e os genes alterados com maior frequência (b).

## JUSTIFICATIVA

Necessidade da caracterização das alterações genéticas presentes frequentemente no CCR e nas suas lesões precursoras em uma amostra da população brasileira para futura implantação no rastreamento molecular do CCR e melhora na conduta médica e terapêutica.

## OBJETIVOS

### 1. Geral

Investigar as alterações genéticas em diferentes estádios do CCR e nas suas lesões precursoras em uma amostra da população brasileira.

### 2. Específicos

- Realizar o perfil de mutações somáticas de um painel de 150 oncogenes e genes supressores tumorais em amostras de câncer colorretal por meio de sequenciamento de nova geração (plataforma Illumina);

- Realizar perfil de mutações somáticas de um painel de 50 genes supressores tumorais e oncogenes (2.800 mutações *hotspots*) em amostras de lesões colorretais por meio de sequenciamento de nova geração utilizando o Ion AmpliSeq Cancer Hotspot Panel v2 (plataforma Ion Torrent);

- Correlacionar os achados do perfil mutacional com dados clínicos e anátomo-patológicos.

A partir desse momento essa tese foi dividida em duas partes, a primeira direcionada à caracterização do perfil mutacional do câncer colorretal e a segunda, à caracterização do perfil mutacional das lesões precursoras do câncer colorretal. A primeira parte foi publicada na recentemente na revista *Scientific Reports* (**Anexo A**).

## PARTE 1: CARACTERIZAÇÃO DO PERFIL MUTACIONAL DO CÂNCER COLORRETAL

### MATERIAL E MÉTODOS

#### 1. População com câncer colorretal (CCR)

As amostras foram selecionadas de forma retrospectiva e incluiu 96 pacientes com CCR atendidos no Hospital de Câncer de Barretos (HCB) entre agosto de 2008 e janeiro de 2016, com diagnóstico de câncer colorretal em diferentes estádios, ressecados por cirurgia, com material biológico (sangue e tecido tumoral) armazenados no Biobanco e que consentiram o uso dessas amostras para estudos futuros.

#### 2. Extração de DNA

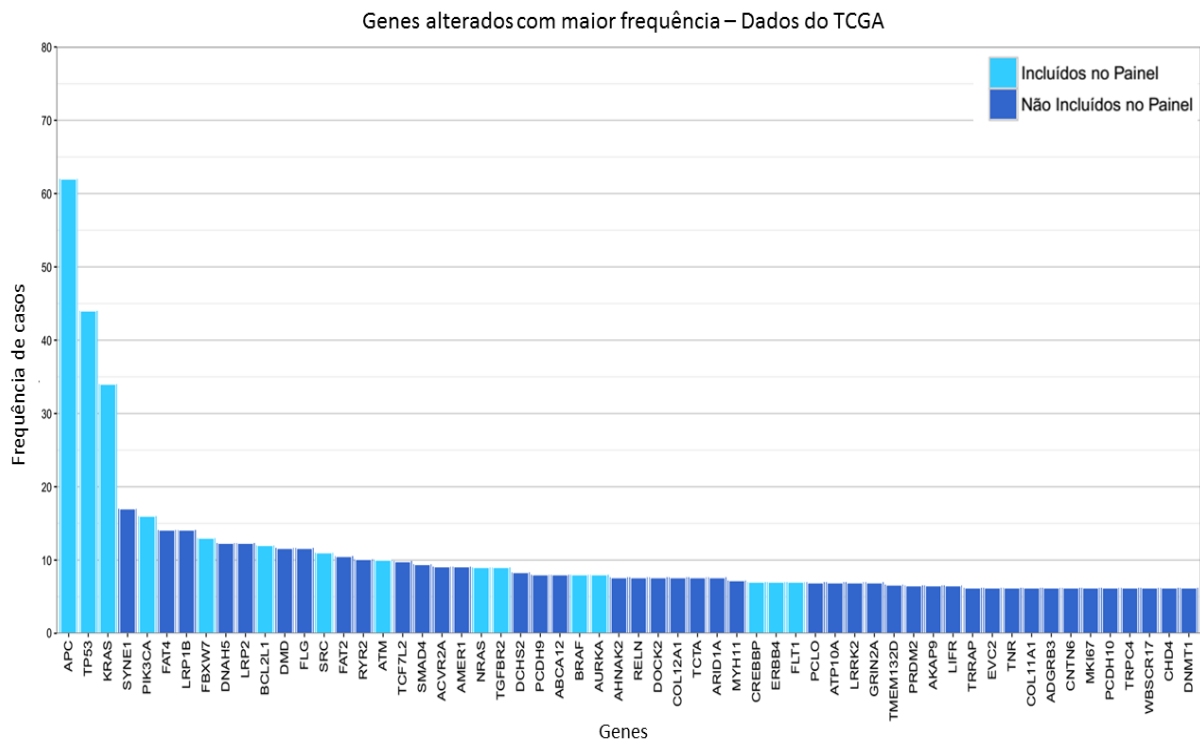
A preparação e extração de DNA das amostras de CCR foram realizadas pelo Biobanco do HCB, utilizando método automatizado através do QIASymphony®, o qual utiliza partículas magnéticas para extração e purificação de ácido nucleicos. A extração de DNA do tecido tumoral foi feita a partir de amostras congeladas. Para as amostras de tecido normal pareado, foi utilizado DNA de leucócitos isolado de sangue periférico (*buffy coat*) de forma automatizada por meio do QIASymphony®. A qualidade e quantidade do DNA extraído foi estabelecida por meio do Nanodrop e Qubit (Thermo Fisher Scientific, Waltham, MA, USA), respectivamente.

#### 3. Análise do perfil mutacional

##### 3.1. Análise das variantes por NGS utilizando a plataforma Illumina

O perfil mutacional das amostras de CCR foi avaliado por um painel de 150 oncogenes e genes supressores tumorais, nomeadamente: *ABL1, AKT1, AKT2, AKT3, ALK, APC, AR, ARAF, ATM, AURKA, AURKB, AXIN1, AXL, BAP1, BARD1, BCL2, BCL2L1, BCL2L2, BLM, BRAF, BRCA1, BRCA2, BRD2, BRD3, BRD4, BRDT, BRIP1, BTK, CBF3, CCND1, CCND2, CCND3, CCNE1, CD22,*

*CD274, CD79A, CD79B, CDH1, CDK12, CDK4, CDKN1A, CDKN2A, CHEK1, CHEK2, CREBBP, CRKL, DDR2, DDX3X, EGFR, EPHA7, ERBB2, ERBB3, ERBB4, ESR1, EZH2, FAM175A, FAS, FBXW7, FGFR1, FGFR2, FGFR3, FGFR4, FHIT, FLT1, FLT3, FLT4, FRS2, GATA2, GNA11, GNAQ, HDAC1, HDAC4, HDAC7, HGF, HRAS, IDH1, IDH2, IGF1R, JAK1, JAK2, JAK3, KDR, KIT, KRAS, MAP2K1, MAP2K2, MAP2K4, MAP3K1, MAPK1, MCL1, MDM2, MET, MLH1, MPL, MRE11A, MS4A1, MSH2, MSH6, MTOR, MUTYH, MYC, MYD88, NBN, NF1, NF2, NFE2L2, NFKBIA, NOTCH1, NOTCH2, NOTCH3, NRAS, NT5C2, NTRK1, PALB2, PDGFRA, PDGFRB, PDK1, PIK3CA, PIK3CB, PIK3CG, PMS2, PTCH1, PTEN, RAC1, RAD51, RAF1, RANBP2, RARA, RB1, RET, RICTOR, ROS1, RRM1, RUNX1, SDHB, SMO, SOX2, SRC, STAT3, STAT5B, STK11, TERC, TERT, TGFBR2, TP53, TSC1, TSC2, VEGFA, WT1, e XPO1.* Este painel abrange os genes alterados com maior frequência em CCR descrito em outras populações (**Figura 4**).



**Figura 4.** Genes frequentemente alterados segundo dados publicados em 2012 pelo TCGA (*The Cancer Genome Atlas*)<sup>23,40</sup>. Os genes incluídos no painel analisado no presente estudo (azul claro) englobam os genes alterados com maior frequência (acima de 6% de casos) em CCR.

Este painel cobre toda a região codificante destes genes (395 Kbases) em um total de 2382 regiões alvo. Além das regiões codificantes, 229 destas regiões possuem sobreposição com regiões promotoras (5000 bp *upstream*). O sequenciamento foi realizado na plataforma *Illumina HiSeq 2500™ System*, pela empresa de biotecnologia Mendelics (São Paulo).

Para cada amostra foi realizado o sequenciamento pareado de tecido tumoral com tecido normal (*buffy coat*), para identificar as mutações somáticas presentes nas amostras tumorais. Para a análise bioinformática, o serviço da *Mendelics* realizou o pré-processamento dos dados (controle de qualidade e alinhamento) com genoma de referência hg19 através de *BWA* (do inglês, *Burrows – Wheeler Aligner*), o qual mapeia e alinha *reads* sequenciadas contra o genoma de referência<sup>41</sup>. As variantes somáticas foram identificadas através do algoritmo *VarScan2*, o qual analisa dados de sequenciamento derivados do tumor e tecido normal correspondente simultaneamente, aplicando métodos estatísticos e heurísticos para detecção e classificação das variantes somáticas de acordo com seu *status* somático (germinativo, somático ou perda de heterozigosidade)<sup>42</sup>.

Para anotação, foram mantidas apenas variantes que apresentavam uma cobertura maior que 10x no sequenciamento, e apresentavam uma frequência alélica na amostra tumoral superior a 5%. A anotação das variantes, determinando seus efeitos nos genes (SNPs, inserções e deleções) foi realizada utilizando a ferramenta *VEP (Variant Effect Prediction)*<sup>43</sup> e posteriormente utilizando a ferramenta “*Cancer Genome Interpreter*”

### **3.2. Critério de patogenicidade para as variantes encontradas no câncer colorretal**

Para identificação das potenciais alterações drivers encontradas, utilizamos o “*Cancer Genome Interpreter*” (CGI)<sup>44</sup>. O CGI é uma ferramenta desenvolvida para automatizar a interpretação de variantes novas ou conhecidas como potenciais drivers para o tumor analisado. Para a caracterização de mutações já conhecidas, o CGI utiliza um catálogo de genes construído com informações derivadas de banco de dados manualmente curados ou genes que possuam evidência experimental ou clínica de capacidade de conduzir a tumorigênese. O Catálogo de Genes do Câncer gerado compreende 837 tumores em 193 tipos tumorais. A classificação das diferentes variantes encontradas nesses genes é feita por meio de um catálogo validado de 5314 mutações oncogênicas ou aplicação do *OncodriveMUT*, um algoritmo que classifica as mutações como drivers ou não de acordo com diversas características das mutações, tais como: o modo de ação do gene, a consequência biológica do tipo de mutação, a posição da mutação dentro do gene, seu impacto funcional predito, sua frequência em populações humanas e se a mutação ocorre em um domínio da proteína que tem poucas variantes germinativas. Com todas as informações integradas, o CGI classifica as

mutações como oncogênicas conhecidas, preditas como driver, preditas como não driver ou polimorfismos<sup>44</sup>. Como critério de patogenicidade, foram utilizadas apenas as mutações classificadas como oncogênicas (conhecidas ou preditas) pelo CGI. Para as definição de mutações *drivers* dos tumores, classificamos as amostras utilizando o CGI e mantivemos mutações que já foram identificadas como *drivers* em tumores em geral ou aquelas mutações que foram classificadas como preditas pelo algoritmo *OncodriveMUT*, excluindo, dessa forma variantes classificadas como polimorfismos (alta frequência alélica em banco de dados públicos), aquelas variantes preditas como neutras ou passageiras para o processo de oncogênese ou, ainda, aquelas variantes fora de regiões codificantes ou regiões de splice.

### 3.3. Análise estatística

As análises estatísticas foram realizadas com o programa *IBM SPSS Statistics*<sup>®</sup> (versão 21) para sistema operacional *Windows*<sup>®</sup>. A associação entre as frequências de alterações moleculares e as variáveis clínicas, histológicas, patológicas e de sangue oculto dos pacientes foi realizado por meio do teste  $\chi^2$  ou teste exato de Fisher.

## RESULTADOS

### 1. Características clínicas e anátomo-patológicas dos pacientes com CCR

Dentre os 96 pacientes com câncer colorretal selecionados, 1 amostra foi excluída por apresentar baixa cobertura no sequenciamento e, 4 amostras foram excluídas por suas amostras não apresentarem nenhuma alteração considerada patogênica de acordo com os critérios definidos nesse estudo, restando 91 pacientes incluídos no final do estudo.

As características clínico-patológicas dos 91 pacientes incluídos estão descritas na **Tabela 2**. A média de idade dos pacientes foi de 61,2 anos (desvio padrão: 14,0), a maioria sendo do gênero masculino. A maior parte dos tumores incluídos no estudo era de estadiamento inicial, 0, I e II (3,3, 33,0 e 41,8%, respectivamente) e estavam, em sua maioria, localizados no cólon distal (78,0%). Invasão angiolímfática foi observada em 18,9% dos tumores, 7,0% apresentaram invasão perineural e 3,4% dos pacientes tiveram recorrência.



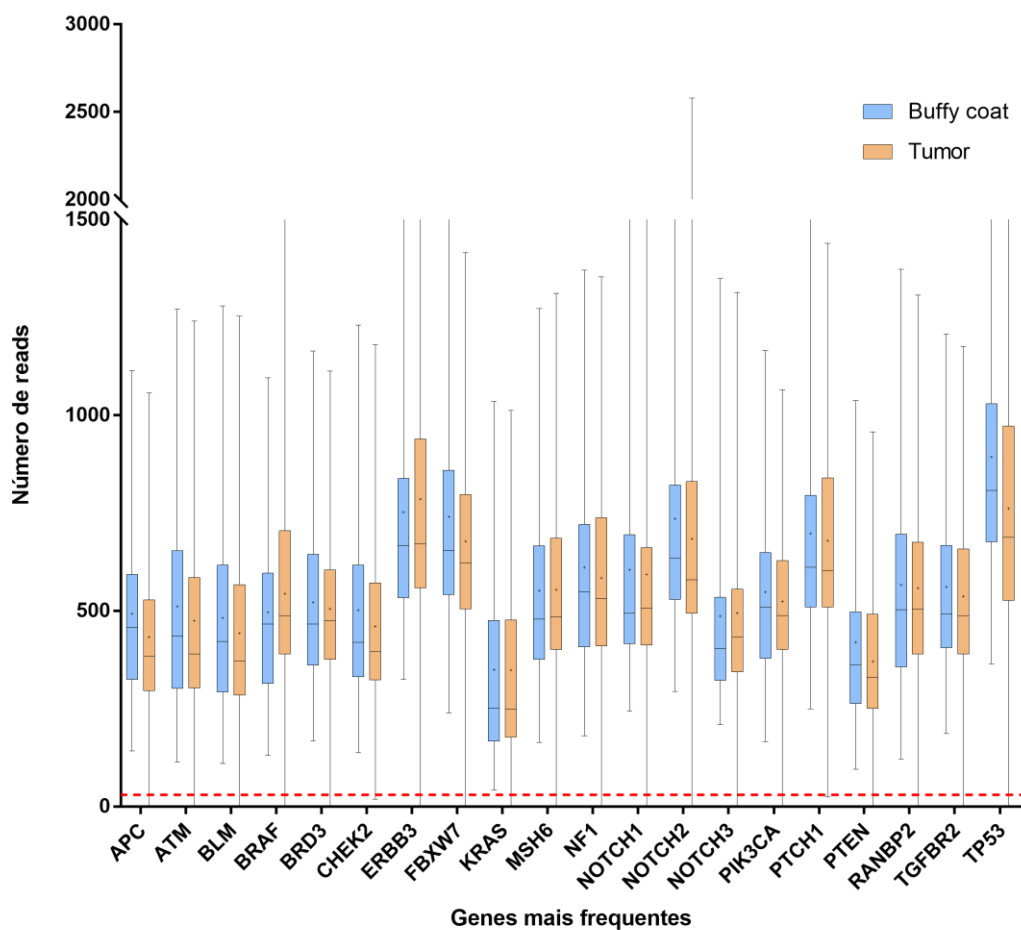
**Tabela 2 - Características clínico-patológicas dos pacientes com câncer colorretal.**

<b>Característica (n = 91)</b>	<b>Frequência</b>	<b>(%)</b>
Idade ao diagnóstico	Média 61,2	Desvio Padrão 14
<b>Gênero (n=91)</b>		
Masculino	49	53,8
Feminino	42	46,2
<b>Cor (n=89)</b>		
Branca	76	85,4
Negra	2	2,2
Parda	10	11,2
Amarelo	1	1,1
<b>Sítio primário do tumor (n=91)</b>		
Cólon proximal	20	22,0
Ceco	4	20,0
Cólon ascendente	9	45,0
Cólon transverso	7	35,0
Cólon distal	71	78,0
Cólon sigmoide	31	43,7
Junção retossigmoide	15	21,1
Reto	25	35,2
<b>Estadiamento (n=91)</b>		
0	3	3,3
I	30	33,0
II	38	41,8
III	16	17,6
IV	4	4,4
<b>TNM (n=91)</b>		
Tumor primário (T)		
Tis	5	5,5
T1	6	6,6
T2	28	30,8
T3	42	46,2
T4 (a,b)	10	11,0
Linfonodos regionais (N)		
N0	69	75,8
N1 (a,b,c)	15	16,5
N2 (a,b)	7	7,7
Metástase (M)		
M0	87	95,6
M1a	3	3,3
M1b	1	1,1
<b>Grau histológico (n=91)</b>		
I	10	11,0
II	77	84,6
III	4	4,4
Invasão angiolinfática (n=90)	17	18,9
Invasão perineural (n=86)	6	7,0
Quimioterapia adjuvante (n=91)	30	33,0
Radioterapia adjuvante (n=89)	4	4,4
Recorrência (n=89)	3	3,4
Instabilidade de microsatélite (n=90)		
MSS	75	83,3
MSI-L	3	3,3
MSI-H	12	13,3

## 2. Caracterização mutacional das amostras de câncer colorretal

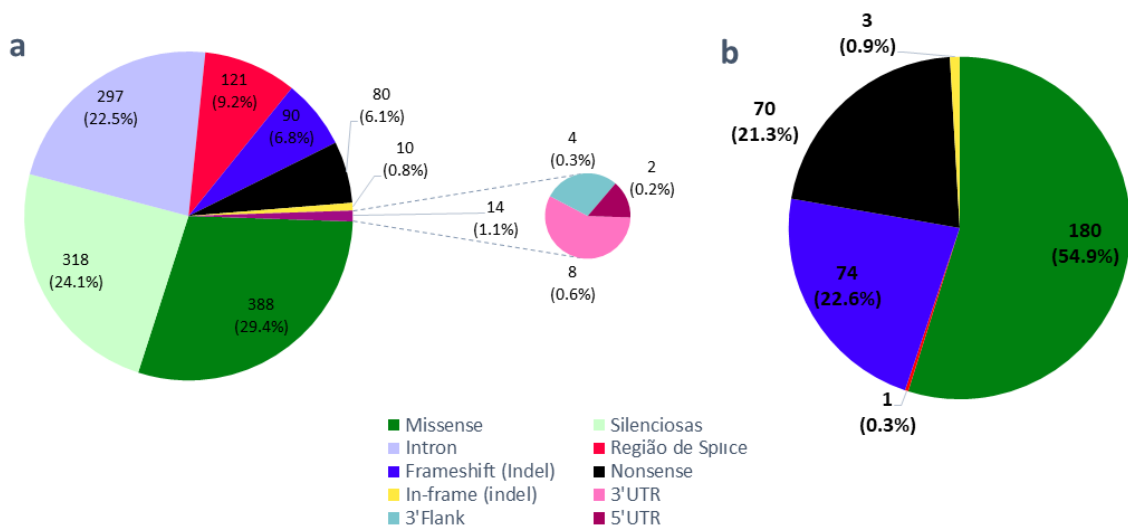
### 2.1. Cobertura do sequenciamento e filtro das variantes das amostras de CCR

A variância de cobertura dos principais genes incluídos no painel das amostras de câncer colorretal e das amostras de *buffy coat* pareadas é mostrada na **Figura 5**. Foi observada grande variância entre as amostras de câncer colorretal sequenciadas, apresentando uma média de cobertura de 625,7x por gene (desvio padrão de 173,7; mínimo: 264,4 e máximo: 1414,9). Uma variância semelhante foi observada para as amostras de *buffy coat* pareadas, as quais apresentaram uma média de cobertura de 632,5x por gene (desvio padrão: 166,2; mínimo: 271,2 e máximo: 1394,9). Uma das amostras sequenciadas foi excluída das análises posteriores por apresentar uma baixa cobertura, com uma média de cobertura de 2,9x.



**Figura 5.** Cobertura média dos genes mais frequentemente alterados nas amostras sequenciadas. Indicado em cada caixa está a mediana (linha) e a média (+) de cobertura por gene das 96 amostras sequenciadas. Podemos observar grande variância da cobertura e uma amostra ficou abaixo de 30x (linha vermelha tracejada) e foi retirada das análises posteriores.

No total, foram identificadas 1318 variantes, incluindo 318 variantes silenciosas, 297 variantes em regiões de *introns*, 388 variantes não silenciosas do tipo *missense*, 121 variantes em regiões de *splice*, 90 *frameshift*, 80 *nonsense*, 10 *in-frame* e 14 variantes em regiões flangeadores dos genes (**Figura 6a**). Uma lista completa das alterações encontradas pode ser vista no **Anexo B**. Após aplicar o filtro de patogenicidade, foram removidas 990 variantes não patogênicas de acordo com os critérios descritos previamente. Entre as 328 mutações mantidas, o tipo mais frequente foi de mutações do tipo *missense* (54,9%), seguidas de mutações *frameshift* (22,6%), *nonsense* (21,3%), *in-frame* (0,9%) e região de *splice* (0,3%, **Figura 6b**).

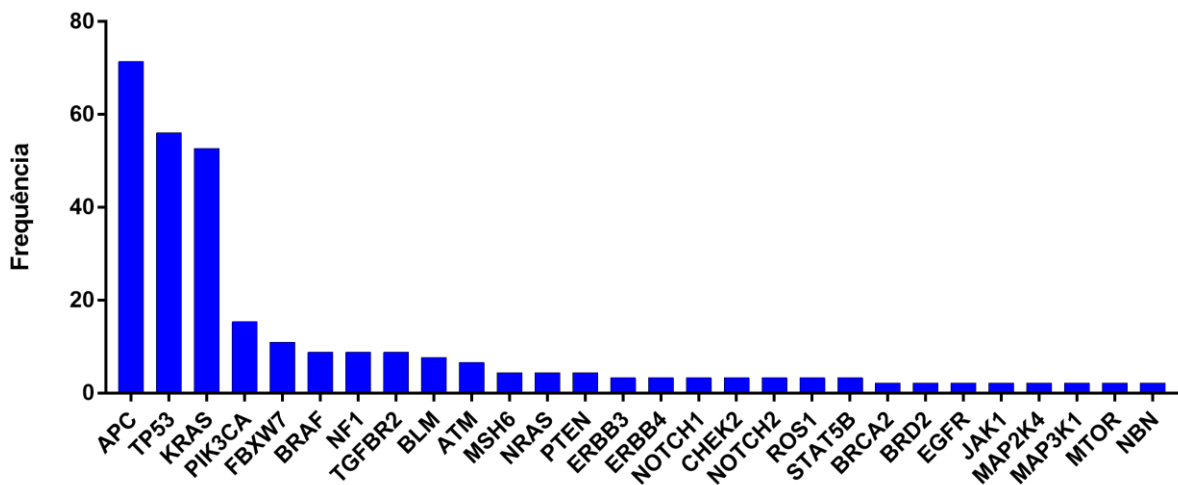


**Figura 6.** Variantes encontradas nas amostras sequenciadas, incluindo todas as variantes alteradas em relação ao tecido germinativo pareado (a) e mutações mantidas após o filtro de patogenicidade (b).

Após o filtro de patogenicidade, 4 amostras não apresentaram mutações patogênicas nos genes analisados de acordo com os critérios definidos previamente. Duas destas amostras apresentaram variantes apenas em regiões não codificantes e duas, além de variantes em regiões não codificantes, apresentaram variantes *missense* sem indícios de patogenicidade nos bancos de dados analisados.

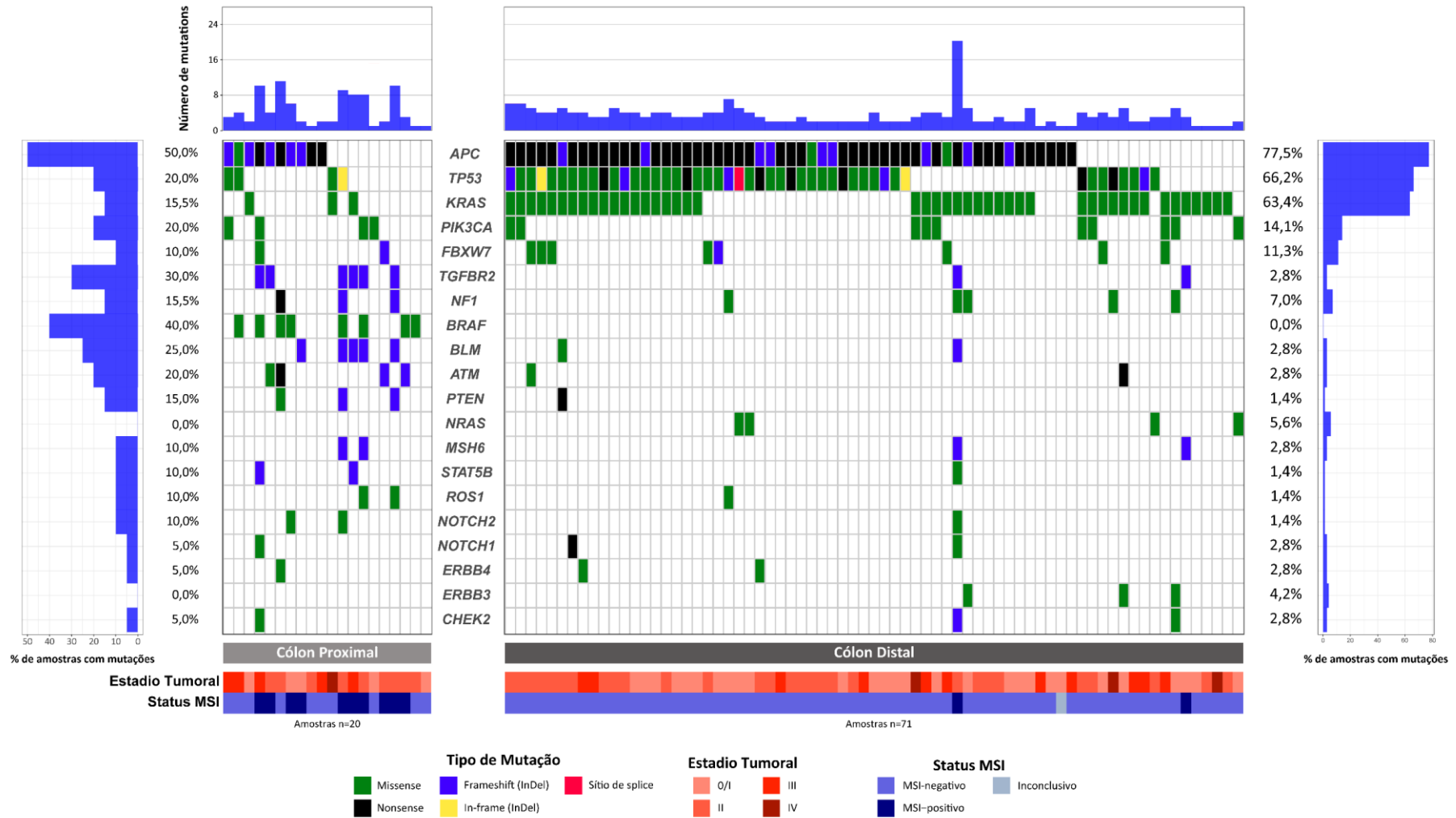
## 2.2. Perfil de mutações das amostras de CCR

Após o filtro de patogenicidade, como descrito anteriormente, 4 amostras foram removidas por não apresentarem nenhuma mutação considerada patogênica, restando 91 tumores que apresentaram pelo menos uma mutação. Encontramos uma média de 3,6 mutações por amostra tumoral, variando entre 1 a 20 mutações. Treze tumores (14,3%) apresentaram mutações em apenas 1 gene, 26 tumores (28,6%) apresentaram mutações em 2 genes, 24 (26,4%) em 3 genes e 28 (30,8%) tumores apresentaram mutações em 4 ou mais genes. Dos 150 genes sequenciados, 49 apresentaram pelo menos uma mutação. Como mostra a **Figura 7**, o gene *APC* foi o gene alterado com maior frequência em nossa população (72,5%), seguido pelo gene *TP53* (56,0%), *KRAS* (52,7%), *PIK3CA* (15,4%) e *FBXW7* (12,1%, **Figura 7**). A lista completa das variantes encontrada pode ser vista no **Anexo B**.



**Figura 7.** Genes alterados com maior frequência em nossa população.

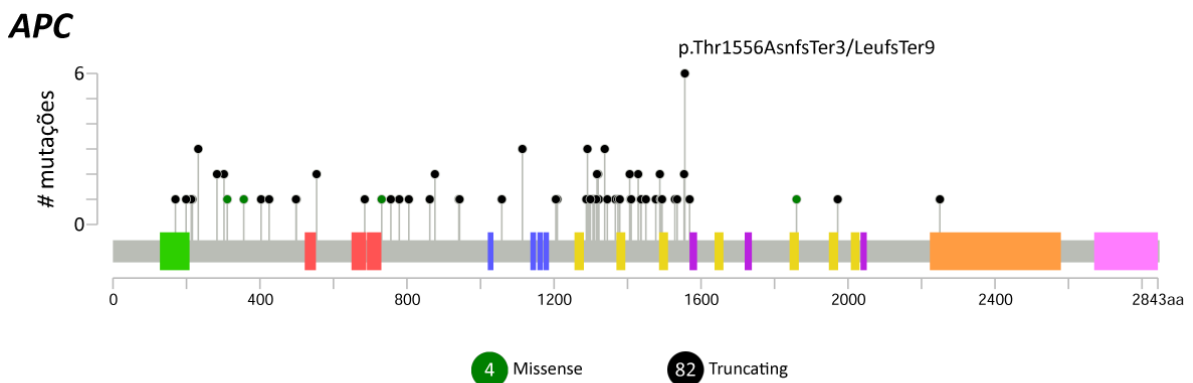
A **Figura 8** mostra os genes alterados com maior na nossa população de acordo com características clínico-patológicas. Podemos observar diferenças no perfil mutacional entre o cólon proximal e distal, como a presença de maior frequência de mutações em genes como *BRAF*, *TGFBR2* no cólon proximal e maior frequência de alteração nos genes *KRAS* e *TP53* no cólon distal.



**Figura 8.** Mutações *drivers* encontradas nos genes analisados em nossa população de estudo classificados por frequência de tumores alterados (apenas os genes com alterações em mais de 3% das amostras são mostrados). A parte central da figura mostra os tipos de mutações de cada gene distribuídas por cada amostra. No topo podemos observar a contagem de mutação por amostra. As laterais mostram a frequência de mutação em cada gene de acordo com a localização (cólón proximal ou distal). Na parte inferior podemos observar o estadiamento tumoral bem como status de instabilidade de microssatélite (MSI).

### 2.2.1 Mutações no gene *APC*

No total, o gene *APC* apresentou 86 mutações distribuídas em 65 tumores, a maioria apresentando apenas uma mutação no gene (44 tumores), e 21 tumores apresentaram duas mutações. A maioria das mutações encontradas eram mutações truncantes de proteínas, 52 (60,5%) *nonsense* e 30 (34,8%) *frameshift*, sendo apenas 4 (4,6%) mutações *missense* (**Figura 9**). Dentre as mutações encontradas, 44 eram mutações já registradas no COSMIC ou dbSNP (34 mutações *nonsense* e 10 *frameshift*) e 22 eram variantes novas e preditas como patogênicas pelo CGI (16 mutações *frameshift*, 4 *missense* e 2 *nonsense*). As mutações encontradas estavam distribuídas ao longo de todo o gene e o sítio proteico mais afetado foi a posição 1556 da proteína, apresentando alterações *frameshift* em 6 tumores (**Figura 9**).

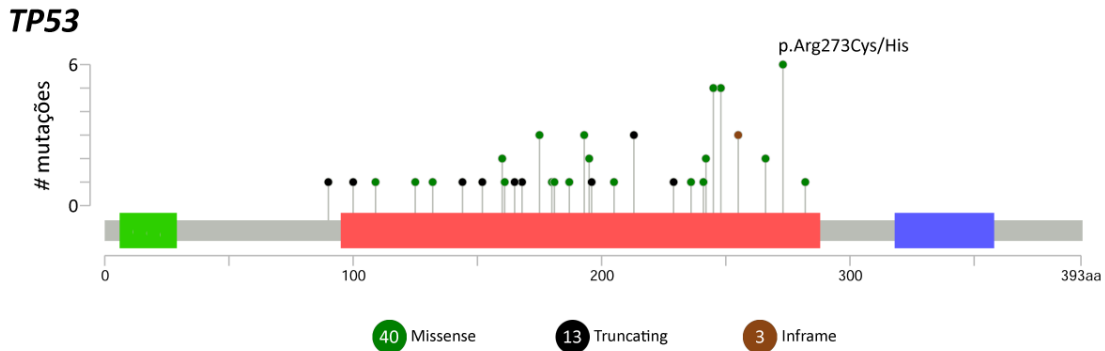


**Figura 9.** Diagrama das mutações no gene *APC*. A maioria das mutações encontradas foram truncantes, incluindo mutações *nonsense* e *frameshift*. Em destaque, podemos observar a posição 1556, que apresentou maior frequência de alteração (6 mutações).

### 2.2.2 Mutações no gene *TP53*

Mutações no gene *TP53* ocorreram em 51 (56,0%) tumores, dos quais 46 apresentaram uma única mutação no gene e 5 tiveram duas mutações, totalizando 56 alterações. Mutações do tipo *missense* foram as mais frequentes (71,4% das mutações), seguidas de mutações *nonsense* (12,5%), *frameshift* (8,9%), *in-frame* (5,4%, **Figura 10**) e sítio de *splice* (1,8%). Entre as mutações identificadas em nossa população, a maioria já estava registrada no COSMIC ou

dbSNP (40 mutações *missense*, 7 *nonsense*, 3 *frameshift*, 2 *in-frame* e 1 sítio de *splice*) e 3 eram novas mutações preditas como patogênicas pelo CGI (2 *frameshift* e 1 *in-frame*). O domínio de ligação ao DNA foi o domínio mais afetado da proteína, com apenas uma mutação encontrada fora dessa região. O sítio proteico mais afetado foi o resíduo 273 da proteína, apresentando 6 mutações (**Figura 10**).

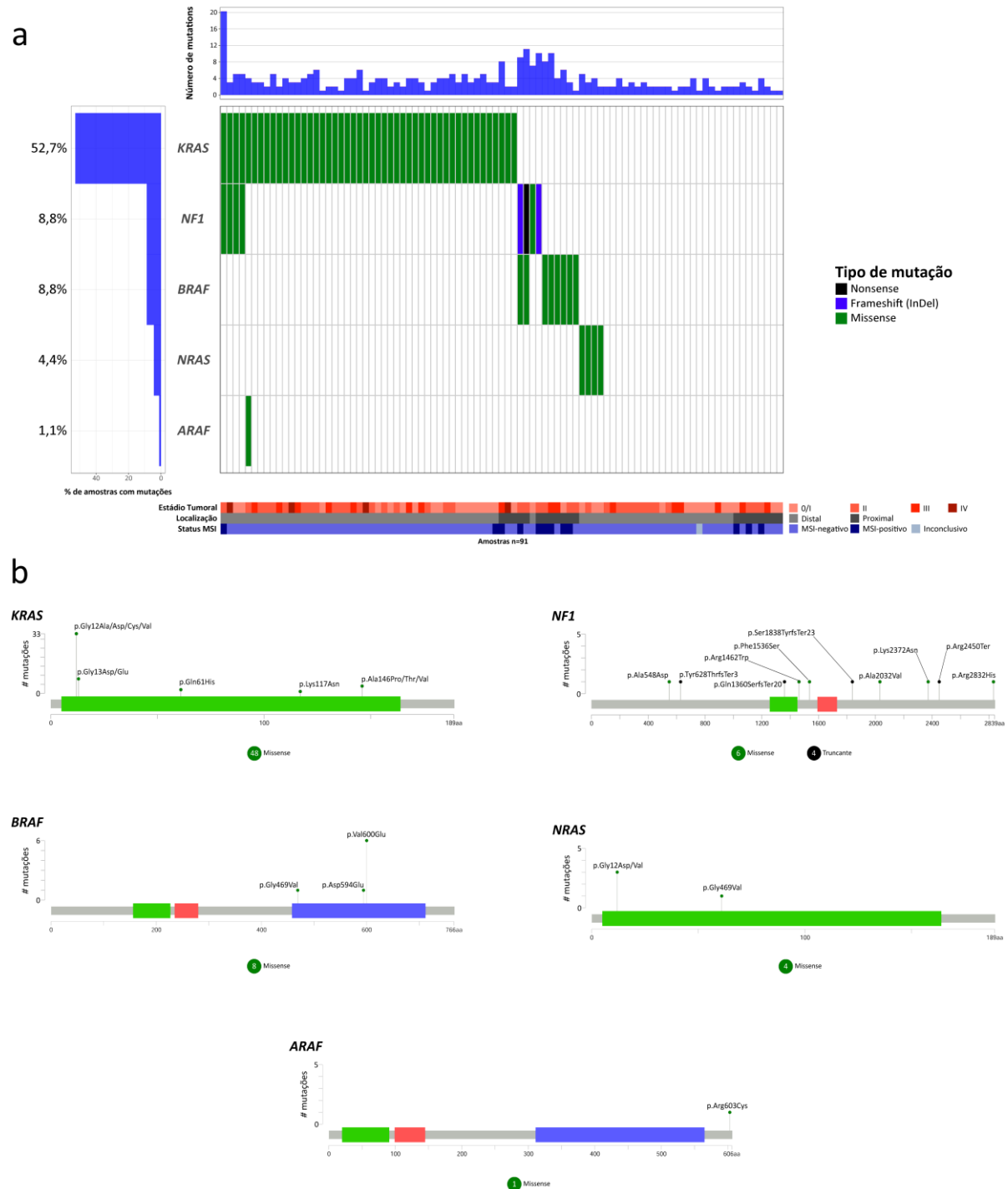


**Figura 10.** Diagrama das mutações no gene *TP53*. A maioria das mutações encontradas no gene foram do tipo *missense* (pontos verdes) e se concentravam na região de ligação ao DNA (caixa em vermelho). Em destaque podemos observar a posição 273, que apresentou a maior frequência de alteração (6 mutações).

### 2.2.3 Alterações na via MAPK-ERK

Dentre as vias alteradas, a via de sinalização MAPK-ERK apresentou pelo menos um gene mutado em 62 tumores (68,0%). Entre os genes alterados, o gene *KRAS* apresentou maior frequência mutações (48 mutações *missense*), todas em *hotspots* já conhecidos do gene (**Figura 11**), afetando os códons 12 e 13 no éxon 2 (68,7 e 16,7% das mutações encontradas no gene, respectivamente), o códon 61 (4,2%) no éxon 3 e os códons 117 e 146 no éxon 4 (2,1 e 8,3%, respectivamente). O gene *NF1* apresentou 10 mutações em 8 amostras tumorais, a maioria (7) sendo mutações novas preditas como patogênicas (4 *missense* e 3 *frameshift*) e 3 mutações já conhecidas (2 *missense* e 1 *nonsense*). Todas as mutações truncantes (*nonsense* e *frameshift*) no gene *NF1* foram mutualmente exclusivas com mutações no gene *KRAS*, ao passo que 4 tumores apresentaram mutações *missense* em ambos os genes (**Figura 11a**). O gene *BRAF* teve 8 mutações *missense* mutualmente exclusiva com o gene *KRAS* e *NRAS*. Todas as mutações encontradas em *BRAF* eram mutações já conhecidas: p.Val600Glu (p.V600E), sendo a mais frequente, afetando 75% (6 tumores) de todas as mutações no gene *BRAF* e duas mutações que também afetam o domínio quinase da proteína (p.Asp594Glu e p.Gly469Val,

**Figura 11b).** O gene *NRAS* apresentou 4 mutações já conhecidas que foram mutuamente exclusivas com o *KRAS*, *NF1* e *BRAF*. Outro gene afetado na via foi o gene *ARAF*, que apresentou uma mutação já conhecida fora dos principais domínios da proteína (**Figura 11**).

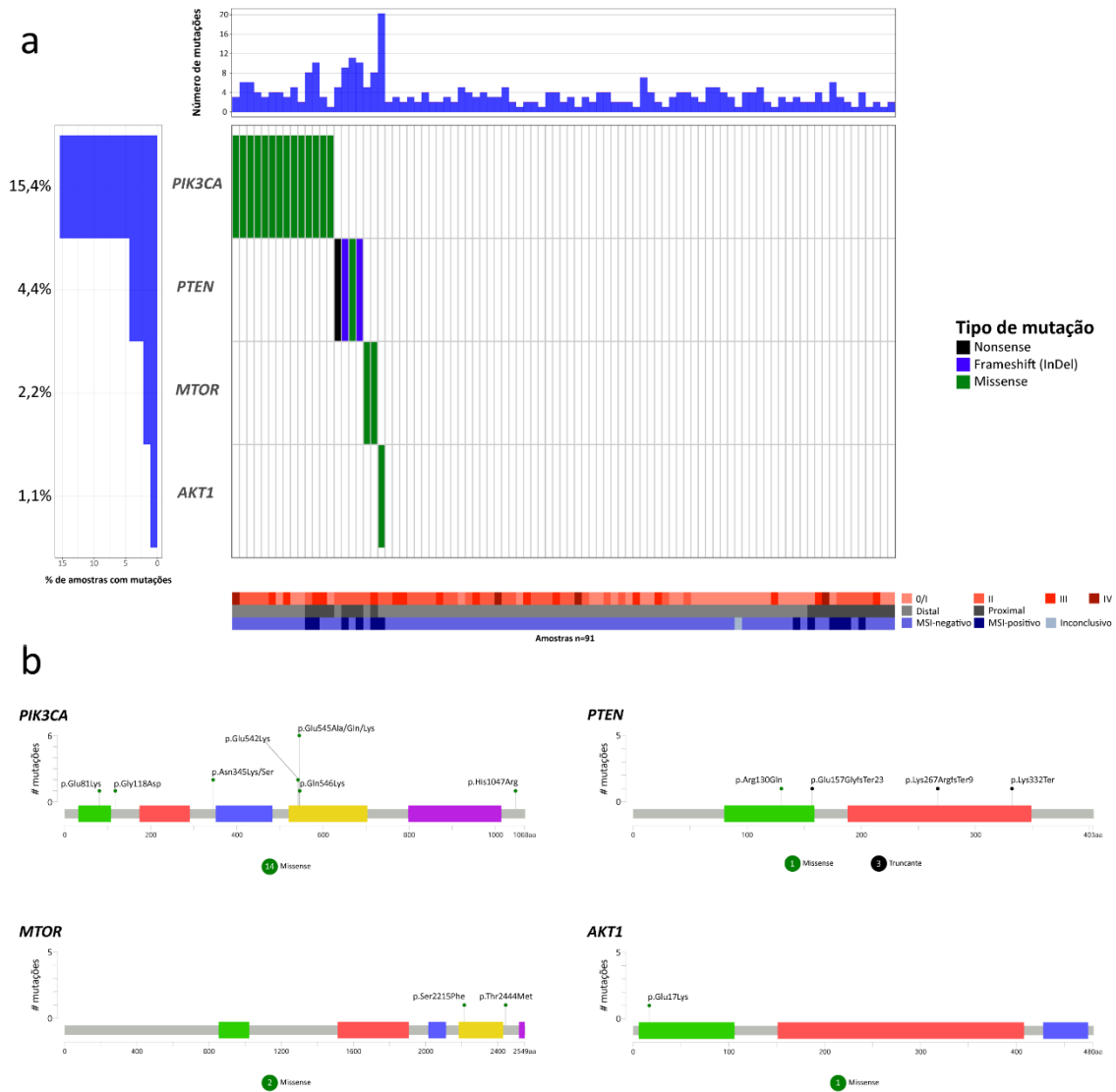


**Figura 11.** Genes alterados na via MAPK (a) e a distribuição de mutações ao longo dos principais domínios proteicos destes genes (b). MSI: instabilidade microsatélite.



#### 2.2.4 Alterações na via PI3K-AKT

Alterações na via PI3K-AKT ocorreram em 21 (23,1%) tumores. O gene com maior frequência de alteração foi *PIK3CA*, apresentando mutações em 14 (15,4%) tumores (**Figura 12a**), sendo todas mutações *missense* já conhecidas. O domínio PIK da proteína apresentou a maioria das mutações encontradas no gene, incluindo 6 mutações no códon 545 (42,9% das mutações no gene), 2 mutações no códon 542 (14,3%) e uma mutação no códon 546 (7,1%, **Figura 12b**). Outras mutações ocorreram fora do domínio PIK, incluindo uma mutação no domínio de ligação p85 (**Figura 12b**) da proteína. O gene *PTEN* apresentou mutações em 4,4% dos tumores, incluindo duas mutações já conhecidas (1 *frameshift* e 1 *missense*) e duas novas mutações preditas como patogênicas (1 *frameshift* e 1 *nonsense*). Outras mutações na via incluem alterações no gene *MTOR*, que apresentou duas mutações *missense* já conhecidas, incluindo uma mutação no domínio P3K/PI4K da proteína e alteração no gene *AKT1*, que apresentou uma mutação *missense* já conhecida (**Figura 12b**). Como mostrado na **Figura 12**, as mutações encontradas nesta via foram mutuamente exclusivas.

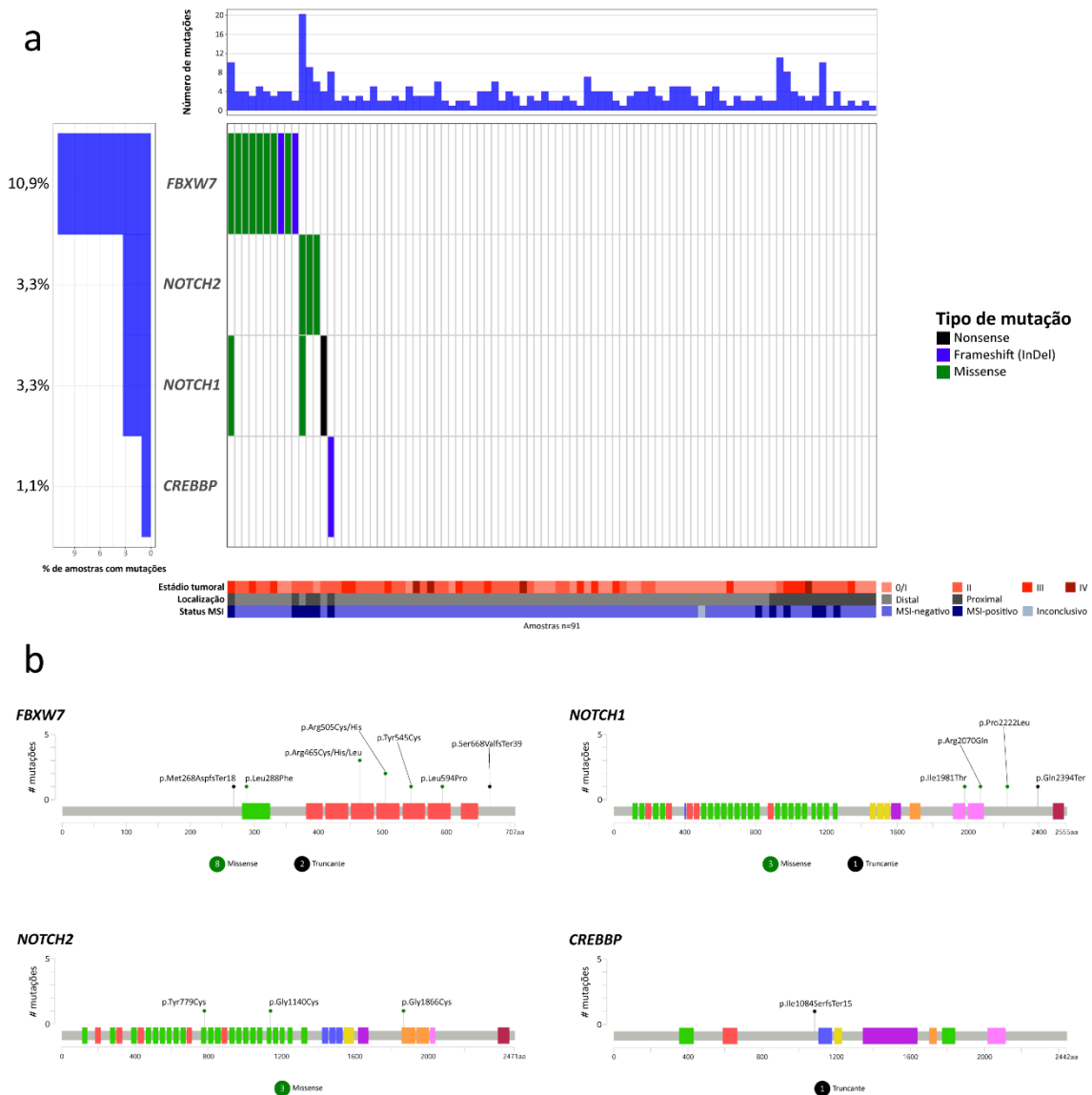


**Figura 12.** Oncoplot dos genes alterados na via PI3K-AKT (a) e a distribuição das mutações ao longo dos principais domínios proteicos destes genes (b). MSI: instabilidade microsatélite.

### 2.2.5 Alterações na via Notch

Mutações nessa via ocorreram em 16,5% dos tumores (**Figura 13a**). O gene mais frequentemente alterado foi o gene *FBXW7*, que apresentou mutações em 10,9% (10) dos tumores, incluindo 9 mutações já conhecidas (7 mutações *missense* e 2 *frameshift*) e 1 nova mutação *missense* predita como patogênica (**Figura 13b, Anexo B**). A maioria das mutações no gene *FBXW7* (70,0%) afetaram um dos domínios repetitivos de transducina beta (**Figura 13b**), uma mutação ocorrendo no domínio F-box-like da proteína e outras duas afetando regiões fora dos domínios proteicos principais. O gene *NOTCH2* estava alterado em 3,3% (3) dos tumores, apresentando novas mutações preditas com patogênicas nos domínios repetitivos EGF-like e no domínio repetitivo *ankyrin* (**Figura 13b**). O gene *NOTCH1* também

apresentou mutações em 3,3% dos tumores, afetando um dos domínios repetitivos *ankyrin* e regiões fora dos principais domínios proteicos (**Figura 13b**), incluindo 3 mutações já conhecidas (2 mutações missense e 1 mutação nonsense) e uma nova mutação predita como patogênica. Outro gene afetado na via foi o gene *CREBBP* (1,1% dos tumores), que apresentou uma mutação já conhecida localizada fora dos principais domínios da proteína (**Figura 13b**).



## 2.2.6 Alterações nos receptores de tirosina quinase (RTKs)

Genes da família dos receptores de tirosina quinases apresentaram mutações em 15,4% das amostras tumorais (**Figura 14**), apresentando mutações em 9 genes, todas mutações do tipo *missense*. O gene *ROS1* apresentou mutações em 3,3% dos tumores, incluindo uma

mutação já conhecida no domínio proteico tirosina quinase e duas novas mutações *missense* preditas com patogênicas fora dos principais domínios da proteína (**Figura 14b**). O gene *ERBB4* apresentou mutações em 3 (3,3%) tumores, incluindo uma nova mutação no domínio *furin-like* rico em cisteína e duas mutações já conhecidas, uma afetando o domínio tirosina quinase e uma fora dos principais domínios proteicos (**Figura 14b**). O gene *ERBB3* também estava alterado em 3 (3,3%) amostras tumorais, apresentando uma nova mutação no domínio *furin-like* rico em cisteína e, duas mutações já conhecidas, uma afetando o domínio tirosina quinase, e duas novas mutações no domínio receptor de fator de crescimento IV (**Figura 14b**). O gene *EGFR* estava alterado em 2 (2,2%) tumores, incluindo duas novas mutações no domínio de tirosina quinase (**Figura 14b**). O gene *RET* estava alterado em 1 tumor (1,1%), com duas mutações no domínio tirosina quinase (**Figura 14b**). O gene *NTRK1* também estava alterado em 1 (1,1%) tumor, com uma nova mutação fora dos domínios principais da proteína (**Figura 14b**). O gene *FLT3* apresentou uma nova mutação no domínio tirosina quinase e o gene *FLT4* apresentou uma mutação já conhecida no domínio *immunoglobulin I-set* (**Figura 14b**). O gene *ALK* também apresentou uma mutação já conhecida no domínio tirosina quinase (**Figura 14b**).

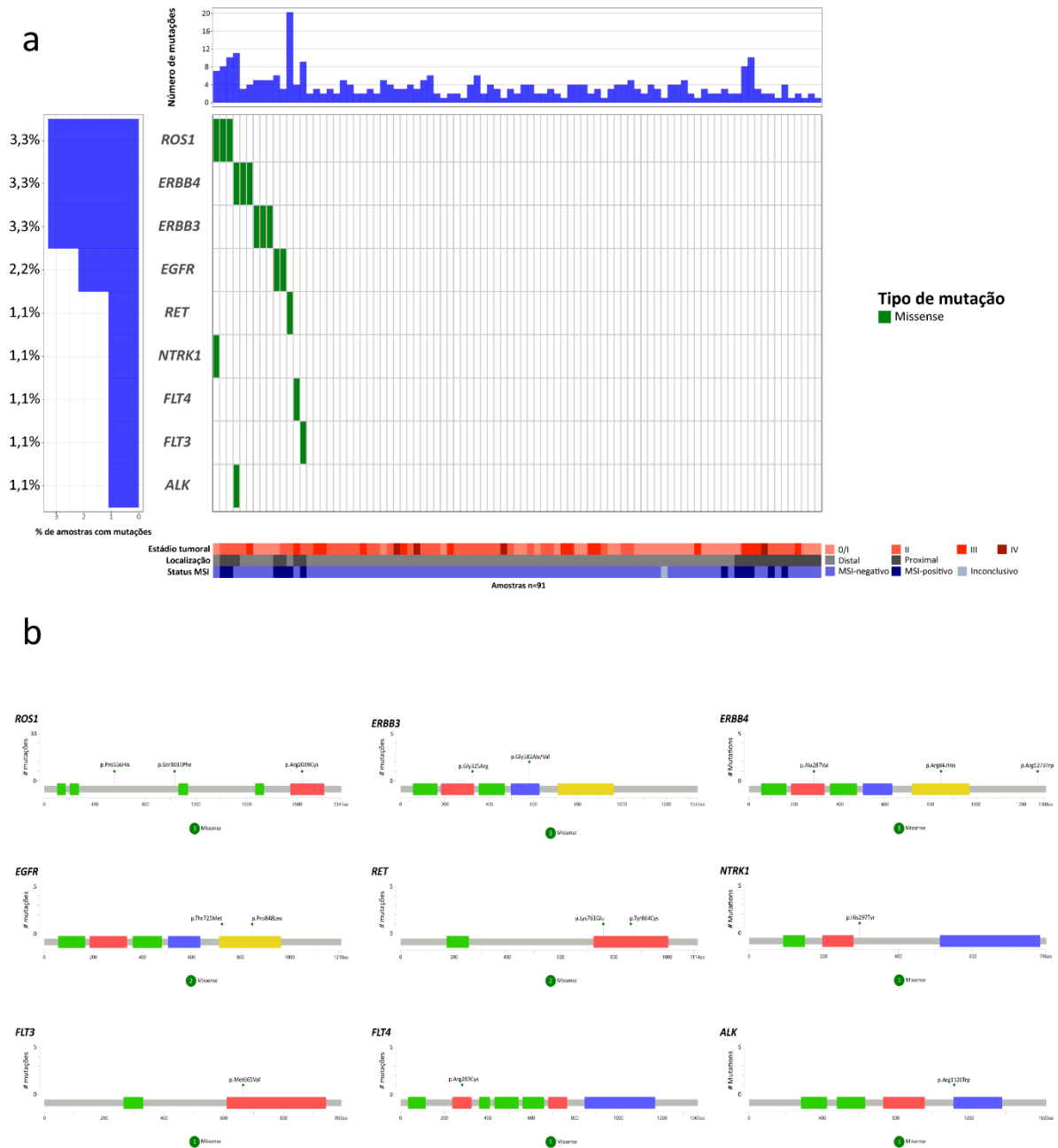


Figura 14. Genes alterados na família de tirosina quinases. MSI: instabilidade microsatélite.

### 2.3 Associação entre mutações e dados clínico-patológicos

As Tabela 3 a Tabela 4 apresentam as associações entre os dados sociodemográficos (sexo e cor da pele), as características clínico-patológicas dos pacientes incluídos e a presença de mutações nos genes com maior frequência de alteração (acima de 3%) na nossa população.

**Tabela 3.** Associação entre os genes com maior frequência de mutações e as características sociodemográficas

	Gênero			Cor				<i>p</i>
	Feminino		<i>p</i>	Amarelo	Branco	Negro	Pardo	
	n (%)	n (%)						
<i>APC</i>	29 (69,0)	36 (73,5)	0,816*	1 (100,0)	53 (69,7)	2 (100,0)	8 (80,0)	0,869
<i>TP53</i>	21 (50,0)	31 (63,3)	0,288*	0 (0,0)	42 (55,3)	2 (100,0)	7 (70,0)	0,371
<i>KRAS</i>	21 (50,0)	27 (55,1)	0,677*	1 (100,0)	40 (52,6)	1 (50,0)	5 (50,0)	0,999
<i>PIK3CA</i>	5 (11,9)	9 (18,4)	0,562*	0 (0,0)	13 (17,1)	0 (0,0)	0 (0,0)	0,596
<i>FBXW7</i>	3 (7,1)	7 (14,3)	0,331	0 (0,0)	8 (10,5)	0 (0,0)	2 (20,0)	0,532
<i>NF1</i>	4 (9,5)	4 (8,2)	0,999	0 (0,0)	7 (9,2)	1 (50,0)	0 (0,0)	0,284
<i>TGFBR2</i>	4 (9,5)	4 (8,2)	0,999	0 (0,0)	7 (9,2)	0 (0,0)	0 (0,0)	0,999
<i>ATM</i>	3 (7,1)	3 (6,1)	0,999	0 (0,0)	5 (6,6)	0 (0,0)	1 (10,0)	0,624
<i>BRAF</i>	6 (14,3)	2 (4,1)	0,137	0 (0,0)	7 (9,2)	0 (0,0)	0 (0,0)	0,999
<i>BLM</i>	5 (11,9)	2 (4,1)	0,242	0 (0,0)	6 (7,9)	0 (0,0)	0 (0,0)	0,999
<i>MSH6</i>	3 (7,1)	1 (2,0)	0,332	0 (0,0)	3 (3,9)	0 (0,0)	0 (0,0)	0,999
<i>NOTCH1</i>	1 (2,4)	2 (4,1)	0,999	0 (0,0)	3 (3,9)	0 (0,0)	0 (0,0)	0,999
<i>NRAS</i>	2 (4,8)	2 (4,1)	0,999	0 (0,0)	4 (5,3)	0 (0,0)	0 (0,0)	0,999
<i>PTEN</i>	3 (7,1)	1 (2,0)	0,332	0 (0,0)	4 (5,3)	0 (0,0)	0 (0,0)	0,999
<i>CHEK2</i>	0 (0,0)	3 (6,1)	0,246	0 (0,0)	3 (3,9)	0 (0,0)	0 (0,0)	0,999
<i>ERBB3</i>	1 (2,4)	2 (4,1)	0,999	0 (0,0)	3 (3,9)	0 (0,0)	0 (0,0)	0,999
<i>ERBB4</i>	3 (7,1)	0 (0,0)	0,094	0 (0,0)	2 (2,6)	1 (50,0)	0 (0,0)	0,130
<i>NOTCH2</i>	1 (2,4)	2 (4,1)	0,999	0 (0,0)	3 (3,9)	0 (0,0)	0 (0,0)	0,999
<i>ROS1</i>	1 (2,4)	2 (4,1)	0,999	0 (0,0)	1 (1,3)	1 (50,0)	0 (0,0)	0,078
<i>STAT5B</i>	1 (2,4)	2 (4,1)	0,999	0 (0,0)	3 (3,9)	0 (0,0)	0 (0,0)	0,999

Valor *p* do Teste Exato de Fisher ou Teste  $\chi^2$  (\*)

**Tabela 4.** Associação entre os genes com maior frequência de mutações e as características clínico-patológicas dos tumores

	Estadiamento			<i>p</i>	Linfonodos regionais		<i>p</i>	Metástase		<i>p</i>	Invasão angiolinfática		<i>p</i>	Invasão perineural		<i>p</i>
	0/I	II	III/IV		N0	N1/N2		M0	M1a/M1b		Não	Sim		Não	Sim	
	n (%)	n (%)	n (%)		n (%)	n (%)		n (%)	n (%)		n (%)	n (%)		n (%)	n (%)	
<i>APC</i>	25 (75,8)	27 (71,1)	13 (65,0)	0,726*	51 (73,9)	14 (63,6)	0,419*	64 (73,6)	1 (25,0)	0,069*	51 (69,9)	13 (76,5)	0,769*	58 (72,5)	5 (83,3)	0,999*
<i>TP53</i>	17 (51,5)	24 (63,2)	11 (55,0)	0,608*	39 (56,5)	13 (59,1)	0,999*	50 (57,5)	2 (50,0)	0,999*	44 (60,3)	8 (47,1)	0,415*	48 (60,0)	3 (50,0)	0,684*
<i>KRAS</i>	17 (51,5)	19 (50,0)	12 (60,0)	0,798*	34 (49,3)	14 (63,6)	0,328*	44 (50,6)	4 (100,0)	0,119*	39 (53,4)	9 (52,9)	0,999*	41 (51,3)	5 (83,3)	0,209
<i>PIK3CA</i>	4 (12,1)	5 (13,2)	5 (25,0)	0,427*	7 (10,1)	7 (31,8)	<b>0,036</b>	13 (14,9)	1 (25,0)	0,494*	12 (16,4)	2 (11,8)	0,999	11 (13,8)	2 (33,3)	0,585
<i>FBXW7</i>	2 (6,1)	5 (13,2)	3 (15,0)	0,519	7 (10,1)	3 (13,6)	0,699	10 (11,5)	0 (0,0)	0,999	7 (9,6)	3 (17,6)	0,392	8 (10,0)	1 (16,7)	0,496
<i>NF1</i>	3 (9,1)	4 (10,5)	1 (5,0)	0,897	7 (10,1)	1 (4,5)	0,674	7 (8,0)	1 (25,0)	0,312	6 (8,2)	1 (5,9)	0,999	6 (7,5)	0 (0,0)	0,999
<i>TGFBR2</i>	1 (3,0)	5 (13,2)	2 (10,0)	0,357	6 (8,7)	2 (9,1)	0,999	8 (9,2)	0 (0,0)	0,999	6 (8,2)	2 (11,8)	0,643	6 (7,5)	0 (0,0)	0,999
<i>ATM</i>	1 (3,0)	5 (13,2)	0 (0,0)	0,143	6 (8,7)	0 (0,0)	0,329	6 (6,9)	0 (0,0)	0,999	4 (5,5)	1 (5,9)	0,999	5 (6,3)	0 (0,0)	0,999
<i>BRAF</i>	1 (3,0)	5 (13,2)	2 (10,0)	0,357	6 (8,7)	2 (9,1)	0,999	8 (9,2)	0 (0,0)	0,999	6 (8,2)	1 (5,9)	0,999	6 (7,5)	0 (0,0)	0,999
<i>BLM</i>	1 (3,0)	5 (13,2)	1 (5,0)	0,322	6 (8,7)	1 (4,5)	0,999	7 (8,0)	0 (0,0)	0,999	6 (8,2)	1 (5,9)	0,999	5 (6,3)	0 (0,0)	0,999
<i>MSH6</i>	1 (3,0)	3 (7,9)	0 (0,0)	0,537	4 (5,8)	0 (0,0)	0,569	4 (4,6)	0 (0,0)	0,999	4 (5,5)	0 (0,0)	0,999	3 (3,8)	0 (0,0)	0,999
<i>NOTCH1</i>	0 (0,0)	2 (5,3)	1 (5,0)	0,437	2 (2,9)	1 (4,5)	0,569	3 (3,4)	0 (0,0)	0,999	2 (2,7)	1 (5,9)	0,471	3 (3,8)	0 (0,0)	0,999
<i>NRAS</i>	3 (9,1)	1 (2,6)	0 (0,0)	0,344	4 (5,8)	0 (0,0)	0,569	4 (4,6)	0 (0,0)	0,999	4 (5,5)	0 (0,0)	0,999	4 (5,0)	0 (0,0)	0,999
<i>PTEN</i>	0 (0,0)	4 (10,5)	0 (0,0)	0,075	4 (5,8)	0 (0,0)	0,569	4 (4,6)	0 (0,0)	0,999	3 (4,1)	0 (0,0)	0,999	2 (2,5)	0 (0,0)	0,999
<i>CHEK2</i>	1 (3,0)	1 (2,6)	1 (5,0)	0,999	2 (2,9)	1 (4,5)	0,569	3 (3,4)	0 (0,0)	0,999	2 (2,7)	1 (5,9)	0,471	3 (3,8)	0 (0,0)	0,999
<i>ERBB3</i>	3 (9,1)	0 (0,0)	0 (0,0)	0,054	3 (4,3)	0 (0,0)	0,999	3 (3,4)	0 (0,0)	0,999	3 (4,1)	0 (0,0)	0,999	3 (3,8)	0 (0,0)	0,999
<i>ERBB4</i>	0 (0,0)	2 (5,3)	1 (5,0)	0,437	2 (2,9)	1 (4,5)	0,569	3 (3,4)	0 (0,0)	0,999	2 (2,7)	0 (0,0)	0,999	2 (2,5)	0 (0,0)	0,999
<i>NOTCH2</i>	1 (3,0)	2 (5,3)	0 (0,0)	0,794	3 (4,3)	0 (0,0)	0,999	3 (3,4)	0 (0,0)	0,999	3 (4,1)	0 (0,0)	0,999	2 (2,5)	0 (0,0)	0,999
<i>ROS1</i>	1 (3,0)	2 (5,3)	0 (0,0)	0,794	3 (4,3)	0 (0,0)	0,999	3 (3,4)	0 (0,0)	0,999	3 (4,1)	0 (0,0)	0,999	3 (3,8)	0 (0,0)	0,999
<i>STAT5B</i>	0 (0,0)	1 (2,6)	2 (10,0)	0,165	1 (1,4)	2 (9,1)	0,144	3 (3,4)	0 (0,0)	0,999	1 (1,4)	2 (11,8)	0,090	2 (2,5)	0 (0,0)	0,999

Valor p do Teste Exato de Fisher ou Teste  $\chi^2$  (\*)

A presença de linfonodos comprometidos estava associada com mutações no gene *PIK3CA*, onde 31,8% de tumores com N1/2 estavam mutados para este gene quando comparado com 10,1% dos tumores N0 ( $p=0,036$ ). Não encontramos associação entre a presença de mutação nos genes mais frequentemente alterados com sexo e cor de pele (**Tabela 3**), invasão angiolinfática, invasão perineural e metástase à distância e estadiamento tumoral (**Tabela 4**).

As principais diferenças encontradas nas frequências dos genes estão relacionadas à localização do tumor e são descritas a seguir.

### 2.3.1 Diferença entre cólon proximal e distal

Dentre os 91 tumores analisados, 71 (78%) estavam localizados no cólon distal e 20 (22%) estavam localizados no cólon proximal. Foram encontradas diferenças significantes entre tumores localizados no cólon proximal e distal (**Tabela 5**).

**Tabela 5.** Associações entre a presença de mutações e localização do tumor

	Localização do Tumor		Valor <i>p</i>
	Proximal	Distal	
	n (%)	n (%)	
<i>APC</i>	10 (50,0)	55 (77,5)	<b>0,025*</b>
<i>TP53</i>	4 (20,0)	48 (67,6)	<b>&lt;0,001*</b>
<i>KRAS</i>	3 (15,0)	45 (63,4)	<b>&lt;0,001*</b>
<i>PIK3CA</i>	4 (20,0)	10 (14,1)	0,499
<i>FBXW7</i>	2 (10,0)	8 (11,3)	0,999
<i>NF1</i>	3 (15,0)	5 (7,0)	0,367
<i>TGFBR2</i>	6 (30,0)	2 (2,8)	<b>0,001</b>
<i>ATM</i>	4 (20,0)	2 (2,8)	<b>0,020</b>
<i>BRAF</i>	8 (40,0)	0 (0,0)	<b>&lt;0,001</b>
<i>BLM</i>	5 (25,0)	2 (2,8)	<b>0,005</b>
<i>MSH6</i>	2 (10,0)	2 (2,8)	0,209
<i>NOTCH1</i>	1 (5,0)	2 (2,8)	0,530
<i>NRAS</i>	0 (0,0)	4 (5,6)	0,572
<i>PTEN</i>	3 (15,0)	1 (1,4)	<b>0,032</b>
<i>CHEK2</i>	1 (5,0)	2 (2,8)	0,530
<i>ERBB3</i>	0 (0,0)	3 (4,2)	0,999
<i>ERBB4</i>	1 (5,0)	2 (2,8)	0,530
<i>NOTCH2</i>	2 (10,0)	1 (1,4)	0,120
<i>ROS1</i>	2 (10,0)	1 (1,4)	0,120
<i>STAT5B</i>	2 (10,0)	1 (1,4)	0,120

Valor *p* do Teste Exato de Fisher ou Teste  $\chi^2$  (\*)



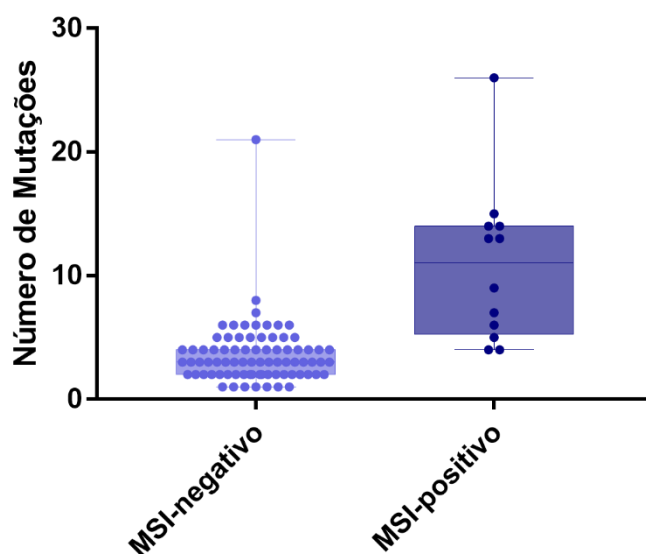
Quando comparados com tumores localizados no cólon distal, tumores localizados no cólon proximal tiveram maior frequência de mutações nos genes *ATM* (20,0% vs 2,8%,  $p=0,02$ ), *BLM* (25,0% vs 4,2%,  $p=0,012$ ), *BRAF* (40,0% vs 0,0%,  $p<0,001$ ), *NOTCH2* (15,0% vs 1,4%,  $p=0,032$ ), *NOTCH3* (15,0% vs 1,4%,  $p=0,032$ ), *PTEN* (15,0% vs 1,4%,  $p=0,032$ ), *RANBP2* (15,0% vs 1,4%,  $p=0,032$ ) e *TGFBR2* (30% vs 2,8%,  $p<0,001$ ). Além disso, alguns genes estavam exclusivamente mutados nos tumores localizados no cólon proximal: *BRAF*, *EGFR*, *JAK1*, *MAP2K4*, *MAP3K1*, *NBN*, *ALK* (cada um mutado em 2 tumores) e *BRCA1*, *BRIP1*, *CREBBP*, *FLT3*, *FLT4*, *MLH1*, *MPL* e *NT5C2* (cada um mutado em 1 tumor).

Em contraste, os genes de maiores frequências no cólon distal comparados ao cólon proximal foram os genes *APC* (77,5% vs 50,0%,  $p=0,025$ ), *TP53* (67,6% vs 20,0%,  $p<0,001$ ) e *KRAS* (63,4% vs 15,0%,  $p<0,001$ ). Assim como no cólon proximal, encontramos genes mutados exclusivamente no cólon distal: *NRAS* (mutado em 4 tumores), *ERBB3* (mutado em 3 tumores), *RET*, *AKT1*, *AR*, *ARAF*, *BARD1*, *CDKN2A*, *GATA2*, *IDH2*, *MRE11A*, *NTRK1*, *PALB2* e *SMO* (cada um mutado em 1 tumor).

Além das diferenças entre as frequências de mutações, tumores localizados no cólon proximal tiveram maior frequência de instabilidade de microssatélite (50% vs 2,0% no cólon distal).

#### **2.4 Associação entre perfil mutacional e o status de MSI**

A presença de instabilidade de microssatélite foi avaliada previamente nos tumores do estudo<sup>45</sup>, por meio de marcadores repetitivos quasi-monomórficos. Dentre as 91 amostras incluídas, 75 (83,3%) não apresentavam instabilidade (MSS), 3 tumores (3,3%) apresentaram baixa instabilidade (MSI-L) e 12 (13,3%) apresentavam alta instabilidade (MSI-H). Como esperado, a maioria dos casos com MSI-H estava localizada no cólon proximal (10; 83,3%,  $p<0,001$ ). Além disso, a quantidade de mutação observada nos tumores MSI-H foi maior do que naqueles com baixa instabilidade ou sem instabilidade, com uma média de mutação de 7,1 mutações por amostra tumoral nos casos com MSI-H (MSI-positivo) comparada a 3,1 mutações por amostra tumoral nos casos de tumores MSS/MSI-L (MSI-negativo,  $p < 0,001$ , **Figura 8 e Figura 15**).



**Figura 15.** Comparação do número de mutação entre os tumores com instabilidade de microssatélite e os tumores sem instabilidade.

Somente tumores MSI-positivos apresentaram mutações nos genes *MSH6* e *MLH1*, que codificam para proteínas do sistema de reparo do DNA que, quando ausentes, levam à instabilidade microssatélite. Ambos os genes apresentaram mutações que truncam a proteína, embora esses genes estivessem afetados em apenas 5 amostras tumorais (4 para *MSH6* e 1 para o *MLH1*), o que representa 41,7% das amostras com instabilidade de microssatélite na nossa população. Outros genes também se encontraram exclusivamente mutados nos casos MSI-positivos e apresentaram diferença estatisticamente significativa quando comparados com os tumores MSI-negativos, tais como *TGFBR2* (66,7% dos casos MSI-positivos,  $p < 0,001$ ), *NOTCH2* (25%,  $p = 0,002$ ), *STAT5B* (25%,  $p = 0,002$ ) (**Tabela 6**). Para os demais genes exclusivamente mutados em tumores MSI-positivos, não houve diferença estatisticamente significativa, sendo esses: *SMO*, *RET*, *MER11A*, *AKT1*, *PALB2*, *FLT3*, *CREBBP*, *MPL*, *JAK1*, *NBN*, *BRCA1*, *EGFR*, *MAP3K1*. Além disso, tumores MSI-positivos estavam mutados com maior frequência para os genes *BLM* (50% vs 1,3%,  $p < 0,001$ ), *BRAF* (41,7% vs 3,8%,  $p < 0,001$ ), *CHEK2* (16,7% vs 1,3%,  $p = 0,046$ ), *NOTCH1* (16,7% vs 1,3%,  $p = 0,046$ ) (**Tabela 6**). Por outro lado, os tumores MSI-negativos apresentaram maior frequência de mutações nos genes *APC* (75,6% vs 41,7%,  $p = 0,034$ ) e *TP53* (65,4% vs 8,3,  $p < 0,001$ ).

**Tabela 6.** Frequência de alteração nos diferentes genes em relação ao status MSI

	<b>Instabilidade de Microssatélite</b>		Valor <i>p</i>
	<b>MSI-negativo</b>	<b>MSI-positivo</b>	
	n (%)	n (%)	
<i>APC</i>	59 (75,6)	5 (41,7)	<b>0,034</b>
<i>TP53</i>	51 (65,4)	1 (8,3)	<b>&lt;0,001*</b>
<i>KRAS</i>	45 (57,7)	3 (25,0)	0,059*
<i>PIK3CA</i>	12 (15,4)	2 (16,7)	0,999
<i>FBXW7</i>	8 (10,3)	2 (16,7)	0,617
<i>NF1</i>	5 (6,4)	3 (25,0)	0,070
<i>TGFBR2</i>	0 (0,0)	8 (66,7)	<b>&lt;0,001</b>
<i>ATM</i>	3 (3,8)	3 (25,0)	<b>0,029</b>
<i>BRAF</i>	3 (3,8)	5 (41,7)	<b>0,001</b>
<i>BLM</i>	1 (1,3)	6 (50,0)	<b>&lt;0,001</b>
<i>MSH6</i>	0 (0,0)	4 (33,3)	<b>&lt;0,001</b>
<i>NOTCH1</i>	1 (1,3)	2 (16,7)	<b>0,046</b>
<i>NRAS</i>	4 (5,1)	0 (0,0)	0,999
<i>PTEN</i>	2 (2,6)	2 (16,7)	0,084
<i>CHEK2</i>	1 (1,3)	2 (16,7)	0,046
<i>ERBB3</i>	3 (3,8)	0 (0,0)	0,999
<i>ERBB4</i>	3 (3,8)	0 (0,0)	0,999
<i>NOTCH2</i>	0 (0,0)	3 (25,0)	<b>0,002</b>
<i>ROS1</i>	1 (1,3)	2 (16,7)	<b>0,046</b>
<i>STAT5B</i>	0 (0,0)	3 (25,0)	<b>0,002</b>

Valor *p* do Teste Exato de Fisher ou Teste  $\chi^2$  (\*)

## DISCUSSÃO

No presente trabalho, foram investigadas mutações em 150 oncogenes e genes supressores tumorais na população brasileira. Até onde podemos verificar, este é o maior perfil mutacional do câncer colorretal realizado em nossa população. De maneira geral, os resultados aqui apresentados mostraram que a o perfil mutacional da população brasileira é similar com o observado em outras populações (**Tabela 7**), com alterações em genes canonicamente mutados no CCR, tais como *APC*, *TP53*, *KRAS*, *PIK3CA* e *FBXW7*. Também foram identificadas diferenças entre os tumores localizados no cólon proximal e distal.

Mutações no gene *APC* são um evento comum no CCR, ocorrendo inicialmente na progressão da doença<sup>10,46,47</sup>. Sua perda acarreta desregulação na via Wnt, com consequências na diferenciação e migração celular<sup>48</sup>. Estudos recentes avaliando evolução clonal e heterogeneidade em tumores colorretais têm confirmado mutações em *APC* como um dos eventos iniciais no processo de diversificação clonal do CCR<sup>49,50</sup>. No presente estudo, mutações no gene *APC* foram as mais frequentes, majoritariamente aquelas que ocasionam interrupção prematura da proteína, similar à diversos estudos em diferentes populações (**Tabela 7**)<sup>23,40,51,52</sup>. Além disso, como reportado previamente<sup>22,52</sup>, mutações no gene foram mais frequentes no cólon distal e em tumores MSI-negativos.

Mutações no gene *TP53* são frequentes em diversos tipos tumorais e é um dos mais alterados no câncer colorretal. Mutações e perda de *TP53* tem sido constatados como um dos eventos que ocorrem durante a transformação de lesões colorretais benignas em tumores avançados<sup>10,47,53</sup>. Mutações no gene *TP53* são observadas em 55-60% dos tumores colorretais (**Tabela 7**)<sup>23,52,54</sup>. Quando comparados com outras populações<sup>23,55</sup>, os dados do presente estudo apresentou uma frequência similar, com o gene *TP53* apresentando maior frequência de mutação em tumores sem hipermutação (MSI-negativos) e no cólon distal<sup>23,40,56,57</sup>.

A via MAPK-ERK está frequentemente alterada no CCR e outros tumores<sup>58</sup>. A via é constitutivamente ativada por alterações em diversos genes drivers, incluindo *KRAS*, *BRAF* e *NRAS*<sup>23</sup>. Foram encontradas mutações em 68% dos casos com mutações em pelo menos um dos genes da via MAPK-ERK analisados, que incluem os genes *KRAS*, *NF1*, *BRAF*, *NRAS* e *ARAF*, o que é semelhante ao observado nos dados do TCGA (59,3%)<sup>23,40</sup>.

**Tabela 7.** Comparação da frequência de mutações em diferentes bancos de dados públicos (disponíveis no cBioPortal) e outros estudos.

Gene	Presente estudo (n=91)		TCGA (n=224)		DFCI (n=619)		Genentech (n=72)		ICGC: COCA-CN (n=321)		Salem et al. (n=2413)	
	N	(%)	N	(%)	N	(%)	N	(%)	N	(%)	N	(%)
APC	65	(71,4)	168	(75,0)	361	(58,3)	29	(40,3)	173	(53,9)	1513	(62,7)
TP53	51	(56,0)	121	(54,0)	320	(51,7)	40	(55,6)	158	(49,2)	1641	(68,0)
KRAS	48	(52,7)	94	(42,0)	173	(27,9)	37	(51,4)	118	(36,8)	1156	(47,9)
PIK3CA	14	(15,4)	45	(20,1)	132	(21,3)	22	(30,6)	60	(18,7)	326	(13,5)
FBXW7	10	(11,0)	37	(16,5)	86	(13,9)	15	(20,8)	59	(18,4)	-	-
BRAF	8	(8,8)	21	(9,4)	127	(20,5)	6	(8,3)	51	(15,9)	168	(7,0)
NF1	8	(8,8)	8	(3,6)	45	(7,3)	4	(5,6)	66	(20,6)	-	-
TGFBR2	8	(8,8)	23	(10,3)	29	(4,7)	2	(2,8)	25	(7,8)	-	-
BLM	7	(7,7)	10	(4,5)	12	(1,9)	0	(0,0)	25	(7,8)	-	-
ATM	6	(6,6)	25	(11,2)	64	(10,3)	13	(18,1)	53	(16,5)	34	(1,4)
MSH6	4	(4,4)	15	(6,7)	27	(4,4)	3	(4,2)	25	(7,8)	-	-
NRAS	4	(4,4)	20	(8,9)	27	(4,4)	2	(2,8)	17	(5,3)	-	-
PTEN	4	(4,4)	8	(3,6)	51	(8,2)	4	(5,6)	30	(9,3)	37	(1,5)
ERBB3	3	(3,3)	14	(6,3)	36	(5,8)	6	(8,3)	31	(9,7)	-	-
ERBB4	3	(3,3)	19	(8,5)	37	(6,0)	8	(11,1)	46	(14,3)	-	-
NOTCH1	3	(3,3)	0	(0,0)	65	(10,5)	6	(8,3)	38	(11,8)	-	-
CHEK2	3	(3,3)	1	(0,4)	36	(5,8)	5	(6,9)	22	(6,9)	-	-
NOTCH2	3	(3,3)	12	(5,4)	25	(4,0)	4	(5,6)	63	(19,6)	-	-
ROS1	3	(3,3)	13	(5,8)	31	(5,0)	2	(2,8)	46	(14,3)	-	-
STAT5B	3	(3,3)	3	(1,3)	15	(2,4)	3	(4,2)	12	(3,7)	-	-
BRCA2	2	(2,2)	10	(4,5)	42	(6,8)	5	(6,9)	35	(10,9)	-	-
BRD2	2	(2,2)	3	(1,3)	7	(1,1)	0	(0,0)	16	(5,0)	-	-
EGFR	2	(2,2)	10	(4,5)	28	(4,5)	4	(5,6)	53	(16,5)	-	-
JAK1	2	(2,2)	4	(1,8)	19	(3,1)	2	(2,8)	25	(7,8)	-	-
MAP2K4	2	(2,2)	11	(4,9)	13	(2,1)	2	(2,8)	16	(5,0)	-	-
MAP3K1	2	(2,2)	5	(2,2)	26	(4,2)	1	(1,4)	28	(8,7)	-	-
MTOR	2	(2,2)	17	(7,6)	50	(8,1)	8	(11,1)	42	(13,1)	-	-
NBN	2	(2,2)	2	(0,9)	20	(3,2)	1	(1,4)	26	(8,1)	-	-

Somente genes com frequência acima de 2% no presentes estudo são exibidos; TCGA: Muzny et al.<sup>23</sup>; DFCI: Giannakis et al.<sup>51</sup>; Genentech: Seshagiri et al.<sup>59</sup>; ICGC: COCA-CN: Colorectal cancer from non-Western samples (China)<sup>60</sup>; Salem et al.<sup>61</sup>.

Individualmente, foi observado alta frequência de mutações no gene *KRAS* (52,7%) na nossa população, a qual foi maior que a reportada no TCGA (42,0%) e aquela encontrada em outras populações (**Tabela 7**)<sup>40,51,55,62</sup>. Além disso, estudos anteriores relatando mutações no gene *KRAS* na população brasileira mostrou uma frequência menor (31,9-49,2%)<sup>25,38</sup>. A alta frequência observada no presente trabalho em comparação àquela observada em outros estudos na população brasileira pode estar associada com a maior sensibilidade e cobertura do NGS utilizada. Em concordância com outros estudos, a maioria das mutações encontradas no gene *KRAS* foi encontrada nas posições 12 e 13 da proteína, com predominância de mutações do Gly12Asp<sup>63,64</sup>. Observações epidemiológicas e moleculares demonstraram que mutações no gene *KRAS* são levemente mais frequentes em tumores do cólon proximal<sup>61,65,66</sup>, diferentemente do observado em nossos dados, que apontou predominância de mutações no gene do cólon distal. Semelhantemente, nós também observamos uma maior frequência de mutações no gene *NF1* (**Tabela 7**).

Mutações no gene *BRAF* são relatadas em 8-20,5% dos casos de CCR (**Tabela 7**) e tem sido associadas com tumores MSI-positivos e localizados no cólon proximal em diversos estudos<sup>22,23,40,61</sup>. Aqui foi confirmada essa observação e encontrada frequência similar de mutações no gene *BRAF*, incluindo predominância da mutação ativante V600E<sup>67</sup>. Outras mutações no gene *BRAF* podem ativar a via MAPK-ERK por outros mecanismos<sup>68</sup> como, por exemplo, mutações no sítio D594, os quais apresentam perda completa ou parcial da atividade quinase da proteína<sup>69</sup>, mas pode contribuir com a ativação constitutiva da via por outros mecanismos de ação, tal como deleções do gene *NF1*<sup>68</sup>. Outras mutações não-V600E, como G469V observada no presente estudo, apresentam alta atividade quinase, mas atuam de forma diferente do mecanismo de ação observado por mutações V600E<sup>68</sup>. O gene *NRAS* apresentou baixa frequência comparados com dados do TCGA, mas foi similar com frequências observadas em outros estudos de larga escala (**Tabela 7**). Além disso, o padrão de mutações mutualmente exclusivas entre os genes *KRAS*, *BRAF* e *NRAS* observado tem sido relatado em diversos estudos<sup>23,40,51,70</sup>.

A via PI3K é outra via que está frequentemente alterada no CCR, favorecendo a sobrevivência celular quando ativada<sup>7,23,58</sup>. Dados do TCGA mostraram que genes da via apresentam mutações em aproximadamente 25,9% dos pacientes com CCR<sup>23,40</sup>, similar à

frequência observada no presente estudo (23,1%). O gene *PIK3CA* é um dos genes mais mutados na via<sup>23,51</sup>, apresentando mutações ativantes principalmente no domínio PI3Ka (*helicase*), em adição à outras mutações já conhecidas, como H1047R e N345K/S encontradas no presente estudo<sup>23,71</sup>. O gene *PTEN* é conhecido por reverter as ações das PI3Ks e sua inativação por mutações é observado em baixa frequência no CCR (1,5-11,1%, **Tabela 7**)<sup>23,72</sup>, similarmente ao observado no presente trabalho (4,4%). A associação entre mutações no gene *PTEN* e tumores localizados no cólon proximal já foi observada anteriormente<sup>57</sup>. Embora mutações no gene *MTOR* tenham sido relatadas com variação de 2,8-9,1% em estudos de larga escala (**Tabela 7**), estas mutações tem sido pouco caracterizadas<sup>73</sup>. Mutações no sítio S2215, como a observada no presente estudo, é um dos sítios de *MTOR* mais frequentemente alterado em diversos tipos tumorais e tem sido demonstrado por agir aumentando a atividade da proteína mTOR<sup>73</sup>. Outra mutação observada (T2444M) não é bem caracterizada, mas afeta a região C-terminal da proteína, que está relatada com ativação de mTOR<sup>73</sup>. Baixa frequência de mutações em *AKT1* tem sido observada em estudos de larga escala (**Tabela 7**) e a mutação E17K tem sido relacionada com a ativação constitutiva da proteína<sup>74</sup>.

A via Notch é importante na manutenção da proliferação das criptas intestinais<sup>75</sup> e mutações genéticas nesta via tem sido associada à progressão tumoral<sup>76</sup>. No presente estudo foram observadas mutações em genes desta via, que inclui *FBXW7*, *NOTCH2*, *NOTCH1* e *CREBBP* em 16,5% das amostras tumorais, o que foi levemente menor quando comparado com a frequência observada no TCGA para os mesmos genes (20,1%)<sup>23,40</sup>. O gene *FBXW7* é um dos reguladores das proteínas Notch<sup>77,78</sup>, sua deleção relacionada à ativação da via<sup>79</sup> e mutações do tipo *missense* no gene estão relacionadas com maior atividade tumorigênica no intestino de camundongos<sup>80</sup>. A frequência de mutações em *FBXW7* em nossa população foi semelhante àquelas observada em outras populações (**Tabela 7**)<sup>81</sup>. Foram observadas mutações afetando *hotspots* do gene *FBXW7*, tais como os sítios 465 e 505, que são importantes sítios de ligação de substratos<sup>82</sup>, além de mutações do tipo *frameshift* que levam à interrupção prematura da proteína e também já foram observadas afetando o gene<sup>23,40,51</sup>. A frequência de mutações afetando os genes *NOTCH1* e *NOTCH2* são variáveis (**Tabela 7**) e mutações ativantes do tipo missense em ambos *NOTCH1* e *NOTCH2* já foram relatadas em tumores hematológicos<sup>83,84</sup>. O padrão de mutações mutualmente exclusivas nos genes *NOTCH2* e *FBXW7*, que foi observado no presente estudo, não tem sido observado em estudos de larga escala<sup>23,40,51</sup>.

No presente estudo, 15,4% das amostras tumorais apresentaram mutações do tipo *missense* na família dos receptores de tirosina quinase (RTKs, incluindo os genes *ROS1*, *ERBB4*, *ERBB3*, *EGFR*, *RET*, *NTRK1*, *FLT4*, *FLT3* e *ALK*), a qual foi inferior à observada nos dados do TCGA para os mesmos genes (22,3%)<sup>23,40</sup>. Embora eventos de ativação em algumas proteínas de RTKs podem ser observados por rearranjos cromossômicos, tais como em *ROS1*, *NTRK1*, *ALK*<sup>85</sup> e *RET*<sup>86</sup>, mutações do tipo *missense* nestes genes são encontradas em diversos estudos (**Tabela 7**), mas o potencial oncogênico de tais mutações não é completamente compreendido no CCR. Mutações *missense* no gene *RET* promovem progressão tumoral em tumores de tireoide<sup>87</sup> e já foram demonstradas reduzindo a apoptose em linhagens celulares de CCR<sup>87</sup>. Interessantemente, mutações *missense* nos genes *ALK*, *ROS2* e *NTRK1* foram observadas em tumores sólidos resistentes à inibidores de receptores de tirosina quinase<sup>88-90</sup>. Foram observadas mutações ativantes conhecidas no domínio de quinase de *EGFR*, tais como T725M<sup>40,91</sup> e P848L<sup>92</sup>, além de mutações no gene *FLT3*<sup>40,91</sup>. Outras mutações afetando o domínio quinase dos genes *ROS1*, *ERBB4* e *ALK* também foram observadas, mas o papel de tais mutações na oncogênese é desconhecido<sup>92</sup>. Mutações afetando regiões fora do domínio quinase das RTKs também podem desempenhar um papel na ativação dessas proteínas, tais como mutações no gene *ERBB3*, as quais promovem crescimento de células epiteliais do cólon<sup>93</sup>, e mutações no gene *ERBB4*, as quais promovem aumento da atividade quinase e habilidade de transformação em células de melanoma<sup>94</sup>.

A presença de MSI foi observada em 12% das amostras tumorais, o que está de acordo com o relatado em diversos estudos anteriores<sup>20,22,45</sup>. Adicionalmente, MSI estava associado com tumores no cólon proximal e que apresentavam maior taxa de mutações, o que também é consistente com estudos prévios<sup>22,23,59,95</sup>. Como reportado previamente, também foi observada associação entre mutações nos genes *ATM*<sup>23,40</sup>, *BLM*<sup>23,40</sup> e *TGFRB2*<sup>22,23,40</sup> em tumores com MSI e localizados no cólon proximal.

Diversas alterações moleculares encontradas no presente estudo são potenciais biomarcadores para prognóstico e resposta ao tratamento do CCR<sup>44,96</sup>. A presença de MSI já foi demonstrada por ser fator de melhor prognóstico em pacientes com tumores mais localizados em diversos estudos<sup>95,97,98</sup>, embora o mesmo não seja verdade para pacientes com doença metastática<sup>99</sup>. O uso de quimioterapia adjuvante para pacientes com tumores em estádios II e III e presença de MSI não tem demonstrado melhoras na sobrevida<sup>95,100</sup>. Além disso, mutações envolvendo a família Ras de proteínas são o principal critério de elegibilidade



para terapias direcionadas ao EGFR em pacientes com câncer colorretal metastático<sup>101</sup> e tumores com mutações no gene *BRAF*, *NRAS*, *PIK3CA* e *PTEN* tem apresentado menor sobrevida com este tipo de terapia<sup>102-105</sup>. Ainda, mutações V600E no gene *BRAF* tem demonstrado benefícios em terapias direcionadas combinadas<sup>44,96</sup>.

Apesar dos principais resultados encontrados no presente estudo, uma das limitações observadas é a ausência de representação de genes importantes e que apresentam alta frequência de mutação no CCR, como *SMAD4* e *CTNNB1*<sup>7,23</sup>. Além disso, o sequenciamento do genoma completo ou de todo exoma tumoral poderia revelar novos genes com potencial importância para o processo tumorigênico na população brasileira.

## CONCLUSÕES

O perfil de mutações somáticas dos tumores colorretais da população brasileira é semelhante ao encontrado em diversos estudos, com mutações frequentes nos genes *APC*, *TP53*, *KRAS*, *PIK3CA* e *FBXW7*.

O perfil mutacional dos tumores colorretais variou significativamente em relação à localização tumor, com tumores do cólon proximal apresentando maior frequência de MSI e mutações nos genes *BRAF*, *TGFBR2*, *ATM*, *BLM* e *PTEN* e tumores do cólon distal apresentando maior frequência de mutações nos genes *APC*, *TP53* e *KRAS*.

Tumores com instabilidade de microssatélite tiveram perfil mutacional distinto daqueles sem instabilidade, além de apresentaram maior quantidade de mutações.

## PARTE 2: CARACTERIZAÇÃO DO PERFIL MUTACIONAL DAS LESÕES PRECURSORAS DO CÂNCER COLORRETAL

### MATERIAL E MÉTODOS

#### 1. População com lesões precursoras do câncer colorretal (CCR)

Para as análises de lesões precursoras, foram selecionados pacientes que realizaram colonoscopia diagnóstica no Hospital de Câncer de Barretos, entre maio de 2014 e novembro de 2016. As lesões, após ser realizada ressecção ou biópsia, foram fixadas em formol e incluídas em blocos de parafina e posteriormente armazenadas no Departamento de Patologia do hospital, sendo avaliadas por dois patologistas. Tais lesões colorretais não-câncer incluíram adenomas, lesões serrilhadas sésseis (LSS) e pólipos hiperplásicos (PH).

Características demográficas dos pacientes e características endoscópicas e histológicas das lesões colorretais não câncer, as quais incluem tamanho, localização, morfologia de acordo com a classificação de Paris<sup>106</sup>, foram coletadas e armazenadas para posterior análise. Lesões localizadas no ceco, cólon ascendente, ângulo hepático, cólon transversal ou ângulo esplênico foram agrupados como cólon proximal e as lesões localizadas no cólon descendente, cólon sigmoide e reto foram agrupadas como cólon distal. Adenomas foram classificados como avançado na presença de displasia de alto grau e/ou com componente viloso  $\geq 25\%$ , e/ou com tamanho  $\geq 10\text{mm}$ .

##### 1.1 Critérios de Inclusão

Pacientes adultos  $> 50$  anos, do sexo masculino e feminino, sintomáticos ou assintomáticos, que foram agendados para colonoscopia diagnóstica no departamento de endoscopia do HCB, e concordaram com a coleta de sangue para estudos moleculares.

##### 1.2 Critérios de Exclusão:

Pacientes com CCR familiar ou síndromes hereditárias associadas ao CCR.

#### 2. Extração de DNA

Para extração de DNA, os blocos fixados e incluídos em parafina das lesões foram avaliados por um patologista e a área da lesão foi demarcada. O DNA de lesões precursoras

foi obtido no Centro de Pesquisa em Oncologia Molecular (CPOM) a partir de cortes de 10 micras de tecido parafinizado.

O tecido foi desparafinizado à temperatura de 80°C e posterior lavagem em solução de xilol e reidratação em álcool (série de 100%, 70% e 50%) e água ultrapura. A área contendo a lesão do tecido desparafinizado foi macrodissecada utilizando agulha estéril de 18G x 1<sup>1/2</sup> e armazenada em microtúbulo para extração de DNA com o *QIAmp DNA Micro kit* (Qiagen), seguindo as especificações do fabricante. Este kit se baseia no isolamento do DNA a partir do rompimento da membrana plasmática celular e digestão proteica e posterior purificação a partir de colunas com membranas baseada em sílica.

### 3. Análise do perfil mutacional

#### 3.1. Análise das variantes por NGS utilizando na plataforma Ion Torrent

Diferentemente das amostras de CCR, as amostras de lesões colorretais precursoras do presente estudo são provenientes de tecido fixado em parafina, o qual diminui a qualidade do DNA extraído. Além disso, tais lesões são pequenas (menores que 10mm em média).

O sequenciamento das amostras provenientes de tecido parafinado não é possível pela mesma plataforma adotada para as amostras provenientes de tecido congelado. Assim, a metodologia escolhida para identificação das mutações nas amostras de lesões colorretais foi o sequenciamento através da plataforma Ion Torrent™ (ThermoFisher Scientific). Para este sequenciamento, foi selecionado o painel Ion AmpliSeq Cancer Hotspot Panel v2, o qual utiliza primers para amplificação de 207 regiões cobrindo aproximadamente 2.800 mutações anotadas no COSMIC, distribuídas ao longo de 50 oncogenes e genes supressores tumorais, nomeadamente: *ABL1, AKT1, ALK, APC, ATM, BRAF, CDH1, CDKN2A, CSF1R, CTNNB1, EGFR, ERBB2, ERBB4, EZH2, FBXW7, FGFR1, FGFR2, FGFR3, FLT3, GNA11, GNAQ, GNAS, HNF1A, HRAS, IDH1, IDH2, JAK2, JAK3, KDR, KIT, KRAS, MET, MLH1, MPL, NOTCH1, NPM1, NRAS, PDGFRA, PIK3CA, PTEN, PTPN11, RB1, RET, SMAD4, SMARCB1, SMO, SRC, STK11, TP53* e *VHL*. Esse painel permite a utilização de quantidades pequenas de DNA (10 ng) e, por se tratar de *hotspots* ao invés de genes inteiros, possibilita a utilização de amostras com um grau de degradação maior, como nossas amostras de lesões colorretais.

As amostras de lesões colorretais foram sequenciadas a partir de 10 ng de DNA totais. As bibliotecas foram construídas utilizando o kit Ion Ampliseq™ Kit for Chef DL8, o qual é realizado de maneira automatizada no sistema Ion Chef™. Após a preparação da biblioteca de sequenciamento, a preparação do chip de foi realizado também de forma automatizada no sistema Ion Chef™. O sequenciamento foi realizado na plataforma PGM (Personal Genome Machine), seguindo recomendações do fabricante (ThermoFisher Scientific). O alinhamento das reads sequenciadas, chamada, e anotação das variantes foram realizadas através da plataforma online Ion Reporter (disponível em: <https://thermofisher.com/myso/loginDisplay>). Foram mantidas variantes que apresentavam uma cobertura acima de 200x e uma frequência alélica acima de 10%.

### **3.2. Critério de patogenicidade para as variantes encontradas nas lesões colorretais**

Assim como realizado para as variantes encontradas no CCR, a identificação de potenciais alterações *drivers* foi realizada utilizando o CGI<sup>44</sup>, embora de maneira mais restritiva. Para as definições de mutações *drivers* das lesões precursoras, as variantes foram classificadas como *drivers* quando identificadas como *drivers* especificamente no CCR, diferentemente da classificação utilizada para as amostras de tumores, nos quais foram mantidas variantes conhecidas como *drivers* em tumores em geral. Também foram mantidas variantes não conhecidas como *drivers* mas que ocorreram em genes com efeito patogênico em CCR e foram preditas como *drivers* pelo algoritmo *OncodriveMUT*.

## RESULTADOS

### 1. Características dos pacientes com lesões colorretais precursoras do CCR

Foram selecionadas 90 lesões colorretais provenientes de 87 pacientes. As características clínico-patológicas desses pacientes estão descritas na **Tabela 8**. A média de idade dos pacientes foi de 62,9 anos (desvio padrão: 27,5), 50,6% sendo do gênero feminino. A maior parte das lesões estava localizada no cólon distal (53,3%) e era menor que 10mm (77,8%). Dentre as 90 lesões, 67 (74,4%) eram adenomas e 23 (25,6%) pólipos serrilhados. Dentre os adenomas, 51 (56,7%) apresentavam histologia tubular e 16 (17,8%) apresentavam componente túbulo-viloso, sendo 28 (41,8%) classificados como adenomas avançados e 39 (58,2%) classificados como adenomas precoces. Dentre os pólipos serrilhados, 30,4% (7) eram lesões serrilhadas sésseis e 16 (69,6%) pólipos hiperplásicos.

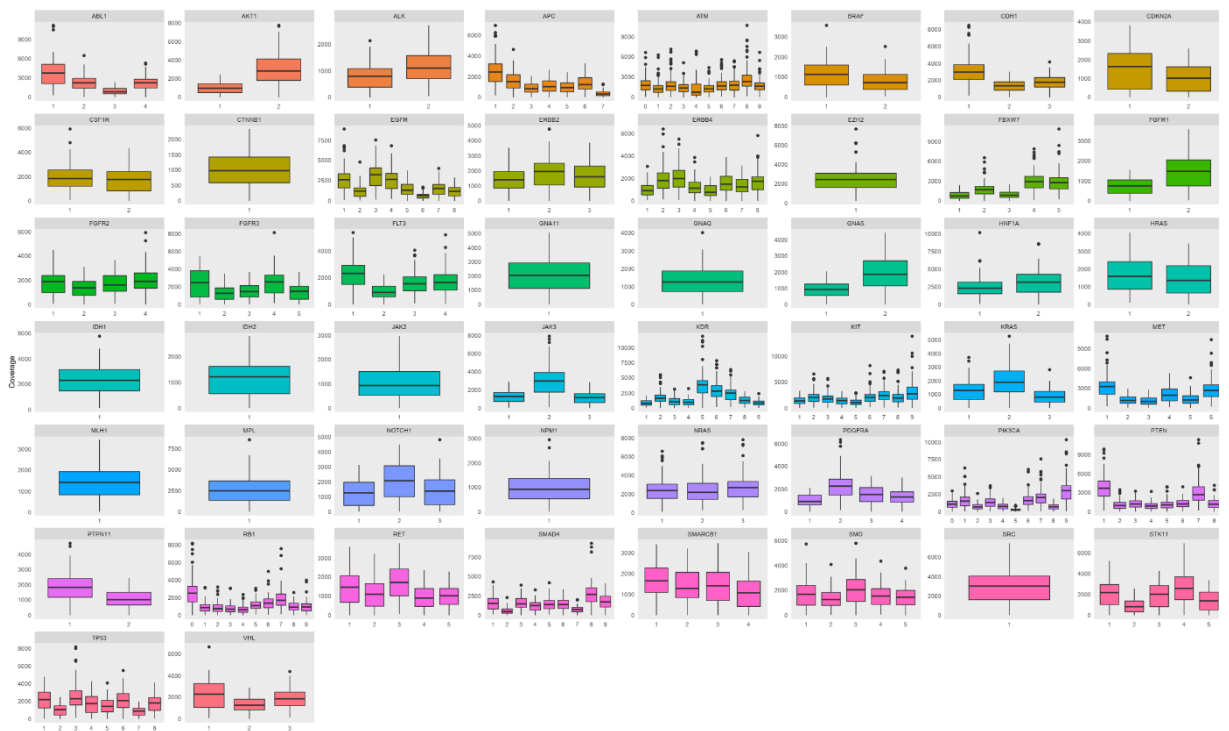
**Tabela 8** – Características sócio-demográficas dos pacientes e morfológicas e histopatológicas das lesões colorretais não-câncer.

<b>Características dos pacientes (n=87)</b>	<b>Frequência</b>	<b>(%)</b>
Idade ao Diagnóstico	Média	(desvio padrão)
	62,92	8,1
<b>Gênero (n = 87)</b>		
Feminino	44	50,6
Masculino	43	49,4
<b>Cor da pele auto-declarada (n = 87)</b>		
Branca	53	60,9
Morena	19	21,8
Parda	9	10,3
Negra	2	2,3
Amarela	4	4,6
<b>Características das lesões (n=90)</b>	<b>Frequência</b>	<b>(%)</b>
<b>Localização da Lesão (n = 90)</b>		
Cólon Proximal	40	44,4
Cólo ascendente	15	16,7
Cólon transverso	17	18,9
Ângulo hepático	8	8,9
Cólon Distal	50	55,6
Cólon descendente	15	16,7
Cólon sigmóide	16	17,8
Transição retossigmóide	2	2,2
Reto	17	18,9
<b>Tamanho (n = 90)</b>		
<10mm	69	76,7
≥10mm	21	23,3
<b>Tipo de lesão (n = 90)</b>		
Adenomas	67	74,4
Tubular	51	56,7
Túbulo-viloso	13	14,4
Viloso	3	3,3
Pólipos Serrilhados	23	25,6
Pólipo Hiperplásico	16	17,8
MVHP	9	10,0
GCHP	7	7,8
Lesão serrilhada séssil	7	7,8

## 2 Caracterização mutacional das lesões precursoras do câncer colorretal

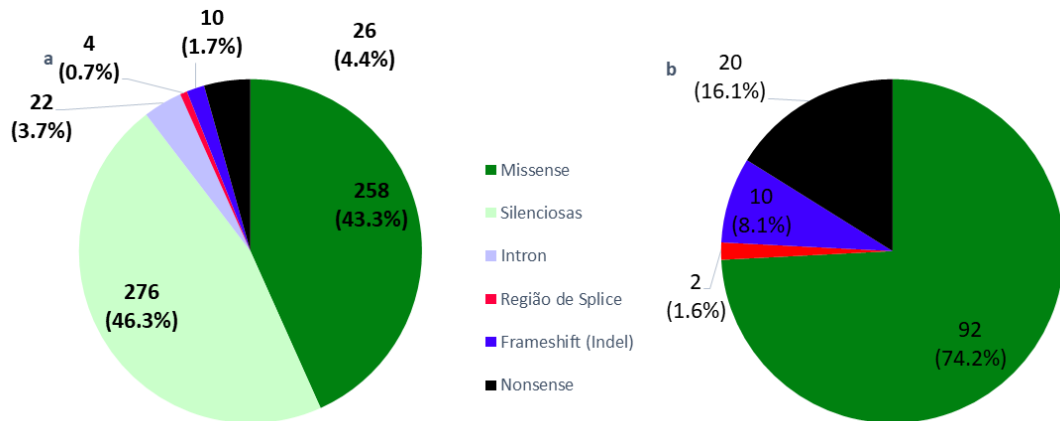
### 2.1 Cobertura e qualidade do sequenciamento das lesões colorretais não câncer

A variação de cobertura dos amplicons nos 50 genes analisados para as lesões não câncer é mostrada na **Figura 16**. A cobertura média de todos amplicons foi de 1631,5x por amostra (desvio padrão 695,9), variando de 293,2x à 3839,6x. Obtivemos uma cobertura média de 1100x, com uma taxa de ontarget (reads alinhadas com o alvo de interesse) de 98%.



**Figura 16.** Cobertura dos 207 amplicons sequenciados pela plataforma Ion Torrent, distribuídos pelos 50 genes analisados. Foi obtida uma cobertura média de 1631,5x.

Identificamos um total de 596 variantes nas 90 amostras de lesões analisadas, sendo 276 (46,3%) variantes consideradas silenciosas, 258 (43,3%) *missense*, 26 (4,4%) *nonsense*, 22 (3,7%) variantes em regiões intrônicas, 10 (1,7%) *frameshift* e 4 (0,7%) em regiões de *splice*. Após aplicar o filtro de patogenicidade, foram removidas 472 variantes não patogênicas de acordo com os critérios descritos previamente, restando 124 mutações em 60 lesões de 58 pacientes. O tipo mais frequente de mutações classificadas como *drivers* foi *missense* (74,2%), seguidas de mutações *nonsense* (16,1%), *frameshift* (8,1%), e região de *splice* (1,6%) (**Figura 17**).

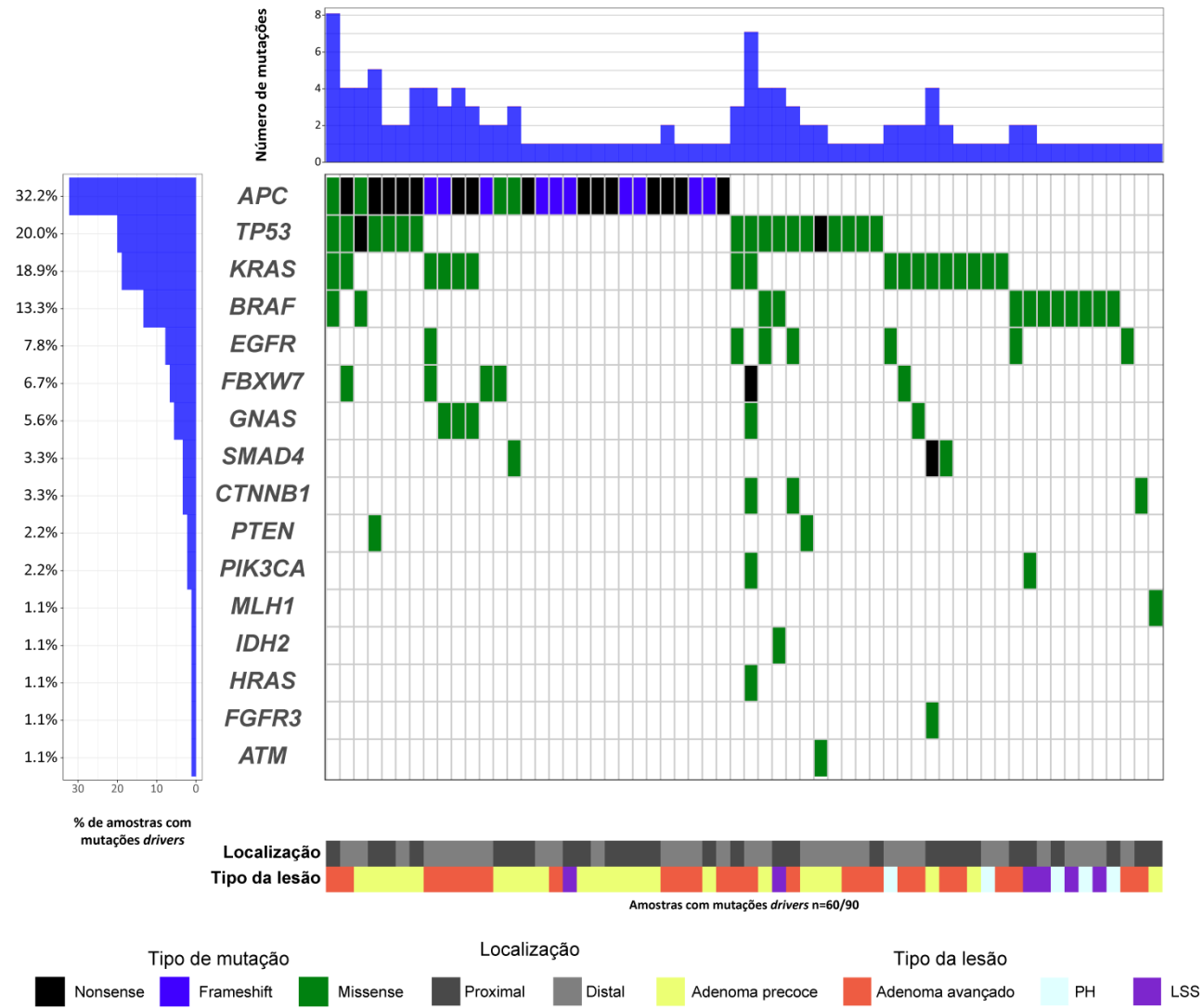


**Figura 17.** Variantes encontradas nas amostras de lesões sequenciadas, incluindo todas as variantes alteradas nas 90 amostras (a) e mutações mantidas após o filtro de patogenicidade (b) em 60 amostras.

## 2.2 Perfil de mutações nas amostras de lesões precursoras do câncer colorretal

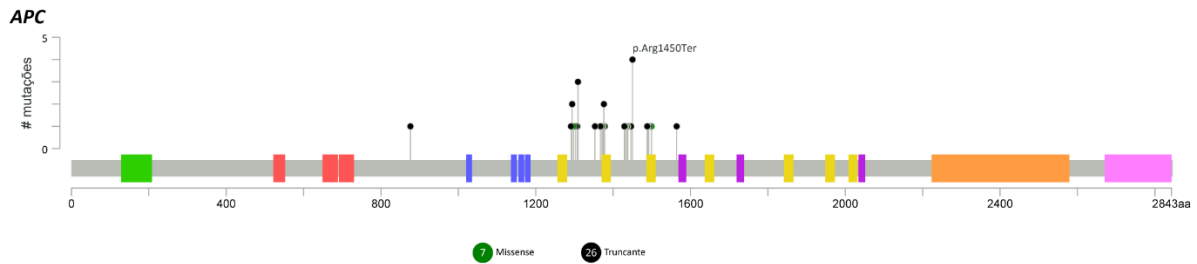
As 60 amostras com mutações *drivers* apresentaram um média de 2,1 mutações, variando de 1-8 mutações, com 32 lesões apresentando apenas um gene mutado, 14 lesões com mutações em 2 genes, 10 lesões com mutações em 3 genes e 4 lesões com mutações em 4 ou mais genes. Entre os 50 genes sequenciados, 16 apresentaram ao menos uma mutação *driver* (**Anexo B**). Os genes com maior frequência de alteração foram *APC* (alterado em 32,2% das amostras), *TP53* (20,0%), *KRAS* (18,9%), *BRAF* (13,3%) e *EGFR* (7,8%). A **Figura 18** mostra as mutações encontrados nos 16 genes ao longo das 60 lesões com mutações *drivers*.





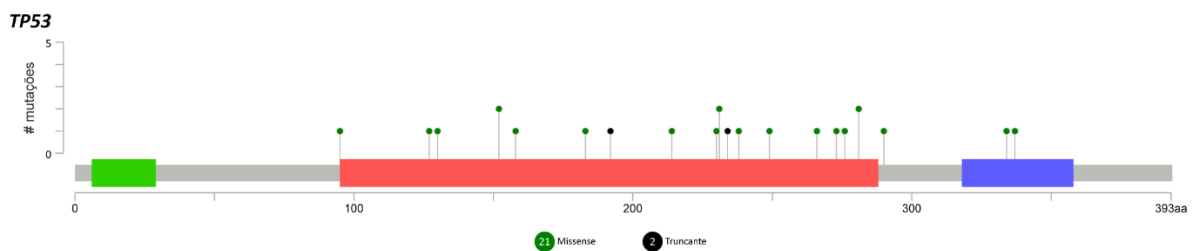
**Figura 18.** Mutações *drivers* encontradas nas lesões colorretais. Os genes são exibidos de acordo com a frequência (esquerda). Amostras sem nenhuma mutação *driver* foram omitidas da figura. PH: pólipo hiperplásico; LSS: lesão serrilhada séssil

O gene *APC* foi o gene com maior frequência de alteração, apresentando 33 mutações em 32,2% das lesões (29/90, **Figura 19**). Mutações truncantes foram as mais frequentes no gene, sendo 48,5% delas do tipo *nonsense* e 30,3% do tipo *frameshift*. Mutações *missense* foram as menos frequentes, totalizando 21,2%. Dentre as mutações encontradas, 78,8% eram conhecidas no COSMIC ou dbSNP (16 *nonsense*, 7 *frameshift* e 3 *missense*) e 21,2% foram mutações preditas como *drivers*.



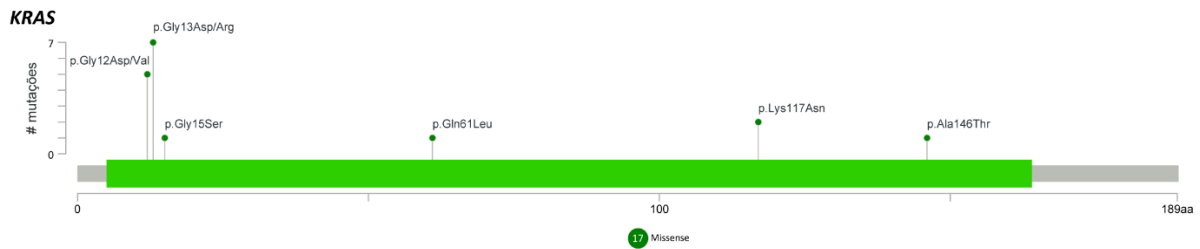
**Figura 19.** Diagrama das mutações no gene *APC*. A maioria das mutações encontradas foram truncantes, incluindo mutações *nonsense* e *frameshift*. Em destaque, podemos observar a mutação p.Arg1450Ter, que apresentou maior frequência (4 mutações).

O gene *TP53* estava alterado em 20,0% das lesões (18/90), apresentando 25 mutações. Cinco lesões apresentaram mais de uma mutação no gene, com o restante apresentando mutações únicas. Dentre as mutações encontradas, todas eram conhecidas no dbSNP e no COSMIC, a maioria de mutações do tipo *missense* (84,0%), seguida de mutações *nonsense* (8,0%) e mutações em regiões de *splice* (8,0%, **Figura 20**).



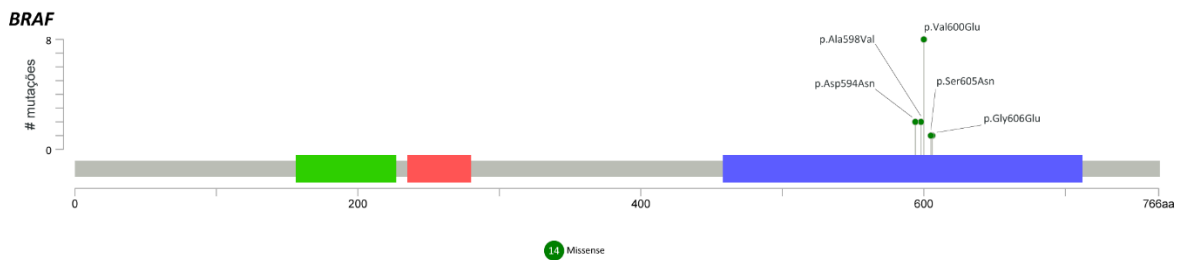
**Figura 20.** Diagrama das mutações no gene *TP53*, incluindo mutações do tipo *missense* (pontos verdes) e mutações *nonsense* (pontos pretos). A maioria das mutações encontradas no gene concentrava-se na região de ligação ao DNA (caixa vermelha).

O gene *KRAS* foi o terceiro gene mais alterado nas lesões, apresentando mutações em 18,9% das amostras (17/90). Todas as mutações encontradas foram de mutações *missense* conhecidas no COSMIC ou dbSNP, afetando os códons 12, 13, 15, 61, 117 e 146 (**Figura 21**). O sítio mais afetado foi a resíduo 13, com 41,2% das mutações que ocorreram no gene.



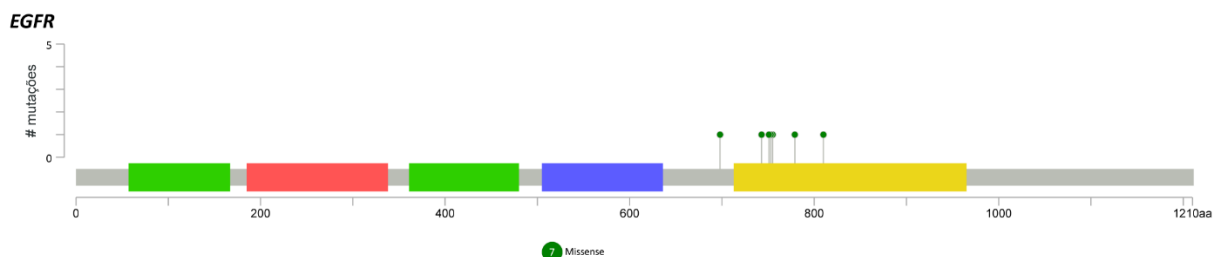
**Figura 21.** Diagrama das mutações no gene *KRAS*, todas mutações do tipo missense (pontos verdes) afetando hotspots do gene no domínio Ras da proteína (caixa verde). O resíduo proteico mais afetado teve 41,2% das mutações na proteína (códon 13).

O gene *BRAF* foi o quarto mais alterado nas lesões, apresentando mutações em 13,3% das amostras (12/90). Todas as mutações foram do tipo *missense* já conhecidas no COSMIC e dbSNP e afetavam o domínio tirosina quinase da proteína. A mutação mais frequente foi a p.Val600Glu (p.V600E), contabilizando 57,1% das mutações no gene. Duas amostras apresentaram mais de uma mutação no gene, uma com as mutações p.Asp594Asn e p.Ala598Val e uma apresentando as mutações p.Val600Glu e p.Ala598Val.



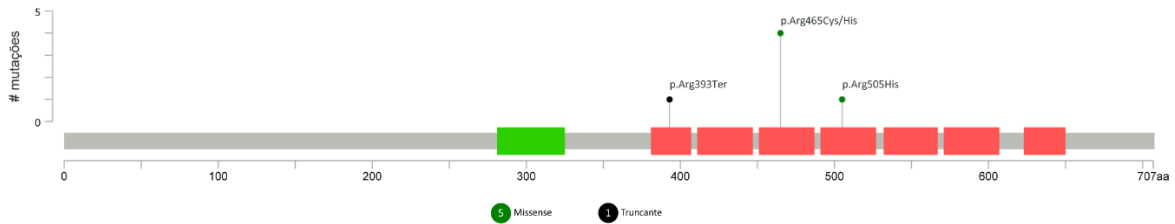
**Figura 22.** Diagrama das mutações no gene *BRAF*, todas mutações do tipo missense (pontos verdes) no domínio tirosina quinase da proteína (caixa azul). A mutação mais frequente foi a p.Val600Glu, com 8 mutações. Duas amostras apresentaram duas mutações no gene: p.Val600Glu e p.Ala598Val em uma amostra e p.Asp594Asn e p. Ala598Val em outra.

O gene *EGFR* estava alterado em 7,8% das amostras (7/90), apresentando mutações *missense* já conhecidas no dbSNP ou COSMIC. As mutações afetando o gene estavam localizadas majoritariamente no domínio tirosina quinase da proteína.



**Figura 23.** Diagrama das mutações no gene *EGFR*, todas mutações do tipo missense (pontos verdes) afetando majoritariamente o domínio no domínio tirosina quinase da proteína (caixa azul). A mutação mais frequente foi a p.Val600Glu, com 8 mutações. Duas amostras apresentaram duas mutações no gene: p.Val600Glu e p.Ala598Val em uma amostra e p.Asp594Asn e p. Ala598Val em outra.

Mutações no gene *FBXW7* ocorreram em 6,7% das amostras (6/90), incluindo mutações do tipo *missense* (83,3%) e *nonsense* (16,7%). Todas as mutações eram conhecidas no dbSNP ou COSMIC e afetavam um dos domínios repetitivos de transducina beta, com o códon 465 com maior número de mutações (66,7% das mutações afetando o gene, **Figura 24**).



**Figura 24.** Diagrama das mutações no gene *FBXW7*, que apresentou mutações do tipo *missense* e *nonsense* nos domínios repetitivos de transducina beta (caixa vermelhas). O sítio 465 foi o códon mais afetado do gene, afetando 66,7% das amostras.

Alterações de genes com menor frequência de alterações (menos que 6% das amostras) podem ser observadas na **Figura 25**. O gene *GNAS* estava alterado em 5,6% das lesões (5/90), e apresentou mutações *missense* conhecidas no dbSNP ou COSMIC. O gene *SMAD4* estava alterado em 3,3% das lesões (3/90), apresentando 4 mutações *missense*, sendo 2 já conhecidas e 2 preditas como *drivers*, e uma mutação *nonsense* já conhecida. O gene *CTNNB1* também estava alterado em 3,3% das amostras (3/90), apresentando mutações do tipo *missense*, sendo duas já conhecidas e uma predita como *driver*. *PIK3CA* estava alterado em 2,2% das amostras (2/90), apresentando duas mutações *missense* já conhecidas. O gene *PTEN* também estava alterado em 2,2% das amostras e apresentou duas mutações *missense* já conhecidas. Os genes com menor frequência de alteração foram os genes *ATM*, *FGFR3*, *HRAS*, *IDH2* e *MLH1*, todos alteradores em 1,1% (1/90) das amostras e apresentando mutações já conhecidas (**Figura 25**).

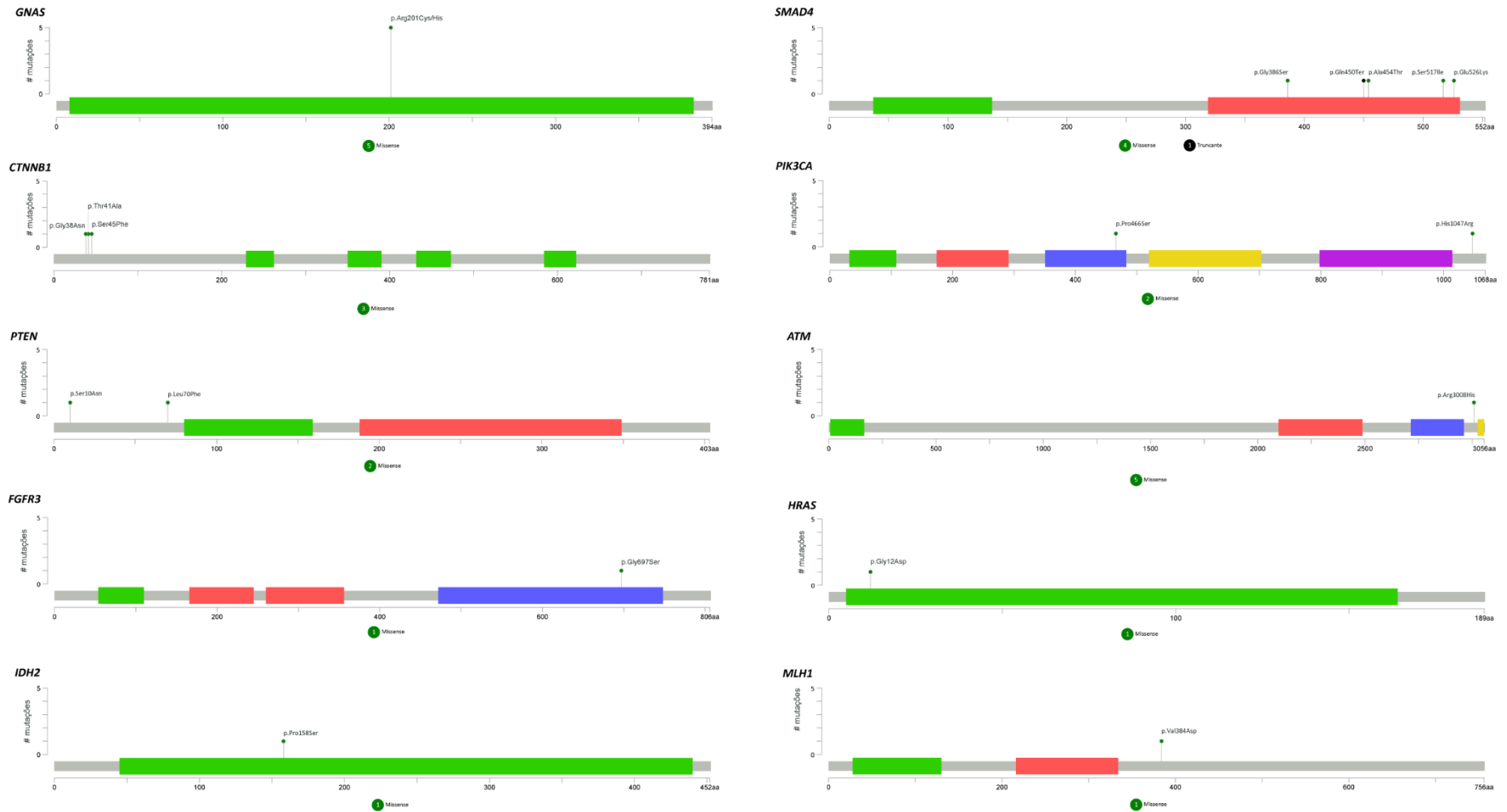


Figura 25. Diagrama das mutações nos genes com frequência abaixo de 6% das amostras de lesões colorretais não câncer.

### 2.3 Perfil de mutações *drivers* de acordo com o tipo histológico nas amostras de lesões precursoras do câncer colorretal

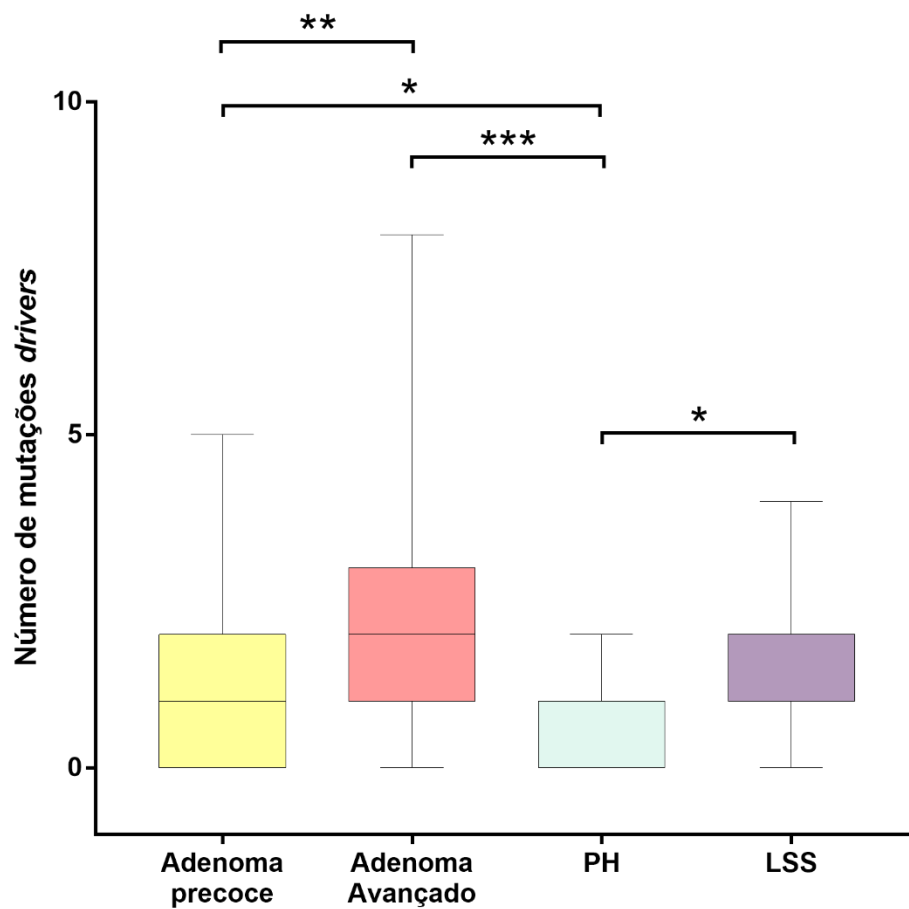
Observando os tipos de lesões, os pólipos hiperplásicos apresentaram a menor proporção de mutações *drivers*, com 68,8% (11/16) sem mutações *drivers*, seguido de adenomas precoces (41,0%, 16/39), LSSs (14,3%, 1/7) e adenomas avançados (7,1%, 2/28,  $p < 0,001$ ). A presença de mutações *drivers* em adenomas foi maior (73,1% com pelo menos uma mutação *driver*) quando comparados com a presença de mutações *drivers* em pólipos serrilhados (PH e LSS que apresentaram 47,8% das lesões com mutações *drivers*,  $p = 0,039$ , **Tabela 9**). O número médio de mutações *drivers* também foi maior entre os adenomas (1,6 vs 0,7,  $p = 0,01$ ).

Adenomas avançados e adenomas precoces diferiram significativamente entre o número de lesões com mutações *drivers* (92,9% vs 59,0%,  $p = 0,002$ , **Tabela 9**), com os adenomas avançados apresentando um número médio de mutações maior do que os precoces (2,2 vs 1,2,  $p = 0,005$ , **Figura 26**). Da mesma forma, entre os pólipos serrilhados, as LSSs foram as lesões que apresentaram maior proporção de mutações *drivers* comparadas com PHs (85,7% vs 31,3%,  $p = 0,027$ , **Tabela 9**), bem como um número médio maior de mutações (1,4 vs 0,4,  $p = 0,014$ , **Figura 26**). A presença de mutações *drivers* nas lesões não câncer não apresentaram diferença em relação ao gênero, cor da pele, localização da lesão, consumo de álcool, tabaco ou ao tipo de alimentação (consumo de fibras, vegetais, carne vermelha, alimentos gordurosos ou peixe).

**Tabela 9.** Presença de mutações drivers nas lesões colorretais

Tipo da lesão	Presença de mutações drivers		Valor <i>p</i>
	Não	Sim	
	n (%)	n (%)	
Adenoma	18 (26,9)	49 (73,1)	0,039*
Pólipo serrilhado	12 (52,2)	11 (47,8)	
<b>Classificação da lesão</b>			
Adenomas			
Adenoma precoce	16 (41,0)	23 (59,0)	0,002*
Adenoma avançado	2 (7,1)	26 (92,9)	
Pólipos serrilhados			
Pólipo hiperplásico	11 (68,8)	5 (31,2)	0,027
Lesão serrilhada séssil	1 (14,3)	6 (85,7)	

Valor *p* do Teste Exato de Fisher ou Teste  $\chi^2$  (\*)



**Figura 26.** Número de mutações drivers diferem entre os tipos de lesões (teste de Kruskal-Wallis  $p < 0,001$ ). Linhas horizontais representam a média de mutações drivers em adenomas precoces, adenomas avançados, pólipos hiperplásicos (PH) e lesões serrilhadas séssis (LSS). Comparações par a par baseadas no teste Mann-Whitney: \* $p < 0,05$ ; \*\* $p < 0,01$ ; \*\*\* $p < 0,001$ .

## 2.4 Diferenças entre os tipos de lesões e frequência de mutações *drivers* em cada gene

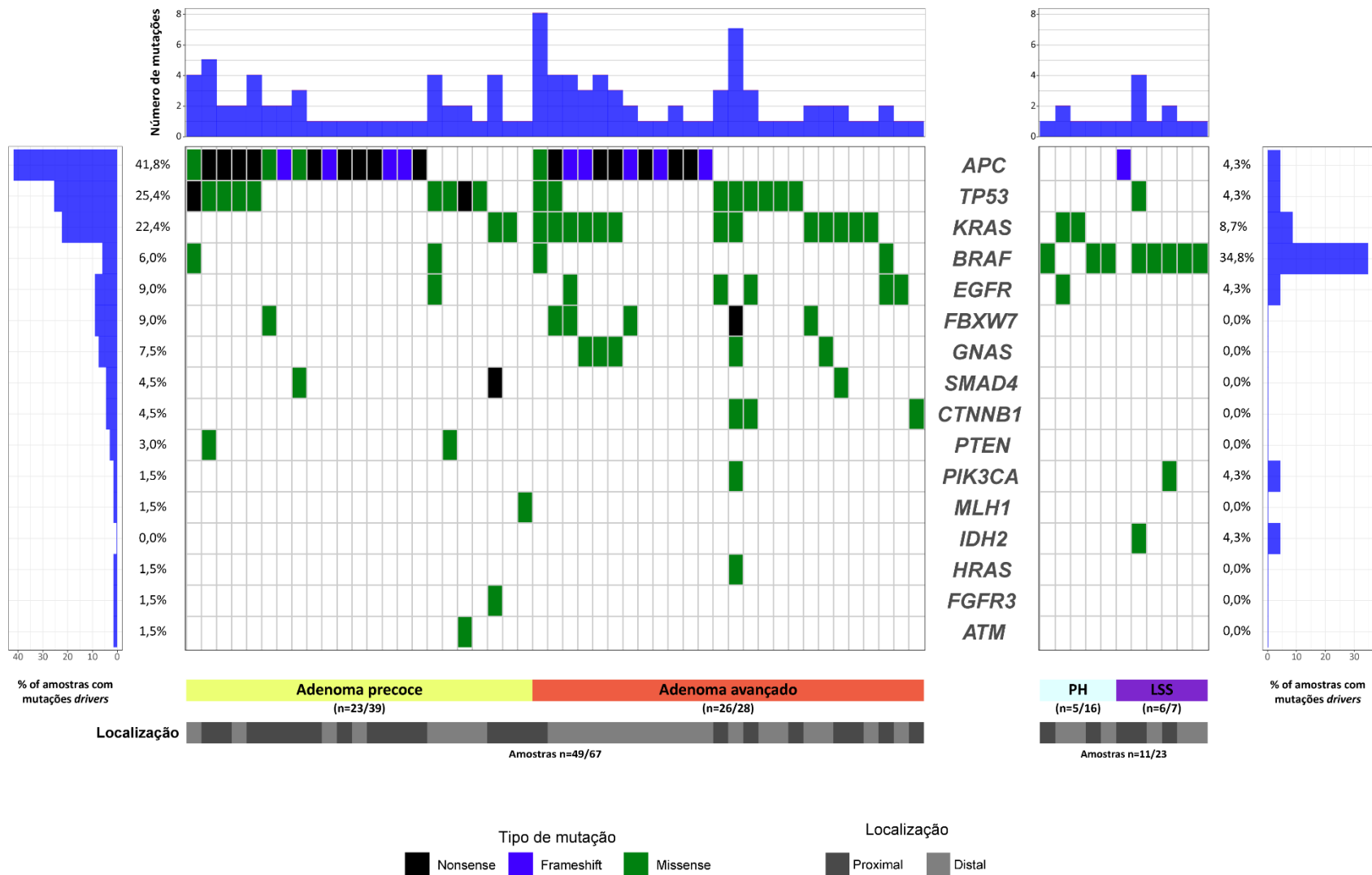
Além de apresentarem diferenças na presença de mutações *drivers*, adenomas e pólipos serrilhados também tiveram perfil mutacional distinto. Mutações no gene *APC* foram mais frequentes em adenomas quando comparados com pólipos serrilhados (41,8% vs 4,3%,  $p = 0,001$ , **Tabela 10, Figura 27**). O gene *TP53* também estava alterado com maior frequência em adenomas (25,4% vs 4,3%,  $p=0,034$ ). Além disso, algumas mutações ocorreram exclusivamente em adenomas, como nos genes *FBXW7* (alterado em 9,0% dos adenomas), *GNAS* (7,5%), *SMAD4* (4,5%), *CTNNB1* (4,5%), *PTEN* (3,0%), *MLH1*, *HRAS* e *ATM* (alterados individualmente em 1,5% dos adenomas), embora a diferença na frequência não tenha sido estatisticamente significativa em comparação com os pólipos serrilhados (**Tabela 10, Figura 27**). De forma contrária, o gene *BRAF* apresentou maior frequência de mutação nos pólipos serrilhados (34,8% vs 6,0%,  $p=0,002$ , **Tabela 10, Figura 27**), e todas mutações p.Vall600Glu no gene foram encontradas exclusivamente em pólipos serrilhados. O único gene exclusivamente mutado em pólipos serrilhados foi o gene *IDH2*, alterado em 4,3% das amostras.

**Tabela 10.** Comparação entre a frequência de mutações em pólipos serrilhados e adenomas

	Tipo de lesão		Valor <i>p</i>
	Pólipo serrilhado (n=23)	Adenoma (n=67)	
	n (%)	n (%)	
<i>APC</i>	1 (4,3)	28 (41,8)	<b>0,001*</b>
<i>TP53</i>	1 (4,3)	17 (25,4)	<b>0,034</b>
<i>KRAS</i>	2 (8,7)	15 (22,4)	0,219
<i>BRAF</i>	8 (34,8)	4 (6,0)	<b>0,002</b>
<i>EGFR</i>	1 (4,3)	6 (9,0)	0,673
<i>FBXW7</i>	0 (0,0)	6 (9,0)	0,332
<i>GNAS</i>	0 (0,0)	5 (7,5)	0,323
<i>SMAD4</i>	0 (0,0)	3 (4,5)	0,567
<i>CTNNB1</i>	0 (0,0)	3 (4,5)	0,567
<i>PTEN</i>	0 (0,0)	2 (3,0)	0,999
<i>PIK3CA</i>	1 (4,3)	1 (1,5)	0,448
<i>MLH1</i>	0 (0,0)	1 (1,5)	0,999
<i>IDH2</i>	1 (4,3)	0 (0,0)	0,256
<i>HRAS</i>	0 (0,0)	1 (1,5)	0,999
<i>FGFR3</i>	0 (0,0)	1 (1,5)	0,448
<i>ATM</i>	0 (0,0)	1 (1,5)	0,999

Valor *p* do Teste Exato de Fisher ou Teste  $\chi^2$  (\*)





**Figura 27.** Mutações *drivers* encontradas em adenomas (à esquerda) e pólipos serrilhados (à direita). Genes diferencialmente mutados incluem *APC* e *TP53* com maior frequência em adenomas e *BRAF* com maior frequência em pólipos serrilhados. Amostras sem mutações *drivers* foram omitidas da figura. PH: pólipo hiperplásico; LSS (lesão serrilhada séssil).

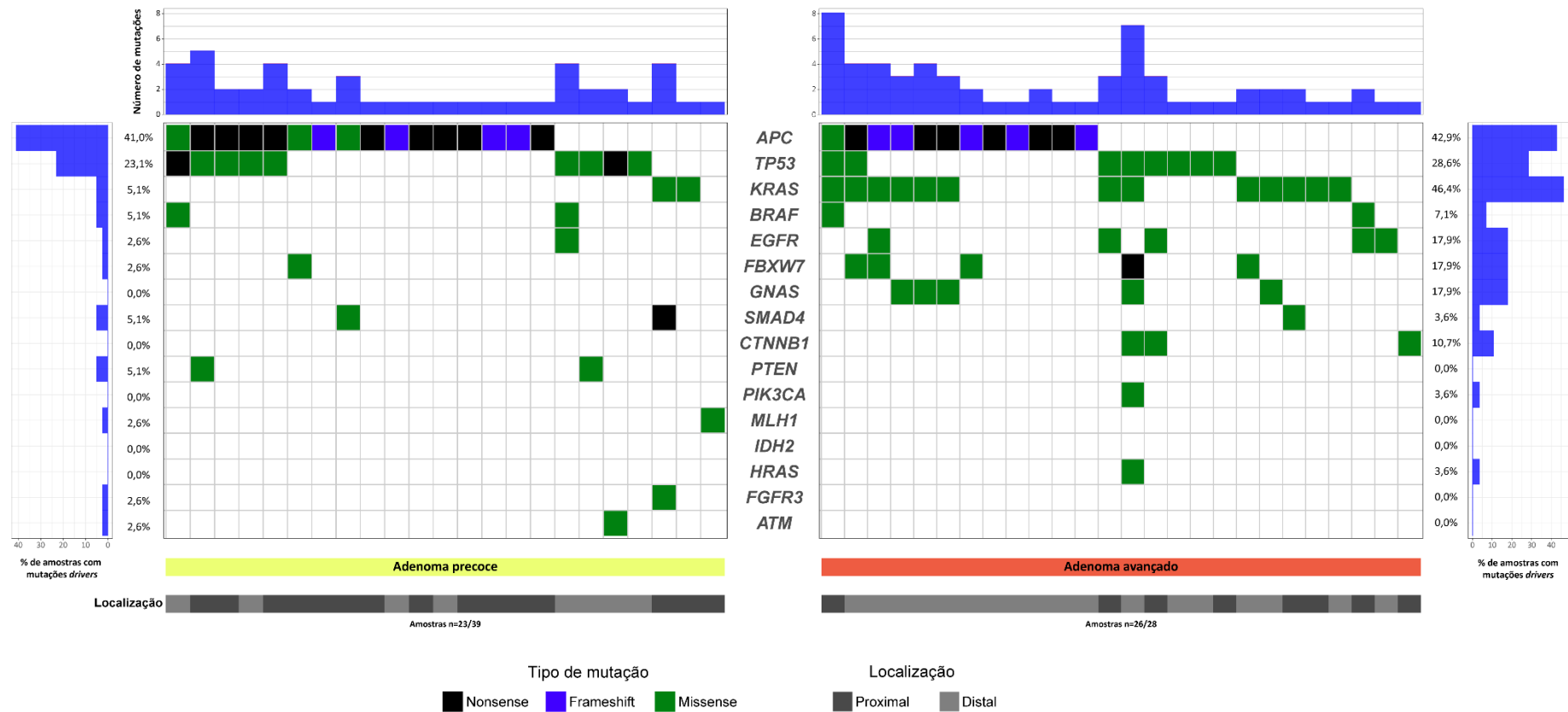
Em seguida, associamos a frequência de mutação *drivers* de cada gene entre os adenomas precoce e avançado (**Figura 28**). Além das diferenças na presença ou não de mutações *drivers*, adenomas avançados apresentaram maior frequência de mutação nos genes *KRAS* (46,4% vs 5,1%,  $p < 0,001$ ) e *GNAS* (17,9% vs 0,0%,  $p = 0,010$ , **Tabela 11**). Mutações nos genes *CTNNB1*, *PIK3CA* e *HRAS* ocorreram exclusivamente nos adenomas avançados (10,7%, 3,6%, 3,6%, respectivamente), embora não tenha apresentado significância estatística (**Tabela 11**). Mutações exclusivas em adenomas precoces aconteceram nos genes *PTEN* (5,1%), *MLH1*, *FGFR3* e *ATM* (todos com frequência de 2,6%), também sem significância estatística.

Além das diferenças entre a presença de mutações *drivers* nos pólipos serrilhados, mutações no gene *BRAF* foram encontradas com maior frequência em lesões serrilhadas sésseis (71,4% vs 18,8%,  $p = 0,026$ ). Algumas mutações ocorreram exclusivamente em LSSs, no gene *APC*, *TP53*, *PIK3CA* e *IDH2* (cada um presente em 14,3% das LSSs). Pólipos hiperplásicos apresentaram mutações exclusivas nos genes *KRAS* (apresentando mutações em 12,5% dos PH), *EGFR* e *FGFR3* (6,3%, **Figura 29**).

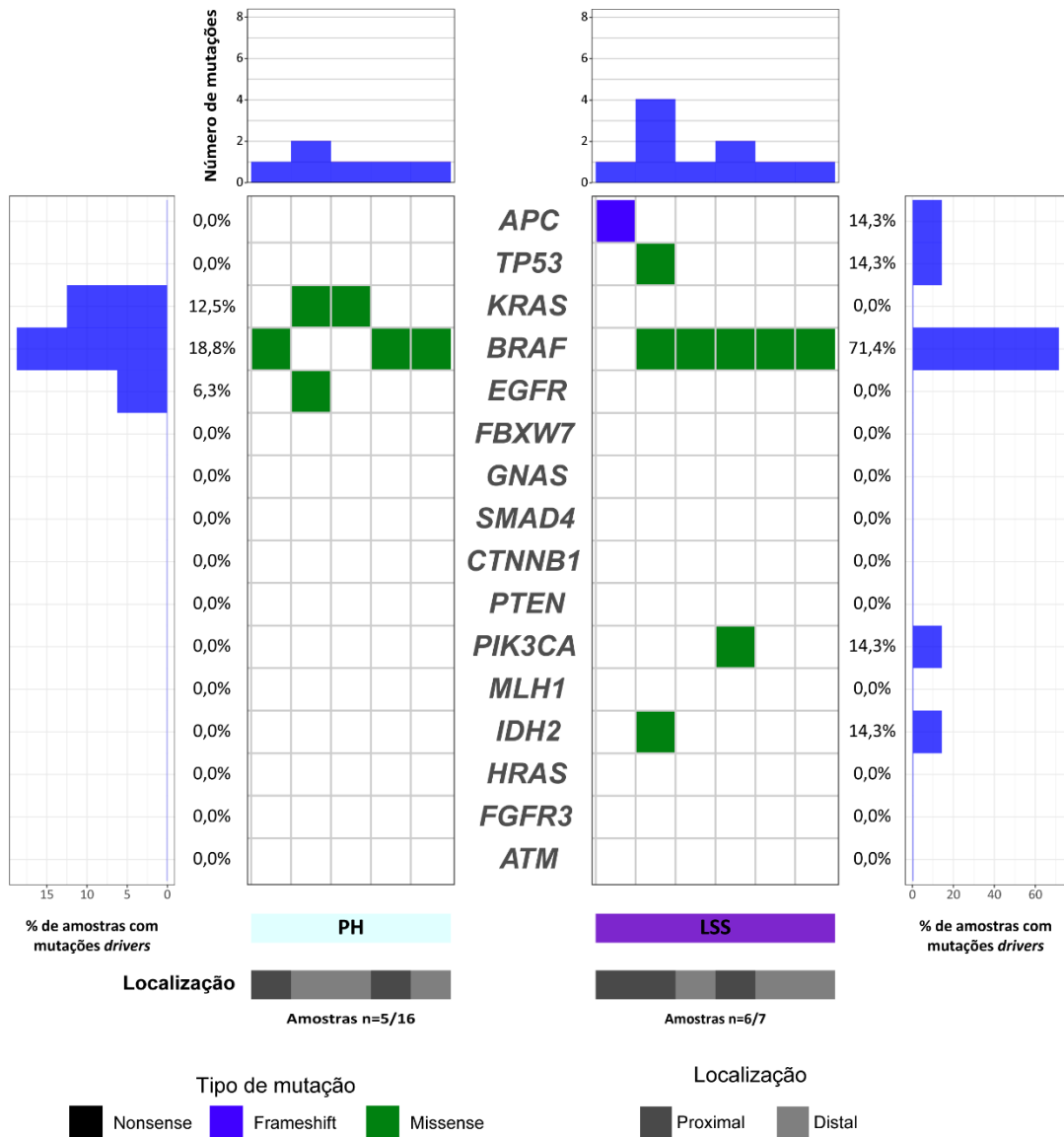
**Tabela 11.** Comparação entre a frequência de mutações em adenomas precoces e adenomas avançados e entre lesões serrilhadas sésseis e pólipos hiperplásicos

	Adenoma		Valor <i>p</i>	Pólipo Serrilhado		Valor <i>p</i>
	Precoce	Avançado		PH	LSS	
	n (%)	n (%)		n (%)	n (%)	
<i>APC</i>	16 (41,0)	12 (42,9)	>0,999*	0 (0,0)	1 (14,3)	0,304
<i>TP53</i>	9 (23,1)	8 (28,6)	0,777*	0 (0,0)	1 (14,3)	0,304
<i>KRAS</i>	2 (5,1)	13 (46,4)	<0,001*	2 (12,5)	0 (0,0)	>0,999
<i>BRAF</i>	2 (5,1)	2 (7,1)	>0,999	3 (18,8)	5 (71,4)	<b>0,026</b>
<i>EGFR</i>	1 (2,6)	5 (17,9)	0,075	1 (6,3)	0 (0,0)	>0,999
<i>FBXW7</i>	1 (2,6)	5 (17,9)	0,075	0 (0,0)	0 (0,0)	-
<i>GNAS</i>	0 (0,0)	5 (17,9)	<b>0,010</b>	0 (0,0)	0 (0,0)	-
<i>SMAD4</i>	2 (5,1)	1 (3,6)	>0,999	0 (0,0)	0 (0,0)	-
<i>CTNNB1</i>	0 (0,0)	3 (10,7)	0,068	0 (0,0)	0 (0,0)	-
<i>PTEN</i>	2 (5,1)	0 (0,0)	0,506	0 (0,0)	0 (0,0)	-
<i>PIK3CA</i>	0 (0,0)	1 (3,6)	0,418	0 (0,0)	1 (14,3)	0,304
<i>MLH1</i>	1 (2,6)	0 (0,0)	>0,999	0 (0,0)	0 (0,0)	-
<i>IDH2</i>	0 (0,0)	0 (0,0)	-	0 (0,0)	1 (14,3)	0,304
<i>HRAS</i>	0 (0,0)	1 (3,6)	0,418	0 (0,0)	0 (0,0)	-
<i>FGFR3</i>	1 (2,6)	0 (0,0)	>0,999	0 (0,0)	0 (0,0)	-
<i>ATM</i>	1 (2,6)	0 (0,0)	>0,999	0 (0,0)	0 (0,0)	-

Valor *p* do Teste Exato de Fisher ou Teste  $\chi^2$  (\*)



**Figura 28.** Mutações *drivers* encontradas em adenomas precoces (à esquerda) e adenomas avançados (à direita). Apesar de apresentarem algumas variações, como mutações exclusivas em alguns genes, apenas mutações nos genes *KRAS* e *GNAS* variaram significativamente entre os adenomas, com maior frequência em adenomas avançados. Amostras sem mutações *drivers* foram omitidas.

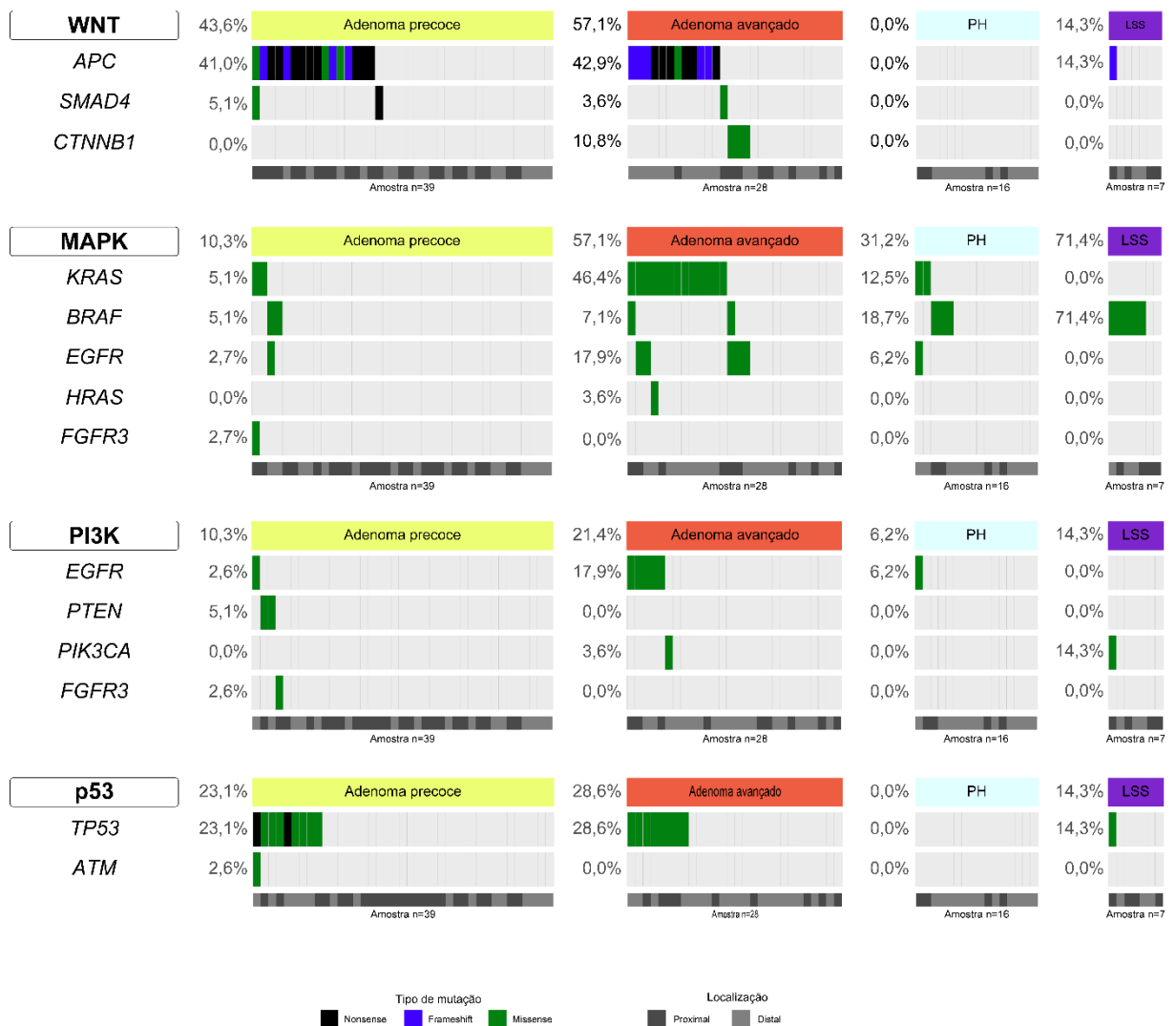


**Figura 29.** Mutações *drivers* encontradas em pólipos hiperplásicos (PH, à esquerda) e lesões serrilhadas sésseis (LSS, à direita). Alterações em genes relacionados à progressão adenoma-carcinoma, como *APC* e *TP53*, foram encontrados com baixa frequência. Mutações no gene *BRAF* diferiram significativamente entre PH e LSS e mutações em alguns genes foram específicas para cada tipo de lesão. Amostras sem mutações *drivers* foram omitidas da figura

## 2.5 Alterações em vias moleculares associadas ao CCR nas lesões não câncer

Alterações em vias moleculares relacionadas com câncer colorretal são mostradas na **Figura 30**. Alterações na via WNT foram encontradas em 37,8% das amostras, com alterações

nos genes *APC*, *SMAD4* e *CTNNB1*. Com exceção de um adenoma precoce, que apresentou mutações no gene *APC* e *SMAD4*, alterações nos genes dessa via foram mutualmente exclusivas. A maior frequência de alteração nesta via foi encontrada entre os adenomas (49,3% vs 4,3%,  $p < 0,001$ ), com maior frequência em adenomas avançados, seguido de adenomas precoces, LSSs e pólipos hiperplásicos (**Figura 30**,  $p < 0,001$ ). A via MAPK estava alterada em 33,3% das amostras, com alterações nos genes *KRAS*, *BRAF*, *EGFR*, *HRAS* e *FGFR3*. Os genes *BRAF* e *KRAS* apresentaram mutações mutualmente exclusivas com exceção de um adenoma avançado com mutações em ambos os genes. Pólipos serrilhados apresentaram maior frequência de alteração na via MAPK, embora sem significância estatística (43,5% vs 29,9%,  $p = 0,306$ ), com maior frequência entre as LSSs, seguido de adenomas avançados, pólipos hiperplásicos e adenomas precoces (**Figura 30**,  $p < 0,001$ ). Alterações na via PI3K foram encontradas em 13,3% das amostras, com mutações nos genes *EGFR*, *PTEN*, *PIK3CA* e *FGFR3*, as quais foram mutualmente exclusivas. Os adenomas apresentaram maior frequência de alteração na via PI3K, embora sem significância estatística (14,9% vs 8,7%,  $p = 0,359$ ). Por último, alterações na via p53, com os genes *TP53* e *ATM*, foram encontradas em 20,0% das lesões, majoritariamente com mutações no gene *TP53*, com uma mutação em *ATM* em uma mesma amostra de adenoma precoce.



**Figura 30.** Alterações em vias relacionadas ao câncer colorretal encontradas nas lesões precursoras. PH: pólipio hiperplásico; LSS: lesão serrilhada sésil

## DISCUSSÃO

Embora tenha sido observado um declínio na incidência do CCR em algumas regiões, como EUA, devido principalmente a implementação de métodos de rastreamento por colonoscopia<sup>3,107</sup>, a doença ainda é o terceiro tipo de câncer mais incidente no mundo, evidenciando a importância do conhecimento do perfil molecular das lesões colorretais tanto para detecção precoce quanto para prevenção da doença. Nesta parte do estudo, demonstramos a presença de mutações *drivers* em vias importantes relacionados ao CCR em lesões precursoras da doença, as quais apresentam um papel importante na proliferação e desenvolvimento tumoral. Além de detectar mutações em genes frequentemente alterados

no CCR, como *APC*, *TP53* e *KRAS*, observarmos diferenças no perfil mutacional de acordo com o tipo de lesão analisada. Adenomas apresentaram mutações em genes já conhecidos no envolvimento da progressão adenoma-carcinoma, tais como *APC*, *KRAS*, *TP53* e *FBXW7*. Pólipos serrilhados, por outro lado, apresentaram baixa frequência em genes como *APC* e *TP53*, estando relacionados com alterações na via MAPK, com mutações nos genes *BRAF* e *KRAS*, corroborando as diferenças moleculares já relatadas previamente entre essas duas vias de carcinogênese<sup>7</sup>.

Segundo o modelo clássico de progressão adenoma-carcinoma, proposto por Fearon e Vogelstein<sup>10</sup>, o acúmulo progressivo de alterações genéticas leva ao desenvolvimento do carcinoma a partir da mucosa normal. Estudos mais recentes têm adicionado complexidade à esse modelo, demonstrando a presença de heterogeneidade molecular em estágios iniciais do desenvolvimento das lesões colorretais<sup>49,50,108,109</sup> e mutações em diversos genes considerados *drivers* para o CCR<sup>49,50,110</sup>. A taxa de aquisição de mutações é aumentada em adenomas quando comparadas ao tecido normal<sup>111</sup>, e a carga mutacional em adenomas avançados têm sido reportada como semelhantes à de tumores<sup>49,50</sup>, mesmo quando apenas mutações *drivers* são analisadas<sup>49</sup>. De fato, encontramos uma média maior de mutações *drivers* em adenomas avançados. Embora tenha sido relatada associação de pólipos serrilhados ao subtipo molecular CMS1<sup>112,113</sup>, o qual frequentemente apresenta MSI e hipermutação, poucos estudos avaliaram a quantidade de mutações *drivers* neste tipo de lesão<sup>110,114</sup>, não identificando diferenças entre a quantidade de mutações em LSSs e adenomas<sup>110</sup>. Nós também não observamos diferenças na quantidade média de mutações entre LSSs e adenomas, o que pode estar associado à ausência de MSI nas LSSs analisadas (artigo submetido<sup>115</sup>).

Alterações na via WNT é um evento inicial na progressão do CCR, principalmente englobando alterações no gene *APC*<sup>10,46,47,49,50</sup>. A perda de *APC* está relacionada com o desenvolvimento de adenomas no intestino de camundongos<sup>116-118</sup> e sua restauração promove diferenciação e reestabelecimento das criptas no cólon, mesmo com a presença de mutações em *KRAS* e *TP53*<sup>118</sup>, que são eventos mais tardios na progressão adenoma-carcinoma. Mutações somáticas no gene *APC* são encontradas em adenomas de pacientes com tumores hereditários<sup>119</sup> e também em adenomas de pacientes sem tumores<sup>110,120,121</sup>, com uma variação de 40,3%<sup>110</sup> até 80%<sup>122</sup>, similar à encontrada no presente estudo (41,8%),. Mutações no gene *CTNNB1* são menos frequentes (11,9-20%)<sup>55,120,122</sup> mas com frequência

maior do que a observada no presente estudo (4,5%). Embora mutações em genes da via WNT foram relatados previamente em pólipos serrilhados<sup>114,123</sup>, a baixa frequência de mutações pode estar relacionada a outros eventos de desregulação, como mutações inativantes e baixa expressão de reguladores negativos da via<sup>124,125</sup>. Mutações no gene *APC* são menos frequentes nestas lesões quando comparados com adenomas<sup>126</sup>. Corroborando com estes dados, adenomas (principalmente aqueles classificados como de alto risco<sup>113</sup>) são associados majoritariamente ao subtipo molecular CMS2, o qual apresenta acentuada ativação da via WNT, ao passo que pólipos serrilhados estão associados com o subtipo CSM1<sup>112</sup>.

Mutações e perda de *TP53* são um evento que geralmente ocorre durante a transição adenoma para carcinoma<sup>10,47</sup> e trabalhos mais recentes com diversificação clonal têm confirmado a presença de mutações em *TP53* em clones iniciais de CCR<sup>49,127,128</sup>. Embora mutações em *TP53* não sejam um evento inicial no desenvolvimento de adenomas<sup>49,129</sup>, estão associadas à adenomas em regiões adjacentes à tumores<sup>122,130</sup>. A frequência de mutação em *TP53* parece aumentar entre adenomas precoces e avançados<sup>110</sup>. Adenomas precoces ou com baixa taxa de crescimento apresentam baixa frequência de mutações de *TP53*<sup>131</sup>. Nossos dados diferem em relação aos encontrados nestes trabalhos, pois não encontramos diferença significativa entre adenomas precoces e avançados, com frequências maiores do que 20% em ambos os casos. Uma vez que a presença de mutações em genes *drivers* do câncer têm sido relatada em diversos tecidos normais<sup>132-136</sup>, incluindo mutações em *TP53*, não podemos descartar a hipótese de que a alta frequência encontrada seja devido a presença de tais mutações no tecido normal destes pacientes. Embora nossos resultados vão de encontro ao modelo de Vogelstein na sequência adenoma carcinoma, onde o *TP53* é um dos responsáveis pela transição adenoma-carcinoma, a frequência de mutação de adenomas foi bem inferior à do CCR estudada neste trabalho (25,4% nos adenomas vs 56,0% no câncer). Nos pólipos serrilhados, a presença de mutações em *TP53* está geralmente associada com LSSs<sup>114,137</sup>, principalmente àquelas sem instabilidade de microssatélite<sup>114</sup>. Concordando com estes dados, mutações em *TP53* nos pólipos serrilhados foram encontradas apenas em LSSs, as quais não apresentavam instabilidade de microssatélite.

Ativação da via MAPK é frequentemente encontrada no CCR<sup>22,58</sup>, principalmente mutações nos oncogenes *KRAS* e *BRAF*. Encontramos uma frequência de 33,3% de mutações nesta via, com mutações no gene *KRAS* ligeiramente mais frequente em adenomas (22,4%), que é superior à frequência relatada previamente em nossa população (14,0%)<sup>25</sup>, mas dentro



da variação observada em outros estudos (10,7-60,0%)<sup>110,122,131,138</sup>. Similarmente ao que observamos, mutações *drivers* no gene *KRAS* apresentam maior frequência em adenomas avançados quando comparados com adenomas precoces<sup>25,110,138</sup>, correspondente ao modelo de progressão adenoma-carcinoma proposto por Fearon e Vogelstein<sup>10,47</sup>. A presença de mutações no gene *BRAF* está relacionada aos pólipos serrilhados<sup>25,114,138,139</sup>, principalmente em LSS (onde é encontrada com alta frequência, 50-67%)<sup>110,120,121</sup>, apresentando associação com o subtipo molecular CMS1<sup>112</sup>, o qual apresenta ativação na via MAPK. Confirmando esses dados, encontramos maior frequência de mutações nos genes da via MAPK nos pólipos serrilhados (43,5% vs 29,8%), principalmente por mutações ativantes do gene *BRAF* nas LSSs. Interessantemente, a principal mutação ativante de *BRAF*, V600E<sup>67</sup>, foi encontrada apenas em pólipos serrilhados, como já observado anteriormente<sup>131</sup>. Mutações no gene *EGFR*, *HRAS* e *FGFR3* são encontrados em baixas frequências em adenomas<sup>120,131</sup> e pólipos serrilhados<sup>114,123,131,140</sup>.

A via PI3K-AKT também está frequentemente ativada no CCR<sup>7,23,58</sup> e genes envolvidos com sua ativação podem ser observados enriquecidos em adenomas classificados como CMS2<sup>113</sup>. A via pode adquirir mutações acumulativas durante a progressão da doença<sup>141</sup>, com mutações observadas em adenomas que apresentam maior taxa de crescimento ao longo do tempo<sup>120</sup> e em adenomas classificados como alto risco para desenvolvimento do CCR<sup>142</sup>. Além disso, a via pode apresentar mutações ou alterações estruturais em tumores com foco de adenomas avançados ou adenomas tradicionais sésseis<sup>143</sup>, sugerindo um papel na progressão adenoma-carcinoma. Concordando com estes dados, nós observamos frequência de mutações ligeiramente maior em adenomas avançados comparados com adenomas precoces. A perda da *PTEN* promove ativação da via por meio da perda de controle de PI(3)K<sup>144,145</sup>, e é capaz de induzir a formação de pólipos no intestino de camundongos<sup>146</sup>. Mutações no gene *PTEN* são observadas em 12,5% de lesões serrilhadas sésseis com displasia<sup>114</sup> e em baixa frequência em adenomas (3,2%)<sup>120</sup>. Nós observamos mutações no gene *PTEN* apenas em adenomas precoces. Mutações no gene *PIK3CA* são encontradas em regiões precursoras em tumores *in situ*, embora não seja um evento inicial durante o processo de diversificação clonal na tumorigenese<sup>50</sup>, como observado em estudos de evolução clonal no CCR<sup>49,127,147</sup>. Além disso, mutações em *PIK3CA* são encontradas em regiões de adenomas associadas a tumores (20,0-30,0%)<sup>50,122</sup>, ou em baixa frequência em adenomas avançados (3,2%)<sup>148</sup>, semelhante ao observado em nossa população de estudo (3,6% de mutações em adenomas avançados).

A presença de mutações no gene *GNAS* é frequente em CCR<sup>7</sup>, e tem sido relatada em adenomas avançados<sup>49,121,149</sup>. Corroborando com estes dados, nós identificamos mutações neste gene apenas em adenomas avançados, não sendo identificada nem em adenomas precoces nem em pólipos serrilhados. Embora a frequência de alterações em *GNAS* em pólipos serrilhados já ter sido relatada, sua frequência não é alta e está relacionado com lesões mais avançadas<sup>150</sup>.

Dentre as limitações do estudo das lesões não câncer, podemos ressaltar como importantes a ausência de análise de tecido normal pareado e os diferentes painéis utilizados para os estudos do CCR e das lesões não câncer, dificultando assim as comparações entre frequências de mutações no contexto dos diferentes estágios anatomo-patológicos da carcinogênese colorretal. Além disso, a escassez de estudos nessa área, prejudica as avaliações minuciosas e comparações dos nossos resultados com os de outros autores.

Coletivamente, embora a ausência de tecido normal pareado para traçar o perfil de mutações somáticas das lesões, nossos dados são consistentes com o que tem sido observado em relação às alterações moleculares de lesões precursoras do CCR, com alterações distintas em adenomas e pólipos serrilhados. Adenomas apresentam maior frequência de mutações em genes já relacionados com a progressão adenoma-carcinoma, como *APC*, *TP53*, com adenomas avançados mostrando um número médio maior de mutações e com maior frequência em genes como *KRAS* e *GNAS*. Pólipos serrilhados apresentam maior frequência de alterações em *BRAF*. Também confirmamos que mutações no gene *KRAS* estão associadas à adenomas mais avançados.

## CONCLUSÕES

Lesões colorretais apresentam mutações em genes conhecidos como *drivers* no CCR, tais como *APC*, *TP53* e *KRAS*, apresentando diferenças de acordo com o tipo de lesão analisada.

Adenomas apresentaram número médio maior de mutações *drivers*, com maior frequência de mutações em *APC* e *TP53*. Adenomas precoces e adenomas avançados diferiram principalmente em relação ao número médio de mutações e frequência de mutações no gene *KRAS* e *GNAS*.

Pólipos serrilhadas apresentaram menor taxa de mutações *drivers* com frequência baixa de mutações na maioria dos genes analisados, com perfil mutacional relacionado com alterações na via MAPK, principalmente mutações no gene *BRAF*.

## CONCLUSÃO GERAL

Este é o primeiro estudo a realizar um perfil mutacional extenso nos diferentes estágios da carcinogênese colorretal, desde as lesões precursoras ao câncer em diferentes estádios da doença. Este perfil mutacional deverá ser mais profundamente estudado com potencial de utilização no rastreamento molecular da doença, prognóstico e tratamento do CCR.

Respondendo aos objetivos, especificamente este estudo mostrou:

1. Os genes mais frequentemente mutados no CCR, na população estudada, foram o *APC* (72,5%), *TP53* (56,0%), *KRAS* (52,7%), *PIK3CA* (15,4%) e *FBXW7* (12,1%).

2. Tumores do cólon proximal em relação ao cólon distal, apresentaram maior frequência de MSI e mutações nos genes *ATM*, *BLM*, *BRAF*, *NOTCH2*, *NOTCH3*, *PTEN*, *RANBP2* e *TGFBR*.

3. Tumores do cólon distal, em relação ao cólon proximal, apresentaram maior frequência de mutações nos genes *APC*, *TP53* e *KRAS*.

4. Tumores com instabilidade de microssatélite localizaram preferencialmente no cólon direito e apresentaram número médio de mutações maior do que os tumores sem instabilidade.

5. Dentre as lesões precursoras, os adenomas e as lesões serrilhadas sésseis apresentaram maiores frequências de mutações *drivers*. Dentre os adenomas, os adenomas avançados apresentaram maior frequência de mutações *drivers* (92,9%) do que os precoces (59,0%).

6. Nos adenomas, os genes mais frequentemente mutados foram *APC* (41,8%), *TP53* (25,4%), *KRAS* (22,4%); Nas LSSs os genes mais frequentemente mutados foram *BRAF* (71,4%), *APC* (14,3%), *TP53* (14,3%), *PIK3CA* (14,3%), e *IDH2* (14,3%); Em contraste, nos PHs, os únicos genes mutados foram *BRAF* (18,8%), *KRAS* (12,5%) e *EGFR* (6,3%);

## **CONSIDERAÇÕES ÉTICAS**

O presente estudo foi aprovado pelo Comitê de Ética e Pesquisa da Fundação PIO XII – Hospital do Câncer de Barretos (Projeto 1060/2015, Parecer 1.423.640, de 25 de Fevereiro de 2016) e foi realizado obedecendo critérios éticos em conformidade com a resolução CNS 466/12. Os avaliados foram selecionados retrospectivamente, obedecendo à anuência pelo Termo de Consentimento Livre e Esclarecido. Os dados serão coletados e utilizados somente para o objetivo do projeto, sendo as informações apresentadas de forma coletiva, sem qualquer prejuízo para as pessoas envolvidas, principalmente no que diz respeito à menção de nomes de pacientes ou profissionais de saúde envolvidos.

## REFERÊNCIAS

1. Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, et al. Cancer incidence and mortality worldwide : Sources , methods and major patterns in GLOBOCAN 2012. **Int J Cancer**. 2015;386:E359–86.
2. Bray F, Ferlay J, Soerjomataram I, Siegel R, Torre L, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. **CA A J Clin**. 2018;
3. Ferlay J, Colombet M, Soerjomataram I, Mathers C, Parkin DM, Piñeros M, et al. Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods. **Int J Cancer**. 2019;144(8):1941–53.
4. MINISTÉRIO DA SAÚDE: Instituto Nacional de Câncer José Alencar Gomes da Silva. Estimativa 2018-Incidência de câncer no Brasil. Rio de Janeiro; 2017. 130 p.
5. Fearon ER. Molecular genetics of colorectal cancer. **Annu Rev Pathol**. 2011;6:479–507.
6. IARC. World Cancer Report 2014. Stewart B, Wild C, editors. 2014. 630 p.
7. Kuipers EJ, Grady WM, Lieberman D, Seufferlein T, Sung JJ, Boelens PG, et al. Colorectal cancer. **Nat Rev Dis Prim**. 2015;1.
8. Bosman FT, Carneiro F, Hruban RH, Theise ND. WHO Classification of Tumours of the Digestive System. 4th ed. Bosman FT, Carneiro F, Hruban RH, Theise ND, editors. Lyon: IARC; 2010. 417 p.
9. Lambert R, Provenzale D, Ectors N, Vainio H, Dixon MF, Atkin W, et al. Early diagnosis and prevention of sporadic colorectal cancer. **Endoscopy**. 2001;33(12):1042–64.
10. Fearon ER, Vogelstein B. A genetic model for colorectal tumorigenesis. **Cell [Internet]**. 1990;61(5):759–67. Available from: <http://linkinghub.elsevier.com/retrieve/pii/0092867490901861>
11. Yamane L, Scapulatemp-Neto C, Reis RM, Guimarães DP. Serrated pathway in colorectal carcinogenesis. **World J Gastroenterol [Internet]**. 2014;20(10):2634. Available from: <http://www.wjgnet.com/1007-9327/full/v20/i10/2634.htm>
12. Leggett B, Whitehall V. Role of the Serrated Pathway in Colorectal Cancer Pathogenesis. **Gastroenterology [Internet]**. Elsevier Inc.; 2010;138(6):2088–100. Available from: <http://dx.doi.org/10.1053/j.gastro.2009.12.066>
13. Crockett SD, Nagtegaal ID. Terminology, Molecular Features, Epidemiology, and Management of Serrated Colorectal Neoplasia. **Gastroenterology [Internet]**. Elsevier, Inc; 2019;157(4):949-966.e4. Available from: <https://doi.org/10.1053/j.gastro.2019.06.041>
14. Mäkinen MJ. Colorectal serrated adenocarcinoma. **Histopathology [Internet]**. 2007;50(1):131–50. Available from: <http://doi.wiley.com/10.1111/j.1365-2559.2006.02548.x>
15. WHO. WHO classification of Tumours. Digestive System Tumours. 5th ed. WHO C of TEB, editor. WHO Press; 2019. 635 p.
16. Markowitz SD, Bertagnolli MM. Molecular Basis of Colorectal Cancer. **N Engl J Med [Internet]**. 2009;361(25):2449–60. Available from: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2843693/pdf/nihms-177087.pdf>
17. O’Connell JB, Maggard MA, Ko CY. Colon cancer survival rates with the new American Joint Committee on Cancer sixth edition staging. **J Natl Cancer Inst**. 2004;96(19):1420–

- 5.
18. Nawa T, Kato J, Kawamoto H, Okada H, Yamamoto H, Kohno H, et al. Differences between right- and left-sided colon cancer in patient characteristics, cancer morphology and histology. **J Gastroenterol Hepatol**. 2008;23(3):418–23.
19. Ghazi S, Lindfors U, Lindberg G, Berg E, Lindblom A, Papadogiannakis N. Analysis of colorectal cancer morphology in relation to sex, age, location, and family history. **J Gastroenterol**. 2012;47(6):619–34.
20. Missiaglia E, Jacobs B, D’Ario G, Di Narzo AF, Sonesson C, Budinska E, et al. Distal and proximal colon cancers differ in terms of molecular, pathological, and clinical features. **Ann Oncol**. 2014;25(10):1995–2001.
21. Drewes JL, Housseau F, Sears CL. Sporadic colorectal cancer: Microbial contributors to disease prevention, development and therapy. **Br J Cancer [Internet]**. Nature Publishing Group; 2016;115(3):273–80. Available from: <http://dx.doi.org/10.1038/bjc.2016.189>
22. Guinney J, Dienstmann R, Wang X, de Reyniès A, Schlicker A, Sonesson C, et al. The consensus molecular subtypes of colorectal cancer. **Nat Med [Internet]**. 2015;21(11):1350–6. Available from: <http://www.nature.com/doifinder/10.1038/nm.3967>
23. Muzny DM, Bainbridge MN, Chang K, Dinh HH, Drummond JA, Fowler G, et al. Comprehensive molecular characterization of human colon and rectal cancer. **Nature [Internet]**. Nature Publishing Group; 2012;487(7407):330–7. Available from: <http://dx.doi.org/10.1038/nature11252>
24. Lao VV, Grady WM. Epigenetics and colorectal cancer. **Nat Rev Gastroenterol Hepatol [Internet]**. 2011;8(12):686–700. Available from: <http://www.nature.com/doifinder/10.1038/nrgastro.2011.173>
25. Yamane LS, Scapulatempo-Neto C, Alvarenga L, Oliveira CZ, Berardinelli GN, Almodova E, et al. KRAS and BRAF mutations and MSI status in precursor lesions of colorectal cancer detected by colonoscopy. **Oncol Rep [Internet]**. 2014;1419–26. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/25050586>
26. Nelson S, Näthke IS. Interactions and functions of the adenomatous polyposis coli (APC) protein at a glance. **J Cell Sci [Internet]**. 2013;126(Pt 4):873–7. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23589686>
27. Sieber OM, Tomlinson IP, Lamlum H. The adenomatous polyposis coli (APC) tumour suppressor-genetics, function and disease. **Mol Med Today**. 2000;6(12):462–9.
28. Logan CY, Nusse R. THE WNT SIGNALING PATHWAY IN DEVELOPMENT AND DISEASE. **Annu Rev Cell Dev Biol [Internet]**. 2004;20(1):781–810. Available from: <http://www.annualreviews.org/doi/10.1146/annurev.cellbio.20.010403.113126>
29. Fearon ER, Wicha MS. KRAS and cancer stem cells in APC-mutant colorectal cancer. **J Natl Cancer Inst**. 2014;106(2):11–2.
30. Vousden KH, Prives C. Blinded by the Light : The Growing Complexity of p53. **Cell**. 2009;137:413–31.
31. Laghi L, Malesci A. Microsatellite Instability and Therapeutic Consequences in Colorectal Cancer. **Dig Dis [Internet]**. 2012;30(3):304–9. Available from: <http://www.karger.com/doi/10.1159/000337003>
32. Vilar E, Gruber SB. Microsatellite instability in colorectal cancer—the stable evidence. **Nat Rev Clin Oncol [Internet]**. 2010;7(3):153–62. Available from: <http://www.nature.com/doifinder/10.1038/nrclinonc.2009.237>

33. Goldstein NS. Serrated Pathway and APC (Conventional)-Type Colorectal Polyps. **Am J Clin Pathol [Internet]**. 2006;125(1):146–53. Available from: <http://ajcp.ascpjournals.org/cgi/doi/10.1309/87BD0C6UCGUG236J>
34. Rajagopalan H, Bardelli A, Lengauer C, Kinzler KW, Vogelstein B, Velculescu VE. RAF/RAS oncogenes and mismatch-repair status. **Nature**. 2002;418.
35. Pawa N, Arulampalam T, Norton JD. Screening for colorectal cancer : established and emerging modalities. **Nat Rev Gastroenterol Hepatol [Internet]**. Nature Publishing Group; 2011;8(12):711–22. Available from: <http://dx.doi.org/10.1038/nrgastro.2011.205>
36. Imperiale TF, Ransohoff DF, Itzkowitz SH, Levin TR, Lavin P, Lidgard GP, et al. Multitarget Stool DNA Testing for Colorectal-Cancer Screening. **N Engl J Med [Internet]**. 2014;370(14):1287–97. Available from: <http://www.nejm.org/doi/abs/10.1056/NEJMoa1311194>
37. Xie T, d' Ario G, Lamb JR, Martin E, Wang K, Tejpar S, et al. A Comprehensive Characterization of Genome-Wide Copy Number Aberrations in Colorectal Cancer Reveals Novel Oncogenes and Patterns of Alterations. Corvalan AH, editor. **PLoS One [Internet]**. 2012;7(7):e42001. Available from: <http://dx.plos.org/10.1371/journal.pone.0042001>
38. Gil Ferreira C, Aran V, Zalcborg-Renault I, Victorino AP, Salem JH, Bonamino MH, et al. KRAS mutations: Variable incidences in a Brazilian cohort of 8,234 metastatic colorectal cancer patients. **BMC Gastroenterol [Internet]**. BMC Gastroenterology; 2014;14(1):1–8. Available from: BMC Gastroenterology
39. Rapozo DCM, Grinmann AB, Carvalho ATP, Souza HSP, Soares-Lima SC, Almeida Simão T, et al. Analysis of mutations in TP53, APC, K-ras, and DCC genes in the non-dysplastic mucosa of patients with inflammatory bowel disease. **Int J Colorectal Dis**. 2009;24(10):1141–8.
40. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, et al. Integrative Analysis of Complex Cancer Genomics and Clinical Profiles Using the cBioPortal. **Sci Signal [Internet]**. 2013;6(269):pl1–pl1. Available from: <http://stke.sciencemag.org/cgi/doi/10.1126/scisignal.2004088>
41. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. **Bioinformatics**. 2009;25(14):1754–60.
42. Koboldt DC, Zhang Q, Larson DE, Shen D, Mclellan MD, Lin L, et al. VarScan 2 : Somatic mutation and copy number alteration discovery in cancer by exome sequencing. **Genome Res**. 2012;22(3):568–76.
43. McLaren W, Pritchard B, Rios D, Chen Y, Flicek P, Cunningham F. Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor. **Bioinformatics**. 2010;26(16):2069–70.
44. Tamborero D, Rubio-Perez C, Deu-Pons J, Schroeder MP, Vivancos A, Rovira A, et al. Cancer Genome Interpreter annotates the biological and clinical relevance of tumor alterations. **Genome Med**. Genome Medicine; 2018;10(1):1–8.
45. Berardinelli GN, Scapulatempo-Neto C, Durães R, de Oliveira MA, Guimarães D, Reis RM. Advantage of HSP110 (T17) marker inclusion for microsatellite instability (MSI) detection in colorectal cancer patients. **Oncotarget**. 2018;9(47):28691–701.
46. Powell SM, Zilz N, Beazer-Barclay Y, Bryan TM, Hamilton SR, Thibodeau SN, et al. APC mutations occur early during colorectal tumorigenesis. **Nature**. 1992;359(6392):235–7.

47. Vogelstein B, Fearon ER, Hamilton SR, Kern SE, Preisinger AC, Leppert M, et al. Genetic alterations during CRC development. **N Engl J Med.** 1988;319(9):525–32.
48. Sansom OJ, Reed KR, Hayes AJ, Ireland H, Brinkmann H, Newton IP, et al. Loss of Apc in vivo immediately perturbs Wnt signaling, differentiation, and migration. **Genes Dev.** 2004;18(12):1385–90.
49. Cross W, Kovac M, Mustonen V, Temko D, Davis H, Baker A-M, et al. The evolutionary landscape of colorectal tumorigenesis. **Nat Ecol Evol [Internet].** 2018;2(10):1661–72. Available from: <http://www.nature.com/articles/s41559-018-0642-z>
50. Saito T, Niida A, Uchi R, Hirata H, Komatsu H, Sakimura S, et al. A temporal shift of the evolutionary principle shaping intratumor heterogeneity in colorectal cancer. **Nat Commun.** 2018;9(1):1–11.
51. Giannakis M, Mu XJ, Shukla SA, Qian ZR, Cohen O, Nishihara R, et al. Genomic Correlates of Immune-Cell Infiltrates in Colorectal Carcinoma. **Cell Rep [Internet].** The Authors; 2016;15(4):857–65. Available from: <http://dx.doi.org/10.1016/j.celrep.2016.03.075>
52. Liu Y, Sethi NS, Hinoue T, Schneider BG, Cherniack AD, Sanchez-Vega F, et al. Comparative Molecular Analysis of Gastrointestinal Adenocarcinomas. **Cancer Cell [Internet].** 2018;33(4):721-735.e8. Available from: <http://linkinghub.elsevier.com/retrieve/pii/S1535610818301144>
53. Baker SJ, Preisinger AC, Jessup JM, Paraskeva C, Markowitz S, Willson JK V, et al. p53 Gene Mutations Occur in Combination with 17p Allelic Deletions as Late Events in Colorectal Tumorigenesis. **Cancer Res [Internet].** 1990;50:7717–22. Available from: <http://cancerres.aacrjournals.org/content/50/23/7717.long>
54. Seshagiri S, Stawiski EW, Durinck S, Modrusan Z, Storm EE, Conboy CB, et al. Recurrent R-spondin fusions in colon cancer. **Nature [Internet].** Nature Publishing Group; 2012;488(7413):660–4. Available from: <http://dx.doi.org/10.1038/nature11282>
55. Liu Z, Yang C, Li X, Luo W, Roy B, Xiong T, et al. The landscape of somatic mutation in sporadic Chinese colorectal cancer. **bioRxiv [Internet].** 2018;9(44):155671. Available from: <https://www.biorxiv.org/content/early/2018/03/15/155671>
56. Russo A, Bazan V, Iacopetta B, Kerr D, Soussi T, Gebbia N. The TP53 colorectal cancer international collaborative study on the prognostic and predictive significance of p53 mutation: Influence of tumor site, type of mutation, and adjuvant treatment. **J Clin Oncol.** 2005;23(30):7518–28.
57. Loree JM, Pereira AAL, Lam M, Willauer AN, Raghav K, Dasari A, et al. Classifying colorectal cancer by tumor location rather than sidedness highlights a continuum in mutation profiles and consensus molecular subtypes. **Clin Cancer Res.** 2018;24(5):1062–72.
58. Bailey MH, Tokheim C, Porta-Pardo E, Sengupta S, Bertrand D, Weerasinghe A, et al. Comprehensive Characterization of Cancer Driver Genes and Mutations. **Cell.** 2018;173(2):371-385.e18.
59. Seshagiri S, Stawiski EW, Durinck S, Modrusan Z, Storm EE, Conboy CB, et al. Recurrent R-spondin fusions in colon cancer. **Nature [Internet].** 2012;488(7413):660–4. Available from: <http://www.nature.com/doi/10.1038/nature11282>
60. Kasprzyk A, Baran J, Wang J, Hsu J, Yao L, Rivkin E, et al. International Cancer Genome Consortium Data Portal--a one-stop shop for cancer genomics data. **Database.** 2011;2011(0):bar026–bar026.
61. Salem ME, Weinberg BA, Xiu J, El-Deiry WS, Hwang JJ, Gatalica Z, et al. Comparative molecular analyses of left-sided colon, right-sided colon, and rectal cancers. **Oncotarget**



- [Internet]. 2017;8(49):86356–68. Available from: <http://www.oncotarget.com/fulltext/21169>
62. Al-Shamsi HO, Jones J, Fahmawi Y, Dahbour I, Tabash A, Abdel-Wahab R, et al. Molecular spectrum of KRAS, NRAS, BRAF, PIK3CA, TP53, and APC somatic gene mutations in Arab patients with colorectal cancer: Determination of frequency and distribution pattern. **J Gastrointest Oncol**. 2016;7(6):882–902.
  63. Loupakis F, Ruzzo A, Cremolini C, Vincenzi B, Salvatore L, Santini D, et al. KRAS codon 61, 146 and BRAF mutations predict resistance to cetuximab plus irinotecan in KRAS codon 12 and 13 wild-type metastatic colorectal cancer. **Br J Cancer [Internet]**. Nature Publishing Group; 2009;101(4):715–21. Available from: <http://dx.doi.org/10.1038/sj.bjc.6605177>
  64. Boeckx N, Peeters M, Van Camp G, Pauwels P, Op De Beeck K, Deschoolmeester V. Prognostic and Predictive Value of RAS Gene Mutations in Colorectal Cancer: Moving Beyond KRAS Exon 2. **Drugs**. Springer International Publishing; 2015;75(15):1739–56.
  65. Rosty C, Young JP, Walsh MD, Clendenning M, Walters RJ, Pearson S, et al. Colorectal carcinomas with KRAS mutation are associated with distinctive morphological and molecular features. **Mod Pathol [Internet]**. Nature Publishing Group; 2013;26(6):825–34. Available from: <http://dx.doi.org/10.1038/modpathol.2012.240>
  66. Samowitz WS, Curtin K, Schaffer D, Study S a P, Robertson M, Leppert M. Relationship of Ki- ras Mutations in Colon Cancers to Tumor Location , Stage , and Survival : A Population-based Study Relationship of Ki- ras Mutations in Colon Cancers to Tumor Location ,. **Cancer Epidemiol Biomarkers Prev**. 2000;9(November):1193–7.
  67. Davies H, Bignell GR, Cox C, Stephens P, Edkins S, Clegg S, et al. Mutations of the BRAF gene in human cancer. **Nature [Internet]**. 2002;417(6892):949–54. Available from: <http://www.nature.com/articles/nature00766>
  68. Yao Z, Yaeger R, Rodrik-Outmezguine VS, Tao A, Torres NM, Chang MT, et al. Tumours with class 3 BRAF mutants are sensitive to the inhibition of activated RAS. **Nature [Internet]**. Nature Publishing Group; 2017;548(7666):234–8. Available from: <http://dx.doi.org/10.1038/nature23291>
  69. Wan PTC, Garnett MJ, Roe SM, Lee S, Niculescu-Duvaz D, Good VM, et al. Mechanism of activation of the RAF-ERK signaling pathway by oncogenic mutations of B-RAF. **Cell**. 2004;116(6):855–67.
  70. Malapelle U, Pisapia P, Sgariglia R, Vigliar E, Biglietto M, Carlomagno C, et al. Less frequently mutated genes in colorectal cancer: evidences from next-generation sequencing of 653 routine cases. **J Clin Pathol [Internet]**. 2016;69(9):767–71. Available from: <http://jcp.bmj.com/lookup/doi/10.1136/jclinpath-2015-203403>
  71. Rosty C, Young JP, Walsh MD, Clendenning M, Sanderson K, Walters RJ, et al. PIK3CA Activating Mutation in Colorectal Carcinoma: Associations with Molecular Features and Survival. **PLoS One**. 2013;8(6):1–9.
  72. Rodon J, Dienstmann R, Serra V, Tabernero J. Development of PI3K inhibitors: lessons learned from early clinical trials. **Nat Rev Clin Oncol [Internet]**. Nature Publishing Group; 2013;10(3):143–53. Available from: <http://dx.doi.org/10.1038/nrclinonc.2013.10>
  73. Grabiner BC, Nardi V, Birsoy K, Possemato R, Shen K, Sinha S, et al. A diverse array of cancer-associated MTOR mutations are hyperactivating and can predict rapamycin sensitivity. **Cancer Discov**. 2014;4(5):554–63.
  74. Carpten JD, Faber AL, Horn C, Donoho GP, Briggs SL, Robbins CM, et al. A transforming

- mutation in the pleckstrin homology domain of AKT1 in cancer. **Nature**. 2007;448(7152):439–44.
75. Zeki SS, Graham TA, Wright NA. Stem cells and their implications for colorectal cancer. **Nat Rev Gastroenterol Hepatol [Internet]**. Nature Publishing Group; 2011;8(2):90–100. Available from: <http://www.nature.com/doi/10.1038/nrgastro.2010.211>
  76. Zhang Y, Li B, Ji ZZ, Zheng PS. Notch1 regulates the growth of human colon cancers. **Cancer**. 2010;116(22):5207–18.
  77. Mu Y, Zou H, Chen B, Fan Y, Luo S. FAM83D knockdown regulates proliferation, migration and invasion of colorectal cancer through inhibiting FBXW7/Notch-1 signalling pathway. **Biomed Pharmacother**. 2017;90:548–54.
  78. Sancho R, Jandke A, Davis H, Diefenbacher ME, Tomlinson I, Behrens A. F-box and WD repeat domain-containing 7 regulates intestinal cell lineage commitment and is a haploinsufficient tumor suppressor. **Gastroenterology [Internet]**. Elsevier Inc.; 2010;139(3):929–41. Available from: <http://dx.doi.org/10.1053/j.gastro.2010.05.078>
  79. Babaei-Jadidi R, Li N, Saadeddin A, Spencer-Dene B, Jandke A, Muhammad B, et al. FBXW7 influences murine intestinal homeostasis and cancer, targeting Notch, Jun, and DEK for degradation. **J Exp Med [Internet]**. 2011;208(2):295–312. Available from: <http://www.jem.org/lookup/doi/10.1084/jem.20100830>
  80. C H, Lewis A, Behrens A, Tomlinson I. Investigation of the atypical FBXW7 mutation spectrum in human tumours by conditional expression of a heterozygous propellor tip missense allele in the mouse intestines. **Gut**. 2014;63(5):792–9.
  81. Korphaisarn K, Morris VK, Overman MJ, Fogelman DR, Kee BK, Raghav KPS, et al. FBXW7 missense mutation: a novel negative prognostic factor in metastatic colorectal adenocarcinoma. **Oncotarget [Internet]**. 2017;8(24):39268–79. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/28424412><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC5503612>
  82. Welcker M, Clurman BE. FBW7 ubiquitin ligase: A tumour suppressor at the crossroads of cell division, growth and differentiation. **Nat Rev Cancer**. 2008;8(2):83–93.
  83. Lee SY, Kumano K, Nakazaki K, Sanada M, Matsumoto A, Yamamoto G, et al. Gain-of-function mutations and copy number increases of Notch2 in diffuse large B-cell lymphoma. **Cancer Sci**. 2009;100(5):920–6.
  84. Weng AP, Ferrando AA, Lee W, Morris IV JP, Silverman LB, Sanchez-Irizarry C, et al. Activating Mutations of NOTCH1 in Human T Cell Acute Lymphoblastic Leukemia. **Science (80- ) [Internet]**. 2004;306(5694):269–71. Available from: <http://www.sciencemag.org/cgi/doi/10.1126/science.1102160>
  85. Pietrantonio F, Di Nicolantonio F, Schrock AB, Lee J, Tejpar S, Sartore-Bianchi A, et al. ALK, ROS1, and NTRK rearrangements in metastatic colorectal cancer. **J Natl Cancer Inst**. 2017;109(12):1–10.
  86. Stransky N, Cerami E, Schalm S, Kim JL, Lengauer C. The landscape of kinase fusions in cancer. **Nat Commun [Internet]**. Nature Publishing Group; 2014;5(1):4846. Available from: <http://dx.doi.org/10.1038/ncomms5846>
  87. Mulligan LM. RET revisited: Expanding the oncogenic portfolio. **Nat Rev Cancer [Internet]**. Nature Publishing Group; 2014;14(3):173–86. Available from: <http://dx.doi.org/10.1038/nrc3680>
  88. Russo M, Misale S, Wei G, Siravegna G, Crisafulli G, Lazzari L, et al. Acquired Resistance to the TRK Inhibitor Entrectinib in Colorectal Cancer. **Cancer Discov [Internet]**. 2016;6(1):36–44. Available from:

- <http://cancerdiscovery.aacrjournals.org/cgi/doi/10.1158/2159-8290.CD-15-0940>
89. Facchinetti F, Loriot Y, Kuo MS, Mahjoubi L, Lacroix L, Planchard D, et al. Crizotinib-resistant ROS1 mutations reveal a predictive kinase inhibitor sensitivity model for ROS1- and ALK-rearranged lung cancers. **Clin Cancer Res.** 2016;22(24):5983–91.
  90. Gainor JF, Dardaei L, Yoda S, Friboulet L, Leshchiner I, Katayama R, et al. Molecular Mechanisms of Resistance to First- and Second-Generation ALK Inhibitors in ALK-Rearranged Lung Cancer. **Cancer Discov [Internet].** 2016;6(10):1118–33. Available from: <http://cancerdiscovery.aacrjournals.org/cgi/doi/10.1158/2159-8290.CD-16-0596>
  91. Chakravarty D, Gao J, Phillips SM, Kundra R, Zhang H, Wang J, et al. OncoKB: A Precision Oncology Knowledge Base HHS Public Access. 2017;1–23. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5586540/pdf/nihms897314.pdf>
  92. Chen H, Wester M, Liang H, Kong K, Meric-Bernstam F, Ip CKM, et al. Systematic Functional Annotation of Somatic Mutations in Cancer. **Cancer Cell [Internet].** Elsevier Inc.; 2018;33(3):450-462.e10. Available from: <https://doi.org/10.1016/j.ccell.2018.01.021>
  93. Jaiswal BS, Kljavin NM, Stawiski EW, Chan E, Parikh C, Durinck S, et al. Oncogenic ERBB3 Mutations in Human Cancers. **Cancer Cell [Internet].** Elsevier Inc.; 2013;23(5):603–17. Available from: <http://dx.doi.org/10.1016/j.ccr.2013.04.012>
  94. Prickett TD, Agrawal NS, Wei X, Yates KE, Lin JC, Wunderlich JR, et al. Analysis of the tyrosine kinome in melanoma reveals recurrent mutations in ERBB4. **Nat Genet [Internet].** Nature Publishing Group; 2009;41(10):1127–32. Available from: <http://dx.doi.org/10.1038/ng.438>
  95. Hutchins G, Southward K, Handley K, Magill L, Beaumont C, Stahlschmidt J, et al. Value of mismatch repair, KRAS, and BRAF mutations in predicting recurrence and benefits from chemotherapy in colorectal cancer. **J Clin Oncol.** 2011;29(10):1261–70.
  96. Sveen A, Kopetz S, Lothe RA. Biomarker-guided therapy for colorectal cancer: strength in complexity. **Nat Rev Clin Oncol [Internet].** Springer US; 2019; Available from: <http://dx.doi.org/10.1038/s41571-019-0241-1>
  97. Popat S, Hubner R, Houlston RS. Systematic review of microsatellite instability and colorectal cancer prognosis. **J Clin Oncol.** 2005;23(3):609–18.
  98. Guastadisegni C, Colafranceschi M, Ottini L, Dogliotti E. Microsatellite instability as a marker of prognosis and response to therapy: A meta-analysis of colorectal cancer survival data. **Eur J Cancer [Internet].** Elsevier Ltd; 2010;46(15):2788–98. Available from: <http://dx.doi.org/10.1016/j.ejca.2010.05.009>
  99. Jin Z, Sanhueza CT, Johnson B, Nagorney DM, Larson DW, Mara KC, et al. Outcome of Mismatch Repair-Deficient Metastatic Colorectal Cancer: The Mayo Clinic Experience. **Oncologist.** 2018;23(9):1083–91.
  100. Sargent DJ, Marsoni S, Monges G, Thibodeau SN, Labianca R, Hamilton SR, et al. Defective mismatch repair as a predictive marker for lack of efficacy of fluorouracil-based adjuvant therapy in colon cancer. **J Clin Oncol.** 2010;28(20):3219–26.
  101. Sepulveda AR, Hamilton SR, Allegra CJ, Grody W, Cushman-Vokoun AM, Funkhouser WK, et al. Molecular biomarkers for the evaluation of colorectal cancer: Guideline from the American society for clinical pathology, college of American pathologists, association for molecular pathology, and American society of clinical oncology. **Arch Pathol Lab Med.** 2017;141(5):625–57.
  102. Roock W De, Vriendt V De, Normanno N, Ciardiello F, Tejpar S. KRAS, BRAF, PIK3CA, and

- PTEN mutations: Implications for targeted therapies in metastatic colorectal cancer. **Lancet Oncol [Internet]**. Elsevier Ltd; 2011;12(6):594–603. Available from: [http://dx.doi.org/10.1016/S1470-2045\(10\)70209-6](http://dx.doi.org/10.1016/S1470-2045(10)70209-6)
103. Sartore-Bianchi A, Martini M, Molinari F, Veronese S, Nichelatti M, Artale S, et al. PIK3CA mutations in colorectal cancer are associated with clinical resistance to EGFR-targeted monoclonal antibodies. **Cancer Res**. 2009;69(5):1851–7.
  104. Peeters M, Oliner KS, Parker A, Siena S, Van Cutsem E, Huang J, et al. Massively parallel tumor multigene sequencing to evaluate response to panitumumab in a randomized phase III study of metastatic colorectal cancer. **Clin Cancer Res**. 2013;19(7):1902–12.
  105. De Roock W, Claes B, Bernasconi D, De Schutter J, Biesmans B, Fountzilias G, et al. Effects of KRAS, BRAF, NRAS, and PIK3CA mutations on the efficacy of cetuximab plus chemotherapy in chemotherapy-refractory metastatic colorectal cancer: A retrospective consortium analysis. **Lancet Oncol**. 2010;11(8):753–62.
  106. Lambert R, Lightdale CJ. The Paris endoscopic classification of superficial neoplastic lesions: Esophagus, stomach, and colon - Paris, France November 30 to December 1, 2002: Foreword. **Gastrointest Endosc**. 2003;58(6 SUPPL.):3–43.
  107. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2018. **CA Cancer J Clin [Internet]**. 2018;68(1):7–30. Available from: <http://doi.wiley.com/10.3322/caac.21442>
  108. Sottoriva A, Kang H, Ma Z, Graham TA, Salomon MP, Zhao J, et al. A big bang model of human colorectal tumor growth. **Nat Genet [Internet]**. Nature Publishing Group; 2015;47(3):209–16. Available from: <http://dx.doi.org/10.1038/ng.3214>
  109. Kang H, Salomon MP, Sottoriva A, Zhao J, Toy M, Press MF, et al. Many private mutations originate from the first few divisions of a human colorectal adenoma. **J Pathol**. 2015;237(3):355–62.
  110. Lin SH, Raju GS, Huff C, Ye Y, Gu J, Chen JS, et al. The somatic mutation landscape of premalignant colorectal adenoma. **Gut**. 2018;67(7):1299–305.
  111. Lugli N, Dionellis VS, Ordóñez-Morán P, Kamileri I, Sotiriou SK, Huelsken J, et al. Enhanced Rate of Acquisition of Point Mutations in Mouse Intestinal Adenomas Compared to Normal Tissue. **Cell Rep**. 2017;19(11):2185–92.
  112. Chang K, Willis JA, Reumers J, Taggart MW, San Lucas FA, Thirumurthi S, et al. Colorectal premalignancy is associated with consensus molecular subtypes 1 and 2. **Ann Oncol**. 2018;29(10):2061–7.
  113. Komor MA, Bosch LJW, Bounova G, Boliijn AS, Delis-van Diemen PM, Rausch C, et al. Consensus molecular subtype classification of colorectal adenomas. **J Pathol**. 2018;246(3):266–76.
  114. Murakami T, Akazawa Y, Yatagai N, Hiromoto T, Sasahara N, Saito T, et al. Molecular characterization of sessile serrated adenoma/polyps with dysplasia/carcinoma based on immunohistochemistry, next-generation sequencing, and microsatellite instability testing: A case series study. **Diagn Pathol**. Diagnostic Pathology; 2018;13(1):1–10.
  115. Sobanski T, Arantes LMRB, dos Santos W, Matsushita M, de Oliveira MA, Costa M, et al. Methylation Profile of Colon Cancer Genes in Colorectal Precursor lesions and Tumor Tissue. **Scientific Reports**. 2020.
  116. Su LK, Kinzler KW, Vogelstein B, Preisinger AC, Moser AR, Luongo C, et al. Multiple intestinal neoplasia caused by a mutation in the murine homolog of the APC gene. **Science (80- )**. 1992;256(5057):668–70.
  117. Cheung AF, Carter AM, Kostova KK, Woodruff JF, Crowley D, Bronson RT, et al. Complete deletion of Apc results in severe polyposis in mice. **Oncogene**. Nature Publishing Group;

- 2010;29(12):1857–64.
118. Dow LE, O'Rourke KP, Simon J, Tschaharganeh DF, van Es JH, Clevers H, et al. Apc Restoration Promotes Cellular Differentiation and Reestablishes Crypt Homeostasis in Colorectal Cancer. **Cell [Internet]**. Elsevier Ltd; 2015;161(7):1539–52. Available from: <http://www.sciencedirect.com/science/article/pii/S0092867415006236>
  119. Borrás E, Lucas FAS, Chang K, Zhou R, Masand G, Fowler J, et al. Genomic landscape of colorectal mucosa and adenomas. **Cancer Prev Res**. 2016;9(6):417–27.
  120. van Lanschot MCJ, Carvalho B, Rausch C, Snaebjornsson P, van Engeland M, Kuipers EJ, et al. Molecular profiling of longitudinally observed small colorectal polyps: A cohort study. **EBioMedicine [Internet]**. The Authors; 2019;39:292–300. Available from: <https://doi.org/10.1016/j.ebiom.2018.12.009>
  121. Nikolaev SI, Sotiriou SK, Pateras IS, Santoni F, Sougioultzis S, Edgren H, et al. A Single-Nucleotide Substitution Mutator Phenotype Revealed by Exome Sequencing of Human Colon Adenomas. **Cancer Res [Internet]**. 2012;72(23):6279–89. Available from: <http://cancerres.aacrjournals.org/cgi/doi/10.1158/0008-5472.CAN-12-3869>
  122. Druliner BR, Wang P, Bae T, Baheti S, Slettedahl S, Mahoney D, et al. Molecular characterization of colorectal adenomas with and without malignancy reveals distinguishing genome, transcriptome and methylome alterations. **Sci Rep**. 2018;8(1):1–10.
  123. Patai Á V., Barták BK, Péterfia B, Micsik T, Horváth R, Sumánszki C, et al. Comprehensive DNA Methylation and Mutation Analyses Reveal a Methylation Signature in Colorectal Sessile Serrated Adenomas. **Pathol Oncol Res**. Pathology & Oncology Research; 2017;23(3):589–94.
  124. Kedrin D, Gala MK. Genetics of the serrated pathway to colorectal cancer. **Clin Transl Gastroenterol [Internet]**. Nature Publishing Group; 2015;6(4):e84-3. Available from: <http://dx.doi.org/10.1038/ctg.2015.12>
  125. Yan HHN, Lai JCW, Ho SL, Leung WK, Law WL, Lee JFY, et al. RNF43 germline and somatic mutation in serrated neoplasia pathway and its association with BRAF mutation. **Gut**. 2017;66(9):1645–56.
  126. Borowsky J, Dumenil T, Bettington M, Pearson SA, Bond C, Fennell L, et al. The role of APC in WNT pathway activation in serrated neoplasia. **Mod Pathol [Internet]**. Nature Publishing Group; 2018;31(3):495–504. Available from: <http://dx.doi.org/10.1038/modpathol.2017.150>
  127. Uchi R, Takahashi Y, Niida A, Shimamura T, Hirata H, Sugimachi K, et al. Integrated Multiregional Analysis Proposing a New Model of Colorectal Cancer Evolution. **PLoS Genet**. 2016;12(2):1–24.
  128. Kim R, Schell MJ, Teer JK, Greenawalt DM, Yang M, Yeatman TJ. Co-evolution of somatic variation in primary and metastatic colorectal cancer may expand biopsy indications in the molecular era. **PLoS One**. 2015;10(5):1–12.
  129. Jones S, Chen WD, Parmigiani G, Diehl F, Beerewinkel N, Antal T, et al. Comparative lesion sequencing provides insights into tumor evolution. **Proc Natl Acad Sci U S A**. 2008;105(11):4283–8.
  130. Wolff RK, Hoffman MD, Wolff EC, Herrick JS, Sakoda LC, Samowitz WS, et al. Mutation analysis of adenomas and carcinomas of the colon: Early and late drivers. **Genes Chromosom Cancer**. 2018;57(7):366–76.
  131. Sievers CK, Zou LS, Pickhardt PJ, Matkowskyj KA, Albrecht DM, Clipson L, et al. Subclonal diversity arises early even in small colorectal tumours and contributes to differential

- growth fates. **Gut**. 2017;66(12):2132–40.
132. Martincorena I. Somatic mutation and clonal expansions in human tissues. **Genome Med**. Genome Medicine; 2019;11(1):11–3.
  133. Martincorena I, Fowler JC, Wabik A, Lawson ARJ, Abascal F, Hall MWJ, et al. Somatic mutant clones colonize the human esophagus with age. **Science (80- )**. 2018;362(6417):911–7.
  134. Martincorena I, Roshan A, Gerstung M, Ellis P, Van Loo P, McLaren S, et al. High burden and pervasive positive selection of somatic mutations in normal human skin. **Science (80- )**. 2015;348(6237):880–6.
  135. Kato S, Lippman SM, Flaherty KT, Kurzrock R. The conundrum of genetic “Drivers” in benign conditions. **J Natl Cancer Inst**. 2016;108(8):1–13.
  136. Lee-Six H, Olafsson S, Ellis P, Osborne RJ, Sanders MA, Moore L, et al. The landscape of somatic mutation in normal colorectal epithelial cells. **Nature [Internet]**. Springer US; 2019;574(7779):532–7. Available from: <http://dx.doi.org/10.1038/s41586-019-1672-7>
  137. Konishi K, Yamochi T, Makino R, Kaneko K, Yamamoto T, Nozawa H, et al. Molecular differences between sporadic serrated and conventional colorectal adenomas. **Clin Cancer Res**. 2004;10(9):3082–90.
  138. Yi C, Huang Y, Yu X, Li X, Zheng S, Ding K, et al. Clinicopathologic distribution of KRAS and BRAF mutations in a Chinese population with colorectal cancer precursor lesions. **Oncotarget**. 2016;7(13):17265–74.
  139. Tanaka Y, Yamano H o., Yamamoto E, Matushita H o., Aoki H, Yoshikawa K, et al. Endoscopic and molecular characterization of colorectal sessile serrated adenoma/polyps with cytologic dysplasia. **Gastrointest Endosc [Internet]**. American Society for Gastrointestinal Endoscopy; 2017;86(6):1131-1138.e4. Available from: <https://doi.org/10.1016/j.gie.2017.05.006>
  140. Kim MJ, Lee EJ, Suh JP, Chun SM, Jang SJ, Kim DS, et al. Traditional serrated adenoma of the colorectum. **Am J Clin Pathol**. 2013;140(6):898–911.
  141. Wu H, Zhang XY, Hu Z, Hou Q, Zhang H, Li Y, et al. Evolution and heterogeneity of non-hereditary colorectal cancer revealed by single-cell exome sequencing. **Oncogene**. 2017;36(20):2857–67.
  142. Komor MA, de Wit M, van den Berg J, Martens de Kemp SR, Delis-van Diemen PM, Bolijn AS, et al. Molecular characterization of colorectal adenomas reveals POFUT1 as a candidate driver of tumor progression. **Int J Cancer**. 2019;00:1–14.
  143. Yamamoto E, Suzuki H, Yamano H, Maruyama R, Nojima M, Kamimae S, et al. Molecular Dissection of Premalignant Colorectal Lesions Reveals Early Onset of the CpG Island Methylator Phenotype. **Am J Pathol [Internet]**. Elsevier Inc.; 2012;181(5):1847–61. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/22995252>
  144. Cantley LC, Neel BG. New insights into tumor suppression: PTEN suppresses tumor formation by restraining the phosphoinositide 3-kinase/AKT pathway. **Proc Natl Acad Sci U S A**. 1999;96(8):4240–5.
  145. Stiles B, Groszer M, Wang S, Jiao J, Wu H. PTENless means more. **Dev Biol**. 2004;273(2):175–84.
  146. He XC, Yin T, Grindley JC, Tian Q, Sato T, Tao WA, et al. PTEN-deficient intestinal stem cells initiate intestinal polyposis. **Nat Genet**. 2007;39(2):189–98.
  147. Kim TM, An CH, Rhee JK, Jung SH, Lee SH, Baek IP, et al. Clonal origins and parallel evolution of regionally synchronous colorectal adenoma and carcinoma. **Oncotarget**. 2015;6(29):27725–35.

148. Whitehall VLJ, Rickman C, Bond CE, Ramsnes I, Greco SA, Umapathy A, et al. Oncogenic PIK3CA mutations in colorectal cancers and polyps. **Int J Cancer**. 2012;131(4):813–20.
149. Chang P-Y, Chen J-S, Chang S-C, Wang M-C, Chang N-C, Wen Y-H, et al. Acquired somatic *TP53* or *PIK3CA* mutations are potential predictors of when polyps evolve into colorectal cancer. **Oncotarget**. 2017;8(42):72352–62.
150. Hashimoto T, Yamashita S, Yoshida H, Taniguchi H, Ushijima T, Yamada T, et al. WNT Pathway Gene Mutations Are Associated With the Presence of Dysplasia in Colorectal Sessile Serrated Adenoma/Polyps. **Am J Surg Pathol**. 2017;41(9):1188–97.

## Anexos

## Anexo A. Artigo publicado referente aos dados de mutações nas amostras tumorais

www.nature.com/scientificreports

SCIENTIFIC  
REPORTS

nature research

## OPEN Mutation profiling of cancer drivers in Brazilian colorectal cancer

Wellington dos Santos<sup>1</sup>, Thais Sobanski<sup>1</sup>, Ana Carolina de Carvalho<sup>1</sup>,  
Adriane Feijó Evangelista<sup>1</sup>, Marcus Matsushita<sup>2</sup>, Gustavo Nóriz Berardinelli<sup>1</sup>,  
Marco Antonio de Oliveira<sup>1</sup>, Rui Manuel Reis<sup>1,3,4</sup> & Denise Peixoto Guimarães<sup>1,5</sup>Received: 10 May 2019  
Accepted: 28 August 2019  
Published online: 23 September 2019

The molecular basis of colorectal cancer (CRC) can guide patient prognosis and therapy. In Brazil, knowledge on the CRC mutation landscape is limited. Here, we investigated the mutation profile of 150 cancer-related genes by next-generation sequencing and associated with microsatellite instability (MSI) and genetic ancestry in a series of 91 Brazilian CRC patients. Driver mutations were found in the *APC* (71.4%), *TP53* (56.0%), *KRAS* (52.7%), *PIK3CA* (15.4%) and *FBXW7* (10.9%) genes. Overall, genes in the MAPK/ERK, PIK3/AKT, NOTCH and receptor tyrosine kinase signaling pathways were mutated in 68.0%, 23.1%, 16.5%, and 15.3% of patients, respectively. MSI was found in 13.3% of tumors, most of which were proximal (52.4%,  $P < 0.001$ ) and had a high mutation burden. European genetic ancestry was predominant (median of 83.1%), followed by Native American (4.1%), Asian (3.4%) and African (3.2%). *NF1* and *BRAF* mutations were associated with African ancestry, while *TP53* and *PIK3CA* mutations were inversely correlated with Native American ancestry. Our study suggests that Brazilian CRC patients exhibit a mutation profile similar to other populations and identify the most frequently mutated genes, which could be useful in future target therapies and molecular cancer screening strategies.

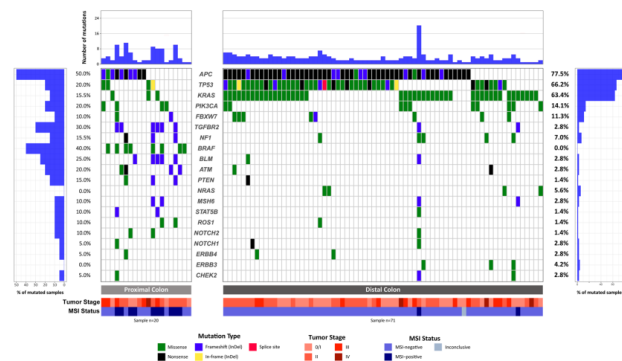
Colorectal cancer (CRC) is the third most common type of cancer worldwide<sup>1,2</sup>. According to GLOBOCAN, over 1.8 million new colorectal cancer cases occurred in 2018. Moreover, the disease is the second cause of death by cancer worldwide<sup>3</sup>. Increases in both incidence and mortality have been observed over the last 10 years in several Europe, Latin America, and Asia countries<sup>4</sup>. The most significant increases in incidence were observed in Brazil, Costa Rica and Bulgaria<sup>5</sup>. In Brazil, the Brazilian National Cancer Institute (INCA) estimated that over 36,000 new cases were expected for 2018, ranking CRC as the third most frequent cancer among men and as the second most frequent cancer among women<sup>6</sup>.

CRC is more frequently observed in the distal colon (left colon, from splenic flexure to rectum) than in the proximal side (right colon, from the cecum to transverse colon)<sup>7</sup>. In addition to incidence differences, the tumors arising from the left and right colon are distinct in their epidemiology, biology, histology and microbial diversity<sup>4,7</sup>. Consequently, it also influences patients' prognosis<sup>8,9</sup>.

Several lifestyles are risk factors associated with CRC, such as red and processed meat consumption, alcohol intake, smoking and body weight<sup>8</sup>. Moreover, the cumulative acquisition of genetic alterations leads to a progressive tumorigenesis process from normal to precursor lesion, culminating in a malignant tumor<sup>10</sup>. The majority (80–85%) of CRC cases are sporadic and emerge from somatic alterations in driver genes<sup>10</sup>. These alterations are linked to three main molecular groups: chromosomal instability, mismatch repair defect and methylator phenotype<sup>10</sup>. In addition to these classic features, the CRC Subtyping Consortium classified four consensus molecular subtypes (CMS) of CRC: CMS1, which are enriched for tumors with MSI, overexpression of DNA damage repair proteins, high immune activation and widespread hypermethylation; CMS2, which have tumors with chromosomal instability and activation of the WNT and MYC signaling pathways; CMS3, with tumors with epithelial and metabolic dysregulation; and CMS4, which have tumors with transforming growth factor- $\beta$  activation, stromal infiltration and angiogenesis<sup>11</sup>. The molecular profile of CRCs is distinct between proximal and distal tumors. Proximal CRCs are associated with microsatellite instability (MSI) and activating mutations on the *BRAF* gene, frequently harboring a high mutation burden, whereas distal CRCs are associated with chromosomal instability

<sup>1</sup>Molecular Oncology Research Center, Barretos Cancer Hospital, Barretos, Brazil. <sup>2</sup>Department of Pathology, Barretos Cancer Hospital, Barretos, Brazil. <sup>3</sup>Life and Health Sciences Research Institute (ICVS), Medical School, University of Minho, Braga, 4710-057, Portugal. <sup>4</sup>ICVS/3B's-PT Government Associate Laboratory, Braga, 4710-057, Portugal. <sup>5</sup>Department of Endoscopy, Barretos Cancer Hospital, Barretos, Brazil. Correspondence and requests for materials should be addressed to R.M.R. (email: [ruielis.hcb@gmail.com](mailto:ruielis.hcb@gmail.com)) or D.P.G. (email: [guimaraes.dp@gmail.com](mailto:guimaraes.dp@gmail.com))





**Figure 1.** Waterfall plot (oncoplot) of the distribution of mutations found in our tumor samples. The upper plot shows the frequency of mutation for each tumor sample. Left and right plots show the frequency of samples mutated for proximal colon (left) and distal colon (right). The central plot shows the types of mutations in each tumor sample (when the sample showed more than one mutation in the same gene, only the most deleterious type is shown). The lower part of the figure shows the tumor stage at diagnosis and microsatellite instability status (MSI) of each sample. Tumor stage: stage of the disease at the diagnosis.

and are more frequently non-hypermethylated and have less frequently the MSI phenotype<sup>67</sup>. In line with this, proximal CRCs are enriched for CMS1 tumors and, inversely, distal CRC are enriched for CMS2 tumors<sup>68</sup>.

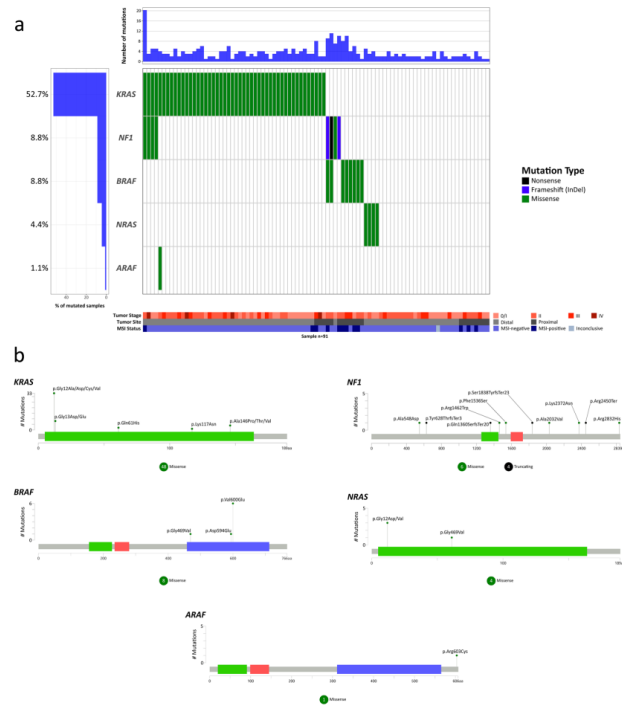
Molecular characterization is crucial for CRC patient care and prognostic and therapeutic assessment<sup>11–13</sup>. Knowledge of the mutation profile of Brazilian CRC patients is limited. Therefore, the present study aimed to interrogate the mutational profile of CRCs by next-generation sequencing mutation analysis of a large panel of 150 cancer-related genes and to identify significant cancer drivers. Moreover, the molecular profile was associated with clinicopathological features and genetic ancestry markers of CRC patients.

## Results

**Description of the mutation profile.** We sequenced 150 cancer-related genes in 91 colorectal tumors. The mean read depth of sequencing was 625.7x per gene and 60.7x per variant. Overall, after filtering out non-driver variants (see Material and Methods), we identified at least one somatic variant for each tumor sample with a mean of 3.6 mutations per patient (range 1–20). Thirteen tumors (14.3%) harbored mutations in single genes, 26 tumors (28.6%) in two genes, 24 tumors (26.4%) in three genes, and 28 tumors (30.8%) harbored alterations in four or more genes. Among the 150 genes sequenced, 46 showed at least one driver somatic mutation, including missense, frameshift, nonsense, in-frame, and splice mutations (54.9%, 22.6%, 21.3%, 0.9%, 0.3%, respectively). The most recurrently altered genes in our study population are shown in Fig. 1 (for a complete list of variants identified, see Supplementary Table S1). *APC* was the most affected gene, altered in 65 tumor samples (71.4%), followed by *TP53* (51 cases, 56.0%), *KRAS* (48 cases, 52.7%), *PIK3CA* (14 cases, 15.4%) and *FBXW7* (10 cases, 10.9%, Fig. 1). Below, we further describe the most affected genes and pathways.

***APC* gene.** A total of 86 somatic mutations in *APC* were found in 65 tumors (71.4%), among which 44 harbored a single mutation, and 21 tumors harbored two mutations in the gene. The majority of the mutations were protein truncating: 52 nonsense (60.5%), 30 frameshift (34.8%), and 4 missense (4.7%) (Supplementary Fig. S1 and Supplementary Table S1). Among the mutations found, 68 were previously reported as somatic mutations in CRC (50 nonsense and 14 frameshift), and 22 were new mutations (16 frameshift, 4 missense and 2 nonsense mutations). The mutations found are spread widely throughout the coding region of the *APC* gene, and the most affected site was the protein position 1556, which showed frameshift mutations in 6 tumors (Supplementary Fig. S1). In addition, nine *APC* mutated cases were selected and PCR-sequenced by direct Sanger sequencing was performed, validating all mutations (Supplementary Fig. S2).

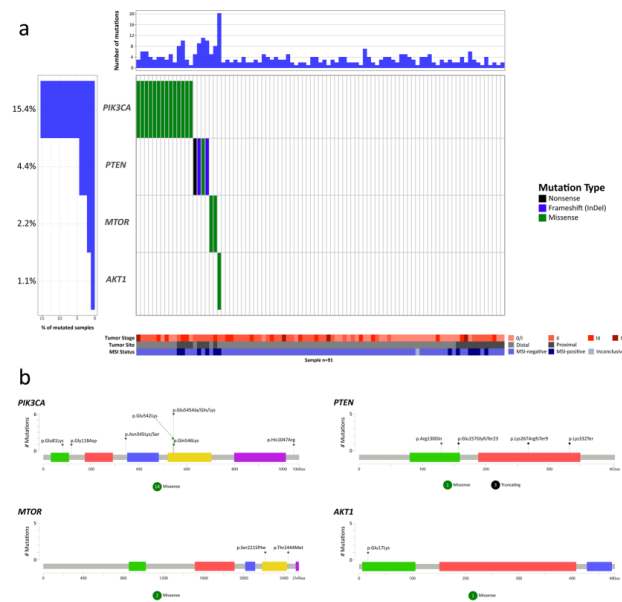
***TP53* gene.** Fifty-six mutations on *TP53* were found affecting 51 tumors (56.0% of the tumors, Fig. 1 and Supplementary Fig. S1), among which 46 harbored a single mutation on the gene and 5 harbored two mutations. The majority of alterations in *TP53* were missense mutations (40, 71.4%), followed by nonsense (7 mutations, 12.5%), frameshift (5 mutations, 8.9%), in-frame (3 mutations, 5.4%) and splice region (1 mutation, 1.8%) (Supplementary Fig. S1 and Supplementary Table S1). Among the mutations found, 53 were previously reported as somatic mutations in CRC (40 missense, 7 nonsense, 3 frameshift, 2 in-frame, and 1 splice region mutations),



**Figure 2.** Waterfall (oncoplot) plot (a) and mutations mapper (b) of MAPK-ERK signaling pathway genes. Tumor stage: stage of the disease at diagnosis. MSI: microsatellite instability.

and 3 were new mutations, including 2 frameshift and 1 in-frame mutation. Most mutations were localized on the DNA-binding domain (only one mutation was found outside this domain), and the most frequently altered site was protein position 273, with 6 missense mutations (Supplementary Fig. S1). In addition, nine *TP53* mutated cases were selected and PCR- followed by direct Sanger sequencing was performed, validating all mutations (Supplementary Fig S2).

**MAPK-ERK signaling pathway alterations.** The MAPK-ERK signaling pathway had at least one mutated gene in 62 tumors (68%). The most frequently altered gene in the MAPK/ERK pathway was *KRAS*, which showed 48 missense mutations, affecting 52.7% of the tumors. All *KRAS* mutations affected hotspots of the gene at protein sites 12 (33 mutations, 68.7% of the *KRAS* mutations), 13 (8 mutations, 16.7%), 146 (4 mutations, 8.3%), 61 (2 mutation, 4.2%), and 117 (1 mutation, 2.1%) (Fig. 2, and Supplementary Table S1). All mutations on codons 12/13 and codon 61 of the *KRAS* gene were confirmed by the cobas® *KRAS* mutation test. The *NF1* gene showed 10 somatic mutations in 8 tumor samples (8.8%), three of which were already known mutations (2 missense and 1 nonsense mutations), four were novel missense mutations, and three were novel frameshift mutations. The mutations were spread throughout the coding region of the gene. All truncating protein mutations in *NF1* were mutually exclusive with *KRAS*, and four tumors had missense mutations on both *KRAS* and *NF1* (Fig. 2a). The *BRAF* gene was mutated in 8 tumor samples (8.8%), showing 8 missense somatic mutations mutually exclusive with *KRAS*, all previously reported as somatic mutations in CRC. The p.Val600Glu mutation

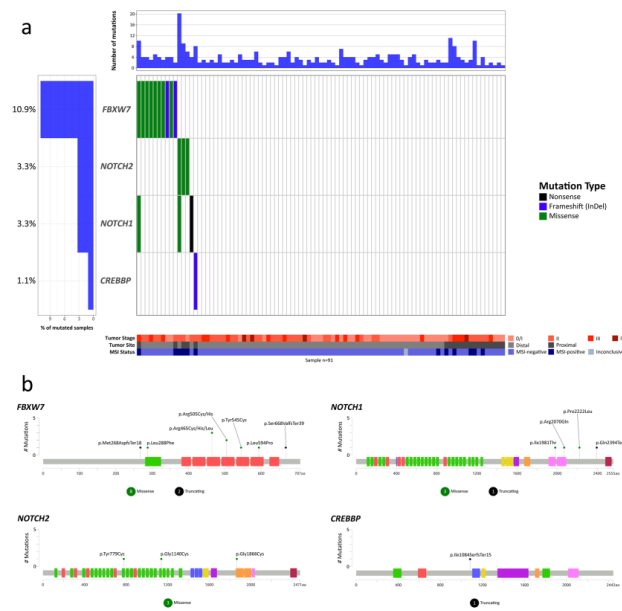


**Figure 3.** Waterfall (oncoplot) plot (a) and mutation mapper (b) of PI3K-AKT signaling pathway genes. Tumor stage: stage of the disease at diagnosis. MSI: microsatellite instability.

(p.V600E) was the most prevalent mutation, occurring in 6 tumors (75% of all mutations on the *BRAF* gene). The two variants, different from V600E mutations, also affected the kinase domain of the *BRAF* protein (p.Asp594Glu and p.Gly469Val). Mutations in the *NRAS* gene were found in 4 tumor samples (4.4%), 3 at codon 12 (75%) and one at codon 61. *NRAS* mutations were mutually exclusive with *KRAS*, *NF1* and *BRAF* mutations. Finally, the *ARAF* gene had one previously known missense mutation that was located outside of the main domains of the protein (Fig. 2 and Supplementary Table S1).

**PI3K-AKT signaling pathway alterations.** Genes in the PI3K-AKT signaling pathway were altered in 21 tumors (23.1%). The most frequently altered gene in the PI3K-AKT signaling pathway was *PIK3CA*, with alterations in 14 tumors (15.4%, Fig. 3a), all harboring already known somatic mutations. The PIK domain of the protein harbored most of the mutations found in *PIK3CA*, 6 mutations at codon 545 (42.9% of all *PIK3CA* mutations), 2 mutations at codon 542 (14.3%), and 1 mutation at codon 546 (7.1%). The p85-binding domain of the protein harbored 1 mutation (7.1%), and the other 4 (28.6%) mutations affected sites outside the main domains of the protein (Fig. 3a). The *PTEN* gene was mutated in 4.4% of the tumors (Fig. 3), showing one previously reported missense mutation and one frameshift mutation, along with two novel frameshift and nonsense mutations. The *MTOR* gene was mutated in 2 tumors (2.2%, Fig. 3), showing two previously described missense mutations in CRC, one affecting the PI3K/PI4K domain and one outside the main domains of the protein. *AKT1* was mutated in 1 tumor (1.1%, Fig. 3) and showed one already known missense mutation affecting the Pleckstrin homology (PH) domain of the protein.

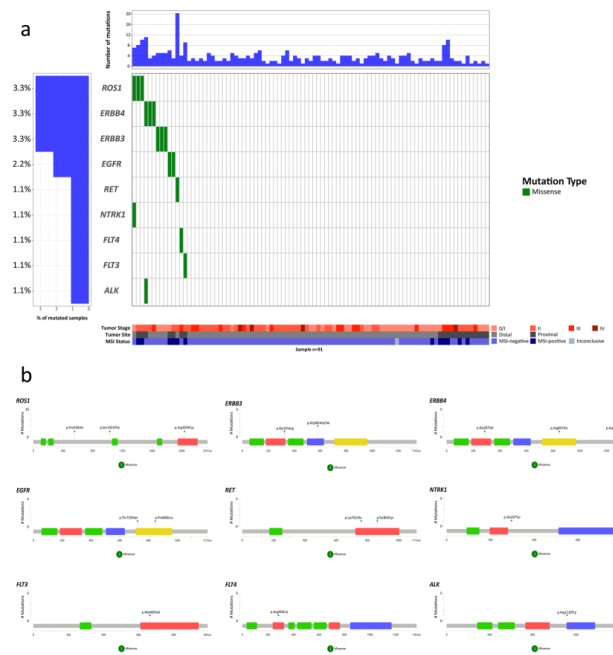
**Notch signaling pathway alterations.** The Notch signaling pathway was altered in 15 tumors (16.5%). *FBXW7* was the most affected gene in this pathway, mutated in 10 tumors (10.9%, Fig. 4), showing nine already known mutations (7 missense and 2 frameshift mutations) and one novel somatic missense mutation. Most of the *FBXW7* mutations (7, 70%) affected one of the repeated beta-transducin domains. One mutation (10%) affected the F-box-like domain, and 2 mutations were localized outside the main domains of the protein (Fig. 4b). The



**Figure 4.** Waterfall (oncoplot) plot (a) and mutation mapper (b) of Notch signaling pathway genes. Tumor stage: stage of the disease at diagnosis. MSI: microsatellite instability.

*NOTCH2* gene was altered in 3 tumors (3.3%, Fig. 4), all harboring novel missense mutations located at one of the repeats of the EGF-like domain (2 mutations) and one at the ankyrin repeat. The *NOTCH1* gene was altered in three tumors (3.3%, Fig. 4) with 4 mutations, three already known mutations (2 missense and one nonsense mutation) and one novel missense mutation. Two *NOTCH1* mutations were localized at one of the ankyrin repeats and two outside of the main domains of the protein. Finally, the *CREBBP* gene was altered in 1 tumor (1.1%, Fig. 4), showing one already known frameshift mutation localized outside the main domains of the protein.

**Receptor tyrosine kinase alterations.** The receptor tyrosine kinase (RTK) family showed 9 altered genes (Fig. 5). *ROS1* was altered in 3 tumors (3.3%, Fig. 5), showing one previously described missense mutation in the protein tyrosine kinase domain and two novel missense mutations outside the main domains of the protein. The *ERBB4* gene was altered in 3 tumors (3.3%, Fig. 5), showing one novel missense mutation in the furin-like cysteine rich region of the protein, one previously described missense mutation outside the main domains of the protein. *ERBB3* was altered in 3 tumors (3.3%, Fig. 5), showing 1 already known missense mutation in the furin-like cysteine rich region of the protein and 2 novel missense mutations in growth factor receptor domain IV. The *EGFR* gene was altered in 2 tumors (2.2%, Fig. 5), showing 2 known missense mutations in the protein tyrosine kinase domain. The *RET* gene was mutated in 1 tumor (1.1%, Fig. 5), showing two novel mutations in the protein tyrosine kinase domain. The *NTRK1* gene was mutated in 1 tumor (1.1%, Fig. 5), showing one novel missense mutation outside the main domains of the protein. The *FLT3* gene was mutated in 1 tumor (1.1%, Fig. 5), showing one novel missense mutation in the protein tyrosine kinase domain. The *FLT4* gene was altered in 1 tumor (1.1%, Fig. 5), showing one already known missense mutation in the immunoglobulin I-set domain. Finally, the *ALK* gene was also altered in 1 tumor (1.1%, Fig. 5) and showed one previously described missense mutation in the protein tyrosine kinase domain.



**Figure 5.** Waterfall (oncoplot) plot (a) and mutation mapper of RTK genes. Tumor stage: stage of the disease at diagnosis. MSI: microsatellite instability.

**MSI status and its association with tumor mutation burden.** We assessed the MSI status of tumor samples by multiplex PCR. Seventy-eight did not exhibit microsatellite instability (MSI-negative, 86.7%), and twelve tumors showed microsatellite instability (MSI-positive, 13.3%). The mean mutation burden of tumors that were MSI-positive was higher than MSI-negative tumors ( $7.1 \pm 1.4$  and  $3.1 \pm 0.2$ , respectively,  $P = 0.002$ , Supplementary Fig. S3). Tumors with MSI-positive status also showed a higher frequency of frameshift mutations (49.4% versus 13.2% in MSI-negative tumors,  $P < 0.001$ ) and a lower frequency of missense (43.5% vs 59.1%,  $P = 0.01$ ) and nonsense mutations (5.9% versus 26.4%,  $P < 0.001$ , Supplementary Fig. S3).

The profile of somatic mutations between tumors with and without MSI was also distinct. Somatic mutations of mismatch repair (MMR) genes were found only in MSI-positive tumors, which showed four frameshift mutations on *MSH6* and one nonsense in *MLH1*. Likewise, *TGFBR2*, *NOTCH2*, and *STAT5B* genes were exclusively mutated in MSI-positive tumors ( $P < 0.001$ ,  $P = 0.002$  and  $P = 0.002$ , respectively, Table 1). Other genes also showed mutations only in MSI-positive tumors; however, there was no significant difference (Supplementary Table S2). MSI-positive tumors also had higher rates of *BRAF* mutations (41.7% vs 3.8%,  $P < 0.001$ ), *ATM* mutations (25% vs 3.8%,  $P = 0.029$ ), *BLM* mutations (50% vs 1.3%,  $P < 0.001$ ), *NOTCH1* mutations (16.7% vs 1.3%,  $P = 0.046$ ), *CHEK2* mutations (16.7% vs 1.3%,  $P = 0.046$ ) and *ROS1* mutations (16.7% vs 0.3%,  $P = 0.046$ , Table 1).

In contrast, MSI-negative tumors showed a higher frequency of *APC* (75.6% vs 41.7%,  $p = 0.034$ ) and *TP53* mutations (65.4% vs 8.3%,  $p < 0.001$ , Table 1), and some genes were exclusively altered in MSI-negative tumors, but without statistical significance (Supplementary Table S2).

**Differences between proximal and distal colon.** Of the 91 tumor samples sequenced, 71 (78%) were from the distal colon, and 20 (22%) were from the proximal colon. The proximal colon had a higher frequency of

	Microsatellite Instability				Tumor Site					
	MSI negative (N=78)		MSI positive (N=12)		P value	Proximal Colon (N=20)		Distal Colon (N=71)		P value
	n	(%)	n	(%)		n	(%)	n	(%)	
APC	59	(75.6)	5	(41.7)	0.034**	10	(50.0)	55	(77.5)	0.025*
TP53	51	(65.4)	1	(8.3)	<0.001*	4	(20.0)	48	(67.6)	<0.001*
KRAS	45	(57.7)	3	(25.0)	0.059*	3	(15.0)	45	(63.4)	<0.001*
PIK3CA	12	(15.4)	2	(16.7)	0.999**	4	(20.0)	10	(14.1)	0.499**
FBXW7	8	(10.3)	2	(16.7)	0.617**	2	(10.0)	8	(11.3)	0.999**
NFI	5	(6.4)	3	(25.0)	0.070**	3	(15.0)	5	(7.0)	0.367**
TGFB2	0	(0.0)	8	(66.7)	<0.001**	6	(30.0)	2	(2.8)	0.001**
ATM	3	(3.8)	3	(25.0)	0.029**	4	(20.0)	2	(2.8)	0.020**
BRAF	3	(3.8)	5	(41.7)	0.001**	8	(40.0)	0	(0.0)	<0.001**
BLM	1	(1.3)	6	(50.0)	<0.001**	5	(25.0)	2	(2.8)	0.005**
MSH6	0	(0.0)	4	(33.3)	<0.001**	2	(10.0)	2	(2.8)	0.209**
NOTCH1	1	(1.3)	2	(16.7)	0.046**	1	(5.0)	2	(2.8)	0.530**
NRAS	4	(5.1)	0	(0.0)	0.999**	0	(0.0)	4	(5.6)	0.572**
PTEN	2	(2.6)	2	(16.7)	0.084**	3	(15.0)	1	(1.4)	0.032**
CHEK2	1	(1.3)	2	(16.7)	0.046**	1	(5.0)	2	(2.8)	0.530**
ERBB3	3	(3.8)	0	(0.0)	0.999**	0	(0.0)	3	(4.2)	0.999**
ERBB4	3	(3.8)	0	(0.0)	0.999**	1	(5.0)	2	(2.8)	0.530**
NOTCH2	0	(0.0)	3	(25.0)	0.002**	2	(10.0)	1	(1.4)	0.120**
ROS1	1	(1.3)	2	(16.7)	0.046**	2	(10.0)	1	(1.4)	0.120**
STAT5B	0	(0.0)	3	(25.0)	0.002**	2	(10.0)	1	(1.4)	0.120**

**Table 1.** Mutation profile of the most frequently altered genes in our study population by microsatellite instability status and tumor site. \*Chi-squared test; \*\*Fisher Exact Test; MSI: microsatellite instability.

MSI-positive tumors than the distal colon (50% versus 2.9%,  $P < 0.001$ ). We found significant differences between the mutation profile of tumors from the proximal and distal colon (Fig. 1 and Table 1). Proximal colon tumors showed higher mutation frequencies than distal tumors on *TGFB2* (30.0% vs 2.8%,  $P < 0.001$ ), *ATM* (20.0% vs 2.8%,  $P = 0.020$ ), *BLM* (25.0% vs 2.8%,  $P = 0.005$ ), and *PTEN* (15.0% vs 1.4%,  $P = 0.032$ ) genes. *BRAF* mutations were exclusively found in tumors from the proximal colon (40.0% vs 0.0%,  $P < 0.001$ ), as well as mutations in the *EGFR*, *JAK1*, *MAP2K4*, *MAP3K1*, *NBN*, *ALK* genes (each of these mutated in 2 tumors), and *BRCA1*, *BRIP1*, *CREBBP*, *FLT3*, *FLT4*, *MLH1*, *MPL* and *NTSC2* genes (each of these mutated in 1 tumor).

In contrast, the distal colon showed a higher frequency than the proximal colon of *APC* (77.5% vs 50.0%,  $P = 0.025$ ), *TP53* (67.6% vs 20.0%,  $P < 0.001$ ) and *KRAS* (63.4% vs 15.0%,  $P < 0.001$ ) gene mutations (Table 1). Mutations in *NRAS* (mutated in 4 tumors), *ERBB3* (mutated in 3 tumors), *RET*, *AKT1*, *AR*, *ARAF*, *BARD1*, *CDKN2A*, *GATA2*, *IDH2*, *MRE11A*, *NTRK1*, *PALB2* and *SMO* genes (each of these mutated in 1 tumor) were exclusively found in the distal colon.

**Mutation profile and genetic ancestry markers.** We further assessed the ancestry of the patients by a panel of ancestry informative markers (AIMs) to estimate the proportion of ancestral background. The proportion of European genetic ancestry was predominant in our study population (median 83.1%, interquartile range 76.6% and 87.8%), followed by Native American (median 4.1%, interquartile range 2.9% and 7.2%), Asian (median 3.4%, interquartile range 2.8% and 5.1%) and African (median 3.2%, interquartile range 2.1% and 7.9%) (Supplementary Fig. S4).

Patients with the highest proportion of African ancestry (those with African ancestry proportions higher than 7.9%) had more frequent *NFI* and *BRAF* mutations when compared with patients with intermediate and lowest proportions of African ancestry ( $P = 0.045$  and  $0.031$ , respectively, Supplementary Table S3). Patients with the highest proportion of Asian ancestry (those with a proportion of Asian ancestry higher than 5.1%) also had more *NFI* mutations than those patients with intermediate and low Asian ancestry ( $P = 0.037$ , Supplementary Table S3). Patients with the highest proportion of Native American ancestry (those with a proportion of Native American ancestry higher than 7.2%) had fewer *TP53* and *PIK3CA* mutations than those patients with the intermediate and the lowest proportion of Native American ancestry ( $P = 0.002$  and  $0.024$ , respectively, Supplementary Table S3).

## Discussion

In this study, we investigated mutations within 150 selected oncogenes and tumor suppressor genes in a Brazilian cohort. To the best of our knowledge, this is the largest mutation profile study in this population. Overall, our results showed that our population had a similar mutation profile to other populations (Table 2), with alterations in canonical driver genes, such as *APC*, *KRAS*, *TP53*, *PIK3CA*, and *FBXW7*. The differences between the proximal and distal colon were also highlighted in Brazilian patients, and the admixed features of our population were revealed by ancestry markers.

Gene	Present Study (n=91)		TCGA (n=224)		DFCI (n=619)		Genetech (n=72)		ICGC: COCA-CN (321)		Salem <i>et al.</i> (2413)	
	N	(%)	N	(%)	N	(%)	N	(%)	N	(%)	N	(%)
APC	65	(71.4)	168	(75.0)	361	(58.3)	29	(40.3)	173	(53.9)	1513	(62.7)
TP53	51	(56.0)	121	(54.0)	320	(51.7)	40	(55.6)	158	(49.2)	1641	(68.0)
KRAS	48	(52.7)	94	(42.0)	173	(27.9)	37	(51.4)	118	(36.8)	1156	(47.9)
PIK3CA	14	(15.4)	45	(20.1)	132	(21.3)	22	(30.6)	60	(18.7)	326	(13.5)
FBXW7	10	(11.0)	37	(16.5)	86	(13.9)	15	(20.8)	59	(18.4)	—	—
BRAF	8	(8.8)	21	(9.4)	127	(20.5)	6	(8.3)	51	(15.9)	168	(7.0)
NFI	8	(8.8)	8	(3.6)	45	(7.3)	4	(5.6)	66	(20.6)	—	—
TGFBR2	8	(8.8)	23	(10.3)	29	(4.7)	2	(2.8)	25	(7.8)	—	—
BLM	7	(7.7)	10	(4.5)	12	(1.9)	0	(0.0)	25	(7.8)	—	—
ATM	6	(6.6)	25	(11.2)	64	(10.3)	13	(18.1)	53	(16.5)	34	(1.4)
MSH6	4	(4.4)	15	(6.7)	27	(4.4)	3	(4.2)	25	(7.8)	—	—
NRAS	4	(4.4)	20	(8.9)	27	(4.4)	2	(2.8)	17	(5.3)	—	—
PTEN	4	(4.4)	8	(3.6)	51	(8.2)	4	(5.6)	30	(9.3)	37	(1.5)
ERBB3	3	(3.3)	14	(6.3)	36	(5.8)	6	(8.3)	31	(9.7)	—	—
ERBB4	3	(3.3)	19	(8.5)	37	(6.0)	8	(11.1)	46	(14.3)	—	—
NOTCH1	3	(3.3)	0	(0.0)	65	(10.5)	6	(8.3)	38	(11.8)	—	—
CHEK2	3	(3.3)	1	(0.4)	36	(5.8)	5	(6.9)	22	(6.9)	—	—
NOTCH2	3	(3.3)	12	(5.4)	25	(4.0)	4	(5.6)	63	(19.6)	—	—
ROSI	3	(3.3)	13	(5.8)	31	(5.0)	2	(2.8)	46	(14.3)	—	—
STAT5B	3	(3.3)	3	(1.3)	15	(2.4)	3	(4.2)	12	(3.7)	—	—
BRCA2	2	(2.2)	10	(4.5)	42	(6.8)	5	(6.9)	35	(10.9)	—	—
BRD2	2	(2.2)	0	(0.0)	0	(0.0)	0	(0.0)	16	(5.0)	—	—
EGFR	2	(2.2)	10	(4.5)	28	(4.5)	4	(5.6)	53	(16.5)	—	—
JAK1	2	(2.2)	4	(1.8)	19	(3.1)	2	(2.8)	25	(7.8)	—	—
MAP2K4	2	(2.2)	11	(4.9)	13	(2.1)	2	(2.8)	16	(5.0)	—	—
MAP3K1	2	(2.2)	5	(2.2)	26	(4.2)	1	(1.4)	28	(8.7)	—	—
MTOR	2	(2.2)	17	(7.6)	50	(8.1)	8	(11.1)	42	(13.1)	—	—
NBN	2	(2.2)	2	(0.9)	20	(3.2)	1	(1.4)	26	(8.1)	—	—

**Table 2.** Comparison of mutation frequency across different public data (cBioPortal) and studies. TCGA: Comprehensive molecular characterization of human colon and rectal cancer<sup>14</sup>; DFCI: Genomic Correlates of Immune-Cell Infiltrates in Colorectal Carcinoma<sup>15</sup>; Genetech: Recurrent R-spondin fusions in colon cancer<sup>16</sup>; ICGC: COCA-CN: Colorectal cancer from non-Western samples (China)<sup>17</sup>; Salem *et al.*: Comparative molecular analyses of left-sided colon, right-sided colon, and rectal cancers<sup>18</sup>.

APC mutations are a common event in CRC, being an early and a general event for tumorigenesis progression<sup>14</sup>. Here, we showed that APC was the most frequently mutated gene, similar to several other studies demonstrated in different populations<sup>7,15–17</sup> (Table 2). Additionally, as previously reported, we found that APC mutations were more frequent in distal colon and MSI-negative tumors<sup>8,9</sup>.

TP53 is a ubiquitous driver gene in several cancer types<sup>19</sup>, and one of the most frequently altered genes in colorectal cancer, ranging from 51.7 to 68% of the cases<sup>7,18,19,20</sup> (Table 2). Our study population showed a similar frequency of TP53 gene alterations when compared with the TCGA frequency and studies in other populations<sup>7,21</sup>, showing the highest frequency in non-hypermutated tumors (MSI-negative tumors) and in distal colon tumors<sup>7,17,22,23</sup>.

The MAPK-ERK signaling pathway is significantly altered in CRCs and other tumors<sup>18</sup>. This pathway is constitutively activated by alterations in several driver genes, including KRAS, BRAF, and NRAS<sup>7</sup>. We observed that 68% of cases had a mutation in at least one gene in the MAPK-ERK pathway, in accordance with TCGA data (59.3%)<sup>7,17</sup>. Individually, we observed a high frequency of KRAS mutations (52.7%) that was slightly higher than the one reported in TCGA data (42%) and those found in other populations<sup>15,17,21,24</sup> (Table 2). Moreover, previous studies on the prevalence of KRAS mutations carried out in Brazilian populations showed a lower frequency (31.9–49.2%) of mutations<sup>25,26</sup>. The higher frequency observed in the present study in comparison to those observed in other Brazilian studies may be associated with the higher sensitivity and coverage of the NGS approach used in our study and the high prevalence of early-stage and localized disease cases included in our study (78% stages 0, I and II). In agreement with other studies, most mutations in the KRAS gene were found at protein residues 12 and 13, with predominant Gly12Asp mutations<sup>11,27</sup>. Epidemiological and molecular observations showed that the proximal and distal colon exhibited distinct features. KRAS mutations are slightly higher in the proximal colon<sup>20,28,29</sup>. However, in our study, a predominance of KRAS mutations was observed in tumors from the distal colon. We also found a high mutation frequency in the NFI gene (Table 2), and interestingly, this frequency was associated with a high proportion of Asian ancestry, corroborating the known higher frequency of NFI mutations in the Asian population<sup>30</sup>.

*BRAF* mutations are reported in less than 10% of cases (Table 2) and have been related to MSI-positive and proximal colon tumors in several studies<sup>6,7,17,20</sup>. Here, we confirmed these observations and found a similar overall frequency of *BRAF* mutations, including the predominance of activating V600E type mutation<sup>31</sup>. Some non-V600E mutations, such as mutations at the D594 site, were shown to have either impaired or complete loss of kinase activity<sup>32</sup>, although constitutive activation of MAPK-ERK by these mutations has been shown to coincide with other mechanisms of activation, such as *NFI* deletions<sup>33</sup>. Non-V600E, such as the G469V mutation found here, has high kinase activity but acts in a different way than the V600E mutation<sup>34</sup>. The *NRAS* gene showed a lower frequency of mutation compared to the TCGA data, but it was similar to other large-scale studies (Table 2). Moreover, the significant pattern of mutually exclusivity of *KRAS*, *BRAF* and *NRAS* mutations found here has been observed in several studies<sup>2,15,17,34</sup>.

The PI3K pathway is another common pathway altered in CRC, leading to cell survival when activated<sup>7,10,18</sup>. Data from TCGA showed that genes from the PI3K signaling pathway were altered in approximately 25.9% of patients<sup>7,17</sup>, similar to the frequency reported in our study (23.1%). The *PIK3CA* gene seems to be the principal gene mutated in this pathway in CRCs<sup>7,15</sup>, showing activating mutations mainly on the PI3Ka domain (helicase), in addition to other already known activating mutations, such as H1047R and N345K/S found in the present study<sup>7,35</sup>. The *PTEN* gene is known to revert the action of PI3Ks<sup>36</sup>, and its inactivation is observed at a lower frequency in CRCs<sup>7</sup> (1.5–9.3%, Table 2) by mutation<sup>7,37</sup>, as reported in our study (4.4%). The association between *PTEN* mutations and proximal colon localization has been observed previously<sup>37</sup>. Although *MTOR* was found to be mutated in CRCs with a range of 7.6–13.1% in large cohort studies (Table 2), these mutations are poorly characterized<sup>38</sup>. Mutations in *MTOR* site S2215, as we found here, seems to be the most frequently altered site across diverse cancer types and has been shown to increase the activity of mTOR protein<sup>38</sup>. The other *MTOR* (T2444M) mutation observed here has not been well characterized, but it affects the C-terminal region related to mTOR activation<sup>38</sup>. A low frequency of *AKT1* has been found across large cohort studies (Table 2), and the E17K mutation observed here has been related to the constitutive activation of *AKT1* protein<sup>39</sup>.

We found that 15.4% of tumors harboring missense mutations in members of the RTK family were lower than the frequency observed in TCGA data (22.3%)<sup>7,17</sup>. Although activating events in some RTK proteins can be observed by rearrangements in CRCs, such as *ROS1*<sup>40</sup>, *NTRK1*<sup>40</sup>, *ALK*<sup>40</sup>, and *RET*<sup>41</sup>, missense mutations in these genes are found across several studies (Table 2), but the oncogenic potential of such mutations is not completely understood in CRC. *RET* missense mutations promote tumor progression in thyroid cancer<sup>42</sup> and are shown to reduce apoptosis in colorectal cancer cell lines<sup>43</sup>. Interestingly, *ALK*, *ROS1*, and *NTRK1* missense mutations have been observed in solid tumors resistant to tyrosine receptor kinase inhibitors<sup>44–47</sup>.

We found known activating mutations present in the protein kinase domain *EGFR* T725M<sup>17,46</sup> and P848L<sup>47</sup>, as well as in *FLT3*<sup>48,49</sup>. Other mutations affecting the protein kinase domain of the *ROS1*, *ERBB4*, and *ALK* genes were also observed, but their role in oncogenesis is unknown<sup>47</sup>. Mutations that occur outside the protein kinase domain also play a role in the activation of RTKs, such as *ERBB3* mutations, which promote the growth of colonic cells<sup>48</sup> and *ERBB4* mutations, which promote an increase in the kinase activity and transformation ability in melanoma cells<sup>49</sup>.

The Notch signaling pathway is important in maintaining the proliferative compartment of intestinal crypts<sup>50</sup> and has been linked to tumor progression<sup>51</sup>. Here, we observed gene alterations in this pathway in 16.5% of the tumors, which was slightly lower when compared to the frequency observed on TCGA data for the same pathway (20.1%)<sup>7,17</sup>. *FBXW7* is one of the regulators of Notch protein<sup>52,53</sup>, and its loss seems to lead to the activation of this pathway<sup>54</sup>. The frequency of *FBXW7* mutations in our population was compared with those reported in other populations (Table 2)<sup>55</sup>. We observed mutations affecting hotspots of *FBXW7*, such as residues 465 and 505, that are important sites for substrate binding<sup>56</sup>, and frameshift mutations that lead to premature termination of the protein and have also been found affecting this gene<sup>7,15,17</sup>. The frequency of alterations in *NOTCH1* and *NOTCH2* genes is variable (Table 2), and missense activating mutations in both *NOTCH1* and *NOTCH2* have been observed in hematological tumors<sup>57,58</sup>. The pattern of mutual exclusivity between mutations in the *NOTCH2* and *FBXW7* genes that we found here has not been observed in large cohort studies<sup>7,15,17</sup>.

The presence of MSI was found in 13.3% of cases, in accordance with previous studies<sup>16,59</sup>. In addition, MSI was associated with tumors in the proximal colon and with those exhibiting a higher mutation burden<sup>67,19</sup>. As previously reported, we observed an association between *ATM*<sup>7,17</sup>, *BLM*<sup>7,17</sup>, and *TGFRB2*<sup>67,17</sup> mutations and MSI and the proximal colon.

Epidemiological evidence suggests that ethnicity has a high impact on cancer incidence and mortality<sup>60</sup>, and differences in the prevalence of cancer driver mutations have been observed in several cancer types, including CRC<sup>61</sup>. Our population is known to be diverse<sup>62</sup>, and our results are in agreement with these observations, showing differences in the proportions of each of the main four population origins compared with previous studies<sup>62,63</sup>. Our patients with a higher African ancestry proportion had a higher frequency of *BRAF* mutations, contrasting with a previous study by Staudacher *et al.*<sup>64</sup>. We did not observe an association between the frequency of *KRAS* mutations and the African ancestry that was reported previously<sup>64</sup>. These disparities may reside in the fact that, although we have an admixture population study, we have a low frequency of self-declared African ancestry and a low frequency of proportion of African ancestry than observed in other population studies<sup>62,63</sup>. In contrast, the association of *NFI* mutations and African ancestry has been observed previously<sup>65</sup>. Moreover, the observation that *NFI* mutations were associated with the high Asian ancestry proportion was consistent with the data from ICGC<sup>30</sup>. Although ethnicity association with *PIK3CA* mutations has been observed in different populations<sup>66–68</sup>, this is the first study to demonstrate an inverse correlation between *PIK3CA* mutations and a high Native American ancestry background.

Despite the major findings, one limitation of the present study lies in the preselection of a panel of known cancer genes, and known genes related to CRC carcinogenesis were not included in the analysis, such as *SMAD4* (altered in 11.6–12.9% of CRC), *TTN* (35.7–48.1%) and *SYNE1* (21.0–48.1%)<sup>3,45,15</sup>. Moreover, using whole-genome or exome approaches, other putative CRC cancer-related genes could have been identified in the Brazilian context.



Characteristics (N=91)	Frequency	
	Mean	(%) (range)
Age at diagnostic	61.18	(29–88)
Gender		
Male	49	53.8
Female	42	46.2
Self-Assessed Skin Color (n = 89)		
White	76	85.4
Black	2	2.2
Brown	10	11.2
Yellow	1	1.1
Primary disease site		
Right-sided	20	22.0
Cecum	4	20.0
Ascending colon	9	45.0
Transverse colon	7	35.0
Left-sided	71	78.0
Sigmoid colon	31	43.7
Rectosigmoid junction	15	21.1
Rectum	25	35.2
Stage (at diagnosis)		
0	3	3.3
I	30	33.0
II	38	41.8
III	16	17.6
IV	4	4.4
TNM		
Primary Tumor (T)		
Tis	5	5.5
T1	6	6.6
T2	28	30.8
T3	42	46.2
T4 (a,b)	10	11.0
Regional Lymph Nodes (N)		
N0	69	75.8
N1 (a,b,c)	15	16.5
N2 (a,b)	7	7.7
Distant Metastasis (M)		
M0	87	95.6
M1a	3	3.3
M1b	1	1.1
Histologic Grade		
I	10	11.0
II	77	84.6
III	4	4.4
Angiolymphatic invasion (n = 90)	17	18.9
Perineural invasion (n = 86)	6	7.0
Adjuvant chemotherapy	30	33.0
Radiotherapy (n = 90)	4	4.4
Tumor Recurrence (n = 89)	3	3.4
MSI (n = 90)		
MSI-negative	78	86.7
MSI-positive	12	13.3

**Table 3.** Characteristics of the study population.

In conclusion, our study constitutes the largest mutation landscape of Brazilian colorectal cancer patients. This study paves the way for a better comprehension of the major alterations identified and could guide better-tailored therapy for colorectal cancer in the Brazilian population.

## Materials and Methods

**Tissue samples.** Ninety-one patients diagnosed with colorectal adenocarcinomas and admitted for a surgical procedure at Barretos Cancer Hospital, Barretos, SP, Brazil, were evaluated. The main clinicopathological features are summarized in Table 3. The mean age of patients was 61 years old, and the majority of cases had early-stage and localized disease (0, I and II).

All included patients signed an informed consent form. Both tumors and blood were obtained and immediately processed and stored at  $-80^{\circ}\text{C}$  at Barretos Cancer Hospital Biobank. The present study was approved by the Barretos Cancer Hospital Institutional Review Board (Project n° 1060/2015, protocol CAAE: 51770115.6.0000.5437). All methods were performed in accordance with the relevant guidelines and regulations.

**DNA isolation.** Tumor DNA was isolated from fresh-frozen tissue using the QIAAsymphony DNA Mini Kit following the Tissue\_200 protocol for automated isolation in the QIAAsymphony (QIAGEN, Hilden, Germany). DNA from leucocytes of peripheral blood was isolated using the QIAamp DNA Blood Mini Kit (QIAGEN, Hilden, Germany) following the manufacturer's instructions. DNA quantity and quality were assessed by Qubit (Thermo Fisher Scientific, Waltham, MA, USA).

**Microsatellite instability analyses.** The MSI status was evaluated using a multiplex PCR comprising six quasi-monomorphic mononucleotide repeat markers (BAT-25, BAT-26, NR-21, NR-24, NR-27 and HSP110) as described previously<sup>49</sup>. The MSI analyses were performed using GeneMapper v4.1 software (Applied Biosystems), and the status was classified as stable (MSS) when none of the markers were unstable. The status was MSI-Low (MSI-L) when one of the markers was unstable and MSI-positive when two or more of the markers were unstable. MSS and MSI-L were considered MSI-negative<sup>49</sup>.

**Mutation profile.** The mutation profile of a commercial panel of 150 cancer-related genes (for a list of genes, see Supplementary Information) was conducted at the Mendelics Genetics company (São Paulo, SP, Brazil, (<https://www.mendelics.com/oncologia/>)). The panel analyzed all coding regions of the 150 cancer-related genes. For sequencing, paired tumor and blood DNA libraries were prepared using the Nextera Rapid Capture Custom Enrichment Kit (Illumina, San Diego, CA, USA). Libraries were quantified using a Qubit fluorometer (Thermo Fisher Scientific, Waltham, MA, USA), and their quality was evaluated using an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA). Cluster generation and sequencing were performed on the Illumina HiSeq 4000 following the manufacturer's instructions. Paired-end reads from Illumina sequencing were processed by script `bc12fastq` (v. 2.17.1.14) and aligned against the human genome reference build GRCh37 using Burrows-Wheeler Aligner (BWA, version 0.7.13)<sup>50</sup>. The somatic variants were called by a VarScan2 algorithm<sup>48</sup>. The variants with artifacts due to indel reads at their position or less than 10% or more than 90% of variant supporting reads on one strand were removed. The variants were further filtered to remove those with fewer than 10 reads covering the variant and with less than 5% variant allele frequency. The annotation of variants was performed using Ensembl Variant Effect Predictor (VEP)<sup>49</sup>.

To identify driver mutations on tumors, we used the Cancer Genome Interpreter – CGI<sup>70</sup>. After the classification of mutations by the CGI, we excluded mutations that were not classified as cancer driver mutations or not predicted as a driver by the OncoMut algorithm that CGI used. Therefore, mutations identified as polymorphisms (high allele frequency) or those predicted as neutral or passenger for oncogenesis, and those found in DNA sequences outside coding regions were excluded.

We further validated mutations on codons 12/13 and codon 61 of the *KRAS* gene using the cobas<sup>®</sup> KRAS Mutation Test (F. Hoffmann-La Roche, Basel, Switzerland) following the manufacturer's recommendations. We also performed PCR-sequencing followed by direct Sanger sequencing for a subset of samples to confirm mutations in the *APC* and *TP53* genes. The PCR and Sanger sequencing conditions were previously reported by our group<sup>71,72</sup>.

**Genetic ancestry markers.** The genetic ancestry of patients was determined by 46 autosomal ancestry informative markers (AIMs), consisting of insertion and deletion polymorphisms and using the genetic data of the Human Genome Diversity Project – center d'Etude du Polymorphisme Humain (HGDP-CEPH1), according to Pereira *et al.*<sup>43</sup>. We classified the proportion of ancestry by tertiles as high, intermediary and low for each of the main populations: African, European, Asian and Native American<sup>73</sup>.

**Statistical analyses.** Chi-square and Fisher's exact tests were used to set differences between mutated and nonmutated genes and patient characteristics. Mutation types and mutation burden differences between MSI-positive and MSI-negative tumors were determined by chi-square and Mann-Whitney tests, respectively. All statistical analyses were performed using SPSS software (v.20) and R 3.5.0 software. Mutation mapper figures were generated using cBioPortal. Oncoplots and ancestry charts were generated using GenVisR and ggplot2 packages (versions 1.14 and 3.1.0, respectively) in R 3.5.0 software.

## Data Availability

Data that support the findings are available from Dr. Denise Peixoto Guimarães and Dr. Rui Manuel Reis and are not publicly available due patient personal information. However, these data are available upon reasonable request and with permission of the Dr. Rui Manuel Reis (Scientific and Executive Director of the Molecular Oncology Research Center, Barretos Cancer Hospital).

## References

1. Arnold, M. *et al.* Global patterns and trends in colorectal cancer incidence and mortality. *Gut* **66**, 683–691 (2017).
2. Bray, F. *et al.* Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA A J Clin. Oncol.* <https://doi.org/10.3322/caac.21492> (2018).
3. MINISTÉRIO DA SAÚDE: Instituto Nacional de Câncer José Alencar Gomes da Silva. *Estimativa 2018-Incidência de câncer no Brasil* (2017).
4. Missiaglia, E. *et al.* Distal and proximal colon cancers differ in terms of molecular, pathological, and clinical features. *Ann. Oncol.* **25**, 1995–2001 (2014).
5. Drewes, J. L., Housseau, F. & Sears, C. L. Sporadic colorectal cancer: Microbial contributors to disease prevention, development and therapy. *Br. J. Cancer* **115**, 273–280 (2016).
6. Guinney, J. *et al.* The consensus molecular subtypes of colorectal cancer. *Nat. Med.* **21**, 1350–1356 (2015).
7. Muzny, D. M. *et al.* Comprehensive molecular characterization of human colon and rectal cancer. *Nature* **487**, 330–337 (2012).
8. Benedix, F. *et al.* Comparison of 17,641 patients with right- and left-sided colon cancer: Differences in epidemiology, perioperative course, histology, and survival. *Dis. Colon Rectum* **53**, 57–64 (2010).
9. Song, M., Garrett, W. S. & Chan, A. T. Nutrients, Foods, and Colorectal Cancer Prevention. *Gastroenterology* **148**, 1244–1260.e16 (2015).
10. Kuipers, E. J. *et al.* Colorectal cancer. *Nat. Rev. Dis. Prim.* **1**, (2015).
11. Boeckx, N. *et al.* Prognostic and Predictive Value of RAS Gene Mutations in Colorectal Cancer: Moving Beyond KRAS Exon 2. *Drugs* **75**, 1739–1756 (2015).
12. Le, D. T. *et al.* PD-1 Blockade in Tumors with Mismatch-Repair Deficiency. *N. Engl. J. Med.* **372**, 2509–2520 (2015).
13. Sargent, D. J. *et al.* Defective mismatch repair as a predictive marker for lack of efficacy of fluorouracil-based adjuvant therapy in colon cancer. *J. Clin. Oncol.* **28**, 3219–3226 (2010).
14. Cross, W. *et al.* The evolutionary landscape of colorectal tumorigenesis. *Nat. Ecol. Evol.* **2**, 1661–1672 (2018).
15. Giannakis, M. *et al.* Genomic Correlates of Immune-Cell Infiltrates in Colorectal Carcinoma. *Cell Rep.* **15**, 857–865 (2016).
16. Liu, Y. *et al.* Comparative Molecular Analysis of Gastrointestinal Adenocarcinomas. *Cancer Cell* **33**, 721–735.e8 (2018).
17. Gao, J. *et al.* Integrative Analysis of Complex Cancer Genomics and Clinical Profiles Using the cBioPortal. *Sci. Signal.* **6**, pii–pii1 (2013).
18. Bailey, M. H. *et al.* Comprehensive Characterization of Cancer Driver Genes and Mutations. *Cell* **173**, 371–385.e18 (2018).
19. Seshagiri, S. *et al.* Recurrent R-spondin fusions in colon cancer. *Nature* **488**, 660–664 (2012).
20. Salem, M. E. *et al.* Comparative molecular analyses of left-sided colon, right-sided colon, and rectal cancers. *Oncotarget* **8**, 86356–86368 (2017).
21. Liu, Z. *et al.* The landscape of somatic mutation in sporadic Chinese colorectal cancer. *bioRxiv* **9**, 155671 (2018).
22. Russo, A. *et al.* The TP53 colorectal cancer international collaborative study on the prognostic and predictive significance of p53 mutation: Influence of tumor site, type of mutation, and adjuvant treatment. *J. Clin. Oncol.* **23**, 7518–7528 (2005).
23. Loree, J. M. *et al.* Classifying colorectal cancer by tumor location rather than sidedness highlights a continuum in mutation profiles and consensus molecular subtypes. *Clin. Cancer Res.* **24**, 1062–1072 (2018).
24. Al-Shamsi, H. O. *et al.* Molecular spectrum of KRAS, NRAS, BRAF, PIK3CA, TP53, and APC somatic gene mutations in Arab patients with colorectal cancer: Determination of frequency and distribution pattern. *J. Gastrointest. Oncol.* **7**, 882–902 (2016).
25. Yamane, L. S. *et al.* KRAS and BRAF mutations and MSI status in precursor lesions of colorectal cancer detected by colonoscopy. *Oncol. Rep.* **1419–1426**, <https://doi.org/10.3892/or.2014.3338> (2014).
26. Gil Ferreira, C. *et al.* KRAS mutations: variable incidences in a Brazilian cohort of 8,234 metastatic colorectal cancer patients. *BMC Gastroenterol.* **14**, 73 (2014).
27. Loupakis, F. *et al.* KRAS codon 61, 146 and BRAF mutations predict resistance to cetuximab plus irinotecan in KRAS codon 12 and 13 wild-type metastatic colorectal cancer. *Br. J. Cancer* **101**, 715–721 (2009).
28. Rosty, C. *et al.* Colorectal carcinomas with KRAS mutation are associated with distinctive morphological and molecular features. *Mod. Pathol.* **26**, 825–834 (2013).
29. Samowitz, W. S. *et al.* Relationship of Ki-ras Mutations in Colon Cancers to Tumor Location, Stage, and Survival: A Population-based Study Relationship of Ki-ras Mutations in Colon Cancers to Tumor Location. *Cancer Epidemiol. Biomarkers Prev.* **9**, 1193–1197 (2000).
30. Kasprzyk, A. *et al.* International Cancer Genome Consortium Data Portal—a one-stop shop for cancer genomics data. *Database* **2011**, bar026–bar026 (2011).
31. Davies, H. *et al.* Mutations of the BRAF gene in human cancer. *Nature* **417**, 949–954 (2002).
32. Wan, P. T. C. *et al.* Mechanism of activation of the RAF-ERK signaling pathway by oncogenic mutations of B-RAF. *Cell* **116**, 855–867 (2004).
33. Yao, Z. *et al.* Tumours with class 3 BRAF mutants are sensitive to the inhibition of activated RAS. *Nature* **548**, 234–238 (2017).
34. Malapelle, U. *et al.* Less frequently mutated genes in colorectal cancer: evidences from next-generation sequencing of 653 routine cases. *J. Clin. Pathol.* **69**, 767–771 (2016).
35. Rosty, C. *et al.* PIK3CA Activating Mutation in Colorectal Carcinoma: Associations with Molecular Features and Survival. *PLoS One* **8**, 1–9 (2013).
36. Vivanco, I. & Sawyers, C. L. The phosphatidylinositol 3-Kinase-AKT pathway in human cancer. *Nat. Rev. Cancer* **2**, 489–501 (2002).
37. Rodon, J., Dienstmann, R., Serra, V. & Tabernero, J. Development of PI3K inhibitors: lessons learned from early clinical trials. *Nat. Rev. Clin. Oncol.* **10**, 143–153 (2013).
38. Grabiner, B. C. *et al.* A diverse array of cancer-associated MTOR mutations are hyperactivating and can predict rapamycin sensitivity. *Cancer Discov.* **4**, 554–563 (2014).
39. Carpten, J. D. *et al.* A transforming mutation in the pleckstrin homology domain of AKT1 in cancer. *Nature* **448**, 439–444 (2007).
40. Pietrantonio, F. *et al.* ALK, ROS1, and NTRK rearrangements in metastatic colorectal cancer. *J. Natl. Cancer Inst.* **109**, 1–10 (2017).
41. Stransky, N., Cerami, E., Schalm, S., Kim, J. L. & Lengauer, C. The landscape of kinase fusions in cancer. *Nat. Commun.* **5**, 4846 (2014).
42. Mulligan, L. M. RET revisited: Expanding the oncogenic portfolio. *Nat. Rev. Cancer* **14**, 173–186 (2014).
43. Russo, M. *et al.* Acquired Resistance to the TRK Inhibitor Entrectinib in Colorectal Cancer. *Cancer Discov.* **6**, 36–44 (2016).
44. Facchinetti, F. *et al.* Crizotinib-resistant ROS1 mutations reveal a predictive kinase inhibitor sensitivity model for ROS1- and ALK-rearranged lung cancers. *Clin. Cancer Res.* **22**, 5983–5991 (2016).
45. Gainor, J. F. *et al.* Molecular Mechanisms of Resistance to First- and Second-Generation ALK Inhibitors in ALK-Rearranged Lung Cancer. *Cancer Discov.* **6**, 1118–1133 (2016).
46. Chakravarty, D. *et al.* OncoKB: A Precision Oncology Knowledge Base. *HHS Public Access.* 1–23 (2017).
47. Chen, H. *et al.* Systematic Functional Annotation of Somatic Mutations in Cancer. *Cancer Cell* **33**, 450–462.e10 (2018).
48. Jaiswal, B. S. *et al.* Oncogenic ERBB3 Mutations in Human Cancers. *Cancer Cell* **23**, 603–617 (2013).
49. Prickett, T. D. *et al.* Analysis of the tyrosine kinome in melanoma reveals recurrent mutations in ERBB4. *Nat. Genet.* **41**, 1127–1132 (2009).
50. Zeki, S. S., Graham, T. A. & Wright, N. A. Stem cells and their implications for colorectal cancer. *Nat. Rev. Gastroenterol. Hepatol.* **8**, 90–100 (2011).

51. Zhang, Y., Li, B., Ju, Z. Z. & Zheng, P. S. Notch1 regulates the growth of human colon cancers. *Cancer* **116**, 5207–5218 (2010).
52. Ma, Y., Zou, H., Chen, B., Fan, Y. & Luo, S. FAM83D knockdown regulates proliferation, migration and invasion of colorectal cancer through inhibiting FBXW7/Notch-1 signalling pathway. *Biomed. Pharmacother.* **90**, 548–554 (2017).
53. Sancho, R. et al. F-box and WD repeat domain-containing 7 regulates intestinal cell lineage commitment and is a haploinsufficient tumor suppressor. *Gastroenterology* **139**, 929–941 (2010).
54. Babaei-Jadidi, R. et al. FBXW7 influences murine intestinal homeostasis and cancer, targeting Notch, Jun, and DEK for degradation. *J. Exp. Med.* **208**, 295–312 (2011).
55. Korpalsarsn, K. et al. FBXW7 missense mutation: a novel negative prognostic factor in metastatic colorectal adenocarcinoma. *Oncotarget* **8**, 39268–39279 (2017).
56. Welcker, M. & Clurman, B. E. FBW7 ubiquitin ligase: A tumour suppressor at the crossroads of cell division, growth and differentiation. *Nat. Rev. Cancer* **8**, 83–93 (2008).
57. Lee, S. Y. et al. Gain-of-function mutations and copy number increases of Notch2 in diffuse large B-cell lymphoma. *Cancer Sci.* **100**, 920–926 (2009).
58. Weng, A. P. et al. Activating Mutations of NOTCH1 in Human T Cell Acute Lymphoblastic Leukemia. *Science* (80-), **306**, 269–271 (2004).
59. Berardinelli, G. N. et al. Advantage of HSP110 (T17) marker inclusion for microsatellite instability (MSI) detection in colorectal cancer patients. *Oncotarget* **9**, 28691–28701 (2018).
60. Wiencke, J. K. Impact of race/ethnicity on molecular pathways in human cancer. *Nat. Rev. Cancer* **4**, 79–84 (2004).
61. Ozdemir, B. C. & Dotto, G. P. Racial Differences in Cancer Susceptibility and Survival: More Than the Color of the Skin? *Trends in Cancer* **3**, 181–197 (2017).
62. Campanella, N. C. et al. Optimization of a pentaplex panel for MSI analysis without control DNA in a Brazilian population: correlation with ancestry markers. *Eur. J. Hum. Genet.* **22**, 875–880 (2014).
63. Pereira, R. et al. Straightforward inference of ancestry and admixture proportions through ancestry-informative insertion deletion multiplexing. *PLoS One* **7**, (2012).
64. Staudacher, J. J. et al. Increased Frequency of KRAS Mutations in African Americans Compared with Caucasians in Sporadic Colorectal Cancer. *Clin. Transl. Gastroenterol.* **8**, e124 (2017).
65. Heath, E. et al. Racial Disparities in the Molecular Landscape of Cancer. *Anticancer Res.* **38**, 2235–2240 (2018).
66. Chong, M.-L. et al. Phosphatidylinositol-3-kinase pathway aberrations in gastric and colorectal cancer: Meta-analysis, co-occurrence and ethnic variation. *Int. J. Cancer* **134**, 1232–1238 (2014).
67. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
68. Koboldt, D. C. et al. VarScan 2: Somatic mutation and copy number alteration discovery in cancer by exome sequencing. *VarScan 2: Somatic mutation and copy number alteration discovery in cancer by exome sequencing*. *Genome Res.* **22**, 568–576 (2012).
69. McLaren, W. et al. The Ensembl Variant Effect Predictor. *Genome Biol.* **17**, 122 (2016).
70. Tamborero, D. et al. Cancer Genome Interpreter annotates the biological and clinical relevance of tumor alterations. *Genome Med.* **10**, 1–8 (2018).
71. de Oliveira, J. C. et al. Genotype-phenotype correlation in 99 familial adenomatous polyposis patients: A prospective prevention protocol. *Cancer Med.* **8**, 2114–2122 (2019).
72. Lacerda, C. F. et al. Mutational profile of TP53 in esophageal squamous cell carcinoma associated with chagasic megaesophagus. *Dis. Esophagus* **30**, 1–9 (2017).
73. Lima-Costa, M. F. et al. Genomic ancestry, Self-rated health and its association with mortality in an admixed population: 10 year follow-up of the Bambuí-EpiGen (Brazil) cohort study of ageing. *PLoS One* **10**, 1–14 (2015).

#### Acknowledgements

We are thankful to Barretos Cancer Hospital. This work was supported by the Brazilian Federal Agency for the Support and Evaluation of Graduate Education (CAPES, Brazil), the National Council for Scientific and Technological Development (CNPq, Brazil), and the Public Ministry of Labor Campinas (Research, Prevention and Education of Occupational Cancer, Brazil).

#### Author Contributions


W.S., D.P.G. and R.M.R. designed and wrote the main manuscript text. D.P.G. and R.M.R. supervised the study. W.S., G.N.B. and T.S. performed the experiments. M.M. performed the histopathological analysis. W.S. and A.F.E. carried out the bioinformatics analysis. M.A.O. and W.S. performed the statistical analysis. W.S., A.C.C., D.P.G., R.M.R. and A.F.E. contributed to the interpretation of the results. The manuscript was reviewed by all authors.

#### Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-019-49611-1>.

**Competing interests:** The authors declare no competing interests.

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019

**Anexo B.** Tabela com todas as mutações encontradas no tumor (removendo mutações silenciosas)

Gene (número de mutações)	N (%)	Anotação Existente
<b>APC (92)</b>	<b>66 (72.5)</b>	
p.Thr1556AsnfsTer3	5 (5.5)	COSM19695
p.Arg1114Ter	3 (3.3)	rs121913331,COSM13125
p.Arg232Ter	3 (3.3)	rs397515734,COSM13130
p.Gln1291Ter	3 (3.3)	COSM19072
p.Gln1338Ter	3 (3.3)	rs121913327,COSM13129
p.Arg283Ter	2 (2.2)	COSM19679
p.Arg302Ter	2 (2.2)	rs137854568,COSM13862
p.Arg554Ter	2 (2.2)	rs137854573,COSM19040
p.Arg876Ter	2 (2.2)	rs121913333,COSM18852
p.Gln1406Ter	2 (2.2)	rs587782518,COSM19087
p.Gln1429Ter	2 (2.2)	rs74535574,COSM18836
p.Glu1317Ter	2 (2.2)	COSM19253
p.Glu1554Ter	2 (2.2)	COSM33818
p.Ala1299LeufsTer6	1 (1.1)	Nova
p.Ala199CysfsTer53	1 (1.1)	Nova
p.Ala731Pro	1 (1.1)	Nova
p.Arg1171Cys	1 (1.1)	rs201830995,COSM1580515
p.Arg1450Ter	1 (1.1)	rs121913332,COSM13127
p.Arg213Ter	1 (1.1)	rs587781392,COSM13134
p.Arg216Gln	1 (1.1)	rs76685252,COSM3826545
p.Arg216Ter	1 (1.1)	rs62619935,COSM98420
p.Arg498Ter	1 (1.1)	Nova
p.Arg499Ter	1 (1.1)	rs137854580,COSM29364
p.Arg805Ter	1 (1.1)	rs587779783,COSM19058
p.Asn1535MetfsTer30	1 (1.1)	Nova
p.Asn862LysfsTer50	1 (1.1)	Nova
p.Asp1058MetfsTer3	1 (1.1)	Nova

Gene	N (%)	Anotação Existente
<b>APC (continuação)</b>		
p.Asp1569GlufsTer2	1 (1.1)	Nova
p.Asp170ValfsTer4	1 (1.1)	COSM1432145
p.Asp1860Tyr	1 (1.1)	Nova
p.Cys1410Ter	1 (1.1)	COSM19592
p.Gln1294AspfsTer8	1 (1.1)	Nova
p.Gln1367Ter	1 (1.1)	rs121913328,COSM13121
p.Gln1477ArgfsTer30	1 (1.1)	Nova
p.Gln1529Ter	1 (1.1)	COSM1183187
p.Glu1209Ter	1 (1.1)	COSM167689
p.Glu1295Ter	1 (1.1)	COSM18961
p.Glu1306Ter	1 (1.1)	rs121913462,COSM18760
p.Glu1309AspfsTer4	1 (1.1)	rs763847228,COSM18764
p.Glu1322Ter	1 (1.1)	COSM18702
p.Glu1345Ter	1 (1.1)	COSM18759
p.Glu1374Ter	1 (1.1)	COSM19085
p.Glu1379Ter	1 (1.1)	rs121913326,COSM18834
p.Glu1408Ter	1 (1.1)	COSM18822
p.Glu1494LysfsTer19	1 (1.1)	COSM19349
p.Glu403LysfsTer51	1 (1.1)	Nova
p.Glu425Ter	1 (1.1)	Nova
p.Glu941Ter	1 (1.1)	COSM18853
p.Gly1120Arg	1 (1.1)	Nova
p.Gly1288Ter	1 (1.1)	COSM19577
p.Gly2250AlafsTer29	1 (1.1)	rs773874693
p.His1490TyrfsTer19	1 (1.1)	Nova
p.Leu1129Ser	1 (1.1)	rs143638171,COSM30751
p.Leu1488PhefsTer26	1 (1.1)	Nova
p.Leu1488TyrfsTer19	1 (1.1)	COSM18999
p.Leu1972GlnfsTer2	1 (1.1)	Nova

Gene	N (%)	Anotação Existente
<b>APC (continuação)</b>		
p.Leu2168Ile	1 (1.1)	COSM1059629
p.Leu356Pro	1 (1.1)	Nova
p.Leu779PhefsTer2	1 (1.1)	Nova
p.Lys1437Ter	1 (1.1)	COSM18818
p.Lys311Asn	1 (1.1)	Nova
p.Lys756AsnfsTer5	1 (1.1)	Nova
p.Phe1491LeufsTer16	1 (1.1)	COSM19623
p.Ser1205AlafsTer60	1 (1.1)	Nova
p.Ser1315Ter	1 (1.1)	COSM19582
p.Ser1436ValfsTer37	1 (1.1)	COSM19607
p.Ser943Ter	1 (1.1)	COSM98439
p.Thr1556LeufsTer9	1 (1.1)	COSM19000
p.Trp685Ter	1 (1.1)	COSM1432191
p.Val1320GlyfsTer95	1 (1.1)	Nova
p.Val1320Ter	1 (1.1)	COSM23635
p.Val1528Gly	1 (1.1)	Nova
<b>TP53 (56)</b>	<b>51 (56.0)</b>	
p.Gly245Ser	5 (5.5)	rs28934575,COSM6932,COSM121035,COSM121036,COSM3356965,COSM1640833
p.Arg248Trp	4 (4.4)	rs121912651,COSM10656,COSM120007,COSM120005,COSM3388183,COSM1640831
p.Arg273His	4 (4.4)	rs28934576,COSM10660,COSM99729,COSM3356963,COSM1645335
p.Arg175His	3 (3.3)	rs28934578,COSM10648,COSM99914,COSM99022,COSM99023,COSM3355994
p.Arg213Ter	3 (3.3)	rs397516436,COSM10654,COSM99618,COSM99615,COSM99616,COSM3378350
p.Arg273Cys	2 (2.2)	rs121913343,COSM10659,COSM99933,COSM3355991,COSM1645518
p.Ile255del	2 (2.2)	COSM1480061,COSM1480062,COSM1480063
p.X224_splice	1 (1.1)	COSM13586,COSM118941,COSM118939,COSM3723947,COSM1638407,COSM118940
p.Ala161Ser	1 (1.1)	COSM43549,COSM1610856,COSM1610857,COSM1610859,COSM3717664,COSM3717663
p.Arg181His	1 (1.1)	rs397514495,COSM10738,COSM1386797,COSM1386798,COSM1386800,COSM3388202
p.Arg196Ter	1 (1.1)	rs397516435,COSM10705,COSM99668,COSM99665,COSM99666,COSM3378446
p.Arg248Gln	1 (1.1)	rs11540652,COSM10662,COSM99602,COSM99020,COSM3356964,COSM1640830

Gene	N (%)	Anotação Existente
<b>TP53 (continuação)</b>		
p.Arg282Trp	1 (1.1)	rs28934574,COSM10704,COSM99925,COSM3378339,COSM1636702
p.Cys229ValfsTer18	1 (1.1)	COSM44323
p.Cys242Arg	1 (1.1)	COSM11738,COSM1386619,COSM1386620,COSM1386621
p.Cys242Phe	1 (1.1)	COSM10810,COSM129834,COSM129835,COSM3378347,COSM1646852,COSM129836
p.Gln100Ter	1 (1.1)	COSM44032,COSM1189391,COSM1189392
p.Gln144Ter	1 (1.1)	rs757274881,COSM11245,COSM318146,COSM318145,COSM318147,COSM3820722
p.Gln165Ter	1 (1.1)	rs730882001,COSM43632,COSM1172488,COSM1172489,COSM1172491,COSM3388213
p.Glu180Lys	1 (1.1)	COSM43772,COSM1522501,COSM1522502,COSM1522504,COSM1522503
p.Gly187Ser	1 (1.1)	rs776167460,COSM44026,COSM437537,COSM437538,COSM437540,COSM3522697
p.Gly266Arg	1 (1.1)	COSM10794,COSM1559475,COSM3388175,COSM1645293
p.Gly266Glu	1 (1.1)	rs193920774,COSM10867,COSM216410,COSM3388174,COSM1679491
p.His168ProfsTer13	1 (1.1)	Nova
p.His193Arg	1 (1.1)	COSM10742,COSM308307,COSM308306,COSM308308,COSM3820719,COSM1740322
p.His193Asp	1 (1.1)	COSM44002,COSM251417,COSM251416,COSM251418,COSM251419
p.His193Leu	1 (1.1)	COSM11066,COSM99919,COSM99916,COSM99917,COSM3732881,COSM2744772
p.Ile195Asn	1 (1.1)	COSM44877,COSM1738250,COSM1738252,COSM1738254,COSM3403267,COSM1738251
p.Ile195Phe	1 (1.1)	COSM44633,COSM129840,COSM129841,COSM129843,COSM3388199,COSM1731912
p.Ile255dup	1 (1.1)	Nova
p.Lys132Arg	1 (1.1)	COSM11582,COSM308311,COSM308310,COSM3388223,COSM1646844,COSM308312
p.Met160AsnfsTer21	1 (1.1)	Nova
p.Met160Ile	1 (1.1)	rs772354334,COSM43891
p.Phe109Ser	1 (1.1)	COSM45169,COSM148181,COSM148180
p.Pro152ArgfsTer18	1 (1.1)	COSM43792,COSM1180849,COSM1180850,COSM1180852,COSM1180851
p.Ser241Phe	1 (1.1)	rs28934573,TP53_g.13359C>T,COSM10812,COSM214171,COSM214170,COSM3522695
p.Ser90LeufsTer59	1 (1.1)	COSM131024,COSM131025,COSM131026
p.Thr125Met	1 (1.1)	COSM44988,COSM168303,COSM168302,COSM3717682,COSM1750374
p.Tyr205His	1 (1.1)	COSM43642,COSM1159830,COSM1159831,COSM1159833,COSM1159832
p.Tyr236Cys	1 (1.1)	COSM10731,COSM116674,COSM116672,COSM3773303,COSM1646851,COSM116673



<b>KRAS (48)</b>	<b>48 (52.7)</b>	
p.Gly12Asp	16 (17.6)	rs121913529,COSM521,COSM1135366
p.Gly12Val	11 (12.1)	rs121913529,COSM520,COSM1140133
p.Gly13Asp	7 (7.7)	rs112445441,COSM532,COSM1140132
p.Gly12Ala	3 (3.3)	rs121913529,COSM522,COSM1140134
p.Gly12Cys	3 (3.3)	rs121913530,COSM516,COSM1140136
p.Ala146Thr	2 (2.2)	rs121913527,COSM19404,COSM1165198
p.Gln61His	2 (2.2)	rs17851045,COSM1135364,COSM1146992,COSM554,COSM555
p.Ala146Pro	1 (1.1)	COSM19905,COSM1140130
p.Ala146Val	1 (1.1)	COSM19900,COSM1360827
p.Gly13Glu	1 (1.1)	COSM87280
p.Lys117Asn	1 (1.1)	COSM28519,COSM1562192
<b>PIK3CA (14)</b>	<b>14 (15.4)</b>	
p.Glu545Lys	4 (4.4)	rs104886003,COSM763,COSM125370
p.Glu542Lys	2 (2.2)	rs121913273,COSM760,COSM125369
p.Asn345Lys	1 (1.1)	rs121913284,COSM754,COSM132748
p.Asn345Ser	1 (1.1)	COSM13476,COSM1420790
p.Gln546Lys	1 (1.1)	rs121913286,COSM766,COSM255876
p.Glu545Ala	1 (1.1)	rs121913274,COSM12458,COSM297145
p.Glu545Gln	1 (1.1)	COSM27133,COSM295672
p.Glu81Lys	1 (1.1)	COSM27502,COSM271871
p.Gly118Asp	1 (1.1)	COSM751,COSM246588
p.His1047Arg	1 (1.1)	rs121913279,COSM775,COSM94986
<b>FBXW7 (11)</b>	<b>11 (12.1)</b>	
p.Arg465Cys	1 (1.1)	COSM22932,COSM170727,COSM1154293,COSM170726,COSM170725
p.Arg465His	1 (1.1)	COSM22965,COSM117310,COSM1149856,COSM117309,COSM117308
p.Arg465Leu	1 (1.1)	COSM33762,COSM673925,COSM673924,COSM673923
p.Arg505Cys	1 (1.1)	rs149680468,COSM22975,COSM108572,COSM1154290,COSM108571,COSM74637
p.Arg505His	1 (1.1)	COSM133117,COSM25812,COSM287374,COSM3127975,COSM287373,COSM287372
p.Glu287Asp	1 (1.1)	Nova
p.Leu288Phe	1 (1.1)	Nova

Gene	N (%)	Anotação Existente
<b>FBXW7</b> (continuação)		
p.Leu594Pro	1 (1.1)	COSM1427660,COSM1427661,COSM1427663,COSM1427664,COSM1427662
p.Met268AspfsTer18	1 (1.1)	COSM1052121,COSM1052122,COSM1594349,COSM1052124,COSM1052123
p.Ser668ValfsTer39	1 (1.1)	COSM34018,COSM1427622,COSM1427624,COSM1427625,COSM1427623
p.Tyr545Cys	1 (1.1)	COSM22989,COSM1052084,COSM1154289,COSM1052086,COSM1052085
<b>NF1</b> (10)		
p.Ala2032Val	1 (1.1)	Nova
p.Ala548Asp	1 (1.1)	Nova
p.Arg1462Trp	1 (1.1)	COSM472532,COSM472531
p.Arg2450Ter	1 (1.1)	COSM24487,COSM303898
p.Arg2832His	1 (1.1)	rs587781523
p.Gln1360SerfsTer20	1 (1.1)	Nova
p.Lys2372Asn	1 (1.1)	Nova
p.Phe1536Ser	1 (1.1)	Nova
p.Ser1838TyrfsTer23	1 (1.1)	Nova
p.Tyr628ThrfsTer3	1 (1.1)	Nova
<b>TGFBR2</b> (9)		
p.Lys153SerfsTer35	6 (6.6)	COSM1180952,COSM1744957
p.Lys153AlafsTer3	2 (2.2)	COSM1645197,COSM1645196
p.Arg218Trp	1 (1.1)	rs61762550
<b>ATM</b> (8)		
p.Lys2811SerfsTer46	2 (2.2)	rs759472682,COSM1351020,COSM1351019
p.Arg2832His	1 (1.1)	rs529296539
p.Arg3008His	1 (1.1)	rs587781894,COSM21626,COSM1139600
p.Gln445Ter	1 (1.1)	Nova
p.Glu376Ter	1 (1.1)	Nova
p.Gly2024Trp	1 (1.1)	Nova
p.His1213Asn	1 (1.1)	Nova
<b>BLM</b> (8)		
p.Asn515MetfsTer16	5 (5.5)	rs772495493,TMP_ESP_15_91304139_91304139,COSM252959

Gene	N (%)	Anotação Existente
<b>BLM (continuação)</b>		
p.Asp757ThrfsTer4	1 (1.1)	rs747341586,COSM1375443
p.Pro707Ser	1 (1.1)	rs146077918,COSM3505421
p.Pro868Leu	1 (1.1)	rs2227935
<b>BRAF (8)</b>	<b>8 (8.8)</b>	
p.Val600Glu	6 (6.6)	rs113488022,COSM476
p.Asp594Glu	1 (1.1)	COSM253330
p.Gly469Val	1 (1.1)	rs121913355,COSM459
<b>NOTCH1(6)</b>	<b>5 (5.5)</b>	
p.Arg1356Cys	1 (1.1)	rs587778567
p.Arg2070Gln	1 (1.1)	rs751048658,COSM1259321,COSM1259320
p.Gln2394Ter	1 (1.1)	COSM13062
p.Ile1981Thr	1 (1.1)	Nova
p.Pro2222Leu	1 (1.1)	rs779125725
p.Thr1379Met	1 (1.1)	Nova
<b>BRD3(5)</b>	<b>5 (5.5)</b>	
p.Pro24ArgfsTer24	2 (2.2)	rs763134487,TMP_ESP_9_136918529_136918529,COSM1181066
p.Asn28Lys	1 (1.1)	Nova
p.Glu509_Lys510del	1 (1.1)	rs569163807
p.Val473CysfsTer13	1 (1.1)	Nova
<b>CHEK2(5)</b>	<b>5 (5.5)</b>	
c.*32C>A (3'UTR)	1 (1.1)	Nova
p.Asn359GlufsTer2	1 (1.1)	Nova
p.Gly486Val	1 (1.1)	Nova
p.Met534Thr	1 (1.1)	Nova
p.Ser207Asn	1 (1.1)	Nova
<b>RANBP2 (5)</b>	<b>4 (4.4)</b>	
p.Arg2076Gln	1 (1.1)	Nova
p.Cys2737LeufsTer5	1 (1.1)	Nova
p.Cys2737ValfsTer35	1 (1.1)	COSM1398708,COSM1398707

Gene	N (%)	Anotação Existente
<b>RANBP2</b> (continuação)		
p.Glu1213Ter	1 (1.1)	rs767158817
p.Thr1004LeufsTer12	1 (1.1)	Nova
<b>ERBB3</b> (4)	4 (4.4)	
p.Gly325Arg	1 (1.1)	COSM1935396
p.Gly582Ala	1 (1.1)	Nova
p.Gly582Val	1 (1.1)	Nova
p.Ser128Arg	1 (1.1)	Nova
<b>MSH6</b> (4)	4 (4.4)	
p.Phe1088SerfsTer2	3 (3.3)	rs770288143,COSM330655
p.Phe1088LeufsTer5	1 (1.1)	rs748452299,COSM308681
<b>NOTCH2</b> (4)	4 (4.4)	
p.Asn853His	1 (1.1)	Nova
p.Gly1140Cys	1 (1.1)	Nova
p.Gly1866Cys	1 (1.1)	Nova
p.Tyr779Cys	1 (1.1)	Nova
<b>NOTCH3</b> (4)	4 (4.4)	
p.Ala1802LeufsTer23	1 (1.1)	Nova
p.Arg1837His	1 (1.1)	rs138265894,COSM3529537,COSM3529538
p.Arg969Trp	1 (1.1)	COSM1391046,COSM1391045
p.Lys2061Asn	1 (1.1)	Nova
<b>NRAS</b> (4)	4 (4.4)	
p.Gly12Asp	2 (2.2)	rs121913237,COSM564
p.Gln61Lys	1 (1.1)	rs121913254,COSM580
p.Gly12Val	1 (1.1)	rs121913237,COSM566
<b>PTCH1</b> (4)	4 (4.4)	
p.Pro1315Leu	3 (3.3)	rs357564,COSM1638394,COSM1638395,COSM1638396,COSM1638397
p.Val1435Met	1 (1.1)	rs587778632,COSM3396768,COSM3396767,COSM3396770,COSM3396769
<b>PTEN</b> (4)	4 (4.4)	
p.Arg130Gln	1 (1.1)	rs121909229,COSM5033

Gene	N (%)	Anotação Existente
<b><i>PTEN</i></b> (continuação)		
p.Glu157GlyfsTer23	1 (1.1)	Nova
p.Lys267ArgfsTer9	1 (1.1)	COSM30622
p.Lys332Ter	1 (1.1)	Nova
<b><i>ALK</i></b> (3)	3 (3.3)	
p.Arg1120Trp	1 (1.1)	rs774951734
p.Pro1620Thr	1 (1.1)	Nova
p.Val351Ala	1 (1.1)	Nova
<b><i>AR</i></b> (3)	3 (3.3)	
p.Arg841His	1 (1.1)	Nova
p.Gln58Leu	1 (1.1)	rs200185441,COSM376477
p.Leu57Gln	1 (1.1)	rs78686797,COSM301687
<b><i>AXL</i></b> (3)	3 (3.3)	
p.Glu483Ter	1 (1.1)	COSM1393945
p.Lys477del	1 (1.1)	Nova
p.Thr568Met	1 (1.1)	rs749923407
<b><i>CREBBP</i></b> (3)	3 (3.3)	
p.Asn199Lys	1 (1.1)	Nova
p.Ile1084SerfsTer15	1 (1.1)	COSM1180829
p.Pro1279Leu	1 (1.1)	rs749189606,COSM4060710
<b><i>ERBB4</i></b> (3)	3 (3.3)	
p.Ala287Val	1 (1.1)	Nova
p.Arg1273Trp	1 (1.1)	COSM1205584
p.Arg847His	1 (1.1)	COSM4681535
<b><i>MAP2K4</i></b> (3)	2 (2.2)	
p.Arg281Ter	1 (1.1)	COSM218507
p.Arg287His	1 (1.1)	COSM180952
p.Val398Ala	1 (1.1)	Nova
<b><i>NBN</i></b> (3)	3 (3.3)	
c.*31delA (3'UTR)	1 (1.1)	Nova

Gene	N (%)	Anotação Existente
<b>NBN</b> (continuação)		
p.Ala46Val	1 (1.1)	Nova
p.Arg551GlyfsTer8	1 (1.1)	COSM1458549
<b>RET</b> (3)	2 (2.2)	
p.Ala1046Val	1 (1.1)	Nova
p.Lys761Glu	1 (1.1)	Nova
p.Tyr864Cys	1 (1.1)	Nova
<b>ROS1</b> (3)	3 (3.3)	
p.Arg2039Cys	1 (1.1)	rs752085591,COSM6038976
p.Pro556His	1 (1.1)	Nova
p.Ser1031Phe	1 (1.1)	Nova
<b>STAT5B</b> (3)	3 (3.3)	
p.Arg200Gln	1 (1.1)	rs779220548
p.Lys164SerfsTer17	1 (1.1)	COSM1383347
p.Lys583AsnfsTer16	1 (1.1)	COSM1383340
<b>BRCA1</b> (2)	2 (2.2)	
p.Asn1509IlefsTer17	1 (1.1)	Nova
p.Thr1196Lys	1 (1.1)	rs80356944
<b>BRCA2</b> (2)	2 (2.2)	
p.Asn1784ThrfsTer7	1 (1.1)	COSM18607,COSM1639329
p.Glu182Gly	1 (1.1)	Nova
<b>BRD2</b> (2)	2 (2.2)	
p.Arg281Trp	1 (1.1)	Nova
p.Pro64ArgfsTer22	1 (1.1)	Nova
<b>BRDT</b> (2)	2 (2.2)	
p.Ala58Val	1 (1.1)	Nova
p.Lys238Asn	1 (1.1)	rs1156281,COSM4144250
<b>CDKN2A</b> (2)	2 (2.2)	
p.Arg58Ter	1 (1.1)	rs121913387,COSM12473,COSM99730,COSM99731,COSM1624870
p.Met154IlefsTer76	1 (1.1)	Nova

Gene	N (%)	Anotação Existente
<b>EGFR(2)</b>	2 (2.2)	
p.Pro848Leu	1 (1.1)	rs148934350,COSM22943
p.Thr725Met	1 (1.1)	rs767505234,COSM53264
<b>FLT3(2)</b>	1 (1.1)	
p.Met665Val	1 (1.1)	Nova
p.Thr227Met	1 (1.1)	rs1933437,COSM5019176
<b>FLT4 (2)</b>	2 (2.2)	
p.Arg119His	1 (1.1)	rs751885535,COSM1471828,COSM1471827
p.Arg1320Trp	1 1.099	rs150279372
<b>HDAC4(2)</b>	2 (2.2)	
p.X997_splice	1 (1.1)	COSM4092775
p.Arg283Cys	1 (1.1)	COSM1018638
<b>HDAC7(2)</b>	2 (2.2)	
p.Arg106Trp	1 (1.1)	rs779766353
p.Glu958del	1 (1.1)	rs769122679
<b>JAK1(2)</b>	2 (2.2)	
p.Ala449Thr	1 (1.1)	rs765489448,COSM1343910
p.Asn339IlefsTer3	1 (1.1)	COSM1343915
<b>MAP3K1 (2)</b>	2 (2.2)	
p.Ala1373Val	1 (1.1)	Nova
p.Ala618SerfsTer28	1 (1.1)	Nova
<b>MRE11A(2)</b>	2 (2.2)	
p.Arg572Ter	1 (1.1)	rs137852761,MRE11Abase_D0061:g.47557C>T,COSM378985,COSM378984
p.Gln677Glu	1 (1.1)	Nova
<b>MS4A1(2)</b>	2 (2.2)	
p.Glu237AsnfsTer6	1 (1.1)	rs779922206
p.Trp216Ter	1 (1.1)	Nova
<b>MTOR(2)</b>	2 (2.2)	
p.Ser2215Phe	1 (1.1)	COSM1686998
p.Thr2444Met	1 (1.1)	rs769849522

Gene	N (%)	Anotação Existente
<b>RRM1</b> (2)	2 (2.2)	
p.Arg499His	1 (1.1)	COSM263457
p.Pro639Ser	1 (1.1)	Nova
<b>ABL1</b> (1)	1 (1.1)	
p.Asn72ThrfsTer28	1 (1.1)	Nova
<b>AKT1</b> (1)	1 (1.1)	
p.Glu17Lys	1 (1.1)	rs121434592,COSM33765
<b>ARAF</b> (1)	1 (1.1)	
p.Arg603Cys	1 (1.1)	COSM2965923
<b>BARD1</b> (1)	1 (1.1)	
p.Ser192LeufsTer20	1 (1.1)	Nova
<b>BRD4</b> (1)	1 (1.1)	
p.Gly732Arg	1 (1.1)	Nova
<b>BRIP1</b> (1)	1 (1.1)	
p.Glu462Ter	1 (1.1)	Nova
<b>CCNE1</b> (1)	1 (1.1)	
p.Met336Ile	1 (1.1)	COSM4076587
<b>CD22</b> (1)	1 (1.1)	
p.Pro746His	1 (1.1)	Nova
<b>CDK12</b> (1)	1 (1.1)	
p.Glu182del	1 (1.1)	Nova
<b>CRKL</b> (1)	1 (1.1)	
p.Arg164Trp	1 (1.1)	rs760471921
<b>DDR2</b> (1)	1 (1.1)	
p.Arg473His	1 (1.1)	rs775098890,COSM1648427,COSM530257,COSM1648428
<b>EPHA7</b> (1)	1 (1.1)	
p.Glu536Ter	1 (1.1)	Nova
<b>ERBB2</b> (1)	1 (1.1)	
p.Ile655Val	1 (1.1)	rs1136201,COSM4000121



Gene	N (%)	Anotação Existente
<b>FAM175A</b> (1)	1 (1.1)	
p.Asp331Gly	1 (1.1)	Nova
<b>GATA2</b> (1)	1 (1.1)	
p.Ala478Thr	1 (1.1)	rs780263343
<b>HGF</b> (1)	1 (1.1)	
p.Arg494Ter	1 (1.1)	Nova
<b>IDH1</b> (1)	1 (1.1)	
c.-16-1G>T_splice	1 (1.1)	Nova
<b>IDH2</b> (1)	1 (1.1)	
p.Arg172Ser	1 (1.1)	COSM133672
<b>MDM2</b> (1)	1 (1.1)	
p.Cys455Tyr	1 (1.1)	Nova
<b>MET</b> (1)	1 (1.1)	
p.Arg1188Ter	1 (1.1)	COSM1214928
<b>MLH1</b> (1)	1 (1.1)	
p.Arg100Ter	1 (1.1)	rs63751221,MMR_c.298C>T,COSM29727
<b>MPL</b> (1)	1 (1.1)	
p.Leu79Met	1 (1.1)	Nova
<b>MYC</b> (1)	1 (1.1)	
p.Asn26Ser	1 (1.1)	rs4645959
<b>NT5C2</b> (1)	1 (1.1)	
p.Phe450LeufsTer5	1 (1.1)	Nova
<b>NTRK1</b> (1)	1 (1.1)	
p.His297Tyr	1 (1.1)	Nova
<b>PALB2</b> (1)	1 (1.1)	
p.Arg1086Ter	1 (1.1)	rs587776527
<b>PIK3CB</b> (1)	1 (1.1)	
p.Glu634Ter	1 (1.1)	Nova

Gene	N (%)	Anotação Existente
<b>PIK3CG(1)</b>	1 (1.1)	
p.Ala85Val	1 (1.1)	Nova
<b>RAF1(1)</b>	1 (1.1)	
p.Pro643Leu	1 (1.1)	Nova
<b>RUNX1(1)</b>	1 (1.1)	
p.Ser362Leu	1 (1.1)	COSM4170266
<b>SMO(1)</b>	1 (1.1)	
p.Arg482Cys	1 (1.1)	rs755295928,COSM3878026
<b>SOX2(1)</b>	1 (1.1)	
p.Lys95Glu	1 (1.1)	Nova
<b>SRC(1)</b>	1 (1.1)	
p.Thr182Ala	1 (1.1)	Nova
<b>VEGFA1)</b>	1 (1.1)	
p.Leu192Ile	1 (1.1)	Nova
<b>XPO1(1)</b>	1 (1.1)	
p.Asp891His	1 (1.1)	Nova

N: número de amostras com determinada mutação

**Anexo C. Tabela das mutações encontradas nas lesões não-câncer após aplicação de filtro de patogenicidade nas amostras sequenciadas**

<b>C</b>	<b>N</b>	<b>(%)</b>	<b>Anotações Existente</b>
<b>APC (33)</b>	<b>29</b>	<b>(48.3)</b>	
p.R1450*	4	(6.7)	rs121913332,CM930030,COSM13127
p.E1309Dfs*4	3	(5.0)	COSM18764
p.Q1294*	2	(3.3)	CD058148,CD132712,CM930027,COSM18960
p.A1446Lfs*27	1	(1.7)	COSM19595
p.A1492Cfs*22	1	(1.7)	Nova
p.E1353*	1	(1.7)	rs1114167568,CD972008,CM130060,COSM19048,COSM6196780
p.E1374K	1	(1.7)	rs876659156,COSM18737,COSM19085
p.G1499K	1	(1.7)	Nova
p.I1307*fs*1	1	(1.7)	Nova
p.L1488Yfs*19	1	(1.7)	COSM18999,COSM32819
p.L1489*	1	(1.7)	COSM19105,COSM291723
p.L1564*	1	(1.7)	CM010762,COSM19039,COSM19478
p.M1431Cfs*42	1	(1.7)	COSM2991032
p.P1373L	1	(1.7)	rs1220647493,CD106455,COSM41619,COSM5981990
p.P1432L	1	(1.7)	CD011101
p.Q1291*	1	(1.7)	CM021065,CM053765,COSM1180985,COSM19072,COSM6973322
p.Q1303*	1	(1.7)	CM994441,COSM13728
p.Q1367*	1	(1.7)	rs121913328,CM940072,COSM13121,COSM3696863
p.Q1378*	1	(1.7)	rs121913329,CD1110399,CM095537,COSM18862
p.Q1429*	1	(1.7)	rs74535574,HM050024,COSM18836,COSM6023917,COSM6957879
p.R876*	1	(1.7)	rs121913333,CM942020,COSM18852
p.S1298F	1	(1.7)	COSM1059592
p.S1371G	1	(1.7)	Nova
p.S1436N	1	(1.7)	Nova
p.T1438Hfs*35	1	(1.7)	COSM18818,COSM4662396
p.Y1376*	1	(1.7)	CD058149,CM106639,COSM18808,COSM18861,COSM19103
p.Y1376Lfs*10	1	(1.7)	Nova

Gene (número de mutações)	N	(%)	Anotações Existente
<b>TP53 (25)</b>	<b>18</b>	<b>(30.0)</b>	
p.D281N	2	(3.3)	rs764146326,CD137626,CM076566,COSM10943,COSM11516,COSM1158315,COSM146336,COSM1646814,COSM1649341
p.T231I	2	(3.3)	TP53_g.13329C>T,COSM44820,COSM45273,COSM6982813,COSM6982814,COSM6982815,COSM6982816,COSM6982817
.	2	(3.3)	rs868137297,CS1313555,COSM1167885,COSM1167886,COSM1167887,COSM18655,COSM2156353,COSM218536,COSM218537
p.A276T	1	(1.7)	TP53_g.13806G>A,COSM308313,COSM3937594,COSM3937595,COSM43663,COSM44114,COSM45467,COSM6964320
p.C238Y	1	(1.7)	rs730882005,CM034930,COSM11059,COSM1649400,COSM179811,COSM179812,COSM179813,COSM249079,COSM249080
p.G266E	1	(1.7)	rs193920774,TP53_g.13777G>A,COSM10867,COSM10958,COSM1646803,COSM1679491,COSM216410,COSM3388173
p.G334E	1	(1.7)	rs1286563734,COSM11514,COSM323935,COSM330687,COSM330688,COSM7325505,COSM7325506
p.H214Y	1	(1.7)	TP53_g.12709C>T,COSM1637584,COSM1637585,COSM1637586,COSM1637587,COSM1637588,COSM44112,COSM44372
p.L130F	1	(1.7)	rs863224683,CM016206,COSM11449,COSM11462,COSM1167680,COSM1167681,COSM1167682,COSM1649369,COSM1728015
p.P152L	1	(1.7)	rs587782705,CM941327,COSM10790,COSM1180849,COSM1180850,COSM1180851,COSM1180852,COSM129856,COSM129857
p.P152S	1	(1.7)	rs767328513,COSM1731754,COSM1731755,COSM1731756,COSM1731757,COSM1731758,COSM3403278,COSM3403279
p.Q192*	1	(1.7)	rs866380588,COSM10733,COSM117946,COSM117947,COSM117948,COSM117949,COSM1649390,COSM3370929,COSM44111
p.R158H	1	(1.7)	rs587782144,CM102353,CM994513,TP53_g.12461G>A,COSM10690,COSM10714,COSM1640853,COSM1649375,COSM1649376
p.R249K	1	(1.7)	rs587782329,COSM129832,COSM129833,COSM1649403,COSM1728798,COSM326723,COSM326724,COSM3388182,COSM3403260
p.R273H	1	(1.7)	rs28934576,CM004342,CM010472,CM920677,TP53_g.13798G>A,COSM10660,COSM10779,COSM1640828,COSM1645335
p.R290C	1	(1.7)	rs770374782,COSM1167879,COSM3522691,COSM3522692,COSM43662,COSM45679,COSM46045,COSM5010282
p.R337H	1	(1.7)	rs121912664,CM012663,CM104660,TP53_g.16901G>A,COSM11411,COSM131485,COSM220135,COSM378685
p.S127F	1	(1.7)	rs730881999,TP53_g.12368C>T,COSM1637542,COSM216412,COSM216413,COSM216414,COSM3378368,COSM3403294
p.S183L	1	(1.7)	TP53_g.12536C>T,COSM10706,COSM11717,COSM1649386,COSM2744812,COSM2744813,COSM2744814,COSM2744815
p.S95F	1	(1.7)	TP53_g.11515C>T,COSM44673,COSM6899671,COSM6899672,COSM6899673,COSM6899674
p.T230I	1	(1.7)	TP53_g.13326C>T,COSM43806,COSM43868
p.Y234*	1	(1.7)	TP53_g.13339C>G,COSM1324800,COSM1324801,COSM1324802,COSM1600318,COSM1600319,COSM1600320
<b>KRAS (17)</b>	<b>17</b>	<b>(28.3)</b>	
p.G13D	6	(10.0)	CM125166,rs112445441,COSM1140132,COSM1152504,COSM3384257,COSM532,COSM533,COSM534
p.G12V	3	(5.0)	rs121913529,COSM1135366,COSM1140133,COSM1140134,COSM12657,COSM49168,COSM520,COSM521,COSM522
p.G12D	2	(3.3)	rs121913529,COSM1135366,COSM1140133,COSM1140134,COSM12657,COSM49168,COSM520,COSM521,COSM522
p.K117N	2	(3.3)	rs770248150,COSM1256061,COSM1562192,COSM19940,COSM28519
p.A146T	1	(1.7)	rs121913527,COSM1140130,COSM1165198,COSM19404,COSM19905,COSM5967494,COSM7325390
p.G13R	1	(1.7)	rs121913535,COSM1152505,COSM3384258,COSM527,COSM528,COSM529

<b>Gene</b> (número de mutações)	<b>N</b>	<b>(%)</b>	<b>Anotações Existente</b>
<b>KRAS</b> (continuação)	29	(48.3)	
p.G15S	1	(1.7)	COSM4384684,COSM538
p.Q61L	1	(1.7)	rs121913240,COSM1140131,COSM1158660,COSM3688142,COSM551,COSM552,COSM553
<b>BRAF</b> (14)	12	(20.0)	
p.V600E	8	(13.3)	CM112509,rs113488022,COSM18443,COSM476,COSM6137
p.A598V	2	(3.3)	COSM21549
p.D594N	2	(3.3)	rs397516896,COSM144576,COSM27639,COSM6928839
p.G606E	1	(1.7)	rs1488384581,COSM1137,COSM144978,COSM144979
p.S605N	1	(1.7)	rs121913373,COSM1136
<b>EGFR</b> (7)	7	(11.7)	
p.A698T	1	(1.7)	COSM41905
p.A743T	1	(1.7)	rs759256622,COSM26509
p.A755V	1	(1.7)	COSM6196863,COSM6954013
p.G779V	1	(1.7)	rs397517120,COSM133202
p.G810S	1	(1.7)	rs121913230,COSM13192
p.P753L	1	(1.7)	rs559717059,COSM1168015
p.T751I	1	(1.7)	rs727504316,COSM13185
<b>FBXW7</b> (6)	6	(10.0)	
p.R465C	3	(5.0)	rs867384286,COSM1154293,COSM170725,COSM170726,COSM170727,COSM22932
p.R393*	1	(1.7)	COSM1052102,COSM1052103,COSM1052104,COSM1154296,COSM22973,COSM6854476
p.R465H	1	(1.7)	rs1057519895,COSM1149856,COSM117308,COSM117309,COSM117310,COSM22965
p.R505H	1	(1.7)	rs1057519896,COSM1155716,COSM133117,COSM1671362,COSM22976,COSM23000
<b>GNAS</b> (5)	5	(8.3)	
p.R201H	4	(6.7)	rs121913495,COSM27895,COSM308279,COSM94388,COSM99221
p.R201C	1	(1.7)	rs11554273,COSM123397,COSM1566192,COSM1757313,COSM27887,COSM27899
<b>SMAD4</b> (5)	3	(5.0)	
p.A454T	1	(1.7)	Nova
p.E526K	1	(1.7)	COSM1151656,COSM14134,COSM6971500
p.G386S	1	(1.7)	rs1057519962,COSM1389070,COSM14147,COSM14249
p.Q450*	1	(1.7)	COSM6475947

Gene (número de mutações)	N	(%)	Anotações Existente
<b>SMAD4</b> (continuação)	29	(48.3)	
p.S517I	1	(1.7)	Nova
<b>CTNNB1</b> (3)	3	(5.0)	
p.G38N	1	(1.7)	Nova
p.S45F	1	(1.7)	rs121913409,COSM5667,COSM5689,COSM5692,COSM6100
p.T41A	1	(1.7)	rs121913412,COSM5664,COSM5688,COSM5716
<b>PTEN</b> (2)	2	(3.3)	
p.L70F	1	(1.7)	rs794727480,CM140349,COSM1684691,COSM6438159
p.S10N	1	(1.7)	COSM1349462,COSM5049
<b>PIK3CA</b> (2)	2	(3.3)	
p.H1047R	1	(1.7)	rs121913279,COSM249874,COSM775,COSM776,COSM94986,COSM94987
p.P466S	1	(1.7)	COSM230020,COSM5939140
<b>MLH1</b> (1)	1	(1.7)	
p.V384D	1	(1.7)	rs63750447,CM981292,COSM1422591,COSM26085,COSM4002663
<b>IDH2</b> (1)	1	(1.7)	
p.P158S	1	(1.7)	COSM86959
<b>HRAS</b> (1)	1	(1.7)	
p.G12D	1	(1.7)	rs104894230,CM053284,CM081305,CM123157,COSM120919,COSM4193615,COSM483,COSM484,COSM485,COSM99915
<b>FGFR3</b> (1)	1	(1.7)	
p.G697S	1	(1.7)	COSM24802,COSM5966954
<b>ATM</b> (1)	1	(1.7)	
p.R3008H	1	(1.7)	rs587781894,CM092589,COSM1139600,COSM21626,COSM356707,COSM6979279