



Ministério da Saúde
Instituto Nacional de Câncer
Coordenação de Pós-graduação

INSTITUTO NACIONAL DE CÂNCER

Pós-Graduação em Oncologia

MARINA CHIANELLO NICOLAU FAGUNDES

**A DESREGULAÇÃO DAS APOBECs E SUA CONTRIBUIÇÃO
PARA O ESTABELECIMENTO DE PERFIS MUTACIONAIS EM
CARCINOMA EPIDERMOIDE DE CABEÇA E PESCOÇO E
ESÔFAGO**

Orientadores: Dra. Sheila Coelho Soares Lima
Dra. Tatiana de Almeida Simão

Rio de Janeiro

2021



Ministério da Saúde
Instituto Nacional de Câncer
Coordenação de Pós-graduação

INSTITUTO NACIONAL DE CÂNCER
Pós-Graduação em Oncologia

MARINA CHIANELLO NICOLAU FAGUNDES

**A DESREGULAÇÃO DAS APOBECs E SUA
CONTRIBUIÇÃO PARA O ESTABELECIMENTO DE PERFIS
MUTACIONAIS EM CARCINOMA EPIDERMÓIDE DE
CABEÇA E PESCOÇO E ESÔFAGO**

Tese apresentada ao Instituto Nacional de Câncer
como parte dos requisitos para a obtenção do título
de Doutor em Oncologia.

Orientadores: Dra. Sheila Coelho Soares Lima
Dra. Tatiana de Almeida Simão

Rio de Janeiro

2021

F156d Fagundes, Marina Chianello Nicolau.

Desregulação das APOBECs e sua contribuição para o estabelecimento de perfis mutacionais em carcinoma epidermoide de cabeça e pescoço e esôfago. / Marina Chianello Nicolau Fagundes. – Rio de Janeiro, 2021. 181 f. : il. Color.

Tese (Doutorado em Oncologia) – Programa de Pós-Graduação em Oncologia, Instituto Nacional de Câncer José Alencar Gomes da Silva, 2021.

Orientadores: Dra. Sheila Coelho Soares Lima.
Profa. Dra. Tatiana de Almeida Simão.

1. Desaminases APOBEC/genética. 2. Neoplasias Esofágicas/genética. 3. Carcinoma de Células Escamosas de Cabeça e Pescoço/genética. 4. Expressão Gênica. I. Lima, Sheila Coelho Soares (Orient.). II. Simão, Tatiana de Almeida (Orient.). III. Instituto Nacional de Câncer José Alencar Gomes da Silva. IV. Título.

CDD 616.994042

Catálogo na fonte
Kátia Simões CRB7/5952



Ministério da Saúde
Instituto Nacional de Câncer
Coordenação de Pós-graduação

INSTITUTO NACIONAL DE CÂNCER
Pós-Graduação em Oncologia

MARINA CHIANELLO NICOLAU FAGUNDES

A DESREGULAÇÃO DAS APOBECS E SUA CONTRIBUIÇÃO PARA O ESTABELECIMENTO DE PERFIS MUTACIONAIS EM CARCINOMA EPIDERMÓIDE DE CABEÇA E PESCOÇO E ESÔFAGO

Orientadores: Dra. Sheila Coelho Soares Lima
Dra. Tatiana Almeida Simão

Aprovada em 30/11/2021.

BANCA EXAMINADORA

Doutor Marcelo Alves Soares- Presidente
Instituto Nacional de Câncer (INCA)

Doutor Martín Hernan Bonamino
Instituto Nacional de Câncer (INCA)

Doutora Lidia Maria Rebolho Batista Arantes
Hospital de Câncer de Barretos

Doutor Renato Santana de Aguiar
Universidade Federal de Minas Gerais (UFMG)

Doutora Mariana Lima Boroni Martins
Instituto Nacional de Câncer (INCA)

Doutor André Luiz Mencialha
Universidade do Estado do Rio de Janeiro (UERJ)

Rio de Janeiro

2021

DEDICATÓRIA

Tudo por eles e para eles, minha família, meu amor verdadeiro.

AGRADECIMENTOS

Quero agradecer primeiramente a Deus por tudo de maravilhoso que tem acontecido em minha vida e por eu estar fechando mais um ciclo profissional;

Às minhas queridas orientadoras, Dra Sheila Coelho Soares Lima e Dra Tatiana de Almeida Simão por toda dedicação, paciência, carinho, orientação e amizade. Estamos juntas desde o início e mais um ciclo sendo cumprido! Muito obrigada pelo conhecimento adquirido, por todos os ensinamentos e por acreditarem em mim. A trajetória não foi fácil, mas nosso trabalho será reconhecido!

Ao Prof.Dr. Luis Felipe Ribeiro Pinto por contribuir na minha carreira científica e por me dar oportunidade de estar nesse grupo com excelentes profissionais;

Aos membros do PCM que sempre estão dispostos a ajudar. A gente não faz pesquisa sozinha e eu só tenho que agradecer por toda discussão científica, pelos cafés na copa, pela diversão garantida, pela amizade. Minhas ANJAS, amo vocês! Amigos, essa conquista é nossa! Pedro, meu primo, obrigada por todo apoio!

Às corujinhas, sem vocês, nada disso seria possível. Obrigada por cada discussão, cada reunião científica. Gostaria de destacar a Paulinha, minha irmã, pelo carinho, pelo apoio, pela amizade; ao Diego pelas análises no R e a Rachele, minha médica preferida;

Aos meus ex companheiros de laboratório e eternos amigos, Bela e Leão, obrigada por tudo! Lion, sem você, metade dessa tese não existiria. Bela, que falta você faz!

Aos membros dessa banca por terem aceitado o convite de avaliar a tese;

À endoscopia do INCA pela coleta das amostras tumorais, em especial a Dra Simone. Sem amostras, não seria possível trabalhar. Aos pacientes, sem vocês, essa tese não existiria!

Às agências de fomento, pois sem ajuda financeira, nada disso teria acontecido;

À minha família (minha mãe, meu pai, meu irmão e minha vó) que é meu porto seguro, minha fortaleza. Obrigada por sempre estarem do meu lado, com apoio incansável, acreditando sempre em mim e no meu potencial. Vocês são a minha vida! Tudo para vocês e por vocês! Agradecer é muito pouco. Não tenho nem palavras.

Ao meu amor, meu aconchego, à minha alma gêmea. Mais um ciclo se encerra e você sempre do meu lado. Você é o meu alicerce, sempre segurando a minha mão em todos os momentos. Obrigada por tudo e por tanto. Só nós sabemos o que vivemos e por tudo que passamos para estar aqui hoje. Mas, o grande dia chegou e eu sou eternamente grata a você. Te amo muito!

Obrigada, meu Paulino, por tudo!

“Ser o melhor que você pode ser é possível para qualquer um que focar nisso. E ela vai voar mais alto e mais longe e mais rápido do que jamais sonhou.”

Capitã Marvel

SUMÁRIO

LISTA DE TABELAS.....	xi
LISTA DE FIGURAS.....	xii
LISTA DE SIGLAS E ABREVIATURAS.....	xvi
LISTA DE SÍMBOLOS.....	xxix
RESUMO.....	xxx
ABSTRACT.....	xxxii
1 INTRODUÇÃO.....	1
1.1 CÂNCER DE CABEÇA E PESCOÇO.....	1
1.1.1 INCIDÊNCIA E MORTALIDADE.....	1
1.1.2 FATORES DE RISCO.....	8
1.1.3 TRATAMENTO.....	9
1.2 CÂNCER DE ESÔFAGO.....	10
1.2.1 INCIDÊNCIA E MORTALIDADE.....	10
1.2.2 FATORES DE RISCO.....	12
1.2.3 TRATAMENTO.....	14
1.3 ALTERAÇÕES MOLECULARES EM CECPE E CEE.....	14
1.4 AID/APOBEC.....	19
1.5 JUSTIFICATIVA.....	25
2 OBJETIVOS.....	26
2.1 GERAL.....	26
2.2 OBJETIVOS ESPECÍFICOS.....	26
3 MATERIAL E MÉTODOS.....	28
3.1 ANÁLISES <i>IN SILICO</i>	28
3.1.1 ANÁLISE DE ASSINATURA MUTACIONAL - BANCO DE DADOS TCGA.....	28
3.1.2 PERFIL DE EXPRESSÃO DAS <i>APOBECs</i> – TCGA e INCA.....	28

3.1.3 IDENTIFICAÇÃO DAS ALTERAÇÕES GENÉTICAS NOS GENES QUE CODIFICAM AS APOBECS - BANCO DE DADOS TCGA.....	30
3.1.4 IDENTIFICAÇÃO DOS miRNAs CUJOS ALVOS SÃO AS APOBECS E CORRELAÇÃO DE SUA EXPRESSÃO COM OS NÍVEIS DE RNAm DAS APOBECS- BANCO DE DADOS TCGA.....	31
3.1.5 METILOMA – TCGA e INCA.....	32
3.1.6 ANÁLISE DE CORRELAÇÃO DO NÍVEL DE METILAÇÃO DE ELEMENTOS TRANSPONÍVEIS COM A EXPRESSÃO DAS APOBECS EM CEE, CEL, CECO, CEOF HPV⁻ e CEOF HPV⁺ - BANCO DE DADOS TCGA.....	33
3.1.7 CARACTERIZAÇÃO DO MICROAMBIENTE TUMORAL E CORRELAÇÃO COM A ASSINATURA MUTACIONAL E EXPRESSÃO DAS APOBECS EM CEE, CEL, CECO, CEOF HPV⁻ e CEOF HPV⁺ - BANCO DE DADOS TCGA.....	34
3.1.8 SINGLE CELL – CEE - BANCO DE DADOS GEO.....	35
3.2 CEE COMO MODELO EXPERIMENTAL - INCA.....	36
3.2.1 EXTRAÇÃO, QUANTIFICAÇÃO E DETERMINAÇÃO DA PUREZA DE RNA.....	36
3.2.1.1 EXTRAÇÃO DE RNA – AMOSTRAS CEE.....	36
3.2.1.2 QUANTIFICAÇÃO E DETERMINAÇÃO DA PUREZA DAS AMOSTRAS DE RNA.....	37
3.2.2 REAÇÃO DE TRANSCRIÇÃO REVERSA (RT).....	38
3.2.3 REAÇÃO EM CADEIA DA POLIMERASE QUANTITATIVA (PCRq).....	38
3.2.4 EXPERIMENTOS <i>IN VITRO</i>.....	41
3.2.4.1 LINHAGENS CELULARES.....	41
3.2.4.2 TRATAMENTO COM AGENTE DESMETILANTE A BAIXAS DOSES.....	43
3.2.4.3 TRATAMENTO COM RNA DE DUPLA FITA.....	44
3.2.4.4 TRATAMENTO COM INTERFERONS.....	45
3.2.4.5 COLETA E VIABILIDADE CELULAR PÓS-TRATAMENTOS.....	47
3.2.4.6 EXTRAÇÃO DO RNA, RT E PCRq PÓS-TRATAMENTOS	47
3.3 ANÁLISE ESTATÍSTICA.....	48

4 RESULTADOS	50
4.1 ASSINATURAS MUTACIONAIS, EXPRESSÃO DE <i>APOBECs</i> E MECANISMOS DE REGULAÇÃO EM CEE, CEL, CECO, CEOF HPV⁻ E CEOF HPV⁺	50
4.1.1 PERFIL MUTACIONAL EM CEE, CEL, CECO, CEOF HPV⁻ E CEOF HPV⁺	51
4.1.2 ASSINATURAS MUTACIONAIS E A EXPRESSÃO DAS <i>APOBECs</i> EM CEE, CEL, CECO, CEOF HPV⁻ e CEOF HPV⁺	60
4.1.3 MECANISMOS DE REGULAÇÃO DAS <i>APOBECs</i> EM CEE, CEL, CECO, CEOF HPV⁻ e CEOF HPV⁺	63
4.1.3.1 ALTERAÇÕES GENÉTICAS	63
4.1.3.2 ALTERAÇÕES EPIGENÉTICAS	64
4.1.3.2.1 MIRNAS	65
4.1.3.2.2 METILAÇÃO DO DNA	69
4.1.3.3 CORRELAÇÃO DO NÍVEL DE METILAÇÃO DE ELEMENTOS TRANSPONÍVEIS COM A EXPRESSÃO DAS <i>APOBECs</i> EM CEE, CEL, CECO, CEOF HPV⁻ e CEOF HPV⁺	72
4.1.3.4 INTERAÇÃO COM O SISTEMA IMUNE	74
4.2 CEE COMO MODELO EXPERIMENTAL	82
4.2.1 VALIDAÇÃO DA EXPRESSÃO DAS <i>APOBECs</i> EM PACIENTES COM CEE DO INCA	82
4.2.2 ENSAIOS <i>IN VITRO</i>	92
4.2.2.1 TRATAMENTO COM AGENTE DESMETILANTE A BAIXAS DOSES	92
4.2.2.2 TRATAMENTO COM RNA DE DUPLA FITA	94
4.2.2.3 TRATAMENTO COM INTERFERONS	95
4.2.2.3.1 TRATAMENTO COM IFNα	96
4.2.2.3.2 TRATAMENTO COM IFNβ	97
4.2.2.3.3 TRATAMENTO COM IFNγ	98
5 DISCUSSÃO	101
6 CONCLUSÕES	116

7 PERSPECTIVAS	119
8 REFERÊNCIA BIBLIOGRÁFICA	120
APÊNDICE A - Termo de consentimento informado - INCA.....	151
APÊNDICE B – Prontuários – INCA.....	159
ANEXO – Carta de aprovação do projeto sob o número 116/11 pelo Comitê de Ética em Pesquisa do INCA.....	161

LISTA DE TABELAS

Tabela 1.1: <i>Membros da família da AID/APOBEC em humanos: localização no genoma, perfil de expressão e função.....</i>	20
Tabela 3.1: <i>Sequências de oligonucleotídeos utilizadas para amplificação dos genes das APOBECs e dos controles positivos dos tratamentos in vitro e o GAPDH (gene - referência) na PCRq.....</i>	39
Tabela 3.2: <i>Caracterização das linhagens celulares TE1 e TE13.....</i>	42
Tabela 4.1: <i>Caracterização dos pacientes com CEE, CEL, CECO, CEOF HPV⁻ e CEOF HPV⁺ do banco de dados do TCGA.....</i>	51
Tabela 4.2: <i>Correlação do nível de metilação dos elementos transponíveis com a expressão das APOBECs em CEE, CEL, CECO, CEOF HPV⁻ e CEOF HPV⁺.....</i>	73
Tabela 4.3: <i>Comparação da expressão das APOBECs nos subtipos celulares em par com seu respectivo valor de p nos dados de single cell de CEE depositado no GEO.....</i>	78
Tabela 4.4: <i>Caracterização dos pacientes com CEE do INCA incluídos neste estudo para validação da expressão de APOBECs por RT-PCRq.....</i>	83
Tabela 4.5: <i>Associação da expressão das APOBEC3A, APOBEC3B e APOBEC3D no tumor e os dados clínico-patológicos dos pacientes com CEE.....</i>	88
Tabela 4.6: <i>Associação da expressão das APOBEC3F, APOBEC3G e APOBEC3H no tumor e os dados clínico-patológicos dos pacientes com CEE.....</i>	90
Tabela Suplementar 1: <i>Nível de metilação das APOBECs entre as mucosas adjacente e tumoral de pacientes com CEE, CEL, CECO, CEOF HPV⁻ e CEOF HPV⁺ do INCA e correlação desse nível com a expressão dessas enzimas nesses tumores no banco de dados do TCGA.....</i>	176

LISTA DE FIGURAS

Figura 1.1: Regiões anatômicas acometidas no câncer de cabeça e pescoço.....	1
Figura 1.2: Distribuição geográfica da incidência de câncer de lábio e cavidade oral no mundo para ambos os sexos.....	2
Figura 1.3: Distribuição proporcional dos dez tipos de câncer mais incidentes estimados para 2020 por sexo, exceto pele não melanoma, no Brasil.....	3
Figura 1.4: Representação espacial das taxas ajustadas ^a de incidência por 100 mil indivíduos, estimadas para o ano de 2020, segundo Unidade da Federação (neoplasia maligna da cavidade oral).....	4
Figura 1.5: Distribuição geográfica da incidência de câncer de orofaringe no mundo para ambos os sexos.....	5
Figura 1.6: Distribuição geográfica da incidência de câncer de laringe no mundo para ambos os sexos.....	6
Figura 1.7: Representação espacial das taxas ajustadas ^a de incidência por 100 mil indivíduos, estimadas para o ano de 2020, segundo Unidade da Federação (neoplasia maligna da laringe).....	7
Figura 1.8: Taxas de incidência e mortalidade dos tumores mais frequentes no mundo, em ambos os sexos.....	10
Figura 1.9: Distribuição geográfica da incidência de câncer de esôfago no mundo para ambos os sexos.....	11
Figura 1.10: Representação espacial das taxas ajustadas ^a de incidência por 100 mil indivíduos, estimadas para o ano de 2020, segundo Unidade da Federação (neoplasia maligna de esôfago).....	12
Figura 1.11: Presença das assinaturas mutacionais de substituição de base única nos diferentes tipos tumorais e a associação com os prováveis fatores responsáveis por tal assinatura.....	17
Figura 1.12: Processos mutacionais que originam as assinaturas 2 e 13 mediadas pela atividade das APOBECs.....	18

Figura 1.13: Processo de desaminação pela família AID/APOBEC conhecido como edição do DNA.....	21
Figura 1.14: Funções biológicas da edição do DNA pela família AID/APOBEC.....	23
Figura 3.1: Exemplo de curva de dissociação de PCRq, mostrando a especificidade dos produtos gerados e sem contaminação da reação.....	40
Figura 3.2: Gráfico de detecção de fluorescência de PCR em tempo real.....	41
Figura 3.3: Esquema do tratamento com 300 nM de 5-AZA-CdR (decitabina) por 6 dias de tratamento.....	44
Figura 3.4: Esquema do tratamento com 10 µg/mL de POLY I:C por 24h.....	45
Figura 3.5: Esquema do tratamento com 100 U/mL de IFN α ou 100 ng/mL de IFN β ou 100 ng/mL de IFN γ por 24h.....	46
Figura 4.1: Número total de variantes de único nucleotídeo por paciente com CEE, CEL, CECO, CEOF HPV $^{-}$ e CEOF HPV $^{+}$ do consórcio TCGA.....	52
Figura 4.2: Assinaturas mutacionais de substituição de base única detectadas em tumores de cabeça e pescoço e de esôfago.....	53
Figura 4.3: Contribuição das assinaturas mutacionais de substituição de base única em tumores de cabeça e pescoço e de esôfago por paciente.....	55
Figura 4.4: Fração de contribuição das assinaturas mutacionais de substituição de base única relacionadas à atividade das APOBECs em tumores de cabeça e pescoço e de esôfago por paciente.....	56
Figura 4.5: Fração de contribuição das assinaturas mutacionais de substituição de base única relacionadas à atividade das APOBECs de acordo com estadiamento tumoral em tumores de cabeça e pescoço e de esôfago por paciente.....	57
Figura 4.6: Fração de contribuição das assinaturas mutacionais de substituição de base única relacionadas à atividade das APOBECs de acordo com status de HPV em tumores de cabeça e pescoço por paciente.....	58
Figura 4.7: O impacto das assinaturas mutacionais relacionadas à atividade das APOBECs (assinaturas 2 e 13) na sobrevida global de pacientes com tumores de cabeça e pescoço e de esôfago.....	59

Figura 4.8: Correlação entre a expressão das <i>APOBECs</i> e as assinaturas mutacionais associadas a elas (2 e 13) em tumores de cabeça e pescoço e de esôfago.....	60
Figura 4.9: Análise da expressão gênica da família <i>AID/APOBEC</i> em tumores de cabeça e pescoço e de esôfago.....	62
Figura 4.10: Alterações genéticas encontradas nos genes que codificam a família <i>AID/APOBEC</i> em tumores de cabeça e pescoço e de esôfago por paciente.....	64
Figura 4.11: Análise dos miRNAs cujos alvos são a família da <i>AID/APOBEC</i> utilizando os bancos de miRNAs Diana, miRDB e miRmap.....	66
Figura 4.12: Análise de correlação entre a expressão dos miRNAs e da família <i>AICDA/APOBEC</i> em tumores de cabeça e pescoço e de esôfago.....	68
Figura: 4.13: Diferença de metilação das <i>APOBECs</i> entre as mucosas adjacentes não tumorais e tumorais de pacientes com CEE, CEL, CECO, CEOF HPV ⁻ e CEOF HPV ⁺ do INCA e correlação desse nível com a expressão dessas enzimas nesses tumores no banco de dados do TCGA.....	71
Figura 4.14: Correlação entre assinatura mutacional relacionada a <i>APOBECs</i> , expressão gênica dessas enzimas e CIBERSORT das células do sistema imune em pacientes com tumores de cabeça e pescoço e de esôfago.....	75
Figura 4.15: Comparação da expressão gênica das <i>APOBECs</i> em células presentes no microambiente tumoral e em células tumorais em CEE.....	77
Figura 4.16: Expressão das <i>APOBECs</i> nas células epiteliais por paciente com CEE.....	80
Figura 4.17: Resumo dos resultados obtidos nos tumores de esôfago e de cabeça e pescoço nas análises <i>in silico</i>	81
Figura 4.18: Expressão das <i>APOBEC3s</i> em pacientes com CEE do INCA.....	85
Figura 4.19: O impacto da expressão das <i>APOBEC3s</i> na sobrevida global de pacientes com carcinoma epidermoide de esôfago do INCA.....	86
Figura 4.20: Viabilidade celular das linhagens de CEE após o tratamento com decitabina.....	92
Figura 4.21: Análise da expressão gênica das <i>APOBECs</i> e dos genes responsivos após o tratamento com decitabina.....	93

Figura 4.22: Viabilidade celular das linhagens de CEE após o tratamento com POLY I:C.....	94
Figura 4.23: Análise da expressão gênica das <i>APOBECs</i> e dos genes responsivos após o tratamento com POLY I:C.....	95
Figura 4.24: Viabilidade celular das linhagens de CEE após o tratamento com IFN α	96
Figura 4.25: Análise da expressão gênica das <i>APOBECs</i> e do gene responsivo após o tratamento com IFN α	97
Figura 4.26: Viabilidade celular das linhagens de CEE após o tratamento com IFN β	97
Figura 4.27: Análise da expressão gênica das <i>APOBECs</i> e do gene responsivo após o tratamento com IFN β	98
Figura 4.28: Viabilidade celular das linhagens de CEE após o tratamento com IFN γ	99
Figura 4.29: Análise da expressão gênica das <i>APOBECs</i> e do gene responsivo após o tratamento com IFN γ	99
FIGURA SUPLEMENTAR 1 - As diferentes assinaturas mutacionais presentes em tumores..	162

LISTA DE SIGLAS E ABREVIATURAS

3'	Extremidade 3', região não transcrita, depois da transcrição
5'	Extremidade 5', região não transcrita, antes da transcrição
5-AZA-CdR	5-AZA-2'-deoxicitidina
5caC	5-carboxilcitosina
5fC	5-formilcitosina
5hmC	5-hidroxi-metilcitosina
5hmU	5-hidroxi-metiluracila
5mC	5-metilcitosina
ADAR1	Desaminase de adenosina específica de RNA, do inglês <i>Adenosine Deaminase RNA Specific</i>
ADE	Adenocarcinoma de esôfago
AID	Desaminase de citidina induzida por ativação, do inglês <i>Activation Induced Cytidine Deaminase</i>
ALL	Leucemia Linfóide Aguda, do inglês <i>Acute lymphocytic leukemia</i>
AML	Leucemia Mieloide Aguda, do inglês <i>Acute Myeloid Leukemia</i>
ALU	Elemento repetitivo pertencente à família SINE, do inglês <i>Arthrobacter luteus (Alu) restriction endonuclease</i>
AMPc	3', 5'- adenosina monofosfato cíclico
ApoB	Apolipoproteína B, do inglês <i>Apolipoprotein B</i>
APOBEC	Enzima de edição de mRNA de apolipoproteína B, tipo polipeptídeo catalítico, do inglês <i>Apolipoprotein B mRNA Editing Enzyme Catalytic Subunit</i>

APS	Adenosina 5' fosfosulfato
ASR	Taxas ajustadas por idade no mundo, do inglês <i>age-standardized rates (World)</i>
ATCC	American Type Culture Collection – Banco de Dados de Linhagens Celulares
ATL	Tampão de Lise de Tecido
ATP	Trifosfato de adenosina
BAX	Regulador de apoptose X associado a BCL2, do inglês BCL2 Associated X, Apoptosis Regulator
BER	Reparo de DNA por excisão de bases, do inglês <i>Base excision repair</i>
BNT	Banco Nacional de Tumores – INCA
BRCA	Câncer de mama início precoce, do inglês BRCA DNA Repair Associated
BSA	Albumina sérica bovina, do inglês Bovine Serum Albumin
Cas9	Endonuclease de duplo corte
CASP8	Caspase 8
CCE	Carcinoma de células escamosas
CCND1	Ciclina D1, do inglês Cyclin D1
CCP	Câncer de cabeça e pescoço
CD45	Glicoproteína expressa na membrana de células hematopoiéticas
CDKN2A	Inibidor 2A de quinase dependente de ciclina, do inglês Cyclin Dependent Kinase Inhibitor 2 ^a
CDH1	Caderina-1, do inglês Cadherin 1
cDNA	Ácido desoxirribonucleico complementar
CE	Câncer de esôfago

Cél	Célula
CEOF	Carcinoma epidermoide de orofaringe
CECO	Carcinoma epidermoide de cavidade oral
CECP	Carcinoma epidermoide de cabeça e pescoço
CEE	Carcinoma epidermoide de esôfago
CEHF	Carcinoma epidermoide de hipofaringe
CEL	Carcinoma epidermoide de laringe
cg	Sonda de metilação
ChIP-seq	Sequenciamento baseado na imunoprecipitação da cromatina
CHO	Grupo aldeído
CI	Intervalo de Confiança
CID	Classificação internacional de doenças
CO₂	Dióxido de carbono (gás carbônico)
COOH	Grupo carboxila
COSMIC	Catálogo de Mutações Somáticas em Câncer
CpG	5'—Citosina—fosfato—Guanina—3'
CRISPR	Repetições Palindrômicas Curtas Agrupadas e Regularmente Interespaçadas, do inglês Clustered Regularly Interspaced Short Palindromic Repeats
CRISPR/Cas9	Técnica de edição do DNA
C_T	Ciclo de <i>threshold</i>
DAPK1	Proteína quinase 1 associada à morte, do inglês Death Associated Protein Kinase 1

ddATP	2',3'-Dideoxiadenosina-5'-trifosfato
ddCTP	2',3'-Dideoxicitidina-5'-trifosfato
ddGTP	2',3'-Dideoxiguanosina-5'-trifosfato
ddTTP	2',3'-Dideoximidina-5'-trifosfato
DDX58	DExH-Box Helicase 58, receptor de resposta imune inata antiviral
DEC	Decitabina
DIPAT	Divisão de Patologia – INCA
DMSO	Dimetilsulfóxido
DNA	Ácido desoxirribonucleico
DNAecc	Elementos de DNA circular extracromossomal
DNMT	DNA metil-transferases, do inglês DNA Methyltransferase
dNTP	Desoxirribonucleotídeo trifosfato
DSB	Quebra dupla do DNA
DTT	Ditiotreitol
E2F1	Fator 1 de transcrição em eucarioto
EBV	Vírus Epstein-Barr, do inglês <i>Epstein-Barr virus</i>
EDTA	Ácido etilenodinitrilotetra acético
ERBB2	Receptor Tirosina Quinase 2 Erb-B2, do inglês Erb-B2 Receptor Tyrosine Kinase 2
ESCC	Carcinoma epidermoide de esôfago, do inglês <i>Esophageal squamous cell carcinoma</i>
ETS1	Fator de transcrição, um proto-oncogene ,do inglês ETS Proto-Oncogene 1, Transcription Factor

FBXW7	F-box com domínio repetido WD 7, do inglês F-Box And WD Repeat Domain Containing 7
FC	<i>Fold Change</i> de expressão (mediana de expressão no tumor/mediana de expressão na mucosa adjacente não tumoral)
FITC	Isotiocianato de fluoresceína
FPKM	Fragments por quilobase de éxon por milhão de reads mapeados, do inglês <i>Fragments Per Kilobase of exon per million reads Mapped</i>
FRC	Células reticulares fibroblásticas, do inglês <i>Fibroblastic reticular cells</i>
GAPDH	Gliceraldeído-3-fosfato desidrogenase, do inglês <i>Glyceraldehyde-3-Phosphate Dehydrogenase</i>
GATA	Fator de transcrição que se liga à sequência de DNA, do inglês GATA Binding Protein
GDC	Dados do NCI, do inglês <i>NCI Genomic Data Commons</i>
GEO	Banco de Dados do NCBI, do inglês <i>Gene Expression Omnibus</i>
H₂O	Água
H3K36	Lisina 36 da histona H3
HBV	Vírus da Hepatite B
HDAC	Histona desacetilase
<i>Heatmap</i>	Mapa de “calor”, gráfico usado na tese para expressão gênica
HIV	Vírus de imunodeficiência humano
HNSCC	Carcinoma epidermoide de cabeça e pescoço, do inglês <i>Head and Neck Squamous cell carcinoma</i>
<i>HOXA9</i>	Homeobox A9, do inglês Homeobox A9
HPV	Papilomavírus Humano, do inglês <i>Human Papillomavirus</i>

HR	Razão de chance, do inglês <i>hazard ratio</i>
HRAS	GTPase, proto-oncogene, do inglês HRas Proto-Oncogene
IARC	Agência Internacional de Pesquisa em Câncer
IC	Intervalo de Confiança
IDH	Índice de desenvolvimento humano
IFNα	Interferon alfa
IFNβ	Interferon beta
IFNγ	Interferon gama
Ig	Imunoglobulina
IL-4	Interleucina 4
IL-6	Interleucina 6
IL-13	Interleucina 13
IL-1β	Interleucina 1 beta
INCA	Instituto Nacional de Câncer
Indels	Inserções ou Deleções
IRF6	Fator regulatório de interferon 6, do inglês Interferon Regulatory Factor 6
IRF7	Fator regulatório de interferon 7, do inglês Interferon Regulatory Factor 7
LINEs	Elementos nucleares longos intercalados, do inglês <i>Long interspersed nuclear elements</i>
LSCC	Carcinoma epidermoide de laringe, do inglês <i>Larynx squamous cell carcinoma</i>
LTR	Região de repetição terminal longa, do inglês <i>long terminal repeat</i>

<i>Machine learning</i>	Algoritmo de aprendizado por máquina (inteligência artificial)
MAF	Formato de anotação de mutação derivado do VCF, do inglês <i>Mutation Annotation Format</i>
M0	Macrófago sem contato com antígeno, não ativado (<i>naive</i>)
M1	Macrófago pró-inflamatório
M2	Macrófago pró-tumoral
MAPK	Proteína quinase ativada por mitógeno, do inglês <i>Mitogen-Activated Protein Kinase</i>
MBDs	Proteínas que possuem domínio de ligação aos sítios CpG metilados
MDA5	Receptor de sensor viral, do inglês <i>Melanoma Differentiation-Associated Protein 5</i>
Meio RPMI	Meio Roswell Park Memorial Institute – É um meio contendo uma série de sais, aminoácidos e outros componentes essenciais para o crescimento celular
MgAc2	Acetato de Magnésio
MGMT	O ⁶ -metil-guanina DNA metil-transferase
MiR/ miRNA	MicroRNA
miR-Seq	Sequenciamento de miRNA
M-MLV	Transcriptase reversa termoestável
MT1M	Metalotioneína 1M
MYC	MYC Proto-Oncogene
NaCl	Cloreto de Sódio
Naive	Sem contato com antígeno (não ativado)
NaOH	Hidróxido de Sódio

NCBI	National Center for Biotechnology Information - Centro Nacional de Informações sobre Biotecnologia
NEFH	Neurofilamento pesado
NER	Reparo por excisão de nucleotídeo, do inglês <i>nucleotide excision repair</i>
NF-κB	Fator nuclear kappa B
NFE2L2	Fator nuclear eritroide 2 tipo 2
NHEJ	<i>Non-Homologous End-Joining</i> - Recombinação das Extremidades não homólogas
NIH	Instituto Nacional de Câncer, do inglês <i>National Cancer Institute</i>
NK	Células <i>natural killer</i>
NKG2D	Um ligante importante para o reconhecimento de células pelo sistema imune, do inglês <i>Killer Cell Lectin Like Receptor K1</i>
NNK	4-(metilnitrosamino)-1-(3-piridil)-1-butanona, do inglês <i>4-(nitrosomethylamino)-1-(3-pyridyl)-1-butanone</i>
NNN	N'-nitrosornicotina, do inglês <i>N'-nitrosornicotine</i>
NOTCH	Receptor NOTCH
O₂	Gás oxigênio
OH	Grupo hidroxila
OPSCC	Carcinoma epidermoide de orofaringe, do inglês <i>Oropharyngeal squamous cell carcinoma</i>
OR	Razão de risco, do inglês <i>odds ratio</i>
Overnight	Incubação feita até o dia seguinte, do inglês durante a noite
p	Braço curto do cromossomo
PBS	Tampão fosfato-salino, do inglês <i>Phosphate buffered saline</i>

PCR	Reação em cadeia da polimerase, do inglês <i>Polymerase Chain Reaction</i>
PCRq	Reação em cadeia da polimerase quantitativa
<i>Pellet</i>	Precipitado de células
pH	Potencial hidrogeniônico
PI3K	Fosfatidil inositol 3 quinase, do inglês <i>Phosphatidylinositol-4,5-Bisphosphate 3-Kinase</i>
PIK3CA	Subunidade catalítica alfa da fosfatidilinositol 3 quinase, do inglês <i>Phosphatidylinositol-4,5-Bisphosphate 3-Kinase Catalytic Subunit Alpha</i>
PKA	Proteína quinase A
PKC	Proteína quinase C
PKR	Proteína quinase ativada por RNA, do inglês <i>RNA-activated protein kinase</i>
POLE	Épsilon DNA polimerase, subunidade catalítica, do inglês <i>DNA Polymerase Epsilon, Catalytic Subunit</i>
POLY I:C	Análogo sintético de RNA de dupla fita, do inglês <i>Polyinosinic-polycytidylic acid</i>
<i>POU2F2</i>	Homeobox, do inglês <i>POU Class 2 Homeobox 2</i>
PPARγ	Receptor ativado por proliferador de peroxissoma gamma, do inglês <i>Peroxisome Proliferator Activated Receptor Gamma</i>
PPi	Pirofosfato
<i>PTEN</i>	Homólogo de tensina e fosfatase, do inglês <i>Phosphatase And Tensin Homolog</i>
q	Braço longo do cromossomo
RAS	GTPases monoméricas - Nome derivado do vírus do sarcoma de rato, do inglês <i>RAt Sarcoma virus</i>
RASSF1	Família de domínio membro 1 associada à Ras, do inglês <i>Ras Association Domain Family Member 1</i>

REN	Reparo por excisão de nucleotídeos
Rho	Coefficiente de correlação
RHOXF1	Membro da família RHOX, do inglês <i>Rhox Homeobox Family Member 1</i>
RLT	Tampão de Lise
RNA	Ácido ribonucleico
RNAg	RNA guia
RNAm	Ácido ribonucleico mensageiro
RNAse	Enzima que degrada RNA
RNAseq	Sequenciamento de nova geração de RNA
RPE	Tampão de lavagem do RNA ligado à membrana
rpm	Rotações por minuto
RPMI	Meio de cultura para linhagem celular, do inglês <i>Roswell Park Memorial Institute</i>
RT	Reação de transcrição reversa
RUNX3	Fator de transcrição, do inglês <i>RUNX Family Transcription Factor 3</i>
RW1	Tampão de lavagem do RNA ligado à membrana
SAM	S-adenosil-metionina
SBS	Assinaturas mutacionais de substituição de base única
SDS	Dodecil sulfato de sódio
<i>Seed match</i>	Sequência de ligação do miRNA-RNA _m
SFB	Soro fetal bovino

SINEs	Elementos nucleares curtos intercalados, do inglês <i>Short interspersed nuclear elements</i>
<i>Single cell</i>	Modelo de célula única
<i>SNAI2</i>	Fator de transcrição, do inglês Snail Family Transcriptional Repressor 2
SOX2	SRY-Box 2
SOX10	SRY-Box 10
SPI1	Proto-oncogene, do inglês <i>Spi-1 Proto-Oncogene</i>
SPIB	Fator de transcrição, do inglês <i>Spi-B Transcription Factor</i>
STAT6	Transdutor de sinal e ativador da transcrição 6, do inglês <i>Signal Transducer And Activator Of Transcription 6</i>
Sting	Estimulador de gene interferon
STP	Segundo tumor primário
<i>SuperScript II</i> ®	Transcriptase reversa
<i>TBX1</i>	T-Box 1
<i>TBX15</i>	T-Box 15
TCGA	Atlas do Genoma Tumoral, do inglês <i>The Cancer Genome Atlas</i>
TCD4⁺	Linfócito T <i>helper</i> (auxiliar)
TCD8⁺	Linfócito T citotóxico
TDG	Enzima timina DNA-glicosilase, do inglês <i>Thymine DNA Glycosylase</i>
TET	Dez-onze translocação metilcitosina dioxygenase, do inglês <i>Tet Methylcytosine Dioxygenase</i>
<i>TFAP2A</i>	Fator de transcrição do inglês <i>Transcription Factor AP-2 Alpha</i>
<i>TFF1</i>	Fator de “trefoil” 1

TGF-β	Fator de crescimento de transformação beta, do inglês <i>Transforming Growth Factor Beta</i>
T helper	Linfócito T auxiliar
Threshold	Linha de base
TNF-α	Fator de necrose tumoral alfa, do inglês <i>Tumor Necrosis Factor alpha</i>
TNM	Sistema onde T= tamanho do tumor, N e M, a presença/ausência de metástase em linfonodos regionais e à distância
TP53	Gene supressor tumoral que codifica uma proteína de massa molecular 53 kDa, do inglês <i>Tumor Protein P53</i>
TP63	Gene que codifica uma proteína de massa molecular 63 kDa, do inglês <i>Tumor Protein P63</i>
TpC	5'—Timina—fosfato—Citosina—3'
TpCpT	Timina seguida de citosina seguida de timina
TPM	Transcritos por milhão, do inglês <i>transcripts per million</i>
TpTpT	Timina seguida de timina seguida de timina
TRAF3	Fator 3 associado ao receptor de fator de necrose tumoral, do inglês <i>TNF Receptor Associated Factor 3</i>
Treg	Célula T regulatória
TrisHCl	Trisma cloridrato – tampão com pH fisiológico
TRIzol®	Reagente para extração de ácidos nucleicos
TSS	Sítio de início da transcrição, do inglês <i>the transcription start site</i>
Tween	Detergente
UCHL1	Hidrolase L1 C-terminal ubiquitina, do inglês <i>Ubiquitin C-Terminal Hydrolase L1</i>

UNG	Uracila DNA glicosilase, do inglês <i>Uracil DNA Glycosylase</i>
UCSC	Instituto Genômico da Universidade da Califórnia Santa Cruz, do inglês <i>University of California Santa Cruz Genomics Institute</i>
UDG	Uracila DNA glicosilase, do inglês <i>Uracil DNA Glycosylase</i>
UTR	Região não traduzida, do inglês <i>untranslated region</i>
VCF	Formato de dado para avaliar mutação, do inglês <i>variant call format</i>
VEGF	Fator de crescimento de endotélio vascular, do inglês <i>Vascular Endothelial Growth</i>
VEGFA	Fator de crescimento de endotélio vascular A, do inglês <i>Vascular Endothelial Growth Factor A</i>
VEP	Ferramenta de predição do efeito da variante, do inglês <i>Ensembl Variant Effect Predictor</i>
VIF	Fator de infectividade viral
ZNF263	Proteína dedo de zinco, do inglês <i>Zinc Finger Protein 263</i>

LISTA DE SÍMBOLOS

%	Porcentagem
Δ	Delta
$\Delta\beta$	variação do valor beta pela mediana (nível de metilação no tumor – nível de metilação na mucosa adjacente ao tumor)
γ	Gama
η	Eta
$\mu\text{M/L}$	Micromolar/litro
h	Hora
nmol/L	Nanomolar/litro
nM	Nanomolar
μM	Micromolar
mM	Milimolar
mg	Miligramma
μL	Micro litro
mL	Mililitro
rpm	Rotação por minuto
$^{\circ}\text{C}$	Grau Celsius
x g	Força centrífuga relativa cuja unidade de medida é grama
s	Segundo
min	Minuto
ng/mL	Nanograma/mililitro
nm	Nanômetro
nM	Nanomolar
μg	Micrograma
ng	Nanograma
M	Molar

U	Unidade
U/mL	Unidade/mililitro
CH₃	Grupamento Metil
CO₂	Dióxido de carbono
Mg⁺²	Magnésio
MgCl₂	Cloreto de Magnésio
NH₂	Amina
<	Menor que
≤	Menor e igual que
±	Mais ou menos
>	Maior que
≥	Maior e igual que
=	Igual
+	Mais



**Ministério da Saúde
Instituto Nacional de Câncer
Coordenação de Pós-graduação**

INSTITUTO NACIONAL DE CÂNCER

**A DESREGULAÇÃO DAS APOBECs E SUA CONTRIBUIÇÃO PARA O
ESTABELECIMENTO DE PERFIS MUTACIONAIS EM CARCINOMA EPIDERMÓIDE DE
CABEÇA E PESCOÇO E ESÔFAGO**

RESUMO

Tese de Doutorado

Marina Chianello Nicolau Fagundes

A família de enzimas AID/APOBEC, envolvida no processo de edição do DNA/RNA, é associada ao estabelecimento de assinaturas mutacionais específicas em câncer. Em carcinoma epidermoide de esôfago (CEE), essas assinaturas estão presentes em cerca de 90% dos casos, sendo responsáveis por 25% da carga mutacional. O carcinoma epidermoide de cabeça e pescoço (CECP) apresenta superexpressão de *APOBEC3B* e assinatura mutacional mediada por APOBEC. Ambos os tumores são frequentes em homens no Brasil, altamente letais e compartilham semelhanças morfológicas e etiológicas. Porém, a causa da desregulação de AID/APOBEC em tumores é desconhecida. Assim, esta tese visou dissecar a desregulação de AID/APOBECs em CEE e em CECP. Nossos dados mostraram que CEE e carcinoma epidermoide de laringe (CEL) apresentam as maiores cargas mutacionais, mas a fração de contribuição das assinaturas de APOBEC é maior em carcinoma epidermoide de orofaringe (CEOF) HPV⁻. Essas assinaturas foram associadas à expressão das *APOBECs* da família 3 nesses tipos de câncer, sendo *APOBEC3A* e *APOBEC3B* em CEE. Essas enzimas também estão superexpressas nos tumores em relação às mucosas adjacentes não tumorais. Em CEOF, tumores HPV⁺ apresentam maior expressão de *APOBEC3s* do que tumores HPV⁻. Em seguida, avaliamos se alterações genéticas e/ou epigenéticas estariam envolvidas na desregulação dessas enzimas, mas esses foram eventos raros nas casuísticas estudadas. A metilação de elementos retrotransponíveis foi avaliada como um possível mecanismo de indução de resposta antiviral e consequente indução da expressão de *APOBECs*. Os níveis de metilação de cerca de 30% dos elementos ALU avaliados foram inversamente correlacionados com a expressão de *APOBEC3A* em CEE. Uma vez que a interação com o sistema imune também pode participar da regulação de APOBECs, avaliamos a correlação entre a proporção dessas células na massa tumoral e as assinaturas e expressão dessas enzimas. Foram observadas correlações significativas com macrófagos (M0, M1, M2) em CECP e CEE. Além disso, correlações diretas entre as assinaturas mutacionais e a proporção de células TCD4 de memória e células TCD8 foi observada em tumores orais. A partir de dados de RNAseq de *single cell* em CEE, observamos que todas as *APOBECs* são detectadas em células epiteliais e que *APOBEC3A* é mais expressa em células epiteliais e mieloides. Prosseguimos com a avaliação dos mecanismos envolvidos na indução dessas enzimas utilizando como modelo experimental linhagens de CEE (TE1 e TE13). Para isso, três abordagens foram utilizadas: tratamentos com agente desmetilante, com interferons e com RNA de dupla fita, nos quais avaliamos a expressão das *APOBEC3A*, *3B* e *3D* e os controles positivos de resposta antiviral mediada por IFN *DDX58*, *MDA5* e *IRF7*. Apesar de não terem sido observadas diferenças estatisticamente significativas, a expressão de *APOBEC3A* foi induzida em níveis semelhantes àqueles observados nos controles positivos. Assim, *APOBEC3A* parece ter um papel importante no estabelecimento da assinatura mutacional em CEE e na interação com o microambiente tumoral e parece ser induzida por agente desmetilante e por interferons.



**Ministério da Saúde
Instituto Nacional de Câncer
Coordenação de Pós-graduação**

INSTITUTO NACIONAL DE CÂNCER

**THE DEREGULATION OF APOBEC AND ITS CONTRIBUTION FOR MUTATIONAL
PROFILE IN HEAD AND NECK AND ESOPHAGEAL SQUAMOUS CELL CARCINOMA
PATIENTS**

ABSTRACT

PhD Thesis

Marina Chianello Nicolau Fagundes

The AID/APOBEC family of enzymes, involved in DNA/RNA editing, is associated with the establishment of specific mutational signatures in cancer. In esophageal squamous cell carcinoma (ESCC), these signatures are present in approximately 90% of the cases, accounting for 25% of the mutational load. Head and neck squamous cell carcinoma (HNSCC) shows *APOBEC3B* overexpression and APOBEC-mediated mutational signature. Both tumors are frequent among men in Brazil, are very lethal and share morphological and etiological similarities. However, the cause of AID/APOBEC dysregulation in tumors is unknown. Thus, this thesis aimed to elucidate the dysregulation of AID/APOBECs in ESCC and in HNSCC. Our data showed that ESCC and laryngeal squamous cell carcinoma (LSCC) show the highest mutational loads, but the contribution fraction of APOBEC signatures is highest in HPV⁻ oropharyngeal squamous cell carcinoma (OPSCC). These signatures were associated with the expression of *APOBEC3s* in these cancers, specifically *APOBEC3A* and *APOBEC3B* in ESCC. These enzymes are also overexpressed in tumors relative to surrounding non-tumor mucosa. In OPSCC, HPV⁺ tumors display higher *APOBEC3s* expression relative to HPV⁻ tumors. Then, we evaluated whether genetic and/or epigenetic alterations would be involved in the dysregulation of these enzymes, but these alterations were rare events in the analyzed cohorts. Methylation of retrotransposable elements was evaluated as a possible mechanism for inducing an antiviral response and consequent induction of *APOBECs* expression. The methylation levels of about 30% of the evaluated ALU elements were inversely correlated with the expression of *APOBEC3A* in ESCC. Since the immune system cells can also participate in the regulation of *APOBECs*, we evaluated the correlation between the proportion of these cells in the tumor mass and the APOBEC mutational signatures and expression. Significant correlations with macrophages (M0, M1, M2) were observed in HNSCC and ESCC tumors. Furthermore, direct correlations between mutational signatures and the proportion of memory CD4 T cells and CD8 T cells were observed in oral tumors. From ESCC single cell RNAseq data, we observed that all *APOBECs* are detected in epithelial cells and that *APOBEC3A* is more expressed in epithelial and myeloid cells than other cells. We investigated the mechanisms involved in the induction of these enzymes using CEE cell lines as an in vitro model (TE1 and TE13). For this, three approaches were used: treatments with demethylating agent, interferons, and double-stranded RNA, in which we evaluated the expression of *APOBEC3A*, *3B* and *3D* and positive controls of antiviral response mediated by IFN, *DDX58*, *MDA5* and *IRF7*. Although no statistically significant differences were observed, *APOBEC3A* expression was induced at similar levels to those observed in positive controls. Thus, APOBEC3A seems to play an important role in establishing the mutational signature in ESCC and its expression seems to be induced by demethylating agent and interferons.

1 INTRODUÇÃO

1.1 CÂNCER DE CABEÇA E PESCOÇO

1.1.1 INCIDÊNCIA E MORTALIDADE

Câncer de cabeça e pescoço é um termo genérico definido por bases anatômico-topográficas para descrever tumores malignos do trato aerodigestivo superior. Estas regiões anatômicas incluem cavidade oral, nasofaringe, orofaringe, hipofaringe, cavidade nasal, seios paranasais e glândulas salivares (**Figura 1.1**) (DOBROSSY, 2005). Esse grupo de tumores representa o sexto tipo de câncer mais frequente mundialmente (FERLAY et al., 2015), sendo considerado um grave problema de saúde pública em muitas partes do mundo (BRAY et al., 2018).

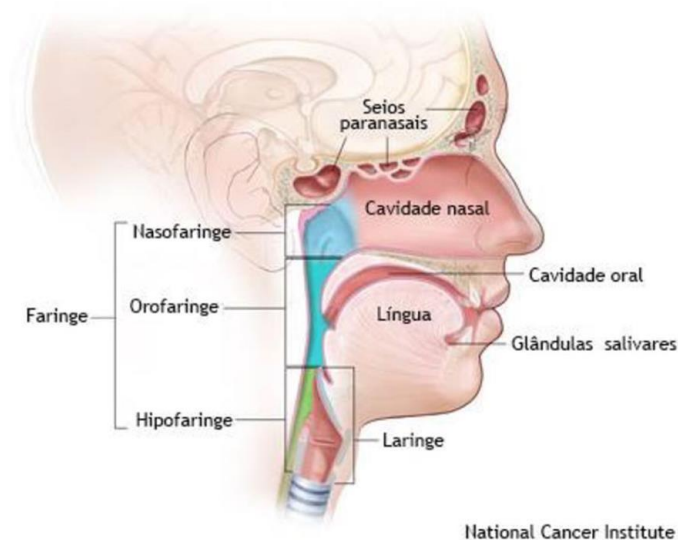


Figura 1.1: Regiões anatômicas acometidas no câncer de cabeça e pescoço. Ilustração da localização dos seios paranasais, cavidade nasal, cavidade oral, língua, glândulas salivares, laringe e faringe (incluindo a nasofaringe, orofaringe e hipofaringe). Fonte: Adaptado de <https://www.cancer.gov/>, último acesso em 17 de setembro de 2021.

Dados de incidência mostram que cerca de 40% dos tumores de cabeça e pescoço ocorrem na cavidade oral, 25% na laringe, 15% na faringe (orofaringe, hipofaringe e nasofaringe) e o restante nos demais sítios anatômicos (glândulas salivares, etc.) (DOBROSSY, 2005). O tipo

histológico mais frequente é o carcinoma de células escamosas, ou epidermoide, diagnosticado em mais de 90% dos casos de tumores malignos da cavidade oral, orofaringe e laringe (DEDIVITIS et al., 2004; BARNES et al., 2005). Este tipo de carcinoma está associado a um alto índice de morbidade e mortalidade. Trata-se de uma neoplasia epitelial invasiva com variados graus de diferenciação e propensa ao aparecimento precoce de metástase em linfonodos (FERREIRA & CASALI, 2010).

O carcinoma epidermoide de cavidade oral (CECO) é um subgrupo de carcinoma epidermoide de cabeça e pescoço (CECP) que surge na boca (língua, assoalho de boca, gengiva, mucosa jugal e palato duro) (DOBROSSY, 2005). Apresenta uma alta incidência em países da América do Norte, como Estados Unidos e Canadá; países da Europa, como Polônia e Alemanha; países da Ásia, como Índia, Paquistão e Afeganistão; a Austrália na Oceania; e países da América do Sul, como o Brasil (Figura 1.2) (GLOBOCAN, 2020).

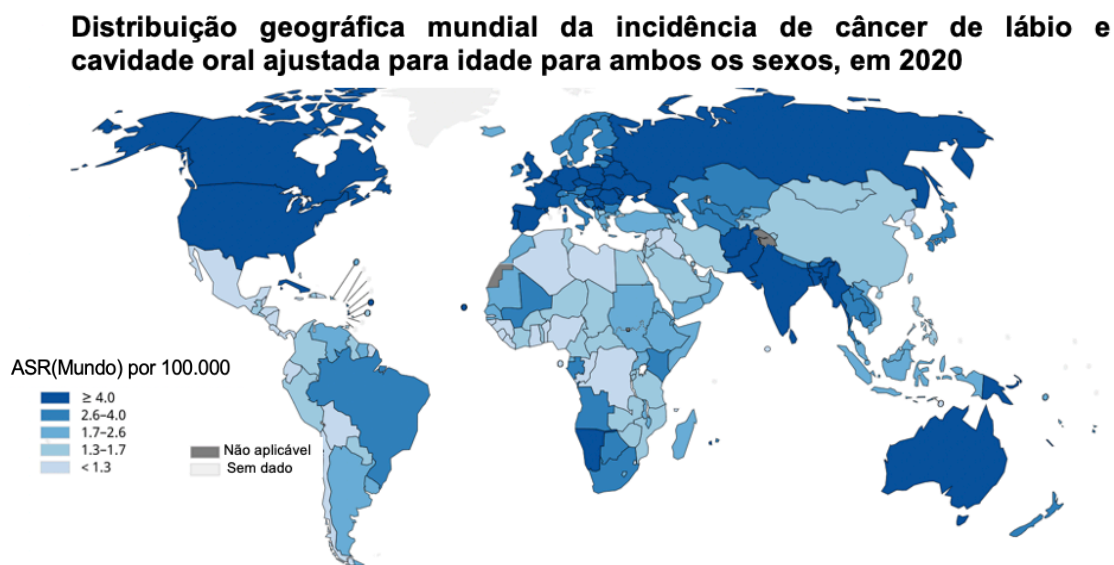


Figura 1.2: Distribuição geográfica da incidência de câncer de lábio e cavidade oral no mundo para ambos os sexos. Tons mais escuros de azul representam regiões com maior incidência enquanto tons mais claros, regiões com menor incidência. ASR: Taxas ajustadas por idade no mundo, do inglês *age-standardized rates (World)*. Fonte: Adaptado de GLOBOCAN 2020 – IARC.

O CECO é o 17º tumor mais incidente no mundo e o 15º mais mortal (BRAY et al., 2018; GLOBOCAN 2020). No Brasil, é o quinto mais incidente entre os homens e o décimo terceiro

entre as mulheres (INCA, 2019) (**Figura 1.3**). Porém, cabe ressaltar que como não há norma ou padronização nas literaturas nacional e internacional sobre quais estruturas anatômicas compõem a cavidade oral, nas estimativas do INCA foram consideradas como neoplasias malignas de lábio e cavidade oral aquelas que tenham como localização primária os lábios, a cavidade oral, as glândulas salivares e a orofaringe (C00-C10), segundo a classificação internacional de doenças (CID) 10 (INCA, 2016). Assim, para o Brasil, estimam-se 11.180 casos novos destes tumores em homens e 4.010 em mulheres para cada ano do triênio 2020-2022.

Distribuição proporcional dos dez tipos de câncer mais incidentes estimados para 2020 por sexo, exceto pele não melanoma*

Localização Primária	Casos	%			Localização Primária	Casos	%
Próstata	65.840	29,2%	Homens	Mulheres	Mama feminina	66.280	29,7%
Cólon e reto	20.520	9,1%			Cólon e reto	20.470	9,2%
Traqueia, brônquio e pulmão	17.760	7,9%			Colo do útero	16.590	7,4%
Estômago	13.360	5,9%			Traqueia, brônquio e pulmão	12.440	5,6%
Cavidade oral	11.180	5,0%			Glândula tireoide	11.950	5,4%
Esôfago	8.690	3,9%			Estômago	7.870	3,5%
Bexiga	7.590	3,4%			Ovário	6.650	3,0%
Linfoma não Hodgkin	6.580	2,9%			Corpo do útero	6.540	2,9%
Laringe	6.470	2,9%			Linfoma não Hodgkin	5.450	2,4%
Leucemias	5.920	2,6%			Sistema nervoso central	5.220	2,3%

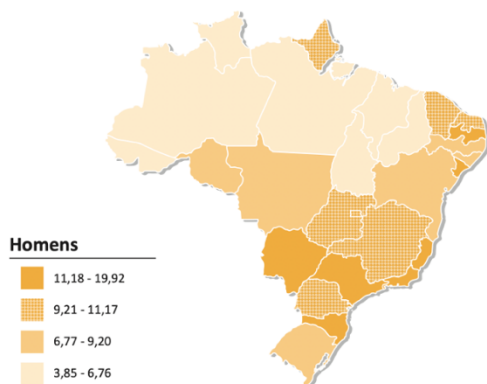
*Números arredondados para múltiplos de 10.

Figura 1.3: Distribuição proporcional dos dez tipos de câncer mais incidentes estimados para 2020 por sexo, exceto pele não melanoma, no Brasil. Fonte: INCA, 2019.

Assim como no mundo, as taxas de incidência de CECO também variam entre as regiões no Brasil, sendo as regiões sul, sudeste e centro-oeste as de maior incidência em homens (**Figura 1.4**) (INCA, 2019). A sobrevida de pacientes acometidos por este tipo de tumor é de cerca de 50% em cinco anos, porém uma proporção significativa de pacientes com CECO desenvolve um segundo tumor primário (STP) no trato aerodigestivo superior, o que reduz bastante essas taxas (JOHNSON, 2011; CHUNG, 2013).

A

Representação espacial das taxas ajustadas^a de incidência por 100 mil homens, estimadas para o ano de 2020, segundo Unidade da Federação (neoplasia maligna da cavidade oral)

**B**

Representação espacial das taxas ajustadas^a de incidência por 100 mil mulheres, estimadas para o ano de 2020, segundo Unidade da Federação (neoplasia maligna da cavidade oral)

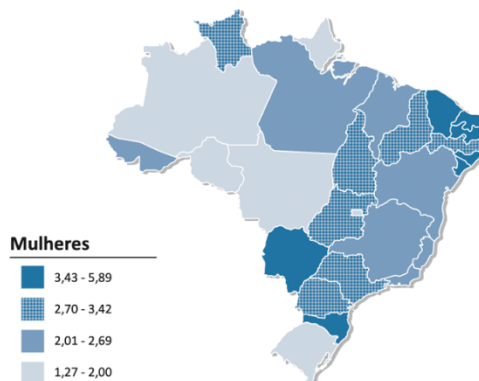


Figura 1.4: Representação espacial das taxas ajustadas^a de incidência por 100 mil indivíduos, estimadas para o ano de 2020, segundo Unidade da Federação (neoplasia maligna da cavidade oral). (A) Representação espacial das taxas brutas de incidência de câncer de cavidade oral em indivíduos do sexo masculino. (B) Representação espacial das taxas brutas de incidência de câncer de cavidade oral em indivíduos do sexo feminino. Os tons mais escuros representam as regiões de maior incidência enquanto os tons mais claros, regiões de menor incidência. ^a População padrão mundial (1960). Fonte: INCA, 2019.

O carcinoma epidermoide de orofaringe (CEOF) é categorizado como CECP, juntamente com os CECO, e carcinomas epidermoides de laringe (CEL) e hipofaringe (CEHF) (GILLISON et al., 2000; BABOCI et al., 2013). O CEOF apresenta cerca de 50.000 casos incidentes por ano (VAN MONSJOU et al., 2010; FERLAY et al., 2015) e, juntamente com CEHF, representa cerca de 1,0% de todas as malignidades do mundo (FERLAY et al., 2015; BRAY et al., 2018). O CEOF é mais comum nos Estados Unidos; em países da Europa, como Alemanha e Polônia; na Austrália; e na América do Sul, no Brasil (**Figura 1.5**) (GLOBOCAN, 2020).

Distribuição geográfica mundial da incidência de câncer de orofaringe ajustada para idade para ambos os sexos, em 2020

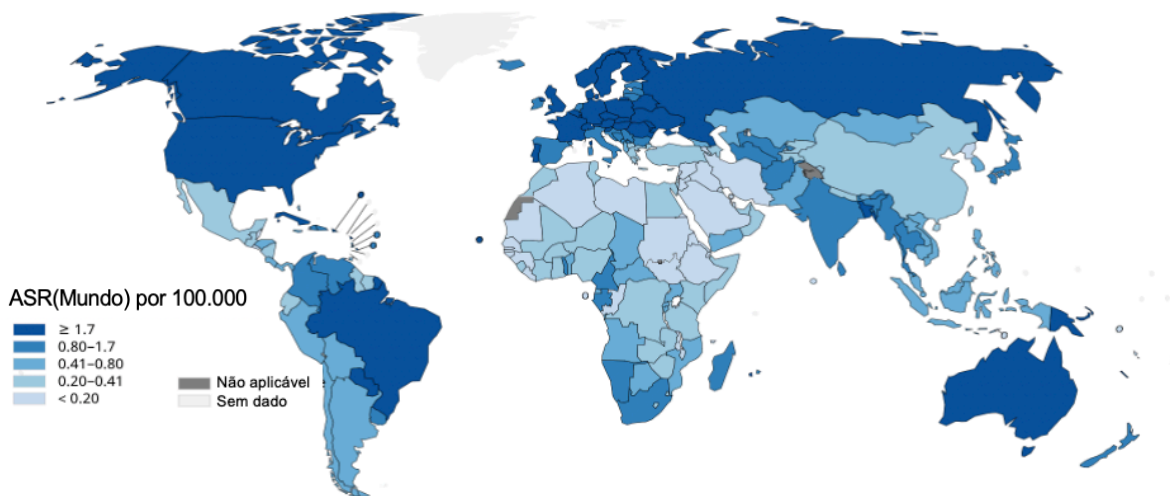


Figura 1.5: Distribuição geográfica da incidência de câncer de orofaringe no mundo para ambos os sexos. Tons mais escuros de azul representam regiões com maior incidência enquanto tons mais claros regiões com menor incidência. ASR: Taxas ajustadas por idade no mundo, do inglês *age-standardized rates (World)*. Fonte: Adaptado de GLOBOCAN 2020– IARC.

O CEOF é o 25º tumor mais incidente no mundo e o 23º mais mortal (BRAY et al., 2018; GLOBOCAN 2020). É mais frequente em homens e a taxa de sobrevivência em cinco anos de pacientes diagnosticados entre 2008 e 2014 foi de aproximadamente 59% nos Estados Unidos, conforme reportado pelo NIH (do inglês *National Cancer Institute*) (NOONE et al., 2018).

Já o câncer de laringe é o segundo tipo tumoral mais frequente do aparelho respiratório, atrás somente do câncer de pulmão (COCA-PELAZ, 2013). A taxa de incidência em indivíduos do sexo masculino é quase oito vezes maior do que em indivíduos do sexo feminino, com idade acima de 40 anos (COCA-PELAZ, 2013). É o 20º mais incidente e o 19º mais mortal no mundo (BRAY et al., 2018). As regiões de alta incidência do câncer de laringe são países da Europa, como a Polônia; países da África, como Líbia e Marrocos, e países da América do Sul, como o Brasil (Figura 1.6) (GLOBOCAN, 2020).

Distribuição geográfica mundial da incidência de câncer de laringe ajustada para idade para ambos os sexos, em 2020

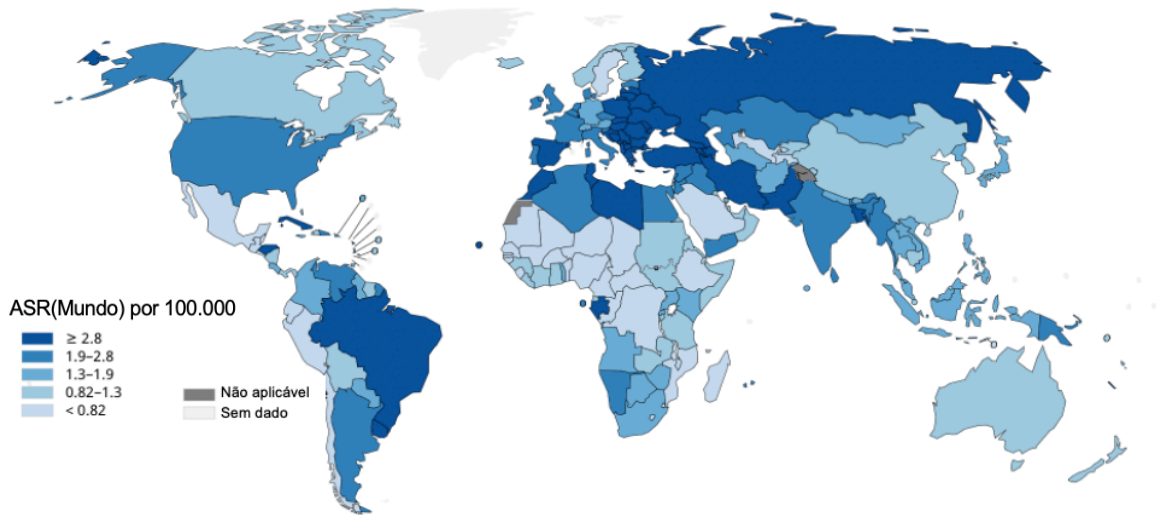
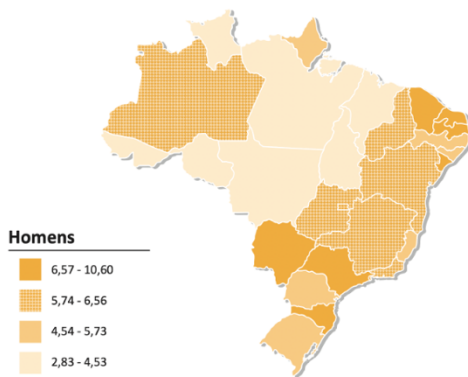


Figura 1.6: Distribuição geográfica da incidência de câncer de laringe no mundo para ambos os sexos. Tons mais escuros de azul representam regiões com maior incidência enquanto tons mais claros, regiões com menor incidência. ASR: Taxas ajustadas por idade no mundo, do inglês *age-standardized rates (World)*. Fonte: Adaptado de GLOBOCAN 2020 – IARC.

No Brasil, o CEL é o nono mais incidente entre os homens e o décimo sexto entre as mulheres (**Figura 1.3**), sendo a região Sul aquela mais afetada por este tumor entre homens (**Figura 1.7**) (INCA, 2019).

A

Representação espacial das taxas ajustadas^a de incidência por 100 mil homens, estimadas para o ano de 2020, segundo Unidade da Federação (neoplasia maligna da laringe)

**B**

Representação espacial das taxas ajustadas^a de incidência por 100 mil mulheres, estimadas para o ano de 2020, segundo Unidade da Federação (neoplasia maligna da laringe)

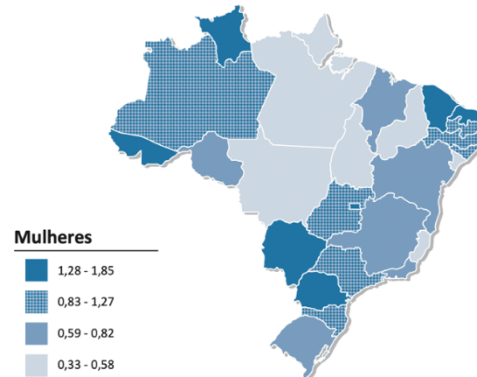


Figura 1.7: Representação espacial das taxas ajustadas^a de incidência por 100 mil indivíduos, estimadas para o ano de 2020, segundo Unidade da Federação (neoplasia maligna da laringe). (A) Representação espacial das taxas brutas de incidência de câncer de laringe em indivíduos do sexo masculino. (B) Representação espacial das taxas brutas de incidência de câncer de laringe em indivíduos do sexo feminino. Os tons mais escuros representam as regiões de maior incidência enquanto os tons mais claros, regiões de menor incidência. ^a População padrão mundial (1960).
Fonte: INCA, 2019.

Com relação ao prognóstico da doença, a taxa de sobrevivência em cinco anos para pacientes com CEL varia de acordo com a localização anatômica: na supraglote (acima das cordas vocais), em torno de 46%; na glote (inclui as cordas vocais), em torno de 77%; e na subglote (abaixo das cordas vocais), em torno de 53%, em pacientes diagnosticados no período entre 2008 e 2014, nos Estados Unidos, de acordo com dados divulgados pelo NIH (NOONE et al., 2018). Quando diagnosticado em estádios iniciais (I e II), o câncer de laringe possui um bom prognóstico com alto percentual de cura (de 80% a 100%) (COCA-PELAZ, 2013).

1.1.2 FATORES DE RISCO

O consumo intenso de álcool e tabaco representam os principais fatores de risco para o desenvolvimento do carcinoma epidermoide de cabeça e pescoço, principalmente para os tumores de cavidade oral, orofaringe, hipofaringe e laringe, podendo atuar de forma sinérgica (WRIGHT & MORGAN, 2013). Já a exposição à radiação solar ultravioleta excessiva e sem proteção ao longo dos anos é fator de risco para o desenvolvimento de câncer de lábio (REZENDE et al., 2007; DAHER et al., 2008). Outro fator de risco para o desenvolvimento de tumores de cabeça e pescoço é a infecção viral pelos Papilomavírus Humano (HPV, do inglês *Human Papillomavirus*) e vírus Epstein-Barr (EBV, do inglês *Epstein-Barr virus*). A exposição crônica ao EBV é fator de risco para o desenvolvimento de carcinoma de nasofaringe, enquanto a infecção pelo HPV vem sendo associada ao desenvolvimento de tumores de orofaringe (FERREIRA & TUNISSIOLLI, 2017).

Os principais fatores de risco para o câncer de cavidade oral estão associados a hábitos como mascar *betel quid*, tabagismo e o etilismo (KUMAR et al., 2016). Estudos já demonstraram associação entre mascar tabaco com o desenvolvimento deste tumor e lesões pré-malignas como leucoplasias, eritoplasias e fibrose submucosa oral (AXELL, 1987). Os carcinógenos presentes na fumaça do cigarro, como as nitrosaminas NNK (4-(metilnitrosamino)-1-(3-piridil)-1-butanona) e NNN (N'-nitrosonornicotina), são capazes de induzir tumores na cavidade oral (KUMAR et al., 2016). Já o álcool foi descrito como fator de risco independente para leucoplasias em população indiana; e a ocorrência de displasia epitelial oral já foi observada em etilistas que não fumam, mostrando o papel do álcool no desenvolvimento desta lesão pré-neoplásica (KUMAR et al., 2016). Dietas pobres em verduras e frutas também são fatores de risco para o desenvolvimento de câncer oral. Assim como o uso de antissépticos bucais que contêm álcool aumenta o risco de desenvolvimento desse tumor, sendo este atribuído à frequência, duração do uso e o conteúdo de álcool (KUMAR et al., 2016).

Já para o CEL, o uso de tabaco é o principal fator de risco e, quando associado ao álcool, esse risco é potencializado. Outros fatores são: histórico familiar, má alimentação, situação socioeconômica desfavorável, inflamação crônica da laringe causada pelo refluxo gastroesofágico, e exposição a produtos químicos, pó de madeira, fuligem ou poeira de carvão e vapores de tinta (COCA-PELAZ, 2013).

Nos últimos anos, a incidência de CECP tem diminuído nos países desenvolvidos devido à redução da exposição ao tabaco. No entanto, há um aumento concomitante na incidência de CEOF como resultado da infecção por tipos de HPV de alto risco (HUBER & TANTIWONGKOSI, 2014; CASTELLSAGUE et al., 2016). Em especial, a detecção de HPV16 em células epiteliais está associada a um risco quase quatro vezes maior de CECP e a um aumento de risco superior a 14 vezes de CEOF. O HPV16 foi identificado em mais de 90% dos CEOF HPV-positivos (EL-NAGGAR et al., 2017).

1.1.3 TRATAMENTO

Com relação ao tratamento, a primeira linha para os tumores de cavidade oral é a cirurgia, com a retirada da área afetada pelo tumor associada à remoção de linfonodos do pescoço e reconstrução, quando necessário. Caso a cirurgia não seja possível, a radioterapia e a quimioterapia são indicadas (INCA, 2019).

O CEOF é frequentemente diagnosticado quando já é sintomático e, portanto, em estádios avançados. Pacientes com estadiamento inicial podem ser tratados com cirurgia primária e/ou radioterapia exclusiva (EVANS *et al.*, 2013). Em estágio avançado, há necessidade de tratamento multimodal, incluindo a quimiorradioterapia ou cirurgia com cisplatina concorrente, geralmente mórbida, seguida de radioterapia adjuvante com ou sem quimioterapia. A decisão sobre o tipo de tratamento para os pacientes com CEOF deve levar em consideração a clínica, o histopatológico e fatores específicos do paciente, além da etiologia (FURLAN *et al.*, 2017).

Já o tratamento do câncer de laringe pode ser dividido em três principais estratégias. A primeira foca em curar o paciente utilizando para tanto um procedimento cirúrgico radical denominado laringectomia total. A segunda almeja a preservação da voz e qualidade de vida do paciente. A terceira, e atual, utiliza uma combinação de abordagens visando a preservação do órgão. Para tanto, radioterapia, quimioterapia e cirurgia são empregadas (INCA, 2019).

1.2 CÂNCER DE ESÔFAGO

1.2.1 INCIDÊNCIA E MORTALIDADE

O câncer de esôfago (CE) é responsável por 7% de todas as neoplasias malignas do trato gastrointestinal, sendo um problema de saúde pública mundial. É o décimo primeiro tipo de câncer mais frequente e o oitavo mais mortal no mundo. Apresenta taxas de incidência e mortalidade muito próximas, demonstrando sua alta agressividade e letalidade (BRAY et al., 2018; GLOBOCAN, 2020) (**Figura 1.8**).

Taxas de incidência e mortalidade no mundo ajustadas por idade para ambos os sexos, todas as idades

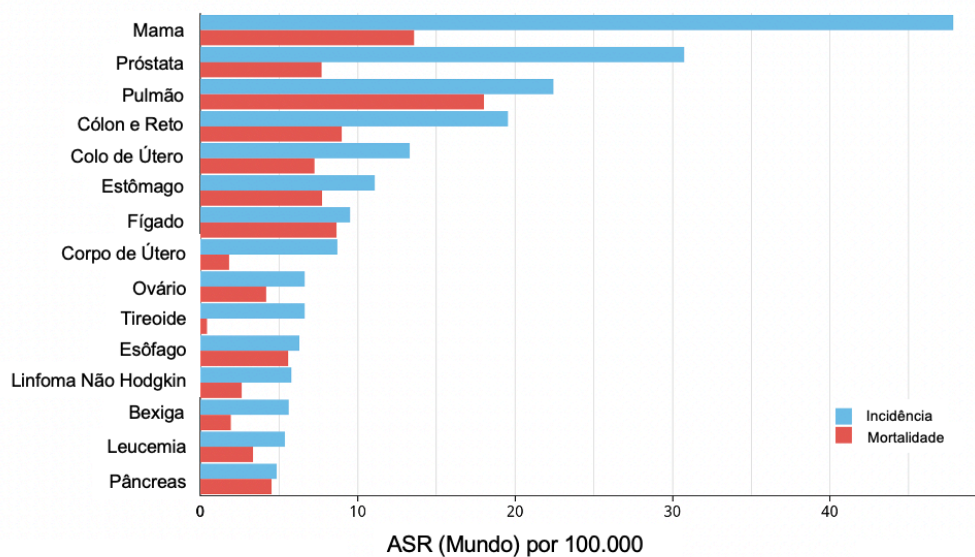


Figura 1.8: Taxas de incidência e mortalidade dos tumores mais frequentes no mundo, em ambos os sexos. As barras azuis mostram as taxas de incidência e as vermelhas mostram as taxas de mortalidade. ASR: Taxas ajustadas por idade no mundo, do inglês *age-standardized rates (World)*. Fonte: Adaptado de GLOBOCAN 2020 – IARC.

Menos de 20% dos pacientes acometidos por esta doença têm uma sobrevida maior que cinco anos após seu diagnóstico (COHEN & AJANI, 2011). Este fato é consequência do diagnóstico tardio, uma vez que os sintomas só aparecem em estágios bem avançados da doença. Somente quando 50 a 75% da luz esofágica está obstruída, o paciente apresenta disfagia

(dificuldade de engolir), sendo normalmente o primeiro sintoma apresentado (STEIN et al., 2001), o que resulta na ineficácia do tratamento (ALLEN et al., 1997; MCLARTY et al., 1997; RIBEIRO PINTO et al., 2003).

O CE apresenta uma ampla distribuição geográfica quanto à sua incidência, com regiões de alta incidência do tumor bem próximas a regiões de baixa incidência (**Figura 1.9**) (ESLICK, 2009). Isso é consequência da forte influência de fatores socioambientais (SMEDS et al., 2002), sendo os países de baixo e médio índice de desenvolvimento humano (IDH) as principais regiões acometidas por esse tumor, com 83% dos casos e 86% dos óbitos decorrentes da doença (PARKIN et al., 2005; ESLICK, 2009; JEMAL et al., 2011).

Distribuição geográfica mundial da incidência de câncer de esôfago ajustada para idade para ambos os sexos, em 2020

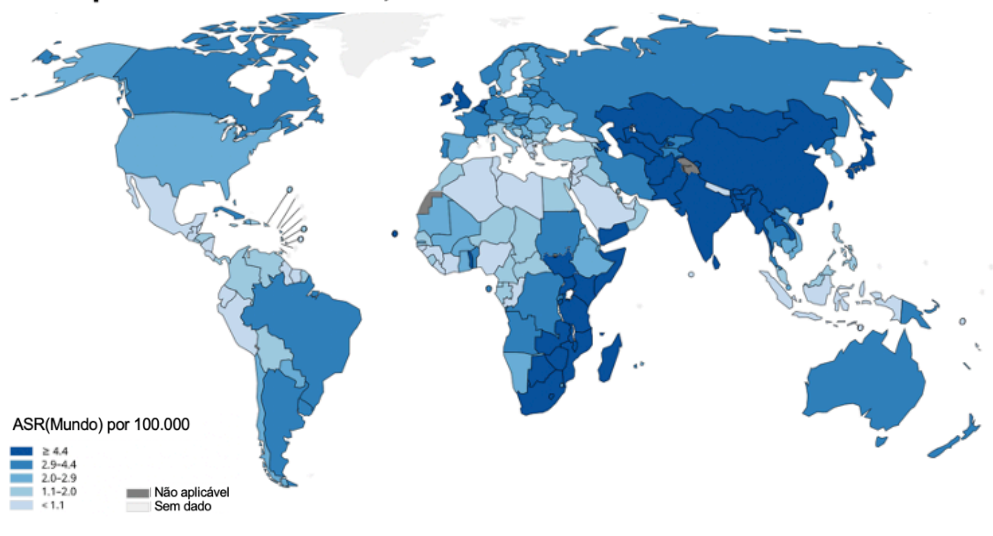


Figura 1.9: Distribuição geográfica da incidência de câncer de esôfago no mundo para ambos os sexos. Tons mais escuros de azul representam regiões com maior incidência enquanto tons mais claros, regiões com menor incidência. ASR: Taxas ajustadas por idade no mundo, do inglês *age-standardized rates (World)*. Fonte: Adaptado de GLOBOCAN 2020 – IARC.

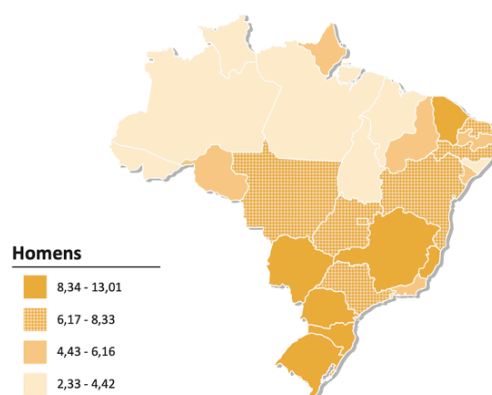
As taxas de incidência variam até 16 vezes entre regiões de alta e baixa incidência. As regiões de alta incidência incluem o sudeste da África (Quênia); leste da Ásia (do norte do Irã até a China, região conhecida como cinturão asiático do câncer de esôfago); o sul da América do Sul (Brasil e Uruguai); alguns países desenvolvidos (França e Inglaterra); e uma baixa incidência é

encontrada, por exemplo, na América Central (PARKIN et al., 2005; ESLICK, 2009; JEMAL et al., 2011) (**Figura 1.9**).

No Brasil, o CE é o sexto tumor mais incidente em homens e o décimo quinto mais incidente em mulheres (**Figura 1.3**), sendo as regiões Sudeste e Sul as mais afetadas por essa neoplasia em ambos os gêneros (**Figura 1.10**) (INCA, 2019).

A

Representação espacial das taxas ajustadas^a de incidência por 100 mil homens, estimadas para o ano de 2020, segundo Unidade da Federação (neoplasia maligna do esôfago)



B

Representação espacial das taxas ajustadas^a de incidência por 100 mil mulheres, estimadas para o ano de 2020, segundo Unidade da Federação (neoplasia maligna do esôfago)

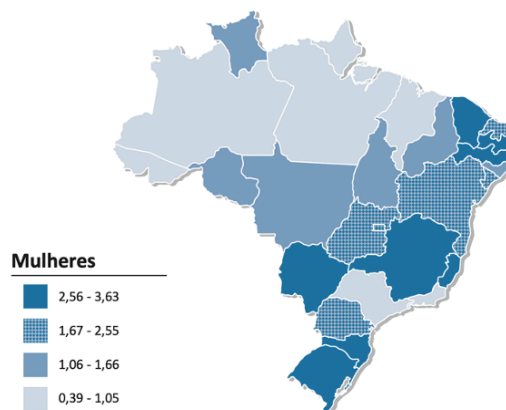


Figura 1.10: Representação espacial das taxas ajustadas^a de incidência por 100 mil indivíduos, estimadas para o ano de 2020, segundo Unidade da Federação (neoplasia maligna de esôfago). (A) Representação espacial das taxas brutas de incidência de câncer de esôfago em indivíduos do sexo masculino. (B) Representação espacial das taxas brutas de incidência de câncer de esôfago em indivíduos do sexo feminino. Os tons mais escuros representam as regiões de maior incidência enquanto os tons mais claros, regiões de menor incidência. ^a População padrão mundial (1960). Fonte: INCA, 2019.

1.2.2 FATORES DE RISCO

O câncer de esôfago apresenta dois subtipos histológicos principais, o adenocarcinoma de esôfago (ADE) e o carcinoma epidermoide de esôfago (CEE). Esses subtipos diferem em características histopatológicas e clínicas, etiologia, distribuição geográfica, população acometida, alterações moleculares envolvidas com a gênese e a progressão da doença, entre outros aspectos (QUEIROGA E PERNAMBUCO, 2006; HONGO et al., 2009).

Mundialmente, o ADE é mais frequentemente observado em regiões com alto IDH, como por exemplo, Europa Ocidental e os Estados Unidos, devido aos fatores de risco associados. Porém, recentemente, tem havido um rápido crescimento da sua incidência em países com baixo e médio IDH, devido principalmente às mudanças nos hábitos. Dentre eles, o consumo de alimentos ricos em gordura e o aumento da média de peso dessas populações, caracterizando obesidade, parecem ser os mais relevantes (FALK, 2009). O principal fator de risco para o ADE é o desenvolvimento do esôfago de Barrett especializado ou intestinal, uma condição pré-maligna na qual o epitélio normal do esôfago é substituído por um epitélio metaplásico colunar (SPECHLER et al., 1994). O esôfago de Barrett se origina em pacientes que sofrem de refluxo gastroesofágico crônico, condição que aumenta em mais de 100 vezes a chance de desenvolver ADE (THOMAS & SOBIN, 1995). O tabagismo foi identificado como fator de risco para o ADE, mas o etilismo mostrou estar pouco relacionado à gênese da doença (TOH et al., 2010). O ADE acomete principalmente o terço inferior do esôfago, enquanto o CEE ocorre mais frequentemente em células escamosas que revestem os terços superior e médio (ENZINGER & MAYER, 2003).

O CEE é o subtipo histológico de CE mais frequente, correspondendo a mais de 80% dos casos de câncer de esôfago no mundo, mesmo com o aumento de incidência de ADE. O CEE apresenta maior incidência em países com baixo e médio IDH, como o Brasil, onde este também é o subtipo mais frequente (PARKIN et al., 2005; INCA, 2019). Diversos estudos epidemiológicos apontam que o etilismo e o tabagismo são os principais fatores de risco para o desenvolvimento de CEE. Em países ocidentais, o consumo concomitante desses dois fatores representa um risco multiplicador (IARC, 1986; CASTELLSAGUÉ et al., 1999). Além disso, há a sugestão de que o fumo tenha um papel importante tanto na iniciação quanto na promoção do tumor, enquanto o álcool teria um papel mais relevante na promoção. Isso porque, após dez anos de abstinência de álcool, o risco retorna aos valores dos abstêmios; enquanto após 10 anos sem tabaco, o risco diminui apenas 50% (CASTELLSAGUÉ et al., 1999).

Todavia, há outros fatores que contribuem para a gênese do CEE. A contaminação de alimentos por nitrosaminas, potentes carcinógenos que apresentam tropismo pelo esôfago (LIJINSKY, 1992), a desnutrição e o consumo de chá quente estão envolvidos na etiologia deste câncer em alguns países, como China e Índia. No Irã, além da desnutrição e da ingestão de chá quente, o consumo de ópio é um importante fator de risco para o desenvolvimento de CEE. No

Japão, apenas o consumo de chá quente parece ter relação direta (TUYNS, 1979; SEPEHR et al., 2001). Já na região sul do Brasil, área de alta incidência, além do consumo de bebidas alcoólicas e tabaco, há também o consumo do chimarrão (chá da erva mate consumido em torno de 70°C), que já foi associado ao desenvolvimento do CEE (CASTELLSAGUÉ et al., 1999; 2000).

1.2.3 TRATAMENTO

A história natural proposta para o desenvolvimento de CEE se baseia no seguinte modelo: depois de uma exposição a longo prazo do indivíduo aos agentes etiológicos, ocorre transformação neoplásica na mucosa esofágica que é raramente detectada nos estádios iniciais. Quando a terapia curativa é possível, a cirurgia (uma mucosectomia para excisão do tumor *in situ*) é o tratamento primário (CHEN et al., 1999). Porém, na maioria dos pacientes, o tumor é detectado em estádios avançados, quando a quimiorradioterapia seguida ou não por esofagectomia são os procedimentos empregados. Entretanto, a maioria dos pacientes não apresenta uma melhora da sobrevida a longo prazo (DA COSTA et al., 2013). Ainda, no momento do diagnóstico, 40% dos pacientes apresentam metástases, inviabilizando o tratamento cirúrgico curativo, reduzindo as taxas de sobrevida (JEMAL et al., 2006).

Além do estadiamento, o grau histopatológico tumoral, baseado na atividade mitótica, atipia nuclear e na diferenciação (BOSMAN et al., 2010), também é utilizado para determinar prognóstico (SOBIN; GOSPODAROWICZ & WITTEKIND, 2009). Tumores bem diferenciados (G1) geralmente apresentam um bom prognóstico, pois se assemelham ao tecido normal. Os tumores moderadamente diferenciados (G2) são as formas intermediárias do tumor podendo apresentar um bom ou um mau prognóstico. E os tumores pouco diferenciados (G3) e indiferenciados (G4) são tumores mais agressivos, que podem promover metástase, tendo assim um pior prognóstico (JOENSUU et al., 2013). O grau de diferenciação é usado quando se avalia a necessidade de tratamento pós-operatório (JOENSUU et al., 2013).

1.3 ALTERAÇÕES MOLECULARES EM CECP E CEE

As alterações genéticas mais comuns em CECP foram caracterizadas tanto por técnicas citogenéticas quanto moleculares. A presença de alterações genéticas, especificamente perdas nas

regiões cromossômicas 3p, 8p, 9p, 15p, 18q, 22q e ganhos nas regiões 1q, 3q, 8q, 11q, 14q, 16q, 20q foram associadas a uma baixa sobrevida dos pacientes (BOCKMUHL et al., 2001; ASHMAN et al., 2003; BERGAMO et al., 2005). Além disso, estudos moleculares que visam encontrar marcadores para esses tumores têm mostrado que proteínas reguladoras do ciclo celular (P53, P16 e Ciclina D1); genes relacionados à angiogênese e à apoptose (*VEGF* e *BAX*, respectivamente); e alguns fatores de transcrição, como o NF-kB e c-MYC, são comumente afetados por alterações genéticas (MCCABE & DLAMINI, 2005).

Em 2015, o maior e mais abrangente estudo envolvendo CEECP foi publicado pelo consórcio TCGA (*The Cancer Genome Atlas*) (CANCER GENOME ATLAS NETWORK et al, 2015). Dentre os principais achados em termos de alterações genéticas, podemos destacar mutações no oncogene *PIK3CA*, perda de *TRAF3* e amplificação de *E2F1* em tumores associados ao HPV. Já em CEECP relacionado ao tabagismo, foram observadas mutações que levam à perda de função em *TP53* e alterações estruturais como deleções em homozigose e fusões cromossômicas que foram associadas à perda de função de *CDKN2A*, além de frequentes alterações no número de cópias incluindo amplificação de 3q26/28 e 11q13/22. Um subgrupo de tumores de cavidade oral com resultados clínicos favoráveis tinha mutações de ganho de função em *HRAS* e *PI3KCA*, e mutações associadas à perda de função de *CASP8*, *NOTCH1* e *TP53*. Em tumores de laringe, foi observada ativação do fator de estresse oxidativo *NFE2L2* (CANCER GENOME ATLAS NETWORK et al, 2015).

Apesar de sua alta incidência e mortalidade no Brasil e no mundo, os mecanismos que levam ao desenvolvimento de CEE ainda são pouco conhecidos. Dentre as alterações genéticas mais frequentes, destacam-se as mutações no gene supressor de tumor *TP53*, que podem ser detectadas em até 80% dos pacientes (ROSSINI et al., 2010; ZHANG et al., 2015; SAWADA et al., 2016).

Em 2017, o consórcio TCGA apresentou um estudo que incluiu análises de transcriptoma, sequenciamento de exoma e metiloma e foram observadas alterações moleculares específicas tanto para CEE quanto para ADE. Os principais achados foram amplificações de *CCND1*, *SOX2* e *TP63* em CEE e amplificações mais frequentes nos genes *ERBB2*, *VEGFA*, *GATA4* e *GATA6* em ADE. Também foi visto que os adenocarcinomas de esôfago se assemelham bastante aos adenocarcinomas gástricos do ponto de vista de perfil mutacional, porém alterações no perfil de

metilação do DNA ocorrem de forma diferenciada nesses tumores. Já CEE se aproxima de CECF do ponto de vista molecular e o principal gene mutado neste subtipo histológico de câncer de esôfago foi *TP53* (91%). Para ADE, a frequência foi de 71% (CANCER GENOME ATLAS RESEARCH NETWORK et al, 2017).

Além de mutações em genes específicos, mais recentemente, foram descritos padrões mutacionais em câncer. Mutações somáticas são encontradas no genoma de tumores (STRATTON et al., 2009), sendo oriundas de erros na maquinaria de replicação do DNA, exposição a agentes mutagênicos exógenos e endógenos, modificações enzimáticas do DNA ou de reparo de DNA deficiente (ALEXANDROV et al 2013). Em alguns tipos tumorais, uma proporção substancial de mutações somáticas é gerada através da exposição ambiental, como por exemplo, à fumaça do cigarro em câncer de pulmão e à radiação ultravioleta no câncer de pele (PFEIFER, 2010); assim como pela não correção de anormalidades no DNA, como é observado em decorrência de defeitos na maquinaria de reparo de mau-pareamento em tumores de cólon (PENA-DIAZ et al., 2012).

Processos mutacionais diferentes frequentemente geram combinações diferentes de tipos de mutação, uma espécie de padrão, designado de assinatura mutacional. Até recentemente, as assinaturas mutacionais em câncer eram exploradas por meio do pequeno número de genes recorrentemente mutados, como *TP53* (PFEIFER et al., 2002; OLIVIER et al., 2010). Hoje, foram criados algoritmos capazes de extrair esse padrão mutacional a partir de análises de sequenciamento de genoma completo ou de exoma. Esses padrões foram demonstrados inicialmente por Alexandrov e colaboradores em 2013 (ALEXANDROV et al., 2013) e, atualmente, existem mais de 90 assinaturas, disponíveis no Catálogo de Mutações Somáticas em Câncer (COSMIC) (**Figura Suplementar 1**). Há assinaturas que ocorrem em bases específicas em qualquer região do genoma, como a assinatura 4 (C>A), associada ao tabagismo. Há também aquelas que ocorrem em bases e regiões específicas do genoma. A assinatura 1 (C>T) acontece preferencialmente em contexto CpG (citosina seguida de guanina através da ligação fosfato) e está associada à idade. As assinaturas mutacionais mediadas por APOBECs (Enzima de edição de mRNA de apolipoproteína B, tipo polipeptídeo catalítico, do inglês *Apolipoprotein B MRNA Editing Enzyme Catalytic Subunit*) são evidentes no genoma de diversos tipos de câncer na forma de hipermutação C>T (processo conhecido como “kataegis”) (Assinatura 2) e de C>G (Assinatura 13) em contexto TpC (timina seguida de citosina através da ligação fosfato). A presença das

assinaturas 2 e 13 geralmente é associada à atividade das APOBEC3s, que fazem a edição do DNA genômico (ALEXANDROV et al., 2013; ZHANG et al., 2015; MORGANELLA et al., 2016). Todas as assinaturas descritas até o presente momento foram apresentadas na **Figura Suplementar 1** e sua distribuição entre as diferentes neoplasias pode ser observada na **Figura 1.11** (ALEXANDROV et al., 2020).

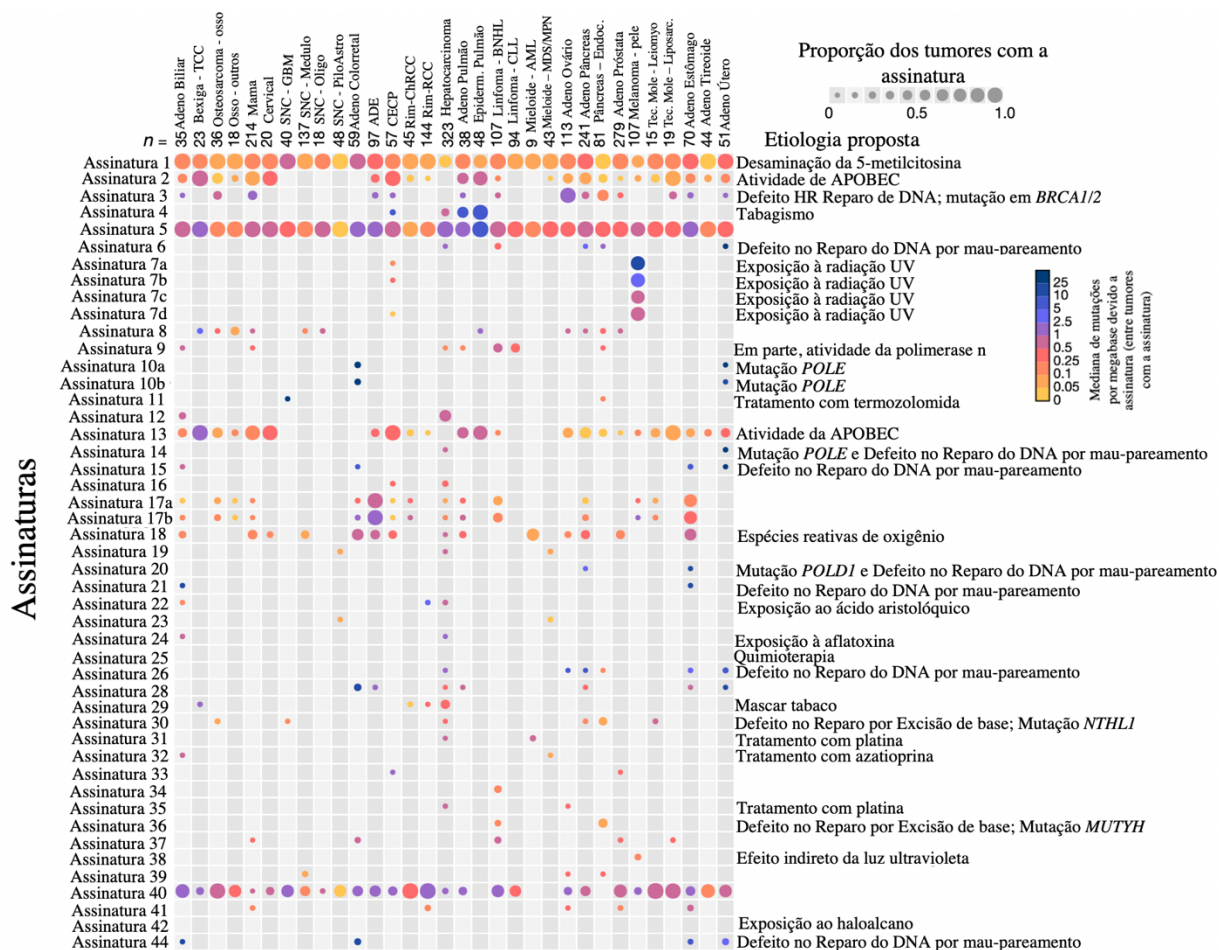


Figura 1.11: Presença das assinaturas mutacionais de substituição de base única nos diferentes tipos tumorais e a associação com os prováveis fatores responsáveis por tal assinatura. O tamanho dos círculos corresponde à proporção dos tumores com a assinatura e a escala de cor à mediana de mutações por megabase. Fonte: Adaptado de ALEXANDROV et al., 2020.

Conforme supracitado, a hipermutação C>T (kataegis) e de C>G em contexto TpC representam as assinaturas 2 e 13, respectivamente (**Figura Suplementar 1**) (COSMIC). Essas assinaturas são geradas da seguinte maneira: a fita simples de DNA, durante a replicação, pode ser alvo da ação de desaminação pelas APOBECs que então convertem citosina em uracila. A uracila é uma base estranha ao DNA, então enzimas de reparo, como UDG (uracila DNA glicosilase, também conhecida como UNG), promovem a remoção dessa base, gerando sítios abásicos. Assim, a DNA polimerase insere a base adenina (A) na fita nascente oposta ao sítio abásico ou oposta à uracila (U) (Regra “A”), gerando a mutação C>T na fita descontínua. Esse é o processo mutacional que gera a Assinatura 2. Já para a Assinatura 13, o mesmo processo acontece, entretanto, uma DNA polimerase de baixa afinidade insere citosina (C) na fita nascente oposta ao sítio abásico. Assim, a mutação C>G na fita descontínua é gerada (MORGANELLA et al., 2016) (**Figura 1.12**).

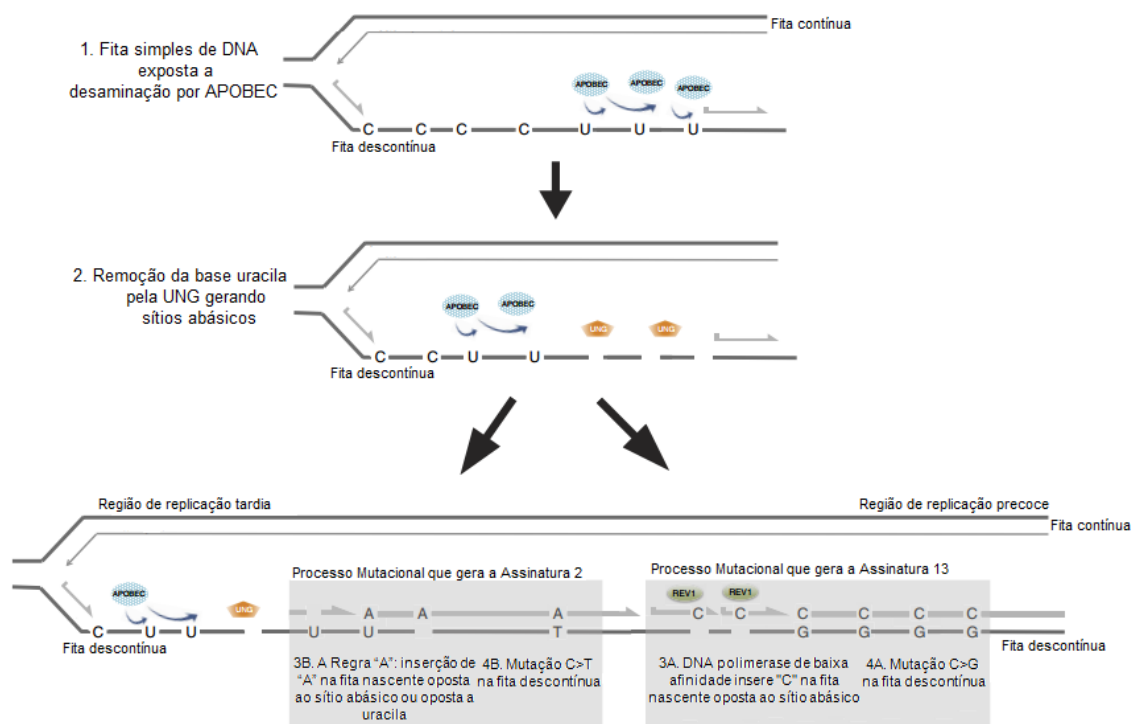


Figura 1.12: Processos mutacionais que originam as assinaturas 2 e 13 mediadas pela atividade das APOBECs. Durante a replicação do DNA, a fita simples de DNA é exposta a desaminação por APOBEC, convertendo citosina em uracila. A uracila é uma base estranha ao DNA, com isso a UNG (enzima de reparo) remove essa base, gerando um sítio abásico. Assim, a DNA polimerase

insere a adenina (A) na fita nascente oposta ao sítio abásico ou oposta à uracila (U) (Regra “A”), gerando a mutação C>T na fita descontínua (geração da assinatura mutacional 2). Já na assinatura 13, uma DNA polimerase de baixa afinidade insere citosina (C) na fita nascente oposta ao sítio abásico. Assim, a mutação C>G na fita descontínua é gerada. Fonte: Adaptado de MORGANELLA et al., 2016.

Em CEE, as assinaturas 2 e 13, mediadas por APOBECs, estão presentes em 88% e 91% dos casos, respectivamente, sendo responsável por 25% da carga de mutação em média, indicando que a ativação de APOBEC é uma etapa crucial no desenvolvimento de CEE (MOODY et al., 2021). Outro estudo mostrou que a elevada taxa de mutação em CEE era de assinaturas em contexto CpG e de APOBEC (SAWADA et al., 2016). Além disso, as transições C:G>T:A em sequências TpC representam o envolvimento da APOBEC3B no desenvolvimento de CEE, consistente com o fato desse gene estar superexpresso nesse tumor (LIN et al., 2014; MATSUMOTO et al., 2015). Além disso, já foi demonstrado para o CECP um elevado nível de RNAm da *APOBEC3B* e da *APOBEC3H* (LIU et al., 2020) e a assinatura mutacional mediada por APOBEC (FADEN et al., 2017). Porém, os mecanismos que levam à desregulação das APOBECs em tumores ainda não são completamente conhecidos.

1.4 AID/APOBEC

A família AID/APOBEC inclui as proteínas AID (desaminase de citidina induzida por ativação, do inglês *Activation Induced Cytidine Deaminase*) e APOBEC1-4 (CONTICELLO, 2008; SEVERI et al., 2011) (**Tabela 1.1**).

Tabela 1.1: *Membros da família da AID/APOBEC em humanos: localização no genoma, perfil de expressão e função.*

Nome	Posição Genômica	Perfil de Expressão	Atividade de edição conhecida	Função Fisiológica Provável
AID	12p13	Células B ativadas, baixos níveis em outros tecidos	Desaminase de DNA	Diversificação de Imunoglobulinas
APOBEC1	12p13.1	Tecidos Gastrointestinais	Desaminase de RNA ou DNA	Edição de RNAm de APOB
APOBEC2	6p21	Músculo esquelético e cardíaco	Desconhecida	Desconhecida
APOBEC3A	22q13.1	Queratinócitos	Desconhecida	Desconhecida
APOBEC3B	22q13.1	Células Sanguíneas Periféricas, células T e queratinócitos	Desaminase de DNA (com menor atividade de edição RNA)	Edição de cDNA retroviral
APOBEC3C	22q13.1	Muitos tecidos e em linhagens celulares tumorais	Desaminase de DNA	Desconhecida
APOBEC3D	22q13.1	Desconhecido	Nenhum	Desconhecida
APOBEC3D-3E	22q13.1	Desconhecido	Nenhum	Desconhecida
APOBEC3E	22q13.1	Desconhecido	Nenhum	Desconhecida
APOBEC3F	22q13.1	Muitos tecidos e provavelmente coexpressa com APOBEC3G	Desaminase de DNA (com menor atividade de edição RNA)	Edição de cDNA retroviral
APOBEC3G	22q13.1	Muitos tecidos e provavelmente coexpressa com APOBEC3F	Desaminase de DNA	Edição de cDNA retroviral
APOBEC3H	22q13.1	Desconhecido	Desconhecida	Desconhecida
APOBEC4	12q23	Desconhecido	Desconhecida	Desconhecida

Fonte: Adaptado de HARRIS & LIDDAMENT, 2004.

Essa família de proteínas altera a sequência de DNA através da conversão de citosina, 5-metilcitosina (5mC) e 5-hidroximetilcitosina (5hmC) em uracila, timina e 5-hidroximetiluracila (5hmU), respectivamente. Essas reações de desaminação são conhecidas como edição do DNA (KNISBACHER et al., 2016) (**Figura 1.13**).

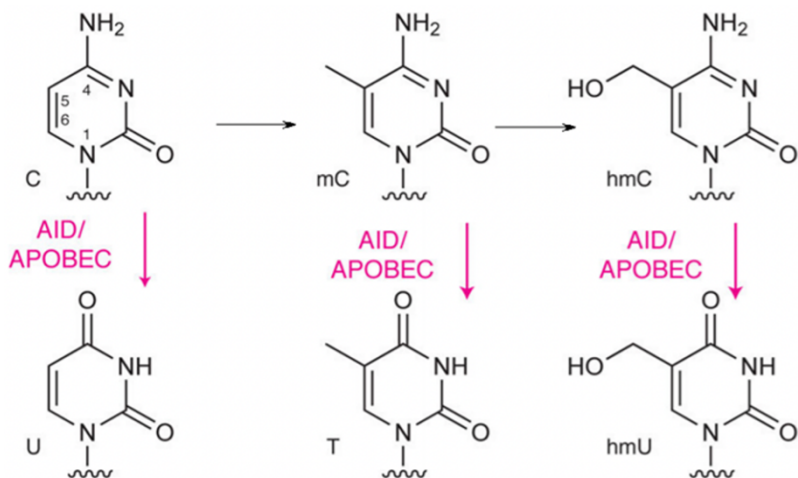


Figura 1.13: Processo de desaminação pela família AID/APOBEC conhecido como edição do DNA. A família AID/APOBEC altera a sequência de DNA através da conversão de citosina (C), 5-metilcitosina (mC) e 5-hidroximetilcitosina (hmC) em uracila (U), timina (T) e 5-hidroximetiluracila (hmU), respectivamente. Fonte: Adaptado de NABEL et al., 2012.

Tanto a uracila, a timina como a 5hmU podem ser excisadas pelas enzimas da via de reparo por excisão de bases (BER, do inglês *base excision repair*), como por exemplo, TDG (timina DNA-glicosilase) (NABEL et al., 2012). Sendo assim, uma vez que o reparo seja eficiente, em ambos os casos, uma citosina não metilada é inserida em substituição a essas bases, levando à desmetilação do DNA (NABEL et al., 2012). Porém, quando o sistema de reparo não é eficiente, estas bases não são reparadas e podem levar ao aparecimento de mutações.

Em humanos, já foram descritas 11 proteínas da família da AID/APOBEC: AID, APOBEC1, APOBEC2, APOBEC3A, APOBEC3B, APOBEC3C, APOBEC3D, APOBEC3F, APOBEC3G, APOBEC3H e APOBEC4 (SALTER et al., 2016). A AID tem função específica na imunidade adaptativa. É altamente expressa em linfócitos B ativados e é essencial na maturação de afinidade e diversificação de anticorpos (MURAMATSU et al., 2000). Por meio da edição do DNA do *locus* genômico das imunoglobulinas (Ig), há a hipermutação somática da cadeia pesada e a recombinação de troca de classe de imunoglobulinas, induzindo quebras de fitas duplas, o que culmina no rearranjo do *locus* de Ig (MURAMATSU et al., 2000; HWANG et al., 2015).

As APOBECs são predominantemente enzimas de edição de DNA, porém a *APOBEC1* foi a primeira descrita com função em humanos de edição de RNA. *APOBEC1* edita a citidina 6666

no RNAm da ApoB, que codifica uma proteína-chave no transporte de lipídios. Essa edição de RNA, no intestino humano, cria um códon de parada prematuro, gerando um peptídeo pequeno ApoB-48, importante para a homeostase lipídica (DANIELS et al., 2009).

A *APOBEC2* é primariamente expressa no coração e no músculo esquelético em mamíferos (LIAO et al, 1999). Não é essencial para o desenvolvimento de camundongos, sobrevivência e fertilidade (MIKL et al., 2005), mas é necessária para o desenvolvimento correto do músculo, por exemplo (ETARD et al., 2010; SATO et al., 2010). Já a função fisiológica da *APOBEC4* ainda é desconhecida (SEVERI et al., 2011), sendo expressa em testículo (ROGOZIN et al., 2005).

As APOBEC3 (A3A-D, A3F-H) são capazes de editar o DNA, participando na imunidade de diferentes maneiras (CASARTELLI, et al., 2010; NORMAN et al., 2011; SHARMA et al., 2015). A APOBEC3G é um fator antiviral intrínseco que inibe a replicação de vírus, como o vírus HIV (vírus da imunodeficiência humana), por meio da desaminação de resíduos de citidina. Isso causa uma hipermutação G>A na fita oposta, resultando na inativação do vírus. APOBEC3G apresenta um papel em “alertar” as células *natural killer* (NK) para presença de patógenos virais. Isso porque a desaminação da citidina induzida por patógenos age como resposta antiviral intrínseca ao regular positivamente a expressão do ligante de ativação de células NK (NKG2D, do inglês *Killer Cell Lectin Like Receptor K1*), aumentando a sensibilidade de células infectadas com HIV à lise por essas células (NORMAN et al., 2011). Na imunidade adaptativa, a edição do DNA no genoma de vírus, como HIV pela APOBEC3G, promove a ativação das células T citotóxicas CD8⁺ e o reconhecimento de células T infectadas (CASARTELLI et al., 2010). Por outro lado, o HIV contrabalança a ação da APOBEC3G através da atividade do fator de infectividade viral (Vif), que promove a degradação da APOBEC3G via proteassomo (NORMAN et al., 2011).

Além disso, as APOBEC3s são potentes inibidores de retrotransposons, induzindo sua degradação (KNISBACHER et al., 2016). Estes elementos têm sua transcrição suprimida pela metilação do DNA. Além disso, quando expressos, a APOBEC3A e a APOBEC3B, que são fatores de resistência antirretroviral inata humana, são capazes de entrar no núcleo e inibir especificamente a retrotransposição de LINE-1 (pertencente à família LINE, do inglês *Long Interspersed Nuclear Elements*) e Alu (pertencente à família SINE, do inglês *Short Interspersed Nuclear Elements*) (BOGERD et al., 2006).

Assim, a edição do DNA mediada pela família AID/APOBEC apresenta diversas funções biológicas (**Figura 1.14**) (KNISBACHER et al., 2016).

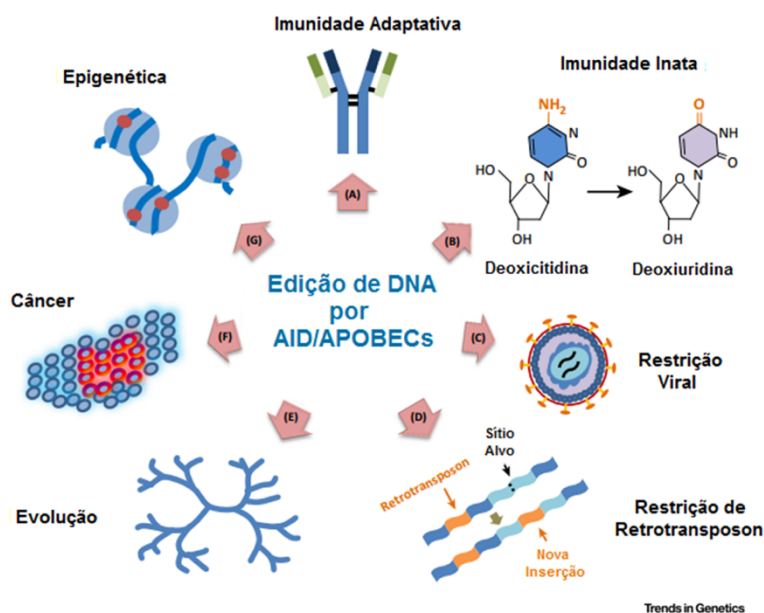


Figura 1.14: Funções biológicas da edição do DNA pela família AID/APOBEC. Papel na imunidade adaptativa: diversificação e maturação de anticorpos do sistema imune; na imunidade inata: reconhecimento de células infectadas pelas células *natural killer*, inibição de vírus e de retrotransposons, induzindo sua degradação; promoção da evolução por gerar variabilidade genética; promoção de mutações no genoma, podendo resultar no desenvolvimento tumoral; participação na desmetilação ativa do DNA, no contexto epigenético. Fonte: Adaptado de KNISBACHER et al., 2016.

Conforme mencionado anteriormente, a hipermutabilidade mediada por APOBECs é frequentemente observada em câncer. As APOBECs, especialmente APOBEC3s, são potentes fontes de mutação endógenas e já foram associadas inclusive ao mau prognóstico de pacientes oncológicos (KNISBACHER et al., 2016). Porém, os mecanismos que levam à superexpressão dessas desaminases em neoplasias não foram completamente elucidados.

Por outro lado, durante infecções virais, já foi mostrado que a expressão de AID/APOBECs pode ser regulada por citocinas. A expressão de *AID*, por exemplo, é regulada por uma série de citocinas pró-inflamatórias. TGF- β , TNF- α e IL-1 β podem estimular sua expressão via sinalização

por NF- κ B em hepatócitos humanos primários e reduzir a infectividade do HBV nas células hospedeiras (ENDO et al., 2007; JIANG et al., 2007; WATASHI et al., 2013). A expressão gênica e proteica de AID também é induzida por IL-4 e IL-6 de maneira STAT6-dependente em linfócitos B e células colônicas humanas (DEDEOGLU et al., 2004; ENDO et al., 2008). Assim, a indução da expressão de AID faz parte de uma resposta pró-inflamatória após a infecção.

APOBEC3A, *APOBEC3F* e *APOBEC3G*, mas não as *APOBEC3B* e *APOBEC3C*, são induzidas por IFN α em diversas células hematopoiéticas, como macrófagos e linfócitos T (PENG et al., 2006; KONING et al., 2009). IFN α aumenta a expressão de *APOBEC3G* em células que são alvo de infecção por HIV (CHEN et al., 2006).

A infecção por HBV em hepatócitos resulta na expressão de citocinas e das APOBEC3s. Em hepatócitos humanos primários ou hepatócitos infectados por HBV, a expressão das *APOBEC3A*, *APOBEC3B*, *APOBEC3F* e *APOBEC3G* é induzida via IFN α (BONVIN et al., 2006; LUCIFORA et al., 2014). Além disso, o tratamento de linhagens celulares de queratinócitos cervicais, que contêm HPV16 episomal, com IFN β resultou no aumento de expressão de *APOBEC3A*, *APOBEC3F* e *APOBEC3G* (WANG et al., 2014). Esses estudos indicam que a expressão de *APOBEC3s* pode ser importante para a resposta inflamatória mediada por citocinas.

Além disso, tumores CECP HPV⁺ apresentam superexpressão das *APOBECs* da família 3, com exceção da *APOBEC3A* quando comparados a tumores CECP HPV⁻ (LIU et al., 2020). A *APOBEC3H* foi identificada como um potencial preditor de prognóstico e alvo terapêutico para oncoimunoterapia em CECP, em consideração ao seu papel essencial na regulação da ativação imune mediada por CXCL10 e infiltração de células T CD8⁺ no microambiente tumoral (LIU et al., 2020).

Assim, infecções virais e a consequente produção de citocinas no microambiente tumoral poderiam explicar, ao menos em parte, a superexpressão de AID/APOBECs observada em câncer. Entretanto, apenas cerca de 15-20% dos tumores humanos são atribuídos a vírus (PARKIN, 2006). Além disso, tumores não relacionados a infecções virais, como o CEE e a maioria dos CECP, apresentam uma importante contribuição da assinatura mutacional relacionada à atividade dessas enzimas.

1.5 JUSTIFICATIVA

CEE e CECP apresentam similaridades em termos de morfologia e fatores etiológicos: são derivados de um epitélio escamoso que, uma vez exposto cronicamente ao etanol e aos componentes do tabaco, pode originar um carcinoma de células escamosas. Entretanto, o CECP compreende tumores de diferentes sítios com algumas especificidades em relação aos fatores de risco associados, como por exemplo, tumores de orofaringe associados à infecção pelo HPV. Apesar disso, o CECP é geralmente estudado como uma entidade única, subestimando as diferenças entre os tumores que abrange.

Até então, nenhum trabalho focou na análise do perfil mutacional em CECP, discriminando por subsítio e associando tal perfil à expressão de *APOBECs*. Além disso, ainda não é conhecida a assinatura mais representativa nesses tumores, aquela relacionada à atividade dessas enzimas ou a relacionada à exposição aos fatores de risco como o tabaco. Finalmente, ainda não está(ão) bem estabelecido(s) o(s) membro(s) da família AID/APOBECs envolvido(s) no estabelecimento de assinaturas mutacionais. Outra questão importante neste contexto seria elucidar os mecanismos que levam à desregulação das APOBECs nessas neoplasias.

Dessa forma, esta tese visa elucidar esses pontos pouco estudados de forma não só a contribuir para a melhor compreensão dos mecanismos moleculares que levam ao desenvolvimento de CECP, separando por subsítio tumoral (CEL, CECO, CEOF HPV-negativo e CEOF HPV-positivo), e CEE, bem como investigar os mecanismos responsáveis pelo estabelecimento das assinaturas mutacionais e pela desregulação da família das APOBECs.

2 OBJETIVOS

2.1 GERAL

Avaliar a associação entre a desregulação de APOBECs e as assinaturas mutacionais induzidas por essas enzimas em carcinoma epidermoide de cabeça e pescoço e de esôfago.

2.2 OBJETIVOS ESPECÍFICOS

- Analisar as assinaturas mutacionais e sua fração de contribuição em carcinoma epidermoide de esôfago (CEE), de laringe (CEL), de cavidade oral (CECO) e de orofaringe HPV-negativo (CEOF HPV⁻) e HPV-positivo (CEOF HPV⁺), correlacionando-as com a expressão das *APOBECs*, utilizando o banco de dados do TCGA;

- Avaliar a presença de alterações genéticas (mutações e alterações do número de cópias) nos genes que codificam as *APOBECs* em CEE, CEL, CECO, CEOF HPV⁻ e CEOF HPV⁺, utilizando o banco de dados do TCGA;

- Avaliar se alterações epigenéticas (miRNAs e metilação do DNA) estão correlacionadas com a expressão das *APOBECs* em CEE, CEL, CECO, CEOF HPV⁻ e CEOF HPV⁺, utilizando o banco de dados do TCGA;

- Avaliar a correlação entre o nível de metilação de elementos repetitivos do genoma e a expressão das *APOBECs* em CEE, CEL, CECO, CEOF HPV⁻ e CEOF HPV⁺, utilizando o banco de dados do TCGA;

- Avaliar a correlação da assinatura mutacional e da expressão das *APOBECs* com a presença de células imunes no microambiente tumoral em CEE, CEL, CECO, CEOF HPV⁻ e CEOF HPV⁺, utilizando o banco de dados do TCGA;

- Avaliar a expressão das *APOBECs* nas células tumorais e do sistema imune em CEE, por análise de *single cell*, em dado depositado no GEO.

- Avaliar a expressão das *APOBECs* em pacientes com CEE do INCA, e sua associação com os dados clínicos patológicos.

- Avaliar a expressão das *APOBEC3A*, *APOBEC3B* e *APOBEC3D* frente ao tratamento com agente desmetilante em baixas doses em linhagens de CEE;

- Avaliar a expressão das *APOBEC3A*, *APOBEC3B* e *APOBEC3D* frente ao tratamento com RNA de dupla fita em linhagens de CEE;

- Avaliar a expressão das *APOBEC3A*, *APOBEC3B* e *APOBEC3D* frente ao tratamento com interferons em linhagens de CEE.

3 MATERIAL E MÉTODOS

3.1 ANÁLISES *IN SILICO*

3.1.1 ANÁLISE DE ASSINATURA MUTACIONAL - BANCO DE DADOS TCGA

Dados processados no formato VCF, do inglês *variant call format* de sequenciamento de exoma de amostras de CEE, CEL, CECO, CEOF HPV⁻ e CEOF HPV⁺ a partir dos estudos *Esophageal Carcinoma* (TCGA, *Firehose Legacy*) para CEE (95 amostras) e *Head and Neck Squamous Cell Carcinoma* (TCGA, *Firehose Legacy*) para CEL (108 amostras), CECO (235 amostras), CEOF HPV⁻ (23 amostras) e CEOF HPV⁺ (48 amostras), depositados no portal de dados GDC (*NCI Genomic Data Commons*) foram baixados utilizando o pacote TCGABiolinks do repositório Bioconductor, em ambiente R (COLAPRICO et al., 2016). As assinaturas mutacionais de variantes de nucleotídeo único (SNVs) para cada tipo tumoral foram inferidas usando o pacote do Bioconductor Sigminer (ALEXANDROV et al., 2020), de acordo com as instruções do pacote. A extração das assinaturas mutacionais biologicamente significativas foi feita pela avaliação da decomposição da fatoração de matriz não negativa pela extração de características interpretáveis dos dados, gerando uma estimativa de assinaturas mutacionais biologicamente distintas presentes nos tumores.

Para cada assinatura, era gerada uma medida de similaridade com as assinaturas já descritas no Catálogo de Mutações Somáticas em Câncer, do inglês *Catalogue of Somatic Mutations In Cancer* (COSMIC) (disponível em <https://cancer.sanger.ac.uk/cosmic>, último acesso em 26 de outubro de 2021). Quando a medida de similaridade entre as assinaturas detectadas era indistinguível, foram representadas juntas.

3.1.2 PERFIL DE EXPRESSÃO DAS *APOBECs* – TCGA e INCA

Os dados processados de expressão gênica das *APOBECs* no formato RNA Seq V2 RSEM, obtidos a partir de sequenciamento de RNA (RNA-seq) realizado pelo consórcio TCGA, foram baixados pelo site do cBioPortal (disponível em <http://www.cbioportal.org>, último acesso em 5 de outubro de 2021) para cada subsítio tumoral CEE, CEL, CECO, CEOF HPV⁻ e CEOF HPV⁺, a

partir dos estudos *Esophageal Carcinoma* (TCGA, *Firehose Legacy*) para CEE (95 amostras) e *Head and Neck Squamous Cell Carcinoma* (TCGA, *Firehose Legacy*) para CEL (115 amostras), CECO (243 amostras) e CEOF HPV⁻ (24 amostras) e CEOF HPV⁺ (52 amostras). Utilizando o ambiente R, foi feita a análise de expressão gênica da família das *APOBECs* comparando a mucosa tumoral e a mucosa adjacente não tumoral para CEL e CECO. Já para CEOF, a análise foi feita entre tumores HPV⁻ e HPV⁺. Os níveis de expressão gênica foram determinados pelo Teste *Wilcoxon matched-pairs signed rank* para CEE, CEL e CECO e Teste *Mann-Whitney* para CEOF em ambiente R, sendo representados por gráfico de *heatmap*, utilizando o z-score.

Para CEE, também foram utilizados dados de RNA-seq previamente gerados pelo grupo a partir da análise de 14 amostras pareadas de RNA (mucosa adjacente não tumoral e tumor) de pacientes do INCA. A técnica de RNA-seq foi realizada durante o desenvolvimento da tese de doutorado do Dr Paulo Thiago de Souza Santos e os dados brutos foram reanalisados para avaliar a expressão desses genes. De forma resumida, o RNA de 14 pares de amostras de CEE (tumor e tecido adjacente não tumoral) foi extraído utilizando o RNeasy Kit (QIAGEN), de acordo com as instruções do fabricante. Esse RNA foi usado para construção de bibliotecas de cDNA utilizando-se o kit TruSeq RNA (Illumina). As leituras de boa qualidade foram alinhadas ao genoma humano pelo alinhador TOPHAT2 (KIM et al. 2013). Uma vez alinhadas, as sequências foram contadas e tabuladas por transcrito pelo programa HTSeq (ANDERS; PYL & HUBER, 2015). A partir da tabela de contagem das leituras, os níveis de expressão gênica entre a mucosa tumoral e adjacente não tumoral foram determinados pelo pacote DESeq2 1.2.5 (LOVE; HUBER & ANDERS, 2014) em ambiente R. Toda essa metodologia foi desenvolvida de acordo com SOUZA-SANTOS e colaboradores (SOUZA-SANTOS et al., 2018). Os genes das *APOBECs* foram considerados diferencialmente expressos quando valor de p ajustado $<0,001$ e $|\log FC| \geq 1$.

Uma vez realizada a análise de expressão gênica, foram feitas análises de correlação entre a contribuição da assinatura mutacional associada à atividade de *APOBECs* e a expressão das *APOBECs* também em ambiente R. A análise de correlação foi realizada pelo teste de correlação de *Spearman* e considerada estatisticamente significativa quando $p < 0,05$.

3.1.3 IDENTIFICAÇÃO DAS ALTERAÇÕES GENÉTICAS NOS GENES QUE CODIFICAM AS APOBECs - BANCO DE DADOS TCGA

Os dados referentes à presença de alterações genéticas (mutações e alterações do número de cópias) nos genes que codificam as enzimas da família AID/APOBECs (*AICDA*, *APOBEC2*, *APOBEC3A-H*) em CEE, CEL, CECO, CEOF HPV⁻ e CEOF HPV⁺ foram obtidos e visualizados a partir dos estudos *Esophageal Carcinoma* (TCGA, *Firehose Legacy*) para CEE e *Head and Neck Squamous Cell Carcinoma* (TCGA, *Firehose Legacy*) para CEL, CECO e CEOF HPV⁻ e CEOF HPV⁺ pelo site cBioPortal (disponível em <http://www.cbioportal.org>, último acesso em 5 de outubro de 2021).

As chamadas de mutação são fornecidas pela própria publicação, que no nosso caso, foi TCGA, *Firehose Legacy* (com o formato MAF, formato de anotação de mutação, derivado do VCF, do inglês *Mutation Annotation Format*), padronizando a anotação das mutações usando *Genome Nexus* (disponível em <https://www.genomenexus.org>, último acesso em 13 de outubro de 2021) (que utiliza VEP, do inglês *Ensembl Variant Effect Predictor* com *canonical UniProt transcript*, disponível em https://github.com/mskcc/vcf2maf/blob/main/data/isoform_overrides_uniprot, último acesso em 13 de outubro de 2021). A ferramenta VEP (predição do efeito da variante, disponível em <https://www.ensembl.org/info/docs/tools/vep/index.html>, último acesso em 13 de outubro de 2021) determina o efeito da variante (substituição de base única, inserção, deleção, variação no número de cópia) no gene, no transcrito, na sequência da proteína, bem como nas regiões regulatórias.

As alterações do número de cópias foram analisadas pelo algoritmo GISTIC. Esse algoritmo tenta identificar regiões significativamente alteradas de amplificação ou deleção em conjuntos de pacientes e gera chamadas específicas de número de cópia de gene/paciente putativas, que são então inseridas no portal. Assim, indica o nível do número de cópias por gene: com deleções profundas, possivelmente com deleção em homozigose; com deleções superficiais, possivelmente com deleção em heterozigose; diploide; ganho de baixo nível (algumas cópias adicionais, sendo amplas) e amplificação de elevado nível (mais cópias, geralmente focal),

utilizando os limiares de +/- 2 para amplificação e deleção de elevado nível e +/- 1 para os de baixo nível. Apenas amplificações e deleções de elevado nível foram consideradas na presente tese.

3.1.4 IDENTIFICAÇÃO DOS miRNAs CUJOS ALVOS SÃO AS APOBECs E CORRELAÇÃO DE SUA EXPRESSÃO COM OS NÍVEIS DE RNAm DAS APOBECs - BANCO DE DADOS TCGA

Foi realizada uma busca por bancos de dados de miRNAs disponíveis no site *tools 4miRs* (um site de predição de alvos, disponível em <https://tools4mirs.org>, último acesso em 5 de outubro de 2021), uma plataforma que agrupa mais de 170 métodos para análises de miRNAs. Foram utilizados os seguintes critérios para selecionar os bancos para as análises posteriores: ser disponível online; ter dados em organismo humano; ter informações sobre evidência experimental da ligação entre RNAm-miRNA; seus dados serem curados a partir de artigos científicos; conter informação sobre a sequência *seed match* (sequência de ligação perfeita com o RNAm) dos miRNAs; utilizar o algoritmo *machine learning* (aprendizado por máquina, inteligência artificial para reconhecer e extrair padrões de um grande volume de dados); e a forma de busca pelos miRNAs ser através de seus alvos (no nosso caso, as *APOBECs*). Assim, foram selecionados três bancos de dados de miRNAs: miRDB (disponível em <http://mirdb.org>, último acesso em 5 de outubro de 2021), Diana versão 5 (disponível em http://diana.imis.athena-innovation.gr/DianaTools/index.php?r=microT_CDS, último acesso em 5 de outubro de 2021) e miRmap (disponível em <https://mirmap.ezlab.org/app/>, último acesso em 5 de outubro de 2021). Para seleção dos miRNAs no banco Diana versão 5, foi avaliado o score de predição de alvo do miRNA, que variava de 0 a 1. Esse score indica a força de ligação do miRNA ao seu alvo, sendo que quanto mais próximo de 1, maior é a confiança de que aquele miRNA interage com o alvo selecionado na busca. A linha de corte do score foi de 0,7. No miRDB, a predição maior do que 80 (escala de 0 a 100) tem maior probabilidade de ser real e inferior a 60 é necessário ter cautela e recomendável ter outras evidências para confirmar tal achado. Assim, nós utilizamos o score de 70 para o miRDB. Abaixo disso, os miRNAs encontrados não eram selecionados para as análises posteriores.

A partir da escolha dos bancos de miRNAs a serem consultados e da seleção de miRNAs que potencialmente tenham as *AID/APOBECs* como alvos, foi avaliada a interseção entre eles. Apenas miRNAs que fossem comuns aos três bancos para cada gene-alvo foram selecionados. Os dados processados do sequenciamento de miRNAs (miR-Seq), contendo a leitura por milhão de cada miRNA para cada indivíduo com CEE, CEL, CECO, CEOF HPV⁻ e CEOF HPV⁺ do consórcio TCGA foram baixados do portal de dados GDC (*NCI Genomic Data Commons*, disponível em <https://portal.gdc.cancer.gov/repository>, último acesso em 5 de outubro de 2021). Com isso, foi possível avaliar a correlação entre a expressão dos miRNAs e das *AID/APOBECs* para todos os tumores estudados. Quando os dados obedeciam à distribuição normal, foi feita a análise de correlação por Pearson, e quando não, por Spearman. Os resultados foram considerados significativos quando $p < 0,05$.

Para esta análise, o nosso foco foi avaliar as correlações inversas, uma vez que seriam indicativas de um mecanismo direto de regulação do gene-alvo pelo miRNA.

3.1.5 METILOMA – TCGA e INCA

Em trabalho prévio do grupo, foi realizada a técnica de microarranjo para avaliação do perfil global de metilação do DNA em CEE, CEL, CECO e CEOF HPV⁻ e CEOF HPV⁺ (SOARES-LIMA et al., 2021). Ao todo, foram incluídas 16 amostras adjacentes não tumorais e 24 amostras tumorais de CEE; 12 amostras adjacentes não tumorais e 20 amostras tumorais de CEL; 7 amostras adjacentes não tumorais e 15 amostras tumorais de CECO e 9 amostras não tumorais (amostras de tonsila de indivíduos sem câncer) e 23 amostras tumorais de CEOF (sendo 18 HPV⁻ e 5 HPV⁺). Resumidamente, o *chip* usado para o microarranjo de metilação foi o Infinium HumanMethylation450 BeadChip Kit (Illumina, Califórnia, Estados Unidos). Após a reação de conversão com bissulfito de sódio, utilizando EZ DNA Methylation Kit® (Zymo Research, Estados Unidos), 250 ng de DNA tratado foram preparados para a aplicação no chip, de acordo com as instruções do fabricante. Todas as análises do microarranjo foram feitas com auxílio do software RStudio com diferentes pacotes do Bionconductor. A metodologia foi desenvolvida de acordo com SOARES-LIMA e colaboradores (SOARES-LIMA et al., 2021). O valor de beta (beta-

value) representa a razão entre a intensidade do alelo metilado e a soma das intensidades dos alelos metilado e não metilado e foi utilizado para as análises posteriores.

Foi realizada a comparação do perfil de metilação de todas as sondas do microarranjo com anotação para as *AID/APOBECs* entre os tecidos tumorais e não tumorais para cada subsítio tumoral. Foram consideradas diferencialmente metiladas sondas com um valor de p ajustado < 0,001 nesta comparação.

Além dos dados gerados pelo grupo, também foram utilizados dados processados de valores de beta para os diferentes subsítios tumorais gerados pela mesma técnica pelo consórcio TCGA, obtidos no portal de dados GDC (*NCI Genomic Data Commons*, disponível em <https://portal.gdc.cancer.gov/repository>, último acesso em 5 de outubro de 2021). As matrizes de valores de beta das casuísticas foram utilizadas para avaliar a correlação entre os níveis de metilação de diferentes regiões dos genes que codificam as enzimas da família *AID/APOBECs* e a expressão gênica dessas enzimas. Quando os dados obedeciam à distribuição normal foi feita a análise de correlação por *Pearson*, e quando não, por *Spearman*. Os resultados foram considerados estatisticamente significativos quando $p < 0,05$.

3.1.6 ANÁLISE DE CORRELAÇÃO DO NÍVEL DE METILAÇÃO DE ELEMENTOS TRANSPONÍVEIS COM A EXPRESSÃO DAS *APOBECs* EM CEE, CEL, CECO, CEOF HPV⁻ e CEOF HPV⁺ - BANCO DE DADOS TCGA

Inicialmente, foi realizada a análise de predição de metilação dos elementos transponíveis a partir dos sítios CpG flanqueadores, pelo pacote REMP (baseado no algoritmo *Random Forest*), em ambiente R (ZHENG et al., 2017). Essa predição foi feita a partir dos dados de microarranjo de metilação para todos os tumores avaliados, CEE, CEL, CECO, CEOF HPV⁻ e CEOF HPV⁺, gerados com a plataforma Infinium HumanMethylation450 pelo consórcio TCGA. O algoritmo identifica padrões de níveis de metilação, predizendo a metilação a partir de sondas que estão dentro ou próximas aos elementos transponíveis. Primeiro, é feita a identificação dos elementos que serão utilizados para predição da metilação através do **INITREMP**. As informações dos elementos foram obtidas a partir do genome browser UCSC (Instituto Genômico da Universidade da Califórnia Santa Cruz, do inglês *University of California Santa Cruz Genomics Institute*) e o

genoma utilizado foi o hg19. Em seguida, os dados foram trabalhados a partir da matriz de beta para fazer a predição por meio do **GROOMETHY**, utilizando os valores de metilação obtidos no portal de dados GDC (*NCI Genomic Data Commons*, disponível em <https://portal.gdc.cancer.gov/repository>, último acesso em 5 de outubro de 2021). Para aquelas amostras que não apresentaram valores para sondas específicas, foi calculado um valor predito, baseado nos valores gerados pelas demais amostras. Assim, com o **REMP**, a predição da metilação desses elementos foi feita por amostra, utilizando uma janela de predição de 1000 pares de base e o algoritmo *Random Forest*. Com o **rempegATE**, foi feita uma média dos valores de metilação para cada elemento com pelo menos dois sítios CpG preditos (2). Por fim, foi feita a análise de correlação da metilação de cada elemento (LINE, ALU e LTR – região de repetição terminal longa, do inglês *long terminal repeat*) com a expressão das *APOBECs* por paciente de cada subsítio tumoral. Quando os dados obedeciam à distribuição normal foi feita a análise de correlação por *Pearson*, e quando não, por *Spearman*. Os resultados foram considerados estatisticamente significativos quando $p < 0,05$.

3.1.7 CARACTERIZAÇÃO DO MICROAMBIENTE TUMORAL E CORRELAÇÃO COM A ASSINATURA MUTACIONAL E EXPRESSÃO DAS *APOBECs* EM CEE, CEL, CECO, CEOF HPV⁻ e CEOF HPV⁺ - BANCO DE DADOS TCGA

CIBERSORT (<http://cibersort.stanford.edu/>) é uma ferramenta analítica do laboratório Alizadeh Lab desenvolvido por NWEMAN et al, 2015 para estimar a abundância de tipos celulares em uma população de células, usando dados de expressão gênica. Nessa tese, foi estimada a proporção dos diferentes subtipos celulares do sistema imune (células T, macrófagos, mastócitos, plasmócitos, etc.) a partir de dados de expressão gênica de cada paciente de CEE, CEL, CECO, CEOF HPV⁻ e CEOF HPV⁺. Os *raw counts* foram obtidos no portal de dados GDC (*NCI Genomic Data Commons*, disponível em <https://portal.gdc.cancer.gov/repository>, último acesso em 5 de outubro de 2021), bem como os dados processados de expressão gênica das *APOBECs* no formato RNA Seq V2 RSEM nesses tumores pelo site cBioPortal (disponível em <http://www.cbioportal.org>, último acesso em 5 de outubro de 2021). Em seguida, foi realizada a análise de correlação da contribuição das assinaturas mutacionais mediadas pelas *APOBECs* (item

3.1) e da expressão das *APOBECs* com os dados estimados de abundância de células do sistema imune. Quando os dados obedeciam à distribuição normal foi feita a análise de correlação por *Pearson*, e quando não, por *Spearman*. Os resultados foram considerados estatisticamente significativos quando $p < 0,05$.

3.1.8 SINGLE CELL – CEE - BANCO DE DADOS GEO

Foram utilizadas contagens normalizadas de dados depositados de CEE no *Gene Expression Omnibus* (GEO), sob o número GSE160269. Para os tumores de CECP esta análise não foi realizada por não haver *single cell* dos subsítios separados. No trabalho utilizado, foi feito o RNA-seq de células únicas (*single cell*) de 60 amostras de CEE e 4 tecidos adjacentes não tumorais de 60 indivíduos utilizando a plataforma 10X Genomics, Illumina NovaSeq 6000 (*Homo sapiens*). Os dados brutos estão disponíveis em <https://bigd.big.ac.cn/gsa-human/>, sob o número de acesso de identificação HRA000195. A metodologia está descrita em ZHANG e colaboradores (ZHANG et al., 2021). Em resumo, a suspensão de célula única foi corada com CD45-FITC (BD Biosciences, 555482, diluição 1: 100) e classificada em células imunes (CD45⁺) ou não imunes (CD45⁻) usando um citômetro de fluxo FACSAria (BD Biosciences, BD FACSDiva (versão 8.0.1)). As células classificadas foram examinadas e contadas antes de serem submetidas à plataforma 10x Genomics Chips. Foram recuperadas 7000 células de cada canal e utilizados kits Chromium Single Cell 5' Reagent (10x Genomics) para preparar bibliotecas de sequenciamento de RNA de transcriptoma total. Os dados de RNA-seq de células únicas foram processados usando Cell Ranger Single-Cell Software Suite (10x Genomics, versão 2.1.0) com parâmetros padrão, alinhados ao genoma de referência GRCh38 e as matrizes de expressão gênica brutas foram geradas para cada amostra. Em média, foram recuperadas 1804 células CD45⁻ e 1841 células CD45⁺ de cada amostra com profundidade de sequenciamento de 143.559 e 80.615 de leituras por célula, respectivamente. O pacote Seurat (versão 2.3.4) foi usado para filtragem de qualidade e análise posteriores. Os tipos celulares foram anotados com base na expressão de marcadores conhecidos para cada célula (epitelial, fibroblasto, endotelial, pericito, células reticulares fibroblásticas – FRC, células T, células B e células mieloides). As células imunes contaminantes no conjunto de dados CD45⁻ e as células não imunes no conjunto de dados CD45⁺ foram removidas

antes do agrupamento com base em seus padrões de expressão (ZHANG et al., 2021). Em seguida, foi comparada a expressão das *APOBECs* nos diferentes tipos celulares através do teste não paramétrico *Kruskal Wallis* e o pós-teste de *Dunn*. Os resultados foram considerados estatisticamente significativos quando $p < 0,05$.

3.2 CEE COMO MODELO EXPERIMENTAL - INCA

Foi incluída nesse estudo uma coorte retrospectiva de 49 pacientes portadores de CEE primário. Foram considerados os seguintes critérios para a inclusão dos pacientes no estudo:

- Casos cujo laudo histopatológico confirmou o diagnóstico de CEE revisto na Divisão de Patologia (DIPAT) do INCA, no período de 2006 a 2015;
- Pacientes que ainda não haviam sido submetidos a qualquer tratamento na ocasião da biópsia;
- Pacientes cadastrados no Banco Nacional de Tumores (BNT) – INCA, onde as amostras eram armazenadas.

Além dos procedimentos vinculados ao BNT-INCA, o projeto de pesquisa foi aprovado pelo Comitê de Ética em Pesquisa da instituição, sob o número 116/11 (Anexo A). Foram coletadas biópsias de CEE e do tecido adjacente não tumoral de cada paciente incluído no trabalho. Todos os pacientes assinaram um termo de consentimento livre e esclarecido (Apêndice A). Além disso, as informações sobre os pacientes foram coletadas a partir de seus prontuários com o preenchimento de uma ficha, incluindo dados sobre tabagismo e etilismo (Apêndice B).

3.2.1 EXTRAÇÃO, QUANTIFICAÇÃO E DETERMINAÇÃO DA PUREZA DE RNA

3.2.1.1 EXTRAÇÃO DE RNA – AMOSTRAS CEE

As amostras de CEE e de tecido adjacente não tumoral pareado coletadas foram processadas no laboratório do BNT-INCA, como determina o regimento da instituição. Para a extração do RNA total, foi utilizado o RNeasy Mini kit (Qiagen), seguindo as orientações do fabricante:

- Um fragmento da biópsia de até 30 mg foi colocado em um microtubo de 1,5 mL junto com 200 μ L do tampão RLT com beta-mercaptoetanol. Após macerar o tecido, com auxílio de um pistão, foram adicionados mais 400 μ L do tampão RLT e posteriormente foram centrifugados em microcentrífuga por 3 min à velocidade de 12.000 x g para sedimentar o material não lisado. Em seguida, o sobrenadante foi transferido para um novo tubo e um volume de etanol 70% foi adicionado. Após homogeneização, 700 μ L da solução foram transferidos para uma coluna de purificação alocada em um tubo coletor de 2 mL e centrifugados por 15 s a 8.000 x g. Ao término, o líquido do tubo coletor foi descartado e o procedimento foi repetido até passar toda a solução pela coluna. Em seguida, 700 μ L do tampão RW1 foram adicionados à coluna de purificação e centrifugados por 15 s a 8.000 x g. Ao término, o líquido do tubo coletor foi descartado e 500 μ L do tampão RPE foram adicionados à coluna de purificação e, posteriormente, centrifugados por 2 min a 8.000 x g. Em seguida, a coluna foi transferida para um novo tubo coletor de 2 mL e centrifugada por 1 min a 12.000 x g. Após essa centrifugação, a coluna foi transferida para um microtubo de 1,5 mL e 30 μ L de água livre de RNase foram adicionados no centro da membrana, sem encostar a ponteira na mesma, a fim de eluir o RNA total. Após 1 minuto à temperatura ambiente, os tubos foram centrifugados por mais 1 min a 8.000 x g. A coluna foi descartada e posteriormente os tubos contendo o RNA total extraído foram armazenados à -80°C.

3.2.1.2 QUANTIFICAÇÃO E DETERMINAÇÃO DA PUREZA DAS AMOSTRAS DE RNA

A quantificação e determinação da pureza das amostras de RNA foram feitas com o auxílio do equipamento NanoDrop (Uniscience), um espectrofotômetro simples, capaz de mensurar a concentração de ácidos nucleicos e proteínas em solução. Para a realização da quantificação, é necessário configurar os parâmetros do equipamento para mensurar ácidos nucleicos, determinados pela absorvância em comprimento de onda de 260 nm. O cálculo da concentração do RNA tomou por base a informação de que uma unidade de absorvância corresponde a uma concentração de 40 μ g / mL de RNA. Em seguida, aplica-se 1 μ L de RNA total extraído no leitor óptico do aparelho. O resultado mostra automaticamente a concentração do RNA em μ g/ μ L e verifica o grau de pureza da amostra a partir da razão entre a absorvância em 260 nm e a absorvância em 280 nm (comprimento de onda necessário para medir a concentração de proteínas

contaminantes na solução). As amostras foram consideradas puras quando esta razão era igual ou superior a 1,7.

3.2.2 REAÇÃO DE TRANSCRIÇÃO REVERSA (RT)

Para a análise da expressão gênica pela reação em cadeia da polimerase (do inglês, *Polymerase Chain Reaction* - PCR) foi necessário primeiramente transcrever o RNAm recém extraído em um DNA complementar (cDNA). O cDNA possui uma sequência de desoxirribonucleotídeos complementar à sequência do RNAm. Este procedimento é essencial para a PCR porque as moléculas de RNA são muito instáveis e pelo fato da DNA polimerase, enzima responsável pela reação, ter a capacidade de polimerizar fitas complementares de DNA a partir, somente, de outra fita de DNA.

A síntese de cDNA foi feita pela reação de transcrição reversa (RT), na qual o RNA total extraído foi incubado com a enzima transcriptase reversa SuperScript II® (Invitrogen). Em um microtubo foram adicionados 250 ng de primer randômico (PROMEGA), 500 ng de RNA alvo, 1 µL do set de dNTPs a 10 mM e água livre de RNase suficiente para completar o volume final de 20 µL de reação e subsequente incubação por 5 minutos a 65°C. Em seguida, foram adicionados 4 µL do tampão da enzima transcriptase reversa (5X) fornecido pelo fabricante (Invitrogen) e 2 µL de DTT 0,1M e subsequente incubação por 2 min a 25°C. Em seguida, foi adicionado 1 µL de SuperScript II® (200 U) e as reações foram incubadas no termociclador por 10 min a 25°C e 50 min a 42°C. A inativação da enzima foi feita incubando a reação a 70°C por 15 min. Após a síntese, o cDNA foi armazenado a -20°C.

3.2.3 REAÇÃO EM CADEIA DA POLIMERASE QUANTITATIVA (PCRq)

A PCRq ou PCR em tempo real é um tipo de PCR em que podemos observar a geração de novas cópias de determinada sequência do DNA ao longo da reação, ou seja, em tempo real. Esta metodologia foi realizada em termociclador equipado com detectores de fluorescência que acompanham o processo de amplificação pela captação de sinal fluorescente emitido por moléculas específicas (fluoróforos), quando estas interagem com a dupla-fita do DNA. O

fluoróforo mais comumente utilizado nas práticas de PCRq é o SYBR Green, que emite fluorescência após interagir fisicamente com a dupla-fita de DNA.

A PCR quantitativa (PCRq) foi otimizada para cada par de oligonucleotídeos a fim de avaliar a expressão gênica das *APOBECs* nas amostras pareadas de CEE. Foi utilizado o sistema Rotor-Gene Q (Qiagen) e o reagente QuantiFast SYBR Green PCR Kit (Qiagen). O gene *GAPDH* foi utilizado como gene-referência. As sequências dos oligonucleotídeos específicos para cada gene encontram-se na **Tabela 3.1**.

Tabela 3.1: Sequências de oligonucleotídeos utilizadas para amplificação dos genes das *APOBECs* e dos controles positivos dos tratamentos *in vitro* e o *GAPDH* (gene - referência) na PCRq.

Genes	Sequências dos Oligonucleotídeos (5'-3')	Referências
<i>APOBEC3A</i>	Senso: CAATGGCACCTCGGTCAAGA Anti-senso: AGAAGATTCTTAGCCTGGTTGTGT	Desenhado pelo grupo
<i>APOBEC3B</i>	Senso: ACCCATCCTCTATGGTCGGA Anti-senso: GCTTGAAATACACCTGGCCTC	Desenhado pelo grupo
<i>APOBEC3D</i>	Senso: CGCAAGCAGGAAGTGAAACC Anti-senso: CGCTCCATCGGATTTCTGATCT	Desenhado pelo grupo
<i>APOBEC3F</i>	Senso: AGGAGATTCTCAGAAACCCGATG Anti-senso: TCTTCCAGGAGACAGGTGAGT	Desenhado pelo grupo
<i>APOBEC3G</i>	Senso: TTGTTGCCCGCCTCTACTAC Anti-senso: CCAACAGTGCTGAAATTCGTCAT	Desenhado pelo grupo
<i>APOBEC3H</i>	Senso: AGCTGTGGCCAGAAGCAC Anti-senso: CGGAATGTTTCGGCTGTT	REFSLAND et al., 2010
<i>DDX58</i>	Senso: CCAGCATTACTAGTCAGAAGGAA Anti-senso: CACAGTGCAATCTTGTCATCC	ROULOIS et al., 2015
<i>MDA5</i>	Senso: CACTTCCTTCTGCCAAACTTG Anti-senso: GAGCAACTTCTTTCAACCACAG	ROULOIS et al., 2015
<i>IRF7</i>	Senso: GGCTGGAAAACCAACTTCCG Anti-senso: GCCTTCTCGCCAGCACA	Desenhado pelo grupo
<i>GAPDH</i>	Senso: CAACAGCCTCAAGATCATCAGCAA Anti-senso: AGTGATGGCATGGACTGTGGTTCAT	DE A SIMÃO et al., 2011

As reações de PCRq foram feitas em duplicata e dois controles negativos foram incluídos em cada reação. A PCRq foi realizada em um rotor de 72 poços e em cada reação foram adicionados 5 µL de tampão QuantiFast SYBR Green (Qiagen), 0,3 µL de oligonucleotídeos

específicos a 10 μ M, 1 μ L de cDNA diluído 1:20 e água deionizada autoclavada para completar o volume final de 10 μ L. A reação de amplificação foi realizada através das seguintes condições:

- Pré-desnaturação: 5 minutos a 95°C;
- 40 ciclos de desnaturação (5 segundos a 95°C) e anelamento e extensão (etapas integradas – 10 segundos a 60°C), com posterior captura de fluorescência;
- Curva de dissociação: 65°C - 95°C.

Após a amplificação dos produtos, foi realizada a análise da curva de dissociação gerada pelo software do termociclador. Essa curva é importante para verificar se os produtos observados na PCRq são específicos para os oligonucleotídeos analisados, para verificar se há formação de estruturas secundárias entre os oligonucleotídeos e contaminação da reação (controle negativo amplificando). Um gráfico representativo da curva de dissociação está representado na **Figura 3.1**.

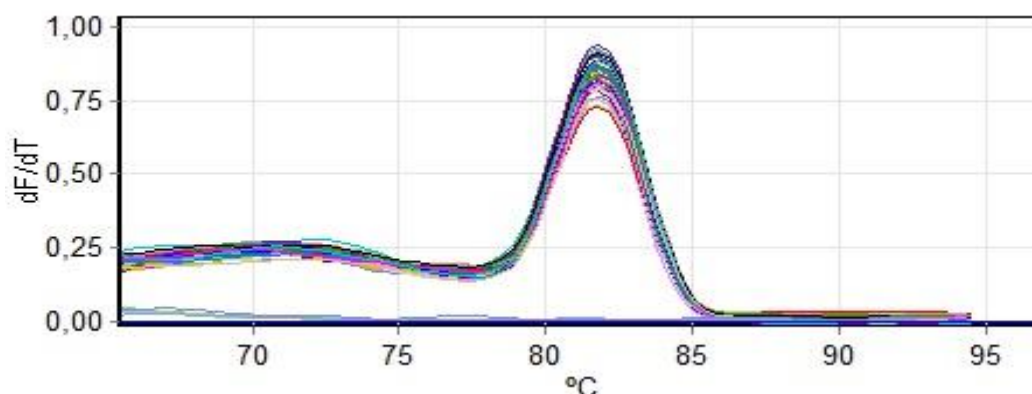


Figura 3.1: Exemplo de curva de dissociação de PCRq, mostrando a especificidade dos produtos gerados e sem contaminação da reação.

Após a reação, foi feita a quantificação relativa do RNAm de cada gene, utilizando o *GAPDH* como controle interno, pelo método de $2^{-\Delta CT}$ (SCHMITTGEN & LIVAK, 2008). Durante a reação de amplificação, é gerado um gráfico em que o eixo Y representa o sinal de fluorescência captado e o eixo X, o número de ciclos da reação. No primeiro ciclo da reação de PCR, há pouca emissão de fluorescência. É justamente isso que define a linha de base do gráfico de amplificação. Ao decorrer da reação, há um aumento do sinal de fluorescência, indicando um acúmulo de produto de PCR. Esse acúmulo de produto de PCR é representado por uma curva exponencial. Durante a

fase exponencial, é determinado um limiar de detecção de fluorescência fixo acima da linha de base (*threshold*). A partir desse limiar, é obtido o parâmetro C_T (ciclo de *threshold*), definido como o número de ciclos, inteiro ou não, em que a fluorescência ultrapassou o limiar previamente estabelecido (**Figura 3.2**). Como o experimento era feito sempre em duplicata, foi calculada a média dos C_{TS} tanto para o gene-alvo (no caso, as *APOBECs*) como para o gene-referência (*GAPDH*). A diferença entre as médias dos C_{TS} (ΔC_T) de cada gene de interesse e a do *GAPDH* foi calculada pelo programa Microsoft Excel e os valores de quantificação relativa foram expressos pelas fórmulas $2^{-\Delta C_T}$ para as *APOBECs* e $1,6^{-\Delta C_T}$ para *APOBEC3H*, também calculados pelo programa Microsoft Excel.

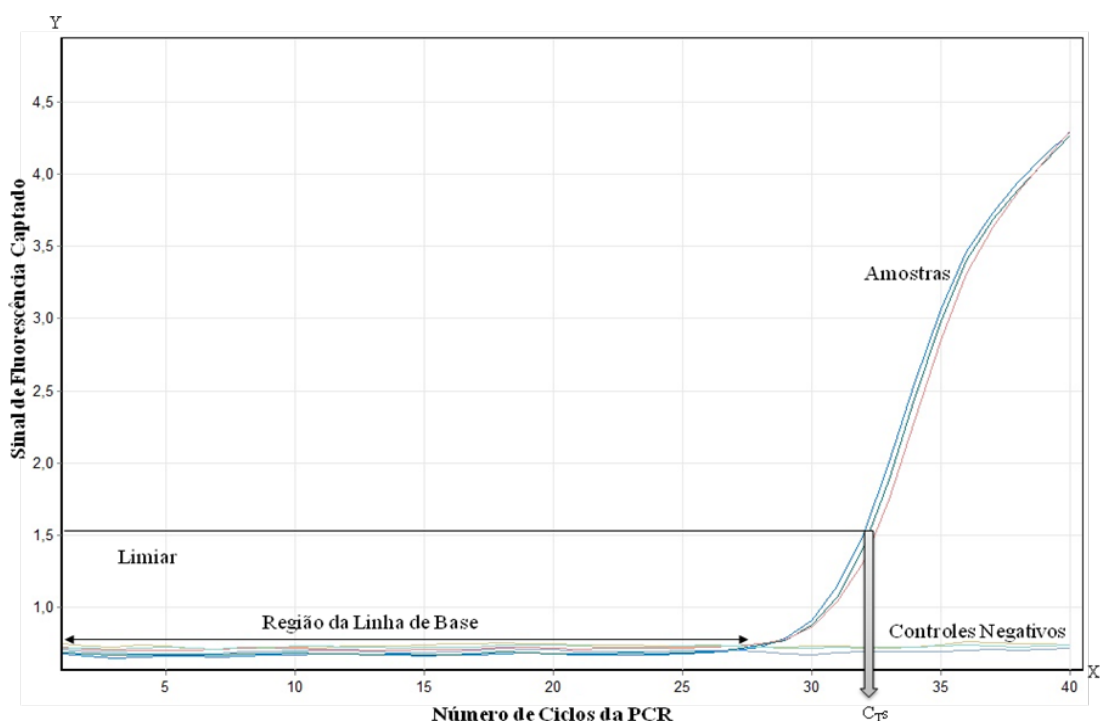


Figura 3.2: Gráfico de detecção de fluorescência de PCR em tempo real.

3.2.4 EXPERIMENTOS *IN VITRO*

3.2.4.1 LINHAGENS CELULARES

As linhagens humanas derivadas de CEE TE1 e TE13, obtidas no banco de linhagens celulares ATCC, foram utilizadas para a realização dos experimentos *in vitro* (**Tabela 3.2**). A

autenticidade das linhagens foi confirmada por Powerplex 18D STR System (Promega). As linhagens celulares foram periodicamente testadas com MycoAlert®Mycoplasma Detection Kit (Cambrex Bio Science, EUA) para a certificação de ausência de infecção por micoplasma.

As linhagens foram cultivadas em meio RPMI 1640 suplementado com 10% de soro fetal bovino (SFB) e 1% do coquetel de penicilina/glutamina/estreptomicina (GIBCO BRL, Gaithersburg, MD, EUA), e foram mantidas em estufa a 37° C sob atmosfera de 5% de CO₂.

Tabela 3.2: *Caracterização das linhagens celulares TE1 e TE13.*

Linhagem celular	Indivíduo			Tumor	
	Idade	Sexo	Histologia	Estágio	Diferenciação
TE1	58	M	CEE	II	Bem
TE13	65	F	CEE	IV	Pouco

Legenda: M: sexo masculino; F: sexo feminino; CEE: Carcinoma Epidermoide de Esôfago.

Fonte: ATCC - <https://www.atcc.org>

Durante o cultivo celular as linhagens foram repicadas e plaqueadas da seguinte forma:

- Retirada do meio de cada linhagem;
- Lavagem com PBS e remoção do mesmo;
- Adição de solução de tripsina a 0,25%, com EDTA e subsequente incubação durante 5 min, a 37°C, 5% CO₂ para soltar as células aderidas;
- Adição do meio de cultura para inibir a tripsina (o meio contém SFB capaz de inibir a tripsina);
- Contagem das células na câmara de Neubauer contendo lamínula utilizando azul de tripan (diluição do corante na proporção de 1:1 com a suspensão celular);
- O número de células foi determinado pela contagem das células não coradas em cada quadrante da câmara de Neubauer e os valores encontrados foram somados e divididos pelo número de quadrantes. Em seguida, este valor foi multiplicado por 2 (fator de diluição) e por 10⁴ para se obter o número de células/mL. As células vivas de membrana intacta não se coram, pois o

azul de tripan não atravessa a membrana e assim não é incorporado. Já nas células não viáveis, esse corante atravessa a membrana e é incorporado. Isso é observado ao microscópio óptico. Assim foi obtida a quantidade de células viáveis;

- Plaqueamento celular de acordo com o tratamento.

3.2.4.2 TRATAMENTO COM AGENTE DESMETILANTE A BAIXAS DOSES

O tratamento com o agente desmetilante a baixas doses foi feito com o objetivo de mimetizar a reativação de elementos transponíveis observada em câncer, conforme já descrito na literatura (ROULOIS et al., 2015). Para tal, a droga 5-AZA-2'-deoxicitidina (5-AZA-CdR), também conhecida como decitabina (Sigma-Aldrich), foi eluída em PBS 1X na concentração de 10 mM, aliquotada e armazenada a -80°C. Foram plaqueadas $6,0 \times 10^4$ de células/poço para a TE1 e $1,0 \times 10^5$ de células/poço para TE13 em placas de 6 poços. Depois de 24h, 5-AZA-CdR foi adicionada ao meio de cultura na concentração final de 300 nM. Após 24h de incubação com a droga, o meio com 5-AZA-CdR foi removido e meio de cultura sem a droga foi adicionado. As células foram mantidas em cultura por 96h na estufa a 37°C sob 5% de CO₂. Em seguida, as células foram coletadas, as células viáveis foram contadas por exclusão por coloração com azul de tripan e posterior análise de expressão gênica das *APOBEC3A*, *APOBEC3B* e *APOBEC3D*, além dos controles positivos do experimento: *DDX58* (do inglês *DExD/H-Box Helicase 58*), *MDA5* (do inglês *Melanoma Differentiation-Associated Protein 5*) e *IRF7* (do inglês *Interferon Regulatory Factor 7*) (conforme descrito no item 3.2.3). *DDX58* e *MDA5* formam receptores de reconhecimento de padrão viral, enquanto o *IRF7* é um fator regulatório da via de interferon (CHIAPPINELLI et al., 2015; ONOMOTO et al., 2021). Nos controles, o solvente da 5-AZA-CdR, PBS 1X estéril, foi adicionado ao meio de cultura, no mesmo volume utilizado para o tratamento. Foram realizados três experimentos independentes em triplicata para cada linhagem celular (**Figura 3.3**).

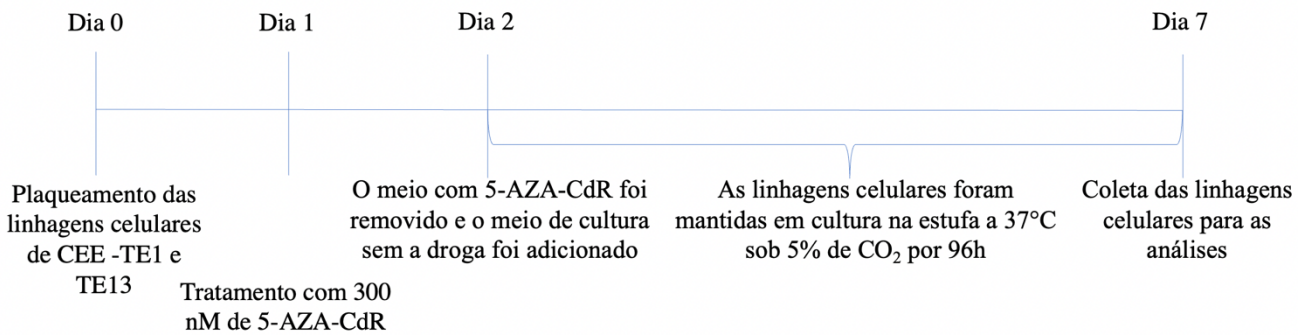


Figura 3.3: Esquema do tratamento com 300 nM de 5-AZA-CdR (decitabina) por 6 dias de tratamento. Inicialmente, as linhagens celulares de CEE TE1 e TE13 foram plaqueadas em placas de 6 poços, com $6,0 \times 10^4$ de células/poço e $1,0 \times 10^5$ de células/poço, respectivamente. Após 24h, a 5-AZA-CdR a 300 nM foi adicionada ao meio de cultura. Após 24h de incubação com a droga, o meio com 5-AZA-CdR foi removido e meio de cultura sem a droga foi adicionado. As linhagens celulares foram mantidas em cultura na estufa a 37°C sob 5% de CO_2 por 96h. Após 96h, as células foram coletadas, as células viáveis foram contadas por exclusão por coloração com azul de tripan e posterior análise de expressão gênica de *APOBEC3A*, *APOBEC3B* e *APOBEC3D* e dos controles positivos do experimento (*DDX58*, *MDA5* e *IRF7*, Tabela 3.1). No controle, foi adicionado apenas o solvente da droga, PBS 1X no meio de cultura.

3.2.4.3 TRATAMENTO COM RNA DE DUPLA FITA

O tratamento com RNA de dupla fita foi feito com POLY I:C, um análogo sintético de RNA de dupla fita (do inglês, *Polyinosinic-polycytidylic acid*), um padrão molecular associado com a infecção viral. Esse análogo ativa os receptores de reconhecimento de padrão antiviral, como por exemplo, DDX58-MDA5, levando à ativação de vias inflamatórias, de resposta a interferon, como por exemplo, IRFs (fatores regulatórios de interferon) (CHIAPPINELLI et al., 2015; ONOMOTO et al., 2021; Invivogen, disponível em <https://www.invivogen.com/polyic-hmw>, último acesso em 22 de setembro de 2021). Para tal, o análogo POLY I:C (Invivogen) foi eluído em água ultrapura, obtendo uma solução estoque de 1 mg/mL, alíquotada e armazenada a 20°C . Foram plaqueadas $1,0 \times 10^5$ de células/poço para a TE1 e $3,0 \times 10^5$ de células/poço para a TE13 em placas de 6 poços. Depois de 24h, as células foram tratadas com 10 $\mu\text{g}/\text{mL}$ de POLY I:C

por 24h, conforme já descrito na literatura (LIM et al., 2009; MATIJEVIĆ, T. & PAVELIĆ, J., 2012). Em seguida, as células foram coletadas, as células viáveis foram contadas por meio da exclusão por coloração com azul de tripan e posterior análise de expressão gênica das *APOBEC3A*, *APOBEC3B* e *APOBEC3D*, além dos controles positivos do experimento, *DDX58*, *MDA5* e *IRF7* (conforme descrito no item 3.2.3). Nos controles, água ultrapura, o solvente do análogo, foi adicionada ao meio de cultura, no mesmo volume utilizado para o tratamento. Foram realizados três experimentos independentes em triplicata para cada linhagem celular (**Figura 3.4**).

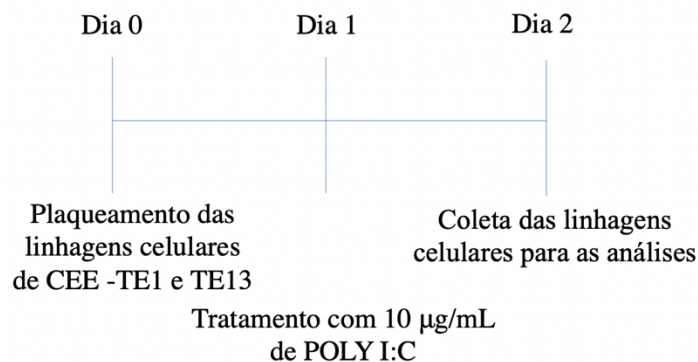


Figura 3.4: Esquema do tratamento com 10 µg/mL de POLY I:C por 24h. Inicialmente, as linhagens celulares de CEE TE1 e TE13 foram plaqueadas em placas de 6 poços, $1,0 \times 10^5$ de células/poço e $3,0 \times 10^5$ de células/poço, respectivamente. Após 24h, 10 µg/mL de POLY I:C foi adicionado ao meio de cultura. Após 24h de incubação com o análogo, as células foram coletadas, as células viáveis foram contadas por exclusão por coloração com azul de tripan e posterior análise de expressão gênica de *APOBEC3A*, *APOBEC3B* e *APOBEC3D* e dos controles positivos do experimento (*DDX58*, *MDA5* e *IRF7*). Nos controles, água ultrapura, o solvente do análogo, foi adicionada ao meio de cultura.

3.2.4.4 TRATAMENTO COM INTERFERONS

O tratamento com interferons foi realizado para mimetizar a resposta à infecção viral e por ser um dos mecanismos de regulação das APOBECs já descrito (BONVIN et al., 2006; CHEN et al., 2006; PENG et al., 2006; KONING et al., 2009; LUCIFORA et al., 2014; WANG et al., 2014). Para tal, os interferons IFN α (R&D Systems), IFN β (Peprotech) e IFN γ (Peprotech) foram

utilizados. IFN β e IFN γ foram reconstituídos em água de injeção, obtendo uma solução de 1 mg/mL. Em seguida, foi necessário diluir essas soluções em PBS + BSA (albumina sérica bovina, do inglês *Bovine Serum Albumin*) a 0,1%, obtendo uma solução estoque de 20 ng/ μ L, sendo aliquotada e armazenada a -80°C. Já para o IFN α , foi adicionado PBS + BSA a 0,1%, obtendo uma solução estoque de 20 U/ μ L, sendo aliquotada e armazenada a -80°C. Para realizar os experimentos, foram plaqueadas $1,0 \times 10^5$ de células/poço para a TE1 e $3,0 \times 10^5$ de células/poço para a TE13 em placas de 6 poços. Depois de 24h, as células foram tratadas com 100 U/mL de IFN α ou 100 ng/mL de IFN β ou 100 ng/mL de IFN γ por 24h, conforme já descrito na literatura (GAN et al., 2016). Após 24h de incubação, as células foram coletadas, as células viáveis foram contadas por meio da exclusão por coloração com azul de tripan e posterior análise de expressão gênica de *APOBEC3A*, *APOBEC3B* e *APOBEC3D*, além do controle positivo dos experimentos, *IRF7*. Nos controles, a solução PBS + BSA a 0,1% foi adicionada ao meio de cultura, no mesmo volume utilizado para os tratamentos. Foram realizados três experimentos independentes em triplicata para cada IFN e para cada linhagem celular (**Figura 3.5**).

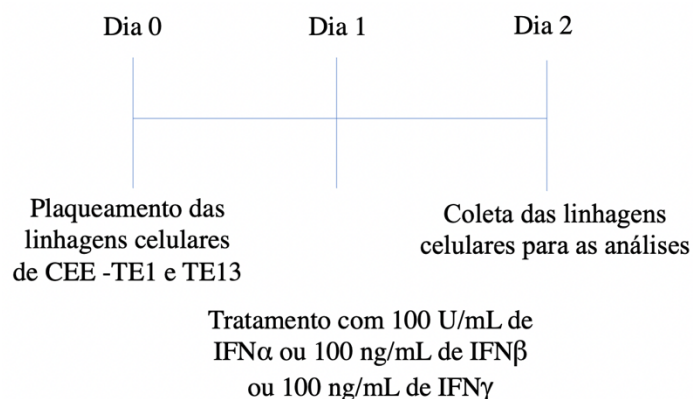


Figura 3.5: Esquema do tratamento com 100 U/mL de IFN α ou 100 ng/mL de IFN β ou 100 ng/mL de IFN γ por 24h. Inicialmente, as linhagens celulares de CEE TE1 e TE13 foram plaqueadas em placas de 6 poços, $1,0 \times 10^5$ de células/poço e $3,0 \times 10^5$ de células/poço, respectivamente. Após 24h, foi adicionado ao meio de cultura 100 U/mL de IFN α ou 100 ng/mL de IFN β ou 100 ng/mL de IFN γ . Após 24h de incubação, as células foram coletadas, as células viáveis foram contadas por exclusão por coloração com azul de tripan e posterior análise de expressão gênica de *APOBEC3A*, *APOBEC3B* e *APOBEC3D* e do controle positivo do experimento (*IRF7*). Nos controles, a solução PBS + BSA a 0,1% foi adicionada ao meio de cultura.

3.2.4.5 COLETA E VIABILIDADE CELULAR PÓS-TRATAMENTOS

Após o tempo determinado para cada tratamento, as células foram coletadas da seguinte maneira:

- Retirada do meio de cada poço das placas;
- Lavagem com PBS e remoção do mesmo;
- Adição de solução de tripsina a 0,25%, com EDTA e subsequente incubação durante 5 min, a 37°C, 5% CO₂ para soltar as células aderidas;
- Adição de meio de cultura para inibir a tripsina (o meio contém SFB capaz de inibir a tripsina);
- Transferência da solução contendo as células para tubos de microcentrífuga de 1,5 mL para realizar a contagem de células, com auxílio da câmara de Neubauer pela técnica de exclusão por azul de tripan, conforme descrito anteriormente (item 3.2.4.1).
- Após a análise de viabilidade celular, os tubos de microcentrífuga de 1,5 mL com a solução contendo as células foram centrifugados por 5 minutos a 10.000 rpm. Em seguida, o sobrenadante foi descartado e o precipitado armazenado em freezer -80°C para posterior extração de RNA.

3.2.4.6 EXTRAÇÃO DO RNA, RT E PCRq PÓS-TRATAMENTOS

O RNA total foi extraído a partir do precipitado de células utilizando o reagente TRIzol® (Invitrogen) seguindo o protocolo do fabricante. O *pellet* de células foi homogeneizado com 500 µL de TRIzol e incubado por cinco minutos à temperatura ambiente. Posteriormente, foram adicionados 100 µL de clorofórmio com subsequente homogeneização em agitador vortex. Após este processo, houve incubação por 15 minutos à temperatura ambiente. As amostras foram centrifugadas à velocidade de 16.128 x g por 15 minutos a 4°C e então a fase aquosa foi transferida para um novo tubo, com descarte do material restante. Ao novo tubo foram adicionados 250 µL de isopropanol com posterior agitação e incubação *overnight* a 4°C, seguida de centrifugação a 16.128 x g por 10 minutos a 4°C. O sobrenadante foi descartado e ao *pellet* foram adicionados 500 µL de etanol a 75%. As amostras foram centrifugadas a 11.200 x g por 5 minutos, com posterior

descarte do etanol. Para eluir o RNA, foram adicionados 15 μ L de água livre de RNase. A quantificação e determinação da pureza das amostras de RNA foram feitas com o auxílio do equipamento NanoDrop (Uniscience), conforme já descrito no item 3.2.1.2.

A síntese de cDNA foi feita com a enzima transcriptase reversa M-MLV (PROMEGA), seguindo o protocolo do fabricante. Cada tubo de reação continha 500 ng a 1 μ g de RNA, 1 μ L de 250 ng de oligonucleotídeos randômicos (GE Healthcare, UK), 1 μ L de 10 mM de desoxirribonucleotídeo na concentração de 10 mM e água ultrapura para completar o volume final de 12 μ L. Após incubação a 65°C por 5 minutos, os tubos contendo os reagentes e amostras foram incubados no gelo. Em seguida, foram adicionados a cada tubo 4 μ L do tampão da enzima 5X (PROMEGA), 2 μ L de DTT e 1 μ L água ultrapura. A mistura foi homogeneizada e incubada a 37°C por 2 minutos. Em seguida, foi adicionado 1 μ L da enzima M-MLV (200 U) (PROMEGA). Em seguida, houve incubação a 25°C por 10 minutos. Depois, incubação a 37°C por 50 minutos e, posteriormente, a 70°C por 15 minutos para inativação da enzima. Após a síntese do cDNA, este foi armazenado a -20°C.

Para a análise de expressão gênica das *APOBEC3A*, *APOBEC3B*, *APOBEC3D* e dos controles positivos dos tratamentos, *DDX58*, *MDA5* e *IRF7*, bem como do gene-referência *GAPDH* (**Tabela 3.1**), foram feitas as reações de PCRq conforme já descritas no item 3.2.3. As condições da PCRq e a ciclagem para os controles positivos dos tratamentos foram as mesmas utilizadas para as *APOBECs* (item 3.2.3). Para as análises de expressão de *APOBEC3A*, foi utilizado 1 μ L de cDNA não diluído. Os valores de quantificação relativa foram expressos pelas fórmulas $2^{-\Delta CT}$, calculados pelo programa Microsoft Excel (item 3.2.3).

3.3 ANÁLISE ESTATÍSTICA

Todas as análises estatísticas foram feitas utilizando o *software* GraphPad Prism versão 5.0 (GraphPad Software, Estados Unidos). O valor de p foi considerado estatisticamente significativo quando $<0,05$.

Para averiguar se os dados obedeciam à distribuição normal, foi utilizado o teste Kolmogorov-Smirnov. Em seguida, na análise de diferença de expressão entre amostras de CEE e de tecidos adjacentes pareados com distribuição normal, foi aplicado o teste t. Quando os dados

não obedeciam à distribuição normal, foi utilizado o *Wilcoxon matched pairs test*. Já para as linhagens de CEE nos ensaios *in vitro*, na análise de viabilidade celular, foi utilizado o Teste *Wilcoxon signed rank* e para a comparação da expressão das *APOBECs* e dos controles positivos após os tratamentos, foi utilizado o *Teste Kruskal Wallis* e pós teste de *Dunn*. Os dados foram considerados estatisticamente significativos quando $p < 0,05$.

O Teste One-Way ANOVA foi usado quando mais de dois grupos foram avaliados. Para os dados paramétricos foi utilizado o Teste *One-Way analysis of variance* e o pós-teste de *Tukey*. Já para os dados não paramétricos foi aplicado o Teste de *Kruskal Wallis* e o pós-teste de *Dunn*.

As correlações entre a expressão de RNAm e contribuição das assinaturas mutacionais de *APOBECs*, bem como com a expressão dos miRNAs, das células do sistema imune e da metilação do DNA para cada subtipo tumoral (CEE, CEL, CECO, CEOF HPV⁻ e CEOF HPV⁺) foram determinadas usando testes de correlação de *Pearson*, quando os dados apresentaram distribuição normal, ou correlação de *Spearman*, quando essa distribuição não era obedecida.

A curva de Kaplan Meier foi utilizada para estimar a sobrevida com análise univariada e as distribuições de sobrevivência foram comparadas com o teste de log-rank. A idade e o estágio do tumor foram selecionados para ajuste na análise multivariada, uma vez que essas variáveis mostraram influenciar na análise de sobrevida. Finalmente, aplicamos a regressão de Cox usando o método *stepwise forward*. As análises de sobrevida dos pacientes com CEE foram realizadas utilizando o pacote de software Survival em ambiente R (THERNEAU, 2015), onde a expressão das *APOBECs* foi dividida em octis e o octil superior definiu o grupo com alta expressão. As análises de sobrevida de acordo com a contribuição da assinatura mutacional relacionada à atividade das *APOBECs* nos diferentes subtipos tumorais (CEE, CEL, CECO, CEOF HPV⁻ e CEOF HPV⁺) a partir dos dados do consórcio TCGA foram realizadas dividindo os casos com alta e baixa contribuição das assinaturas de acordo com a mediana de cada tipo tumoral ($>$ ou \leq mediana). Apesar do acompanhamento, em alguns casos, ter ultrapassado os 60 meses, todos os indivíduos foram censurados nesse período, de modo a tornar os resultados comparáveis.

4 RESULTADOS

4.1 ASSINATURAS MUTACIONAIS, EXPRESSÃO DE *APOBECs* E MECANISMOS DE REGULAÇÃO EM CEE, CEL, CECO, CEOF HPV⁻ E CEOF HPV⁺

Inicialmente, foi feita a caracterização dos pacientes a partir dos estudos *Esophageal Carcinoma* (TCGA, *Firehose Legacy*) para CEE e *Head and Neck Squamous Cell Carcinoma* (TCGA, *Firehose Legacy*) para CEL, CECO e CEOF HPV⁻ e CEOF HPV⁺. A mediana de idade foi de 57 anos para CEE, 62 anos para CEL, 60 anos para CECO, 57,5 anos para CEOF HPV⁻ e 56 anos para CEOF HPV⁺. A maioria dos pacientes é do gênero masculino para todos os tipos tumorais, sendo essa proporção de 85,3% em CEE, 83,5% em CEL, 67,9% em CECO, 79,2% em CEOF HPV⁻ e 92,3% em CEOF HPV⁺. Em sua maioria, todos os grupos de pacientes são compostos por tabagistas ou ex-tabagistas (62,5% para CEE, 92,2% para CEL, 67,5% para CECO, 83,3% para CEOF HPV⁻ e 61,5% para CEOF HPV⁺); etilistas ou ex-etilistas (71,9% para CEE, 64,3% para CEL, 67,9% para CECO, 87,5% para CEOF HPV⁻ e 76,9% para CEOF HPV⁺); e foram diagnosticados em estadiamento avançado, ou seja, III ou IV (75,6% para CEL, 71,2% para CECO, 45,8% para CEOF HPV⁻ e 38,5% para CEOF HPV⁺). A única exceção foi CEE, em que a maioria dos pacientes foi diagnosticada em estadiamento precoce (I ou II, 65,6%) (**Tabela 4.1**).

Tabela 4.1: Caracterização dos pacientes com CEE, CEL, CECO, CEOF HPV⁻ e CEOF HPV⁺ do banco de dados do TCGA.

Critério Avaliado	CEE	CEL	CECO	CEOF HPV ⁻	CEOF HPV ⁺
	TCGA	TCGA	TCGA	TCGA	TCGA
	Número de pacientes (%)	Número de pacientes (%)	Número de pacientes (%)	Número de pacientes (%)	Número de pacientes (%)
Idade (mediana, min-max)	57 anos (36-90)	62 anos (38-83)	60 anos (19-90)	57,5 anos (43-79)	56 anos (35-77)
Gênero					
Masculino	81 (85,3%)	96 (83,5%)	165 (67,9%)	19 (79,2%)	48 (92,3%)
Feminino	15 (15,7%)	19 (16,5%)	78 (32,1%)	5 (20,8%)	4 (7,7%)
Tabagismo					
Nunca	32 (33,3%)	6 (5,2%)	75 (30,9%)	4 (16,7%)	19 (36,5%)
Ex/Ativo	60 (62,5%)	106 (92,2%)	164 (67,5%)	20 (83,3%)	32 (61,5%)
Sem informação	4 (4,2%)	3 (2,6%)	4 (1,6%)	-	1 (2%)
Etilismo					
Nunca	25 (26%)	39 (33,9%)	74 (30,5%)	2 (8,3%)	11 (21,1%)
Ex/Ativo	69 (71,9%)	74 (64,3%)	165 (67,9%)	21 (87,5%)	40 (76,9%)
Sem informação	2 (2,1%)	2 (1,8%)	4 (1,6%)	1 (4,2%)	1 (2%)
Estadiamento do tumor					
I+II	63 (65,6%)	11 (9,6%)	52 (21,4%)	5 (20,8%)	7 (13,5%)
III+IV	31 (32,3%)	87 (75,6%)	173 (71,2%)	11 (45,8%)	20 (38,5%)
Não determinado	2 (2,1%)	17 (14,8%)	18 (7,4%)	8 (33,4%)	25 (48%)

4.1.1 PERFIL MUTACIONAL EM CEE, CEL, CECO, CEOF HPV⁻ E CEOF HPV⁺

A partir de dados de sequenciamento de exoma, nós investigamos o número total de mutações (variantes de único nucleotídeo) por paciente em cada tipo tumoral. Os pacientes com CEE e CEL (mediana de 155 e 153,5, respectivamente) apresentaram, de forma significativa, maior número de mutações quando comparados aos pacientes com CECO (mediana de 101),

CEOF HPV⁻ (mediana de 102) e CEOF HPV⁺ (mediana de 61), porém não houve diferença significativa entre pacientes com CEE e CEL, ou entre os outros tipos tumorais (**Figura 4.1**).

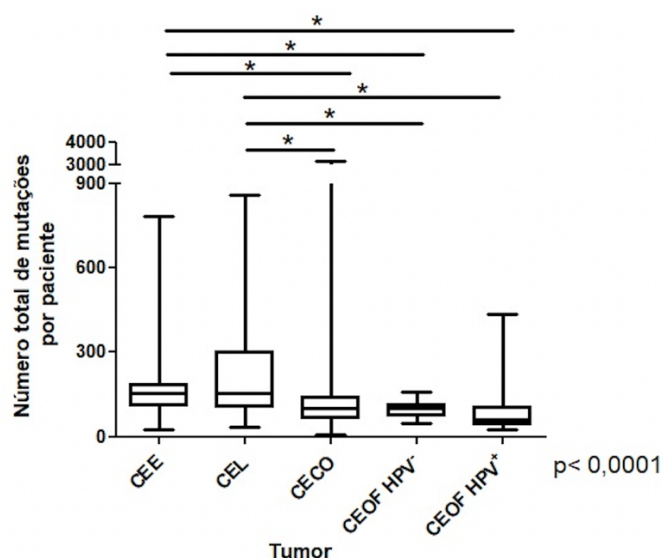


Figura 4.1: Número total de variantes de único nucleotídeo por paciente com CEE, CEL, CECO, CEOF HPV⁻ e CEOF HPV⁺ do consórcio TCGA. Boxplot mostrando o número de mutações por paciente diagnosticado com cada tipo tumoral. Teste de *Kruskal-Wallis*, pós teste de *Dunn*, * $p < 0,0001$. CEE, carcinoma epidermoide de esôfago; CEL, carcinoma epidermoide de laringe, CECO, carcinoma epidermoide de cavidade oral, CEOF HPV⁻, carcinoma epidermoide de orofaringe HPV⁻, CEOF HPV⁺, carcinoma epidermoide de orofaringe HPV⁺.

Em seguida, na análise de assinaturas mutacionais de substituição de base única, foi possível identificar as seguintes assinaturas: assinatura 1/6 (C>T em contexto CpG, associada à idade e C>T, associada a defeitos na maquinaria de reparo de mau-pareamento do DNA); assinatura 2/13 (C>T e C>G em contexto TpC, associadas à atividade das APOBECs); assinatura 4 (C>A, associada à exposição ao benzopireno); assinatura 5 (T>C em contexto ATN – nesse contexto com mais bases adeninas sendo mutadas do que timinas, associadas, com evidências não tão claras, à idade, à deficiência no reparo por excisão de nucleotídeo e ao tabagismo); as assinaturas 6 (C>T), 15 (C>T) e 21 (T>C) (associadas a defeitos na maquinaria de reparo de mau-pareamento do DNA) e assinatura 16 (T>C em contexto ATN – nesse contexto com mais bases adeninas sendo mutadas do que timinas, sem causa conhecida, porém artigos demonstram possível associação ao álcool - LI et al., 2018; MOODY et al., 2021). Para CEE, foram observadas as

assinaturas 1/6, 2/13, 4 e 16 (**Figura 4.2A**). Para CEL, as assinaturas 1/6, 2/13 e 4 (**Figura 4.2B**). Para CECO, as assinaturas 1/6, 2/13, 5 e 6/21/15 (**Figura 4.2C**). Para CEOF HPV⁻, as assinaturas 2/13 e 5 (**Figura 4.2D**) e CEOF HPV⁺, as assinaturas 1/6 e 2/13 (**Figura 4.2E**).

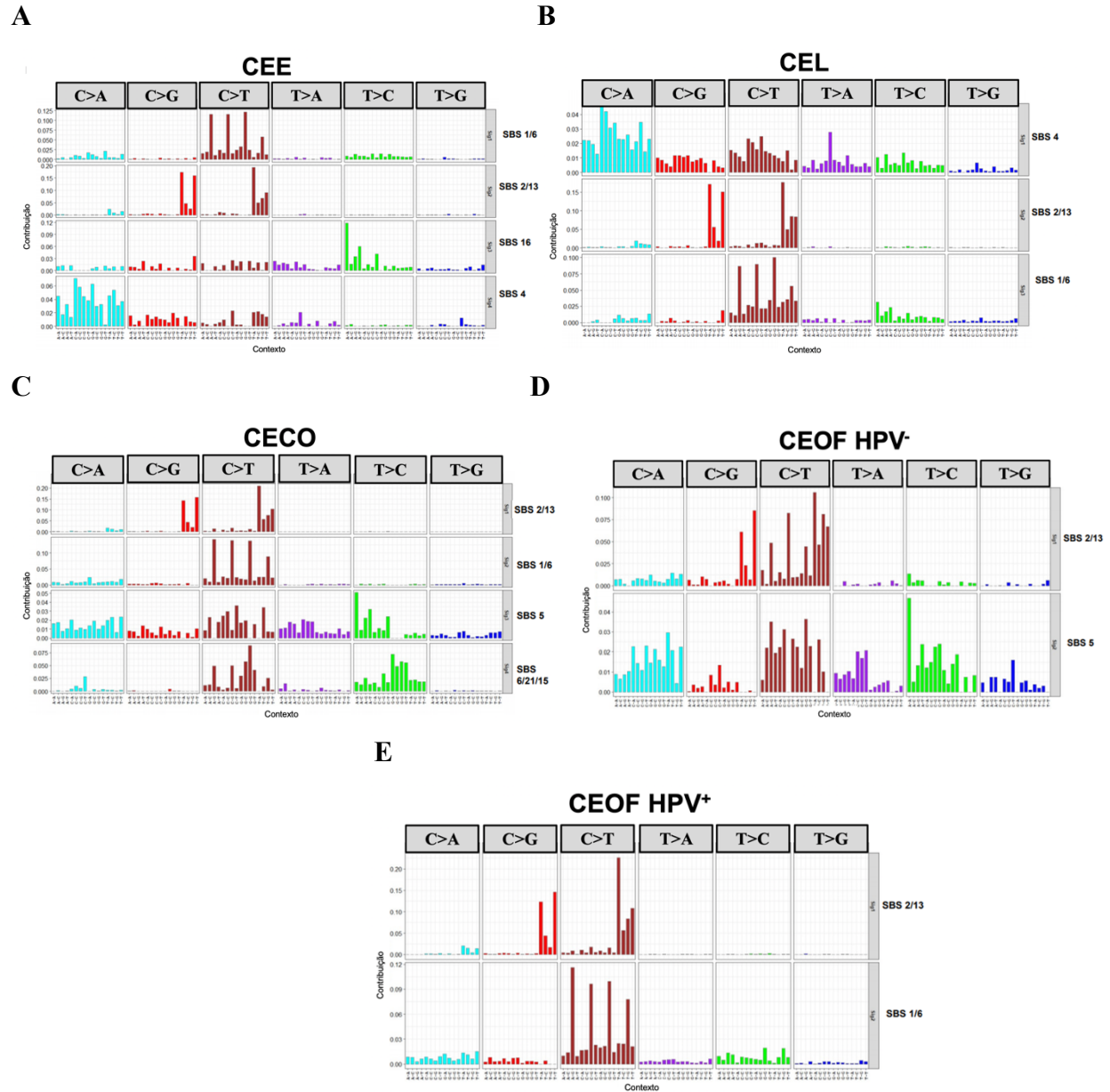


Figura 4.2: Assinaturas mutacionais de substituição de base única detectadas em tumores de cabeça e pescoço e de esôfago. Gráficos de barras mostrando a contribuição de cada troca de bases, estratificadas de acordo com o nucleotídeo que precede e sucede a base alterada (contexto), por sítio tumoral: carcinoma epidermoide de esôfago - CEE (A), carcinoma epidermoide de laringe - CEL (B), carcinoma epidermoide de cavidade oral - CECO (C), carcinoma epidermoide de

orofaringe HPV⁻ - CEOF HPV⁻ (D), carcinoma epidermoide de orofaringe HPV⁺- CEOF HPV⁺(E) e SBS, assinaturas mutacionais de substituição de base única. As assinaturas 1/6, 2/13 e 6/21/15 aparecem juntas, porque para cada assinatura, era gerada uma medida de similaridade com as assinaturas já descritas no COSMIC. Quando a medida de similaridade entre as assinaturas detectadas era indistinguível, foram representadas juntas.

Após identificar quais assinaturas mutacionais estavam presentes em cada tipo tumoral, avaliamos a contribuição delas por paciente (**Figura 4.3**). A assinatura 2/13, associada à atividade das APOBECs e alvo desta tese, foi detectada em todos os tipos tumorais, conforme mencionado anteriormente, e pôde ser observada em quase todos os pacientes, porém com contribuição variável.

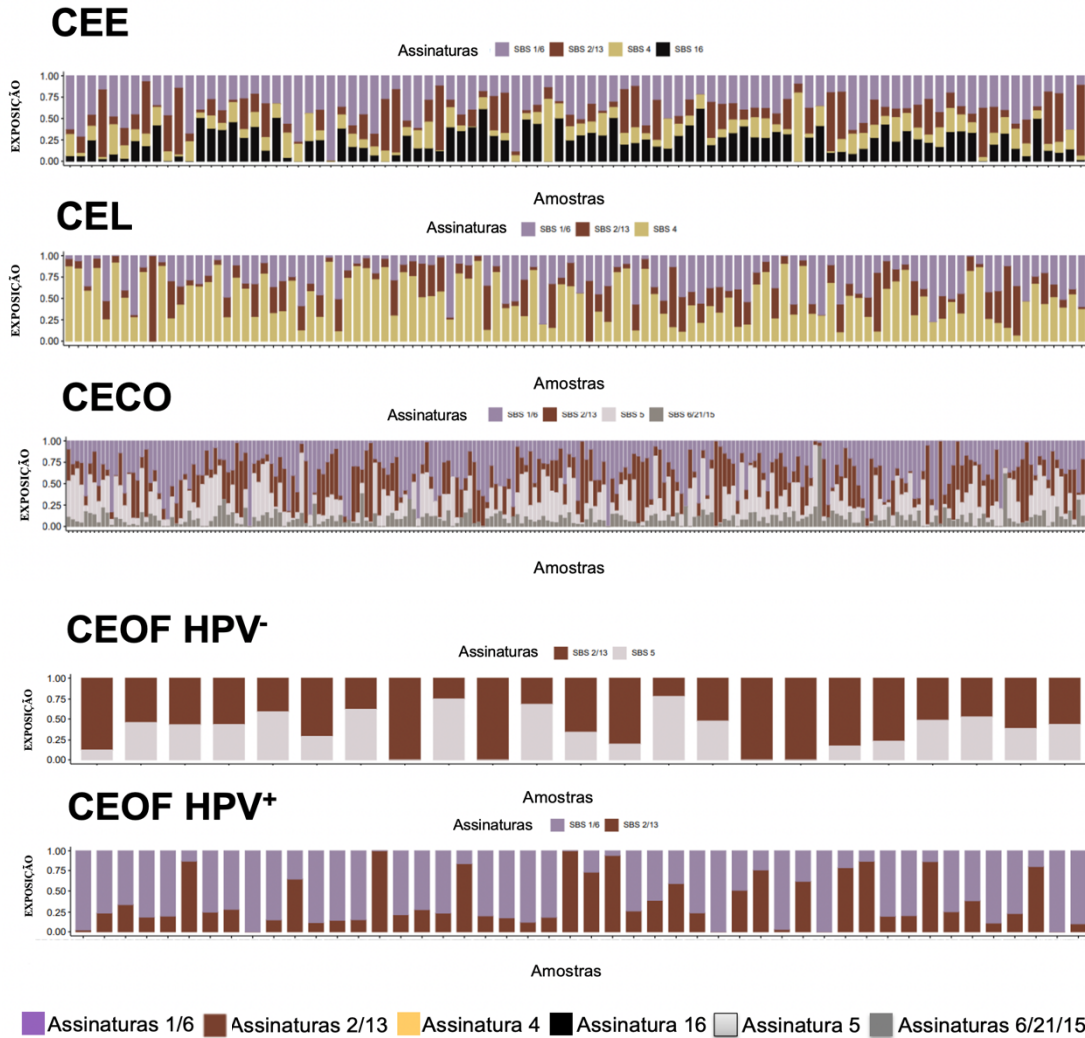


Figura 4.3: Contribuição das assinaturas mutacionais de substituição de base única em tumores de cabeça e pescoço e de esôfago por paciente. Gráfico de barras mostrando a fração de cada assinatura mutacional por paciente diagnosticado com CEE, CEL, CECO, CEOF HPV⁻ e CEOF HPV⁺ do consórcio TCGA. CEE, carcinoma epidermoide de esôfago, CEL, carcinoma epidermoide de laringe, CECO, carcinoma epidermoide de cavidade oral, CEOF HPV⁻, carcinoma epidermoide de orofaringe HPV⁻, CEOF HPV⁺, carcinoma epidermoide de orofaringe HPV⁺, SBS, assinaturas mutacionais de substituição de base única. As assinaturas 1/6, 2/13 e 6/21/15 aparecem juntas, porque para cada assinatura, era gerada uma medida de similaridade com as assinaturas já descritas no COSMIC. Quando a medida de similaridade entre as assinaturas era indistinguível, foram representadas juntas.

Assim, nosso próximo passo foi avaliar se a contribuição da assinatura mutacional associada à atividade das APOBECs diferia de acordo com o tipo tumoral (**Figura 4.4**). Os CEOF HPV⁻ apresentaram maior fração de contribuição da assinatura 2/13 em comparação a todos os outros subsítios tumorais (mediana de 0,7838). Já os CEOF HPV⁺ (mediana de 0,2403) apresentaram maior fração de contribuição da assinatura de APOBECs quando comparados a CEE (mediana de 0,1269) e CEL (mediana de 0,1297). Por fim, pacientes com CECO (mediana de 0,2457) mostraram maior contribuição dessa assinatura em relação a pacientes com CEL (mediana de 0,1297).

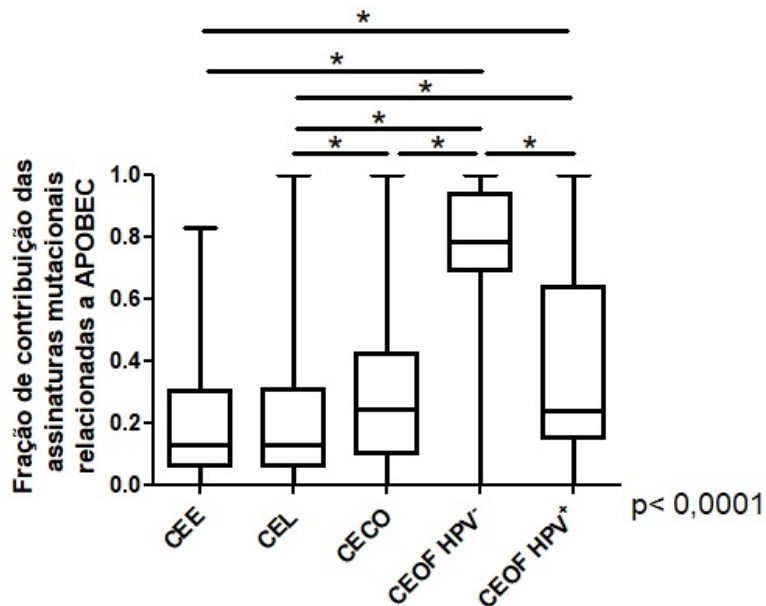


Figura 4.4: Fração de contribuição das assinaturas mutacionais de substituição de base única relacionadas à atividade das APOBECs em tumores de cabeça e pescoço e de esôfago por paciente. Boxplot mostrando a fração de contribuição das assinaturas mutacionais de substituição de base única relacionadas à atividade das APOBECs por paciente diagnosticado com CEE, CEL, CECO, CEOF HPV⁻ e CEOF HPV⁺ do consórcio TCGA. Teste de *Kruskal-Wallis*, pós teste de *Dunn*, * $p < 0,0001$. CEE, carcinoma epidermoide de esôfago, CEL, carcinoma epidermoide de laringe, CECO, carcinoma epidermoide de cavidade oral, CEOF HPV⁻, carcinoma epidermoide de orofaringe HPV⁻, CEOF HPV⁺, carcinoma epidermoide de orofaringe HPV⁺.

Também foi avaliado se a fração de contribuição da assinatura mutacional associada à atividade das APOBECs diferia de acordo com o estadiamento do tumor (tumores precoces e

tumores avançados) para os tumores de cabeça e pescoço e de esôfago (**Figura 4.5**), para averiguar se essa alteração genética estaria presente já no início da carcinogênese ou se teria um papel na progressão do tumor. Não foi encontrada nenhuma diferença estatisticamente significativa entre os tumores precoces e avançados em CEE (**Figura 4.5A**), CEL (**Figura 4.5B**), CECO (**Figura 4.5C**), CEOF HPV⁻ (**Figura 4.5D**) e CEOF HPV⁺ (**Figura 4.5E**), sugerindo ser uma alteração precoce nesses tumores.

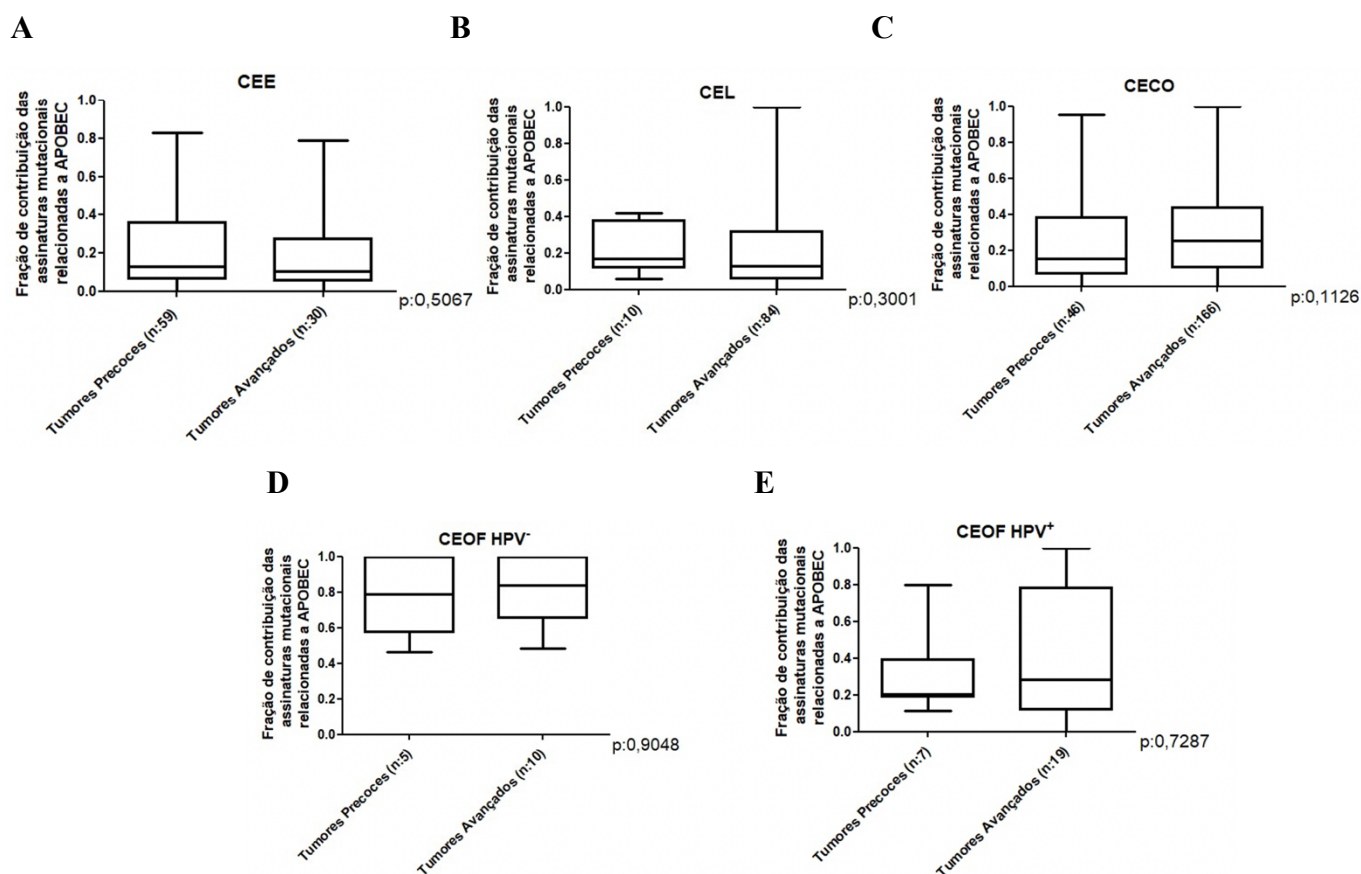


Figura 4.5: Fração de contribuição das assinaturas mutacionais de substituição de base única relacionadas à atividade das APOBECs de acordo com estadiamento tumoral em tumores de cabeça e pescoço e de esôfago por paciente. Boxplot mostrando a fração de contribuição das assinaturas mutacionais de substituição de base única relacionadas à atividade das APOBECs por paciente diagnosticado com CEE (A), CEL (B), CECO (C), CEOF HPV⁻ (D) e CEOF HPV⁺ (E) do consórcio TCGA, de acordo com estadiamento do tumor (tumores precoces *versus* tumores avançados). Teste T não pareado para CEOF HPV⁻ e Teste *Mann-Whitney* para os demais tumores, p>0,05. CEE, carcinoma epidermoide de esôfago, CEL, carcinoma epidermoide de laringe, CECO,

carcinoma epidermoide de cavidade oral, CEOF HPV⁻, carcinoma epidermoide de orofaringe HPV⁻, CEOF HPV⁺, carcinoma epidermoide de orofaringe HPV⁺.

Como as APOBECs são enzimas de resposta antiviral, avaliamos se a fração de contribuição das assinaturas mutacionais relacionadas à atividade dessas enzimas diferia em CEL e CECO de acordo com status de HPV (**Figura 4.6**), conforme foi observado para os tumores CEOF (Figura 4.4). Não houve diferença estatisticamente significativa entre tumores HPV⁻ e HPV⁺ em CEL (**Figura 4.6A**) e em CECO (**Figura 4.6B**).

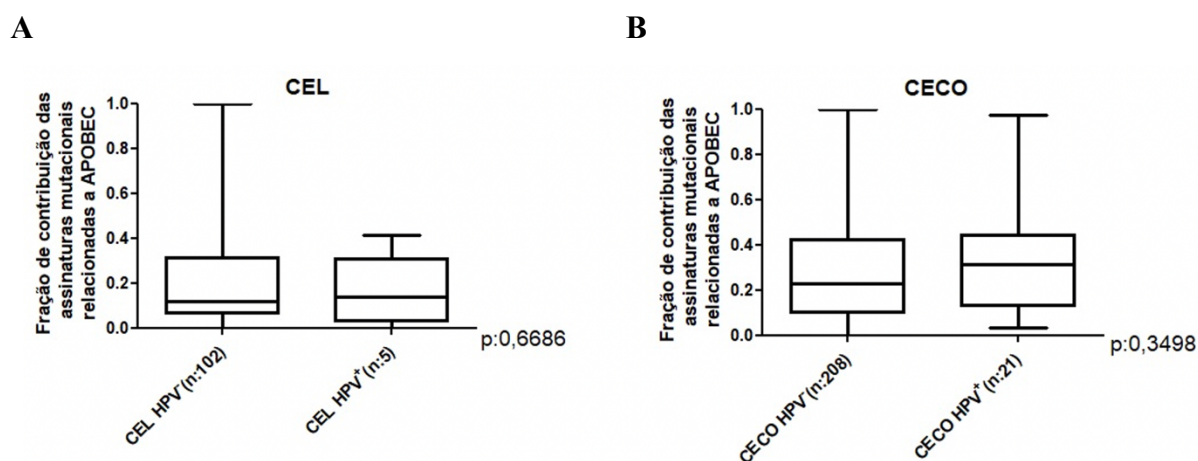


Figura 4.6: Fração de contribuição das assinaturas mutacionais de substituição de base única relacionadas à atividade das APOBECs de acordo com status de HPV em tumores de cabeça e pescoço por paciente. Boxplot mostrando a fração de contribuição das assinaturas mutacionais de substituição de base única relacionadas à atividade das APOBECs por paciente diagnosticado com CEL (A) e CECO do consórcio TCGA, de acordo com status de HPV (HPV⁻ versus HPV⁺). Teste *Mann Whitney*, $p > 0,05$. CEL, carcinoma epidermoide de laringe, CECO, carcinoma epidermoide de cavidade oral.

Em seguida, investigamos se a contribuição das assinaturas mutacionais associadas à atividade das APOBECs (assinaturas 2 e 13) estava associada à sobrevida global dos pacientes com CEE, CEL, CECO, CEOF HPV⁻ e CEOF HPV⁺ do consórcio TCGA (**Figura 4.7**). Não foi observado impacto significativo na sobrevida global dos pacientes com CEE (**Figura 4.7A**), CEL

(Figura 4.7B), CECO (Figura 4.7C), CEOF HPV⁻ (Figura 4.7D) e CEOF HPV⁺ (Figura 4.7E) de acordo com as assinaturas mutacionais associadas a essas enzimas.

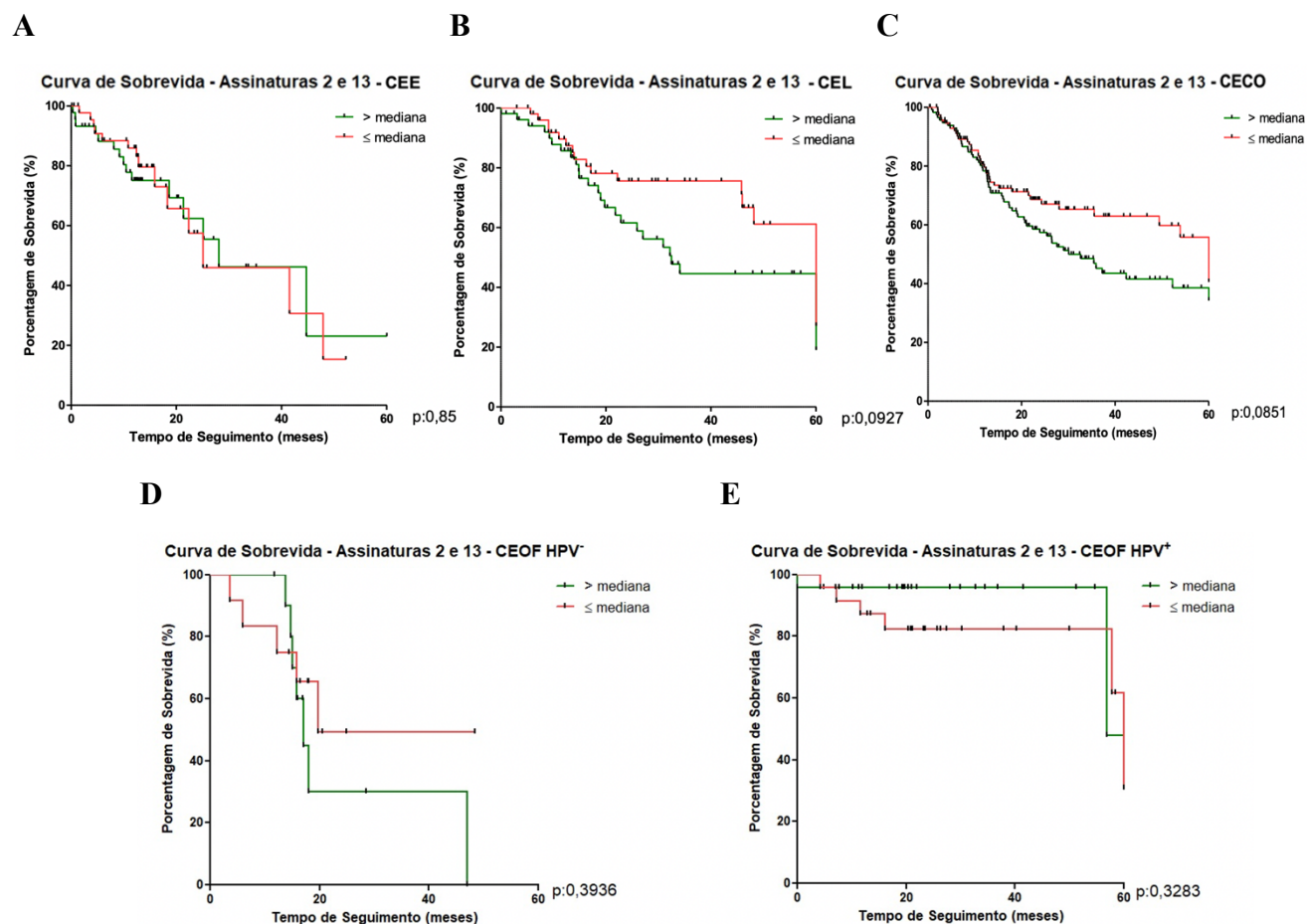


Figura 4.7: O impacto das assinaturas mutacionais relacionadas à atividade das APOBECs (assinaturas 2 e 13) na sobrevida global de pacientes com tumores de cabeça e pescoço e de esôfago. Curva de Kaplan Meier mostrando o impacto das assinaturas mutacionais 2 e 13 na sobrevida global de pacientes com CEE (A), CEL, (B), CECO (C), CEOF HPV⁻ (D) e CEOF HPV⁺ (E). Os casos foram divididos com alta e baixa contribuição das assinaturas de acordo com a mediana de cada tipo tumoral (> ou ≤ mediana). Apesar do acompanhamento, em alguns casos, ter ultrapassado os 60 meses, todos os indivíduos foram censurados nesse período, de modo a tornar os resultados comparáveis. CEE, carcinoma epidermoide de esôfago, CEL, carcinoma epidermoide de laringe, CECO, carcinoma epidermoide de cavidade oral, CEOF HPV⁻, carcinoma epidermoide de orofaringe HPV⁻, CEOF HPV⁺, carcinoma epidermoide de orofaringe HPV⁺.

4.1.2 ASSINATURAS MUTACIONAIS E A EXPRESSÃO DAS *APOBECs* EM CEE, CEL, CECO, CEOF HPV⁻ e CEOF HPV⁺

Foi realizada uma análise de correlação entre a contribuição das assinaturas mutacionais relacionadas à atividade das *APOBECs* e a expressão gênica dessas enzimas em CEE, CEL, CECO, CEOF HPV⁻ e CEOF HPV⁺ no consórcio do TCGA (**Figura 4.8**). Foi observada correlação direta significativa da expressão de *APOBEC3A* ($\rho:0,26$) e *APOBEC3B* ($\rho:0,24$) com as assinaturas mutacionais em CEE. Em CEL, correlação direta significativa da expressão de *APOBEC3B* ($\rho:0,22$), *APOBEC3C* ($\rho:0,36$), *APOBEC3D* ($\rho:0,23$), *APOBEC3F* ($\rho:0,41$), *APOBEC3G* ($\rho:0,32$) com as assinaturas mutacionais. Em CECO, correlação indireta significativa com a expressão de *AICDA* ($\rho:-0,2$) e *APOBEC2* ($\rho:-0,2$) e direta significativa com a expressão de *APOBEC3A* ($\rho:0,26$), *APOBEC3B* ($\rho:0,15$), *APOBEC3C* ($\rho:0,15$) e *APOBEC3D* ($\rho:0,19$). Para CEOF HPV⁺, correlação direta significativa com a expressão da *APOBEC3A* ($\rho:0,35$) e para CEOF HPV⁻, não foram observadas correlações significativas (**Figura 4.8**).

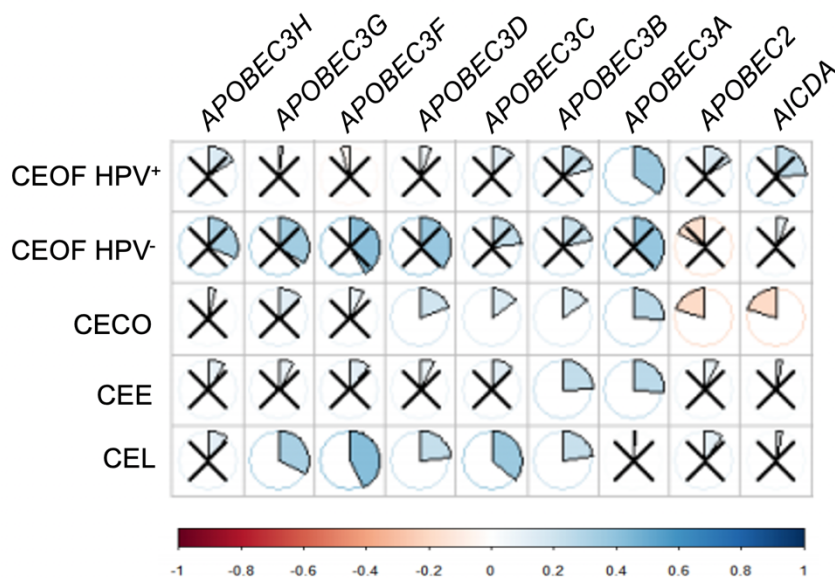


Figura 4.8: Correlação entre a expressão das *APOBECs* e as assinaturas mutacionais associadas a elas (2 e 13) em tumores de cabeça e pescoço e de esôfago. Matriz de correlação entre a expressão das *APOBECs* e as assinaturas mutacionais de substituição de base única relacionadas à atividade das *APOBECs* em CEE, CEL, CECO, CEOF HPV⁻ e CEOF HPV⁺ no consórcio do TCGA.

Correlação de *Spearman*, $p < 0,05$. Azul representa correlação positiva (escala de rho de 0 até 1); Vermelho representa correlação negativa (escala de rho de -1 até 0); X representa sem significância estatística. CEE, carcinoma epidermoide de esôfago, CEL, carcinoma epidermoide de laringe, CECO, carcinoma epidermoide de cavidade oral, CEOF HPV⁻, carcinoma epidermoide de orofaringe HPV⁻, CEOF HPV⁺, carcinoma epidermoide de orofaringe HPV⁺.

Já que foram encontradas correlações significativas entre a contribuição das assinaturas mutacionais de APOBECs e sua expressão, avaliamos se essas enzimas estavam desreguladas nesses tumores (**Figura 4.9**). Em CEE, *AICDA* e todas as *APOBEC3s* são mais expressas no tumor em comparação à mucosa adjacente não tumoral, inclusive aquelas correlacionadas com a assinatura mutacional (*APOBEC3A* e *APOBEC3B*), com exceção da *APOBEC3C*. A *APOBEC2* está menos expressa no tumor quando comparado à mucosa não tumoral (**Figura 4.9A**). Para CEL, *AICDA*, *APOBEC3B*, *APOBEC3F*, *APOBEC3G*, *APOBEC3H* são mais expressas no tumor em comparação à mucosa adjacente não tumoral, inclusive aquelas correlacionadas com as assinaturas mutacionais (*APOBEC3B*, *APOBEC3F* e *APOBEC3G*). A *APOBEC2* está menos expressa no tumor quando comparado à mucosa não tumoral (**Figura 4.9B**). Para CECO, *AICDA*, *APOBEC3B*, *APOBEC3D*, *APOBEC3F* e *APOBEC3G* estão mais expressas no tumor em relação à mucosa adjacente, inclusive aquelas associadas as assinaturas mutacionais (*AICDA*, *APOBEC2*, *APOBEC3B* e *APOBEC3D*). A *APOBEC2* está menos expressa no tumor quando comparado à mucosa não tumoral (**Figura 4.9C**). Já para CEOF, *AICDA* e *APOBEC3s* estão mais expressas nos tumores HPV⁺ quando comparados aos tumores HPV⁻, inclusive a *APOBEC3A* que foi correlacionada com as assinaturas mutacionais (**Figura 4.9D**).

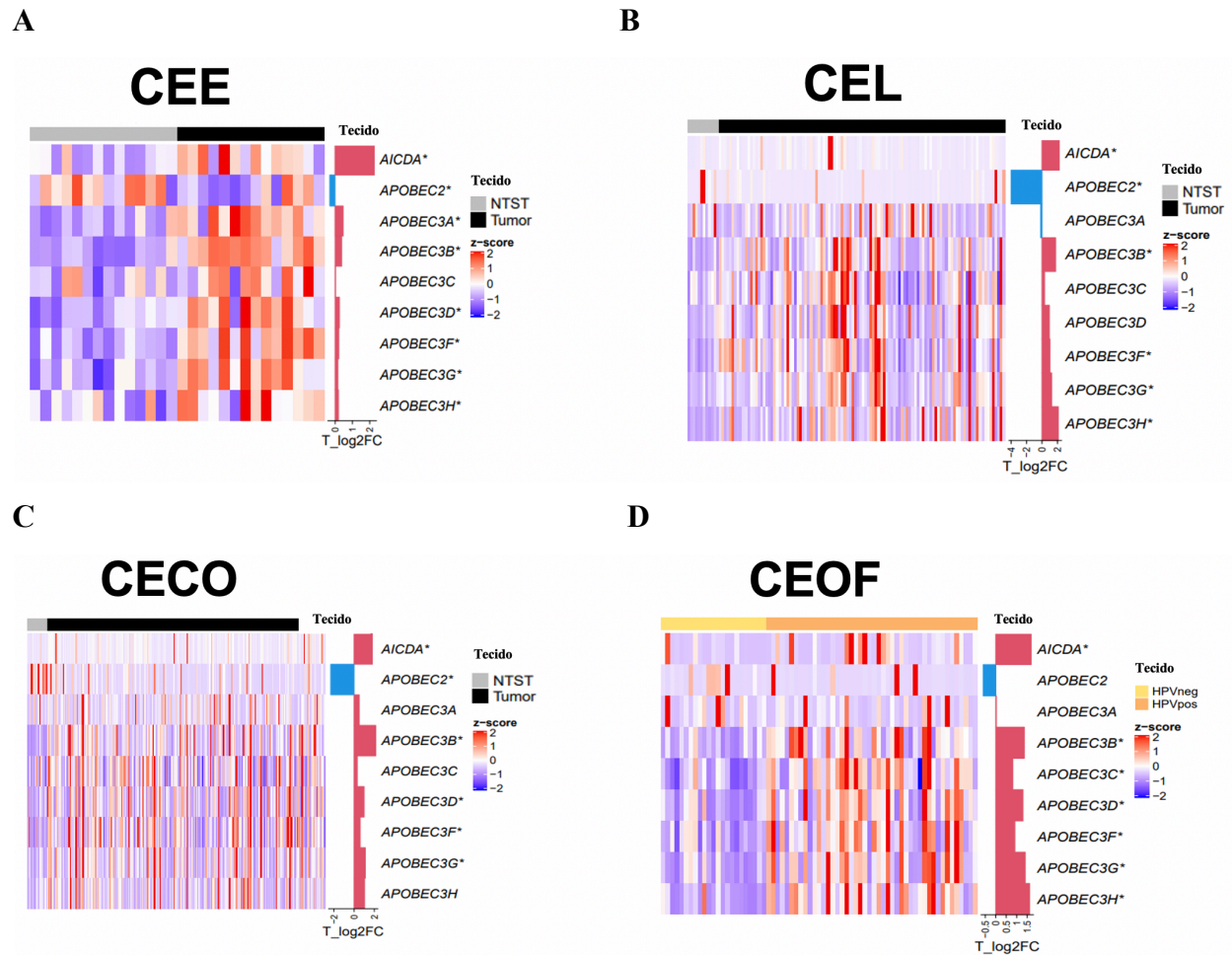


Figura 4.9: Análise da expressão gênica da família *AID/APOBEC* em tumores de cabeça e pescoço e de esôfago. *Heatmap* da expressão do RNAm das *APOBECs* em CEE (A), CEL (B), CECO (C) entre a mucosa tumoral e não tumoral adjacente ao tumor e CEOF (D) entre tumores HPV⁻ e HPV⁺. Para todos os tipos tumorais, foram utilizados dados do consórcio do TCGA, com exceção de CEE, cujos dados foram provenientes de dados gerados pelo grupo com pacientes do INCA. Teste *Wilcoxon matched-pairs signed rank* para CEE, CEL e CECO e Teste *Mann-Whitney* para CEOF, * $p < 0,05$. Vermelho representa maior expressão (escala de 0 até 2) e azul, menor expressão (escala de -2 a 0). CEE, carcinoma epidermoide de esôfago, CEL, carcinoma epidermoide de laringe, CECO, carcinoma epidermoide de cavidade oral, CEOF HPV⁻, carcinoma epidermoide de orofaringe HPV⁻, CEOF HPV⁺, carcinoma epidermoide de orofaringe HPV⁺.

4.1.3 MECANISMOS DE REGULAÇÃO DAS *APOBECs* EM CEE, CEL, CECO, CEOF HPV⁻ e CEOF HPV⁺

Como as assinaturas mutacionais relacionadas à atividade das *APOBECs* estão presentes em todos os tumores analisados, foram encontradas correlações dessas assinaturas mutacionais com a expressão das *APOBECs* e essas enzimas estão desreguladas em CEE, CEL, CECO, CEOF HPV⁻ e CEOF HPV⁺, nós avaliamos os possíveis mecanismos por trás da desregulação dessas enzimas: alterações genéticas (mutações e alterações no número de cópias nos genes que codificam essas enzimas), alterações epigenéticas (miRNAs e metilação do DNA), mecanismos que mimetizam a infecção viral e a interação com o sistema imune.

4.1.3.1 ALTERAÇÕES GENÉTICAS

Foram analisadas alterações genéticas, mutações e alterações do número de cópias, nos genes que codificam as *APOBECs* em tumores de cabeça e pescoço e de esôfago (**Figura 4.10**). Os dados foram obtidos a partir dos estudos *Esophageal Carcinoma* (TCGA, *Firehose Legacy*) para CEE e *Head and Neck Squamous Cell Carcinoma* (TCGA, *Firehose Legacy*) para CEL, CECO e CEOF HPV⁻ e CEOF HPV⁺ pelo site cBioPortal (disponível em <http://www.cbioportal.org>, último acesso em 5 de outubro de 2021). Em CEE, 2,1% dos casos mostraram amplificação de *AICDA* e 1,1% dos casos apresentaram mutações em *APOBEC3F* (**Figura 4.10A**). Em CEL, foi observada amplificação em 0,9% dos casos em *APOBEC3A*, *APOBEC3F*, *APOBEC3G* e *APOBEC3H* e amplificação e mutações em 5% dos casos para *AICDA*, 2,6% dos casos para *APOBEC2*, *APOBEC3B* e *APOBEC3D* e 1,7% dos casos para *APOBEC3C* (**Figura 4.10B**). Em CECO, observamos amplificação em 0,4% dos casos para *AICDA* e em 1,2% dos casos para *APOBEC2* e mutações em 0,8% dos casos em *APOBEC3B* (**Figura 4.10C**). Em CEOF, amplificação em 1,3% dos casos para *AICDA* e *APOBEC2* e mutação em 1,3% dos casos para *APOBEC3D* e *APOBEC3G* (**Figura 4.10D**). Todas as alterações foram encontradas nos tumores HPV⁻, com exceção da alteração em *APOBEC3D* (HPV⁺). Os dados sugerem que mutações e alterações do número de cópias dos genes que codificam as *APOBECs* são raras nos tumores avaliados.

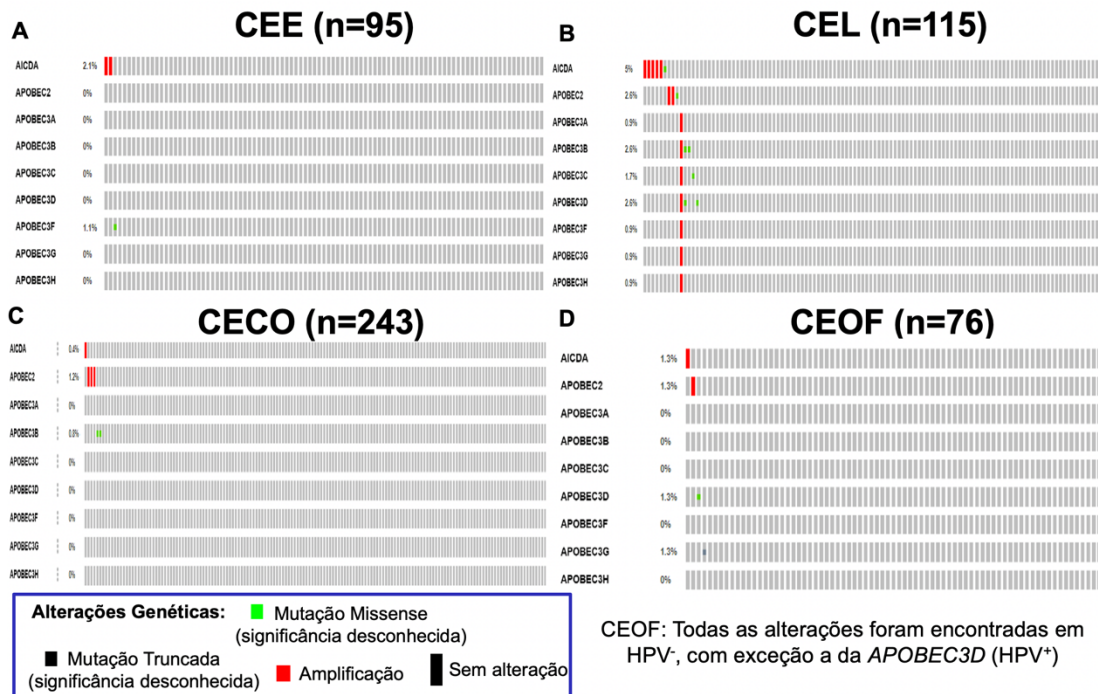


Figura 4.10: Alterações genéticas encontradas nos genes que codificam a família AID/APOBEC em tumores de cabeça e pescoço e de esôfago por paciente. *Oncoprint* de alterações genéticas (mutações e alterações do número de cópias) nos genes que codificam a família AID/APOBEC em CEE (A), CEL (B), CECO (C) e CEOF (D). Em CEOF, todas as alterações foram encontradas nos tumores HPV⁻, com exceção das alterações em *APOBEC3D* (HPV⁺). CEE, carcinoma epidermoide de esôfago, CEL, carcinoma epidermoide de laringe, CECO, carcinoma epidermoide de cavidade oral, CEOF HPV⁻, carcinoma epidermoide de orofaringe HPV⁻, CEOF HPV⁺, carcinoma epidermoide de orofaringe HPV⁺.

4.1.3.2 ALTERAÇÕES EPIGENÉTICAS

Para a análise de alterações epigenéticas, foram avaliados os miRNAs e a metilação do DNA em CEE, CEL, CECO, CEOF HPV⁻ e CEOF HPV⁺.

4.1.3.2.1 MIRNAS

Utilizando três bancos de dados de miRNAs (miRDB, Diana versão 5 e miRmap), foi feito um diagrama de *Venn* dos miRNAs presentes nesses bancos cujo alvo são as enzimas da família AID/APOBECs (**Figura 4.11**). Foram escolhidos os miRNAs comuns aos três bancos para a análise de correlação entre a expressão dos miRNAs e a expressão das *APOBECs* em CEE, CEL, CECO, CEOF HPV⁻ e CEOF HPV⁺, a partir de dados do consórcio TCGA. Para *AICDA*, foram encontrados 28 miRNAs comuns aos três bancos (hsa-mir-155, hsa-mir-4729, hsa-mir-3163, hsa-mir-548c, hsa-mir-5692a, hsa-mir-548x, hsa-mir-548aj, hsa-mir-361, hsa-mir-5002, hsa-mir-6071, hsa-mir-1304, hsa-mir-548ah, hsa-mir-29b, hsa-mir-29a, hsa-mir-502, hsa-mir-29c, hsa-mir-15b, hsa-mir-150, hsa-mir-618, hsa-mir-548at, hsa-mir-5692c, hsa-mir-5692b, hsa-mir-4306, hsa-mir-1208, hsa-mir-548ao, hsa-mir-2115, hsa-mir-548aq e hsa-mir-3680) (**Figura 4.11A**). Para *APOBEC2*, foram sete miRNAs (hsa-mir-509, hsa-mir-3909, hsa-mir-4649, hsa-mir-205, hsa-mir-4267, hsa-mir-4803 e hsa-mir-214) (**Figura 4.11B**). Para *APOBEC3A*, seis miRNAs (hsa-mir-3152, hsa-mir-4297, hsa-mir-4786, hsa-mir-5010, hsa-mir-4775 e hsa-mir-890) (**Figura 4.11C**). Para *APOBEC3B*, 10 miRNAs (hsa-mir-7, hsa-mir-4314, hsa-mir-138, hsa-mir-550b, hsa-mir-505, hsa-mir-4522, hsa-mir-4297, hsa-mir-890, hsa-mir-4646 e hsa-mir-204) (**Figura 4.11D**). Para *APOBEC3C*, 12 miRNAs (hsa-mir-4731, hsa-mir-650, hsa-mir-4443, hsa-mir-3612, hsa-mir-1827, hsa-mir-1291, hsa-mir-4437, hsa-mir-3675, hsa-mir-4283, hsa-mir-6077, hsa-mir-637 e hsa-mir-5589) (**Figura 4.11E**). Para *APOBEC3D*, 16 miRNAs (hsa-mir-4731, hsa-mir-4486, hsa-mir-1291, hsa-mir-340, hsa-mir-371a, hsa-mir-3675, hsa-mir-4524a, hsa-mir-3613, hsa-mir-4437, hsa-mir-3674, hsa-mir-367, hsa-mir-4283, hsa-mir-4482, hsa-mir-3150a, hsa-mir-1285 e hsa-mir-3119) (**Figura 4.11F**). Para *APOBEC3F*, 21 miRNAs (hsa-mir-4731, hsa-mir-940, hsa-mir-340, hsa-mir-660, hsa-mir-6511a, hsa-mir-650, hsa-mir-1291, hsa-mir-3613, hsa-mir-3612, hsa-mir-616, hsa-mir-1911, hsa-mir-4695, hsa-mir-4283, hsa-mir-373, hsa-mir-4437, hsa-mir-371b, hsa-mir-4768, hsa-mir-1827, hsa-mir-6131, hsa-mir-92a e hsa-mir-92b) (**Figura 4.11G**). Para *APOBEC3G*, quatro miRNAs (hsa-mir-4297, hsa-mir-4649, hsa-mir-138 e hsa-mir-1207) (**Figura 4.11H**). E para *APOBEC3H*, oito miRNAs foram selecionados (hsa-mir-4745, hsa-mir-3173, hsa-mir-3163, hsa-mir-3148, hsa-mir-3688, hsa-mir-4773, hsa-mir-3182 e hsa-mir-890) (**Figura 4.11I**).

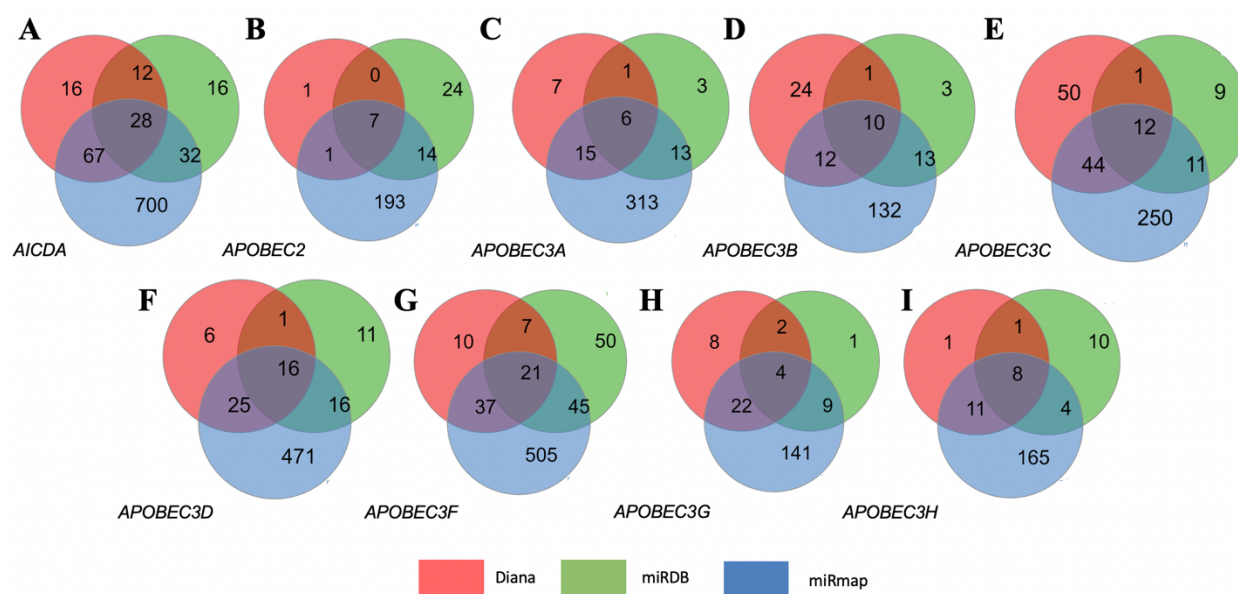


Figura 4.11: Análise dos miRNAs cujos alvos são a família da *AID/APOBEC* utilizando os bancos de miRNAs Diana, miRDB e miRmap. Diagrama de Venn dos miRNAs presentes nos três bancos de dados de miRNAs cujos alvos são *AICDA* (A), *APOBEC2* (B), *APOBEC3A* (C), *APOBEC3B* (D), *APOBEC3C* (E), *APOBEC3D* (F), *APOBEC3F* (G), *APOBEC3G* (H) e *APOBEC3H* (I). Os miRNAs comuns aos três bancos (na intersecção do diagrama) foram selecionados para a análise de correlação com a expressão das *APOBECs*. O diagrama de Venn foi desenhado no site <http://www.interactivenn.net/index2.html>, último acesso em 8 de outubro de 2021.

Em seguida, foi realizada a análise de correlação entre a expressão dos miRNAs e a expressão das *APOBECs* em CEE, CEL, CECO, CEOF HPV⁻ e CEOF HPV⁺. Para esta análise, o nosso foco foi avaliar as correlações inversas, uma vez que seriam indicativas de um mecanismo direto de regulação do gene-alvo pelo miRNA.

Foram observadas correlações inversas significativas da expressão dos miRNAs com *AICDA* em CEE ($\rho: -0,23/\text{hsa-mir-3680-1}$) e CEL ($\rho: -0,202/\text{hsa-mir-548ao}$) e duas em CECO ($\rho: -0,133/\text{hsa-mir1304}$ e $\rho: -0,149/\text{hsa-mir-548x}$) (**Figura 4.12A**). Para a *APOBEC2*, foram encontradas correlações inversas em CEL, CECO e CEE ($\rho: -0,212/\text{hsa-mir-205}$ para CEL; $\rho: -0,252/\text{hsa-mir-205}$ para CECO; e $\rho: -0,231/\text{hsa-mir3909}$ para CEE) (**Figura 4.12B**). A expressão dos miRNAs *hsa-mir-3152* ($\rho: -0,146$) e *hsa-mir-4786* ($\rho: -0,18$) foi inversamente

correlacionada com a expressão da *APOBEC3A* em CECO (**Figura 4.12C**). Para *APOBEC3D*, somente uma correlação inversa significativa em CEL (rho:-0,22/hsa-mir-371a) (**Figura 4.12F**). Para *APOBEC3F*, a maioria das correlações foi inversa e significativa em CEL (rho:-0,306/hsa-mir-92a-2, rho:-0,292/hsa-mir-92a-1, e rho:-0,204/hsa-mir-940), CECO (rho:-0,184/hsa-mir-92a-2, rho:-0,178/hsa-mir-92a-1, e rho:-0,135/hsa-mir-940), CEOF HPV⁻ (rho:-0,492/hsa-mir373) e CEOF HPV⁺ (rho:-0,315/hsa-mir-1911) (**Figura 4.12G**). Para *APOBEC3B* (**Figura 4.12D**), *APOBEC3C* (**Figura 4.12E**) e *APOBEC3H* (**Figura 4.12H**), apenas correlações diretas foram identificadas. Nossos resultados mostraram que esse mecanismo não é comum a todas as *APOBECs* e nem a todos os tumores.

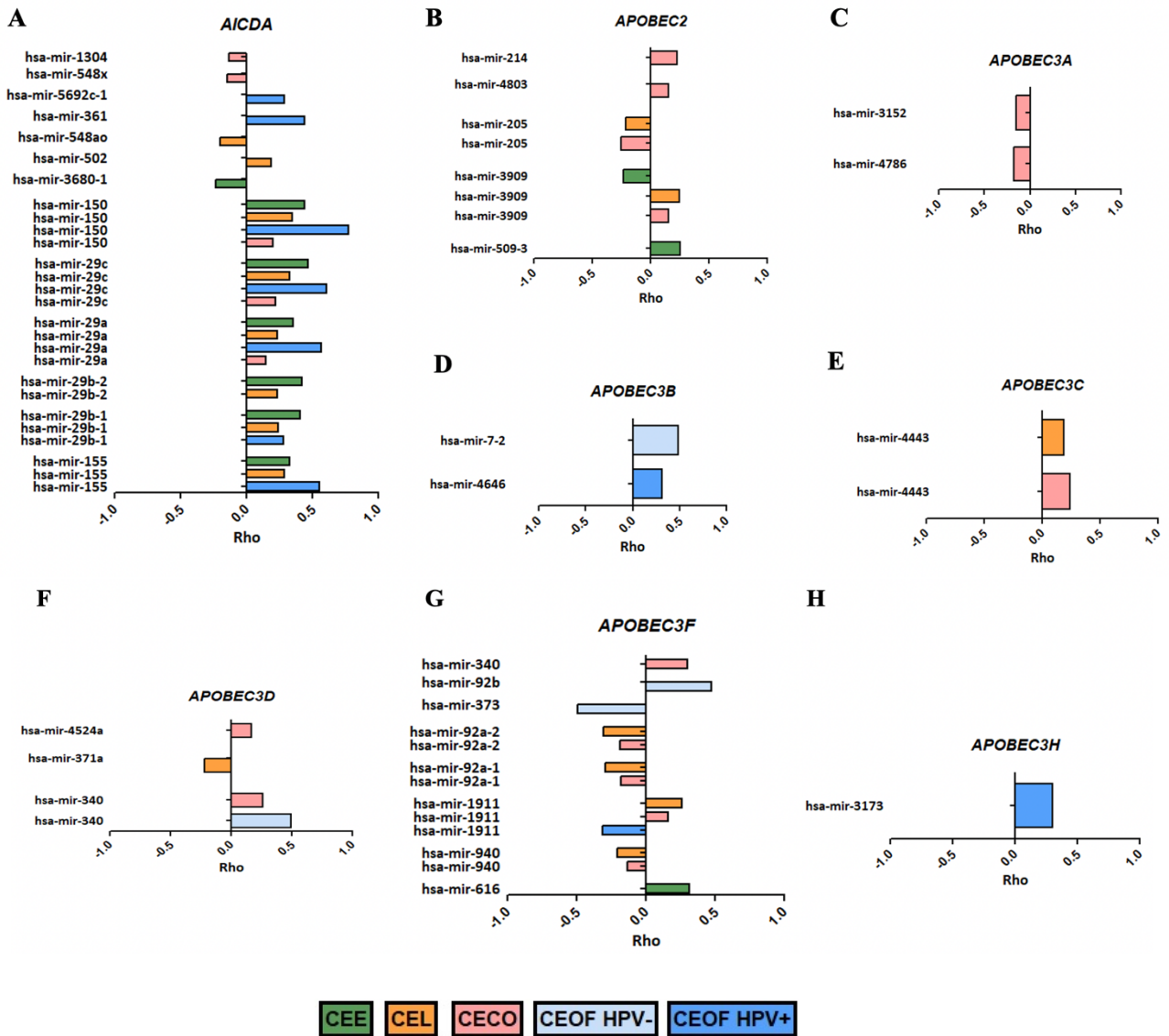


Figura 4.12: Análise de correlação entre a expressão dos miRNAs e da família *AICDA/APOBEC* em tumores de cabeça e pescoço e de esôfago. Análise de correlação entre a expressão dos miRNAs e da *AICDA* (A), da *APOBEC2* (B), da *APOBEC3A* (C), da *APOBEC3B* (D), da *APOBEC3C* (E), da *APOBEC3D* (F), da *APOBEC3F* (G) e da *APOBEC3H* (H) em CEE, CEL, CECO, CEOF HPV⁻ e CEOF HPV⁺ do consórcio do TCGA, utilizando gráfico de barras. Correlação de *Spearman*, $p < 0,05$. Coeficiente de correlação (rho) negativo (-1 a 0) indica correlação inversa e positivo (0 a 1), correlação direta. CEE, carcinoma epidermoide de esôfago, CEL, carcinoma epidermoide de laringe, CECO, carcinoma epidermoide de cavidade oral, CEOF

HPV⁻, carcinoma epidermoide de orofaringe HPV⁻, CEOF HPV⁺, carcinoma epidermoide de orofaringe HPV⁺.

4.1.3.2.2 METILAÇÃO DO DNA

Outro mecanismo epigenético de regulação da expressão gênica bastante estudado em câncer é a metilação do DNA. As *APOBECs* apresentam sítios CpG em sua região promotora, sugerindo que este mecanismo pode impactar a expressão dessas enzimas. Primeiramente, foi analisada a diferença de metilação entre a mucosa tumoral e mucosa adjacente não tumoral das *APOBECs* em CEE, CEL, CECO, CEOF HPV⁻ e CEOF HPV⁺. Depois, foi feita a análise de correlação dos níveis de metilação e expressão dessas enzimas nesses tumores. Na tabela suplementar 1, todas as análises estão descritas por gene, por sonda com localização no genoma em cada subtipo tumoral CEE, CEL, CECO, CEOF HPV⁻ e CEOF HPV⁺ (**Tabela Suplementar 1**). Na **Figura 4.13**, apenas os sítios CpG diferencialmente metilados entre tecido tumoral e não tumoral e correlacionados com a expressão da *APOBEC* analisada são mostrados.

Para *AICDA*, foi observada uma hipometilação no tumor em relação à mucosa adjacente e uma correlação direta significativa entre a expressão e a metilação nesse gene ($\Delta\beta$:-0,29, rho:0,306, cg00245465, enhancer) em CEOF HPV⁺ (**Figura 4.13A**). Para *APOBEC2*, uma hipermetilação no tumor em relação à mucosa adjacente e uma correlação inversa entre a expressão e a metilação em CEL ($\Delta\beta$:0,063, rho:-0,22, cg10567498, enhancer); uma hipometilação no tumor com correlação inversa em CEOF HPV⁻ ($\Delta\beta$:-0,022, rho:-0,45, cg12253732, em região 3'UTR) e em CEOF HPV⁺ ($\Delta\beta$:-0,33, rho:-0,57, cg17548735, enhancer) (**Figura 4.13B**). Para *APOBEC3A*, hipometilação no tumor com correlação inversa em CEL ($\Delta\beta$:-0,07, rho:-0,19, cg27504369, em região 5'UTR), em CECO ($\Delta\beta$:-0,12, rho:-0,14, cg24360617, em região TSS200) e em CEOF HPV⁺ ($\Delta\beta$:-0,09, rho:-0,28, cg24360617, em região TSS200) (**Figura 4.13C**). Para *APOBEC3B*, hipermetilação no tumor em relação à mucosa não tumoral com correlação inversa entre a expressão e a metilação nesse gene em CEE ($\Delta\beta$:0,132, rho:-0,35, cg16045423, enhancer) e em CEL ($\Delta\beta$:0,07, rho:-0,21, cg11816043, enhancer). Foi encontrada também uma hipometilação no tumor com correlações inversas em CEOF HPV⁺ ($\Delta\beta$:-0,15, rho:-0,41, cg16045423, enhancer; $\Delta\beta$:-0,12, rho:-0,38, cg11816043, enhancer) (**Figura 4.13D**). Para *APOBEC3D*, foi observada uma hipometilação no

tumor com correlação inversa em CEE ($\Delta\beta$:-0,27, rho:-0,68, cg16066354, enhancer; $\Delta\beta$:-0,098 rho:-0,51 cg08665930, enhancer), além de uma hipermetilação no tumor com correlação inversa em CEOF HPV⁻ ($\Delta\beta$:0,11, rho: -0,63, cg16066354, enhancer) e em CEOF HPV⁺ ($\Delta\beta$:0,05, rho:-0,55, cg16066354, enhancer) (**Figura 4.13E**). Para *APOBEC3F*, foi observada hipometilação no tumor com correlação direta em CEL ($\Delta\beta$:-0,134, rho:0,27, cg10191799, na região TSS1500; $\Delta\beta$:-0,295, rho: 0,61, cg19333614, na região 3'UTR), em CECO ($\Delta\beta$:-0,235, rho:0,41, cg19333614, na região 3'UTR), em CEOF HPV⁻ ($\Delta\beta$:-0,19, rho: 0,62, cg19333614, na região 3'UTR) e em CEOF HPV⁺ ($\Delta\beta$:-0,21, rho: 0,60, cg19333614, na região 3'UTR). Também foram observadas hipermetilação no tumor com correlação direta em CEE ($\Delta\beta$:0,0785, rho: 0,27, cg09582545, na região 3'UTR) e em CEL ($\Delta\beta$:0,06, rho: 0,29, cg09582545, na região 3'UTR) (**Figura 4.13F**). Para *APOBEC3G*, foram observadas hipometilação no tumor com correlação direta em CECO ($\Delta\beta$:-0,034, rho:0,32, cg24519526, na região 3'UTR) e inversa em CEOF HPV⁺ ($\Delta\beta$:-0,109, rho:-0,29, cg06398754, no corpo) (**Figura 4.13G**). Para *APOBEC3H*, foram observadas hipometilação no tumor com correlação inversa em CEE ($\Delta\beta$:-0,084, rho:-0,25, cg17972162, enhancer) e em CEOF HPV⁺ ($\Delta\beta$:-0,17, rho:-0,43, cg17972162, enhancer) e direta em CEOF HPV⁻ ($\Delta\beta$:-0,11, rho:0,565, cg03582215, na região 3'UTR). Além de uma hipermetilação no tumor com correlação inversa em CEE ($\Delta\beta$:0,063, rho:-0,42, cg24424889, enhancer) (**Figura 4.13H**). Em geral, apesar do elevado número de sítios CpG avaliados para cada gene, apenas sítios CpG individuais apresentaram diferenças de metilação entre os tecidos tumorais e não tumorais e correlação com a expressão gênica. Além disso, a sobreposição desses achados entre sítios tumorais foi muito baixa. Já que regiões diferencialmente metiladas (incluindo mais de um sítio CpG) normalmente estão associadas à regulação da expressão gênica (RAKYAN et al., 2011), estes dados sugerem que a metilação do DNA não seja o principal mecanismo de regulação da expressão das *APOBECs* nos tumores avaliados.

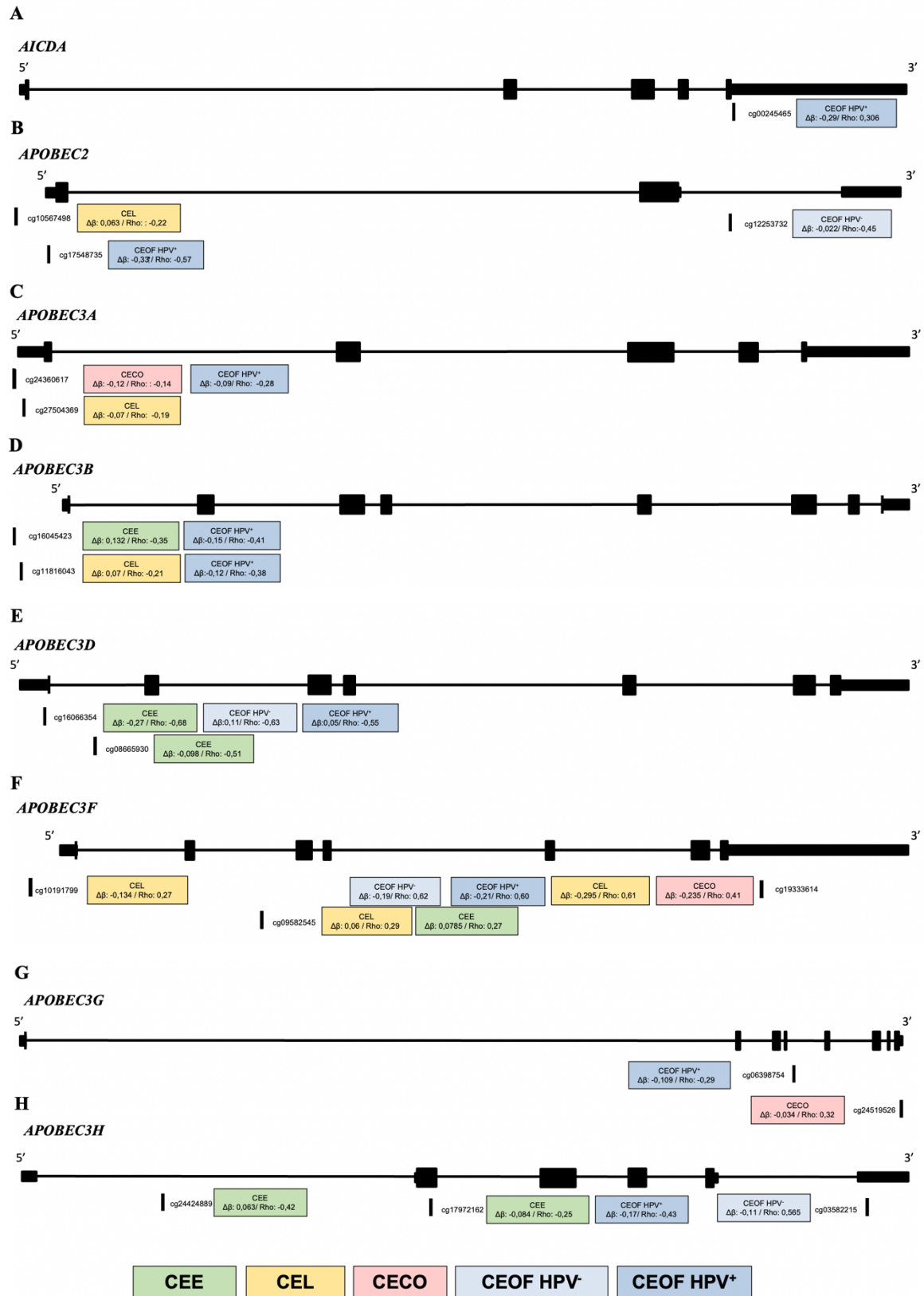


Figura: 4.13: Diferença de metilação das *APOBECs* entre as mucosas adjacentes não tumorais e tumorais de pacientes com CEE, CEL, CECO, CEOF HPV⁻ e CEOF HPV⁺ do INCA e correlação desse nível com a expressão dessas enzimas nesses tumores no banco de dados do TCGA. Esquema representativo dos genes *AICDA* (A), *APOBEC2* (B), *APOBEC3A* (C), *APOBEC3B* (D), *APOBEC3D* (E), *APOBEC3F* (F), *APOBEC3G* (G) e *APOBEC3H* (H) com a localização da sonda de metilação de cada gene, com o valor de $\Delta\beta$ (mediana da diferença de valor de beta de metilação entre tumor e mucosa adjacente não tumoral), bem como o coeficiente de correlação entre o nível de metilação de sítios CpG individuais e a expressão das enzimas em CEE, CEL, CECO, CEOF HPV⁻ e CEOF HPV⁺. Teste de *Wilcoxon matched-pairs* para a análise de metilação e correlação de *Spearman*, $p < 0,05$. CEE, carcinoma epidermoide de esôfago, CEL, carcinoma epidermoide de laringe, CECO, carcinoma epidermoide de cavidade oral, CEOF HPV⁻, carcinoma epidermoide de orofaringe HPV⁻, CEOF HPV⁺, carcinoma epidermoide de orofaringe HPV⁺.

4.1.3.3 CORRELAÇÃO DO NÍVEL DE METILAÇÃO DE ELEMENTOS TRANSPONÍVEIS COM A EXPRESSÃO DAS *APOBECs* EM CEE, CEL, CECO, CEOF HPV⁻ e CEOF HPV⁺

De acordo com os dados obtidos, as alterações genéticas e epigenéticas parecem não ser os principais mecanismos de regulação das *APOBECs*. Por outro lado, um mecanismo já descrito na literatura de regulação dessas enzimas é a infecção viral, conforme demonstrado em CEOF HPV⁺ (Figura 4.9D). Porém, os demais tumores não têm como principal fator de risco a infecção por vírus. Assim, o nosso próximo passo foi avaliar possíveis mecanismos que mimetizariam a infecção viral. Já foi observado na literatura que tumores apresentam uma hipometilação global. Com isso, são ativados elementos retrotransponíveis (ALU, LINE, LTRs - região de repetição terminal longa, do inglês *long terminal repeat*) que podem levar à instabilidade genômica (FEINBERG & TYCKO, 2004; BAYLIN & OHM, 2006; BAYLIN & JONES, 2011; LIMA et al., 2011; LIMA, 2012). Além disso, já foi demonstrado que esses elementos, uma vez ativados, mimetizariam a infecção viral, podendo levar a uma resposta imune antiviral mediada por interferon em células tumorais (CHIAPPINELLI et al., 2015; ROULOIS et al., 2015). Mediante ao exposto, nossa hipótese é que a hipometilação global nesses tumores levaria à ativação dos

elementos retrotransponíveis que mimetizariam a infecção viral, o que poderia levar ao aumento de expressão das *APOBECs* (enzimas de resposta antiviral). Assim, o próximo passo foi correlacionar os níveis de metilação dos elementos transponíveis com a expressão das *APOBECs* em CEE, CEL, CECO e CEOF HPV⁻ e CEOF HPV⁺. Nesta análise, focamos em correlações inversas significativas entre os níveis de metilação desses elementos e a expressão das *APOBECs* (Tabela 4.2), ou seja, uma hipometilação desses elementos com a superexpressão dessas enzimas. Para todos os tipos tumorais, a metilação de elementos ALU foi mais frequentemente correlacionada com a expressão das *AID/APOBECs*. Além disso, nesta análise, podemos destacar o elevado número de correlações significativas entre a metilação de elementos ALU individuais e a expressão de *APOBEC3A* em CEE (6.998 correlações inversas, sendo 28,2% do total de elementos ALU avaliados).

Tabela 4.2: Correlação do nível de metilação dos elementos transponíveis com a expressão das *APOBECs* em CEE, CEL, CECO, CEOF HPV⁻ e CEOF HPV⁺.

Nº Rho Negativo	CEE			CEL			CECO			CEOF HPV ⁻			CEOF HPV ⁺		
	ALU (n:24810)	LINE1 (n:2984)	LTR (n:1036)	ALU (n:25104)	LINE1 (n:3171)	LTR (n:1139)	ALU (n:23718)	LINE1 (n:2815)	LTR (n:927)	ALU (n:21476)	LINE1 (n:2430)	LTR (n:782)	ALU (n:21476)	LINE1 (n:2430)	LTR (n:782)
<i>AICDA</i>	1 (0,004%)	5 (0,2%)	1 (0,1%)	0	0	0	2 (0,008%)	0	0	0	0	0	3 (0,01%)	4 (0,2%)	3 (0,4%)
<i>APOBEC2</i>	0	1 (0,03%)	0	0	0	0	176 (0,7%)	57 (2,0%)	23 (2,5%)	0	0	0	0	0	0
<i>APOBEC3A</i>	6998 (28,2%)	471 (15,8%)	80 (0,8%)	0	0	0	43 (0,2%)	12 (0,4%)	2 (0,2%)	0	0	0	0	0	0
<i>APOBEC3B</i>	1 (0,004%)	0	0	0	0	0	0	1 (0,04%)	0	0	0	0	6 (0,03%)	5 (0,2%)	3 (0,4%)
<i>APOBEC3C</i>	0	0	0	0	0	0	80 (0,3%)	21 (0,75%)	5 (0,5%)	8 (0,04%)	10 (0,4%)	0	146 (0,7%)	26 (1,1%)	11 (1,4%)
<i>APOBEC3D</i>	0	0	0	4 (0,02%)	4 (0,1%)	1 (0,09%)	152 (0,6%)	59 (2,1%)	9 (0,97%)	4 (0,02%)	0	0	1408 (6,6%)	101 (4,2%)	22 (2,8%)
<i>APOBEC3F</i>	0	0	0	186 (0,7%)	26 (0,8%)	5 (0,4%)	67 (0,3%)	17 (0,6%)	2 (0,2%)	0	0	0	666 (3,1%)	34 (1,4%)	15 (1,9%)
<i>APOBEC3G</i>	0	0	0	31 (0,1%)	7 (0,2%)	1 (0,09%)	187 (0,8%)	84 (3,0%)	5 (0,5%)	128 (0,6%)	6 (0,2%)	2 (0,26%)	1304 (6,1%)	112 (4,6%)	20 (2,6%)
<i>APOBEC3H</i>	0	1 (0,03%)	0	0	1 (0,03%)	0	0	72 (2,6%)	9 (0,97%)	0	0	2 (0,26%)	626 (2,9%)	71 (2,9%)	15 (1,9%)

A tabela mostra o número de correlações inversas significativas entre os níveis de metilação de elementos retrotransponíveis individuais e a expressão de *APOBECs* em CEE, CEL, CECO, CEOF HPV⁻, e CEOF HPV⁺. Correlação de *Spearman*, $p < 0,05$. CEE, carcinoma epidermoide de esôfago, CEL, carcinoma epidermoide de laringe, CECO, carcinoma epidermoide de cavidade oral, CEOF HPV⁻, carcinoma epidermoide de orofaringe HPV⁻, CEOF HPV⁺, carcinoma epidermoide de orofaringe HPV⁺, ALU, elemento repetitivo pertencente à família SINE (elementos nucleares curtos intercalados, do inglês *Short interspersed nuclear elements*), LINE, elementos nucleares

longos intercalados, do inglês *Long interspersed nuclear elements*, LTR, região de repetição terminal longa, do inglês *long terminal repeat*.

4.1.3.4 INTERAÇÃO COM O SISTEMA IMUNE

Conforme já foi descrito, os elementos transponíveis ativados podem mimetizar a infecção viral, levando a uma resposta imune antiviral mediada por interferon em células tumorais (CHIAPPINELLI et al., 2015; ROULOIS et al., 2015). Assim, podendo levar à indução das *APOBECs* já que essas enzimas são enzimas de resposta antiviral. Porém, além desse mecanismo, a interação de células tumorais com células imunes presentes no microambiente tumoral também poderia desempenhar um papel no estabelecimento de assinaturas mutacionais mediadas por *APOBECs*. Com intuito de avaliar a plausibilidade desse mecanismo regulatório, avaliamos a correlação entre a proporção de células imunes na massa tumoral e a contribuição das assinaturas mutacionais 2/13 em cada tipo tumoral (**Figura 4.14**). A correlação entre a proporção dos diferentes tipos celulares e a expressão das *APOBECs* associadas à presença das assinaturas mutacionais de acordo com nosso estudo (Figura 4.8) também foi avaliada. Em CEE, a presença de macrófagos M0 foi diretamente correlacionada com a contribuição das assinaturas ($\rho:0,27$, $p=0,009$), enquanto essa correlação foi inversa para macrófagos M2 ($\rho:-0,24$, $p=0,02$). A presença desses tipos celulares também foi correlacionada com a expressão das enzimas *APOBEC3A* e *APOBEC3B* (**Figura 4.14A**), associadas à presença da assinatura mutacional, de acordo com nossos dados. Já em CEL, correlações inversas foram observadas entre a contribuição das assinaturas 2/13 e a presença de plasmócitos ($\rho:-0,26$, $p=0,006$) e células TCD4 naive ($\rho:-0,33$, $p=0,0006$), enquanto para macrófagos M1 essa correlação foi direta ($\rho:0,255$, $p=0,008$). Neste tipo tumoral, a expressão de *APOBEC3B*, *3C*, *3D*, *3F* e *3G* também foi correlacionada à presença destes tipos celulares (**Figura 4.14B**). Para CECO, correlações diretas de células TCD4 de memória ativadas ($\rho:0,207$, $p=0,0016$), células TCD8 ($\rho:0,189$, $p=0,004$) e macrófagos M1 ($\rho:0,244$, $p=0,0002$); e correlações indiretas de células Treg ($\rho:-0,157$, $p=0,017$) com a contribuição das assinaturas 2/13 foram observadas. A presença dessas células também foi correlacionada com a expressão de *APOBEC2*, *3A*, *3C* e *3D* (**Figura 4.14C**). Em CEOF HPV-, células TCD8 ($\rho:-0,528$, $p=0,0096$) e macrófagos M1 ($\rho:-0,421$, $p=0,0452$) foram

inversamente correlacionados com as assinaturas mutacionais, enquanto para macrófagos M0 essa correlação foi direta (rho:0,445, p=0,033) (**Figura 4.14D**). Finalmente, em CEOF HPV⁺, apenas a presença de mastócitos foi correlacionada com a contribuição da assinatura de APOBECs, sendo essa correlação direta para mastócitos em repouso (rho:0,346, p=0,016) e inversa para mastócitos ativados (rho:-0,459, p=0,001) (**Figura 4.14E**).

A		CEE	SBS 2/13	<i>APOBEC3A</i>	<i>APOBEC3B</i>
Macrófago M0		Rho: 0,27 p: 0,009	Rho: 0,27 p: 0,009	Rho: 0,18 p: 0,08	Rho: 0,309 p: 0,003
Macrófago M2		Rho: -0,24 p: 0,02	Rho: -0,24 p: 0,02	Rho: -0,29 p: 0,005	Rho: 0,041 p: 0,7

B	CEL	SBS 2/13	<i>APOBEC3B</i>	<i>APOBEC3C</i>	<i>APOBEC3D</i>	<i>APOBEC3F</i>	<i>APOBEC3G</i>
Plasmócito		Rho: -0,26 p: 0,006	Rho: -0,17 p: 0,068	Rho: -0,25 p: 0,0075	Rho: -0,14 p: 0,127	Rho: -0,19 p: 0,046	Rho: -0,27 p: 0,004
Célula TCD4 Naive		Rho: -0,33 p: 0,0006	Rho: -0,075 p: 0,42	Rho: -0,14 p: 0,135	Rho: -0,216 p: 0,02	Rho: -0,12 p: 0,18	Rho: -0,206 p: 0,027
Macrófago M1		Rho: 0,255 p: 0,0078	Rho: 0,17 p: 0,07	Rho: 0,234 p: 0,012	Rho: 0,28 p: 0,0028	Rho: 0,23 p: 0,012	Rho: 0,39 p: 1,28x10 ⁻⁵

C	CECO	SBS 2/13	<i>AICDA</i>	<i>APOBEC2</i>	<i>APOBEC3A</i>	<i>APOBEC3B</i>	<i>APOBEC3C</i>	<i>APOBEC3D</i>
Célula TCD4 memória ativada		Rho: 0,207 p: 0,0016	Rho: 0,069 p: 0,28	Rho: -0,17 p: 0,0066	Rho: 0,25 p: 9,76x10 ⁻⁵	Rho: 0,099 p: 0,126	Rho: 0,145 p: 0,024	Rho: 0,33 p: 1,05x10 ⁻⁷
Célula TCD8		Rho: 0,189 p: 0,004	Rho: -0,10 p: 0,12	Rho: -0,14 p: 0,029	Rho: 0,22 p: 0,0005	Rho: -0,08 p: 0,216	Rho: 0,064 p: 0,316	Rho: 0,35 p: 2,2x10 ⁻⁸
Célula Treg		Rho: -0,157 p: 0,017	Rho: 0,108 p: 0,09	Rho: 0,21 p: 0,0009	Rho: -0,0026 p: 0,97	Rho: -0,068 p: 0,29	Rho: 0,0327 p: 0,61	Rho: 0,09 p: 0,165
Macrófago M1		Rho: 0,244 p: 0,000189	Rho: -0,09 p: 0,16	Rho: 0,09 p: 0,16	Rho: 0,2 p: 0,0016	Rho: 0,108 p: 0,09	Rho: 0,196 p: 0,0023	Rho: 0,35 p: 2,07x10 ⁻⁸

D	CEOFP HPV ⁻	SBS 2/13
Célula TCD8		Rho: -0,528 p: 0,0096
Macrófago M0		Rho: 0,445 p: 0,033
Macrófago M1		Rho: -0,421 p: 0,0452

E	CEOFP HPV ⁺	SBS 2/13	<i>APOBEC3A</i>
Mastócito em repouso		Rho: 0,346 p: 0,016	Rho: -0,13 p: 0,36
Mastócito ativado		Rho: -0,459 p: 0,001	Rho: -0,027 p: 0,849

Figura 4.14: Correlação entre assinatura mutacional relacionada a APOBECs, expressão gênica dessas enzimas e CIBERSORT das células do sistema imune em pacientes com tumores de cabeça e pescoço e de esôfago. Representação da correlação entre a assinatura mutacional relacionada a APOBECs (SBS2/13), a expressão gênica das *APOBECs* (*AICDA*, *APOBEC2*, *APOBEC3A-APOBEC3G*) e a presença das células do sistema imune em CEE (A), CEL (B), CECO (C), CEOF HPV⁻ (D) e CEOF HPV⁺ (E). Rho – coeficiente de correlação; p: valor de p; em negrito, estatisticamente significativo, p<0,05. Correlação de *Spearman*. CEE, carcinoma epidermoide de esôfago, CEL, carcinoma epidermoide de laringe, CECO, carcinoma epidermoide de cavidade

oral, CEOF HPV⁻, carcinoma epidermoide de orofaringe HPV⁻, CEOF HPV⁺, carcinoma epidermoide de orofaringe HPV⁺, SBS: assinaturas mutacionais de substituição de base única relacionadas à atividade das APOBECs (2/13), célula Treg, célula T regulatória, Macrófago M0, *naive* (sem contato com antígeno, não ativado), *naive*, sem contato com antígeno (não ativado), Macrófago M1, pró-inflamatório, Macrófago M2, pró-tumoral.

Frente a esses achados, nos perguntamos se as diferenças de expressão entre os tumores e tecidos não tumorais, assim como se as correlações entre a expressão das *APOBECs* e a contribuição as assinaturas mutacionais mediadas por essas enzimas seriam, na verdade, um reflexo da composição do infiltrado inflamatório no microambiente tumoral. Para isso, foi analisada a expressão das *APOBECs* em experimentos de *single cell* realizados em CEE cujos dados encontram-se depositados no banco GEO sob o número GSE160269. Estes dados nos mostraram que todas as *APOBECs* são expressas em células de origem epitelial, correspondentes principalmente às células malignas dentro da massa tumoral. Para *APOBEC3A*, as células epiteliais (com 7,0% de células com detecção de expressão) e mieloides (com 15,8% de células com detecção de expressão) são as que mais expressam esse gene quando comparadas aos outros subtipos celulares (células B – 0,5%, endoteliais – 0,1%, fibroblasto – 0,1%, FRC – 0,3%, pericito – 0,1% e células T – 0,1%). O tipo celular que mais expressa *APOBEC3B* é a célula epitelial (11,4%). A *APOBEC3C* é mais expressa em células T (44,6% de células com detecção de expressão), enquanto em 23,1% das células epiteliais a expressão desse gene foi detectada. As células T também são aquelas que mais expressam *APOBEC3D* (9,4% de células com detecção de expressão). As células epiteliais têm 1,1% de detecção desse gene. A *APOBEC3F* é mais expressa nas células T (9,7% de células com detecção de expressão) seguida por células epiteliais (7,9%) e células B (7,2%). A *APOBEC3G* é mais expressa nas células T (50,6% de células com detecção de expressão). As células epiteliais apresentam 13,3% de células com detecção desse gene. Já a *APOBEC3H* não é expressa em fibroblasto (0,0%) e em células reticulares fibroblásticas (FRC – 0,0%), sendo também mais expressa em células T (16,2%). As células epiteliais apresentam 0,8% de células com detecção desse gene (**Figura 4.15**). A análise estatística com os valores de p está demonstrada na **Tabela 4.3**.

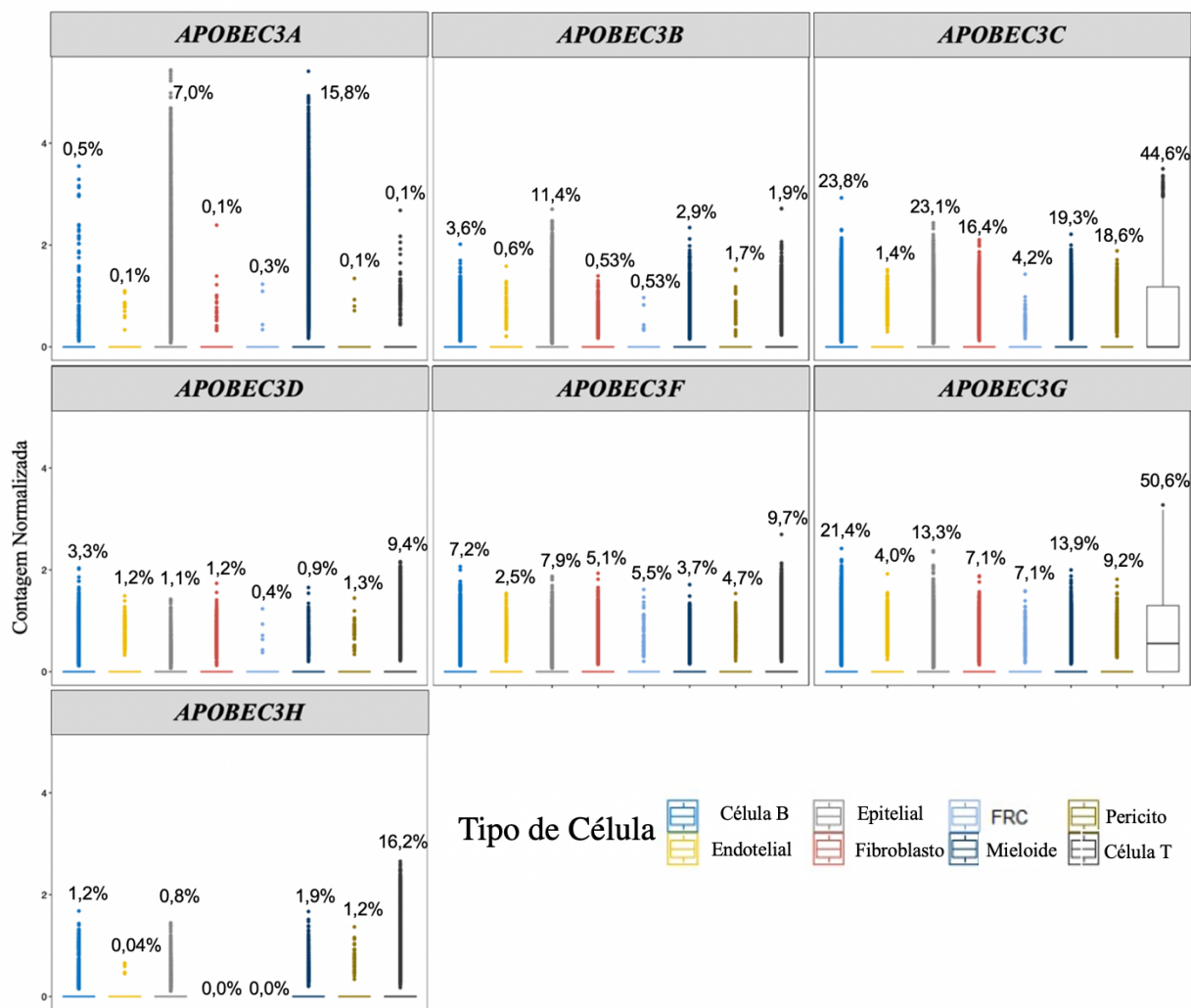


Figura 4.15: Comparação da expressão gênica das *APOBECs* em células presentes no microambiente tumoral e em células tumorais em CEE. Boxplot mostrando a expressão das *APOBEC3A*, *APOBEC3B*, *APOBEC3C*, *APOBEC3D*, *APOBEC3F*, *APOBEC3G* e *APOBEC3H* em células B, células endoteliais, células epiteliais (tumorais), fibroblastos, FRC (células reticulares fibroblásticas), células mieloides, pericitos e células T. Análise realizada a partir de dados de experimentos de *single cell* de CEE depositados no banco GEO, sob o número GSE160269. CEE, carcinoma epidermoide de esôfago, FRC, células reticulares fibroblásticas. Os números indicam a porcentagem de células com detecção da expressão das *APOBECs*. Teste não paramétrico *Kruskal Wallis* e o pós-teste de *Dunn*.

Tabela 4.3: Comparação da expressão das APOBECs nos subtipos celulares em par com seu respectivo valor de p nos dados de single cell de CEE depositado no GEO.

Tipo de Célula	Valor de p						
	<i>APOBEC3A</i>	<i>APOBEC3B</i>	<i>APOBEC3C</i>	<i>APOBEC3D</i>	<i>APOBEC3F</i>	<i>APOBEC3G</i>	<i>APOBEC3H</i>
Cél B vs Cél Endotelial	4,2x10⁻⁸	2,2x10⁻¹⁶	2,2x10⁻¹⁶	2,2x10⁻¹⁶	2,2x10⁻¹⁶	2,2x10⁻¹⁶	2,2x10⁻¹⁶
Cél B vs Cél Epitelial	2,2x10⁻¹⁶	2,2x10⁻¹⁶	4,5x10⁻¹¹	2,2x10⁻¹⁶	0,015	2,2x10⁻¹⁶	8,5x10⁻⁹
Cél B vs Fibroblasto	2,2x10⁻¹⁶	2,2x10⁻¹⁶	2,2x10⁻¹⁶	2,2x10⁻¹⁶	2,2x10⁻¹⁶	2,2x10⁻¹⁶	NA
Cél B vs FRC	0,35	2,2x10⁻⁹	2,2x10⁻¹⁶	9,5x10⁻⁹	0,02	2,2x10⁻¹⁶	NA
Cél B vs Cél Mieloide	2,2x10⁻¹⁶	0,0001	2,2x10⁻¹⁶	2,2x10⁻¹⁶	2,2x10⁻¹⁶	2,2x10⁻¹⁶	4,9x10⁻⁹
Cél B vs Pericito	0,0051	7,6x10⁻⁸	4,9x10⁻¹⁰	2,3x10⁻⁹	2,5x10⁻⁷	2,2x10⁻¹⁶	0,84
Cél B vs Cél T	2,2x10⁻¹⁶	2,2x10⁻¹⁶	2,2x10⁻¹⁶	2,2x10⁻¹⁶	2,2x10⁻¹⁶	2,2x10⁻¹⁶	2,2x10⁻¹⁶
Cél Endotelial vs Cél Epitelial	2,2x10⁻¹⁶	2,2x10⁻¹⁶	2,2x10⁻¹⁶	0,24	2,2x10⁻¹⁶	2,2x10⁻¹⁶	2,2x10⁻¹⁶
Cél Endotelial vs Fibroblasto	0,22	0,23	2,2x10⁻¹⁶	0,89	2,2x10⁻¹⁶	2,2x10⁻¹⁶	NA
Cél Endotelial vs FRC	0,058	0,68	4,8x10⁻¹³	0,012	7,2x10⁻¹⁰	4,4x10⁻⁷	NA
Cél Endotelial vs Cél Mieloide	2,2x10⁻¹⁶	2,2x10⁻¹⁶	2,2x10⁻¹⁶	0,0032	3,5x10⁻⁸	2,2x10⁻¹⁶	2,2x10⁻¹⁶
Cél Endotelial vs Pericito	0,74	2,5x10⁻⁹	2,2x10⁻¹⁶	0,7	3,3x10⁻¹⁰	2,2x10⁻¹⁶	2,2x10⁻¹⁶
Cél Endotelial vs Cél T	0,79	2,2x10⁻¹⁶	2,2x10⁻¹⁶	2,2x10⁻¹⁶	2,2x10⁻¹⁶	2,2x10⁻¹⁶	2,2x10⁻¹⁶
Cél Epitelial vs Fibroblasto	2,2x10⁻¹⁶	2,2x10⁻¹⁶	2,2x10⁻¹⁶	0,12	2,2x10⁻¹⁶	2,2x10⁻¹⁶	NA
Cél Epitelial vs FRC	2,2x10⁻¹⁶	2,2x10⁻¹⁶	2,2x10⁻¹⁶	0,024	0,0031	3x10⁻¹⁰	NA
Cél Epitelial vs Cél Mieloide	2,2x10⁻¹⁶	2,2x10⁻¹⁶	5,2x10⁻¹⁶	0,012	2,2x10⁻¹⁶	0,00078	2,2x10⁻¹⁶
Cél Epitelial vs Pericito	2,2x10⁻¹⁶	2,2x10⁻¹⁶	0,0014	0,26	1x10⁻⁹	1,3x10⁻⁸	0,013
Cél Epitelial vs Cél T	2,2x10⁻¹⁶	2,2x10⁻¹⁶	2,2x10⁻¹⁶	2,2x10⁻¹⁶	2,2x10⁻¹⁶	2,2x10⁻¹⁶	2,2x10⁻¹⁶
Fibroblasto vs FRC	0,0028	0,99	2,2x10⁻¹⁶	0,012	0,45	1	NA
Fibroblasto vs Cél Mieloide	2,2x10⁻¹⁶	2,2x10⁻¹⁶	6x10⁻¹⁴	0,0003	9,4x10⁻¹⁴	2,2x10⁻¹⁶	NA
Fibroblasto vs Pericito	0,25	2,2x10⁻¹⁶	4,6x10⁻⁵	0,62	0,43	5x10⁻⁶	NA
Fibroblasto vs Cél T	0,024	2,2x10⁻¹⁶	2,2x10⁻¹⁶	2,2x10⁻¹⁶	2,2x10⁻¹⁶	2,2x10⁻¹⁶	NA
FRC vs Cél Mieloide	2,2x10⁻¹⁶	3,1x10⁻⁷	2,2x10⁻¹⁶	0,1	0,00074	4,2x10⁻¹²	NA
FRC vs Pericito	0,21	0,0016	2,2x10⁻¹⁶	0,01	0,27	0,017	NA

FRC vs Cél T	0,05	0,0003	2,2x10⁻¹⁶	2,2x10⁻¹⁶	5,9x10⁻⁸	2,2x10⁻¹⁶	NA
Cél Mieloide vs Pericito	2,2x10⁻¹⁶	0,00021	0,64	0,02	0,0052	2,6x10⁻¹¹	0,004
Cél Mieloide vs Cél T	2,2x10⁻¹⁶	2,2x10⁻¹⁶	2,2x10⁻¹⁶	2,2x10⁻¹⁶	2,2x10⁻¹⁶	2,2x10⁻¹⁶	2,2x10⁻¹⁶
Pericito vs Cél T	0,83	0,55	2,2x10⁻¹⁶	2,2x10⁻¹⁶	2,2x10⁻¹⁶	2,2x10⁻¹⁶	2,2x10⁻¹⁶

Cél, célula, FRC, células reticulares fibroblásticas, vs, *versus*, em negrito, estatisticamente significativo, $p < 0,05$, NA, não aplicado porque não houve expressão da *APOBEC3H* em fibroblasto e FRC. Teste não paramétrico *Kruskal Wallis* e o pós-teste de *Dunn*.

Como a expressão de todas as *APOBECs* foi detectada nas células epiteliais, o nosso próximo passo foi investigar se essas enzimas são expressas nas células epiteliais em todos os pacientes com CEE no modelo de *single cell* (**Figura 4.16**). A *APOBEC3A* é detectada com níveis variáveis de expressão nas células epiteliais dos pacientes, com exceção de seis amostras tumorais (10% das amostras). Das quatro amostras pareadas (os quatro pares de amostras finais no gráfico, sendo adjacente não tumoral e tumor, respectivamente, marcados com retângulo vermelho), uma amostra adjacente não tumoral não apresentou detecção de expressão (P128N), porém o RNAm de *APOBEC3A* foi detectado no tumor deste paciente (P128T). A *APOBEC3B* não foi detectada apenas em duas amostras tumorais (3,33% das amostras). Já das quatro amostras pareadas, uma amostra adjacente não apresentou detecção de expressão (P128N), enquanto a expressão de *APOBEC3B* foi detectada no tumor do mesmo paciente (P128T). O RNAm de *APOBEC3C* foi detectado em todos os tumores e tecidos adjacentes não tumorais, mesmo que em poucas células epiteliais. A *APOBEC3D* foi detectada em todas as amostras de CEE, porém não houve detecção em três amostras adjacentes não tumorais (P126N, P128N e P130N). Já os respectivos tumores (P126T, P128T e P130T) apresentaram detecção, mesmo que em uma pequena proporção de células. A expressão da *APOBEC3F* foi detectada com níveis variáveis de expressão nas células epiteliais dos pacientes com CEE, com exceção de duas amostras tumorais (3,33% das amostras). Das quatro amostras pareadas, três amostras adjacentes não tumorais não apresentaram detecção de expressão desse gene nas células epiteliais (P126N, P128N e P130N), sendo detectada nos tumores respectivos (P126T, P128T e P130T). Apenas uma amostra tumoral (1,67% das amostras) não apresentou detecção da expressão de *APOBEC3G*. A *APOBEC3H* é a que apresenta menor ou ausência de detecção de expressão nas células epiteliais da maioria dos pacientes com CEE. Nas

amostras pareadas, somente uma amostra tumoral apresenta uma detecção consistente da *APOBEC3H* (P130T) (Figura 4.16).

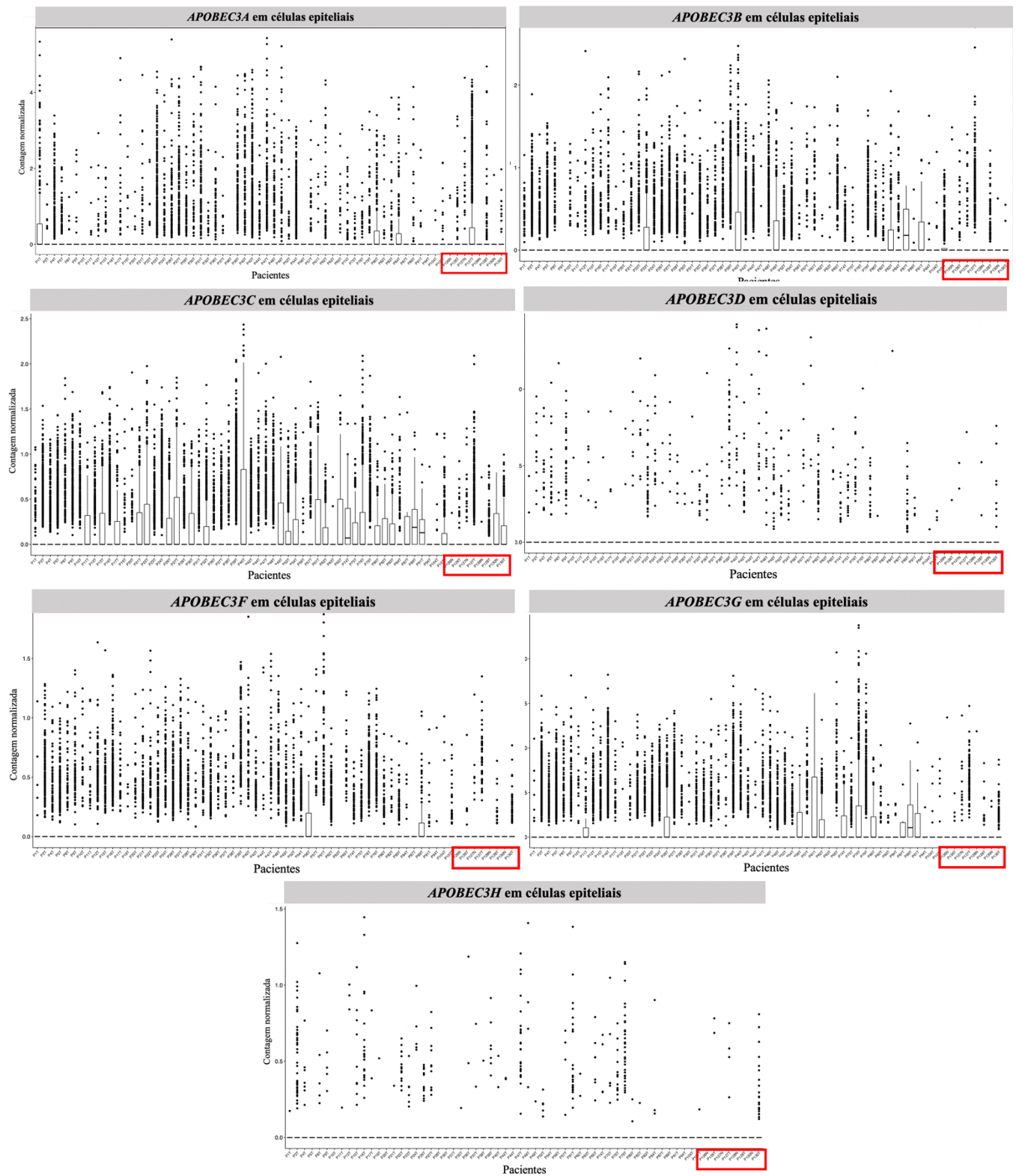


Figura 4.16: Expressão das *APOBECs* nas células epiteliais por paciente com CEE. Boxplot mostrando a expressão das *APOBEC3A*, *APOBEC3B*, *APOBEC3C*, *APOBEC3D*, *APOBEC3F*, *APOBEC3G* e *APOBEC3H* nas células epiteliais por paciente com CEE em dados de experimentos de *single cell* de CEE, depositados no banco GEO sob o número GSE160269. CEE, carcinoma epidermoide de esôfago. Quadrado vermelho representa as 4 amostras pareadas, sendo não tumoral e tumor, respectivamente (P126N e P126T, P127N e P127T, P128N e P128T e P130N e P130T).

A figura 4.17 apresenta um resumo dos resultados obtidos até então. Assim, o nosso próximo passo foi investigar mais a fundo o mecanismo de regulação dessas enzimas.

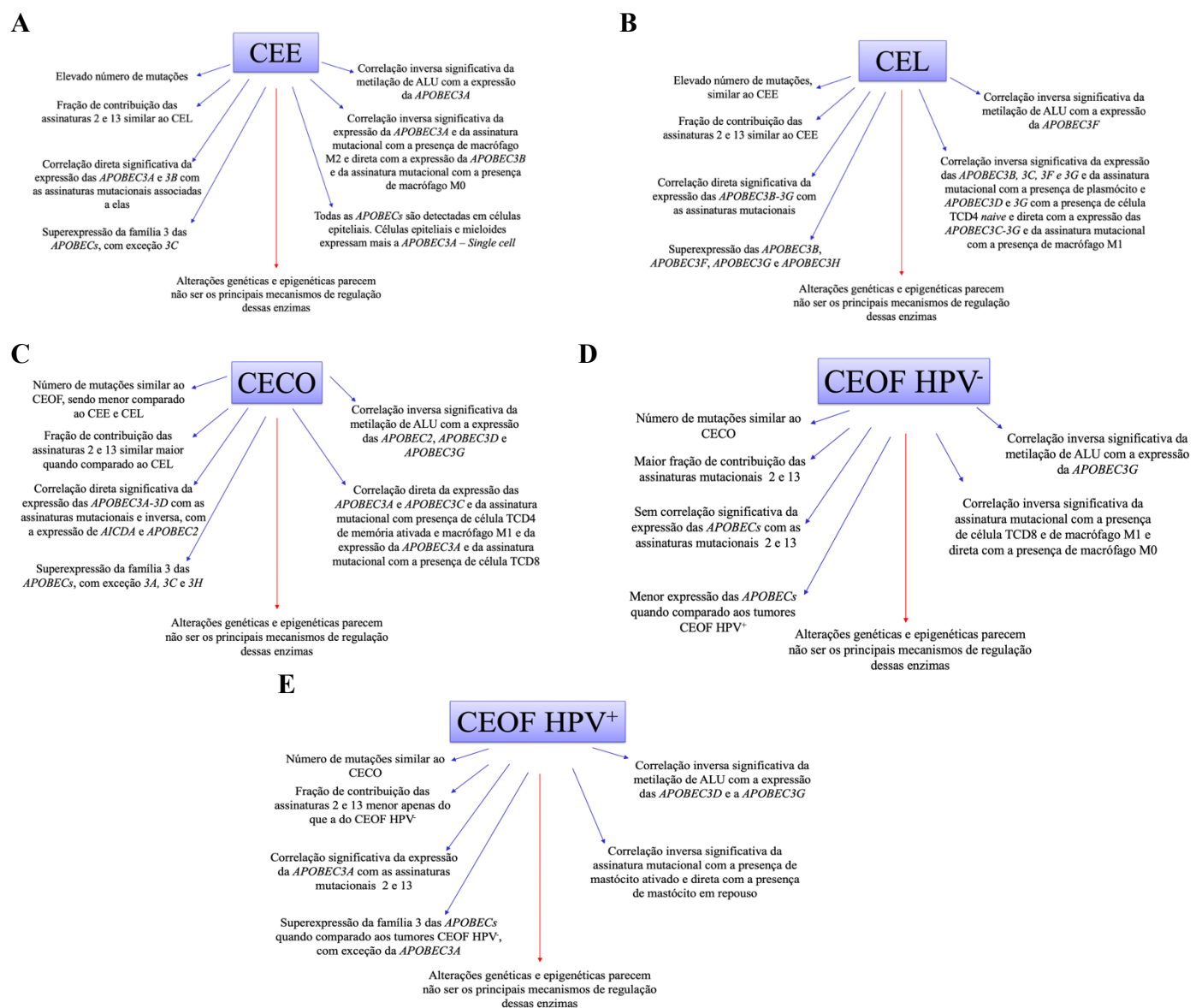


Figura 4.17: Resumo dos resultados obtidos nos tumores de esôfago e de cabeça e pescoço nas análises *in silico*. Esquema dos resultados obtidos em CEE (A), CEL (B), CECO (C), CEOF HPV⁻ (D) e CEOF HPV⁺ (E) estudados nas análises *in silico*.

4.2 CEE COMO MODELO EXPERIMENTAL

4.2.1 VALIDAÇÃO DA EXPRESSÃO DAS *APOBECs* EM PACIENTES COM CEE DO INCA

Foram utilizadas para validação da expressão das *APOBECs* 49 amostras pareadas de tecido tumoral e tecido adjacente não tumoral de pacientes com CEE do INCA, utilizando a técnica de RT-PCRq. A maioria dos pacientes incluídos nesta validação foi do sexo masculino (71%), ex-tabagistas e tabagistas (78%), ex-etilistas e etilistas (67%), com mediana de idade de 59 anos (39-78 anos). Os tumores acometeram principalmente dois ou mais terços do esôfago (59%), eram, em geral, moderadamente diferenciados (78%), e foram detectados em estágios mais avançados, III e IV (51%). A maioria dos pacientes foi a óbito (61%) no tempo de seguimento de 50 meses, no período de 2006 a 2015 (**Tabela 4.4**).

Tabela 4.4: Caracterização dos pacientes com CEE do INCA incluídos neste estudo para validação da expressão de APOBECs por RT-PCRq.

Critério Avaliado	CEE INCA Número de Pacientes (%)
Idade (mediana, min-max)	59 anos (39-78)
Gênero	
Masculino	35 (71%)
Feminino	14 (29%)
Tabagista	
Nunca	1 (2%)
Ex/Ativo	38 (78%)
Sem informação	10 (20%)
Etilismo	
Nunca	4 (8%)
Ex/Ativo	33 (67%)
Sem informação	12 (25%)
Óbito	
Não	19 (39%)
Sim	30 (61%)

	CEE INCA
Critério Avaliado	Número de Pacientes (%)
Localização do tumor	
Terço Médio	12 (25%)
Terço Superior	4 (8%)
Dois ou mais terços	29 (59%)
Sem informação	4 (8%)
Grau de diferenciação do tumor	
Pouco	7 (14%)
Moderadamente	38 (78%)
Bem	1 (2%)
Sem diferenciação	3 (6%)
Estadiamento do tumor	
I+II	9 (18%)
III+IV	25 (51%)
Não determinado	15 (31%)

CEE, carcinoma epidermoide de esôfago; INCA, Instituto Nacional de Câncer; Min, mínimo; Max, máximo.

A mucosa tumoral apresentou uma expressão aumentada das *APOBEC3A* (Figura 4.18A), *APOBEC3B* (Figura 4.18B), *APOBEC3D* (Figura 4.18C), *APOBEC3F* (Figura 4.18D), *APOBEC3G* (Figura 4.18E) e *APOBEC3H* (Figura 4.18F) em relação à mucosa adjacente não tumoral em pacientes com CEE (Figura 4.18).

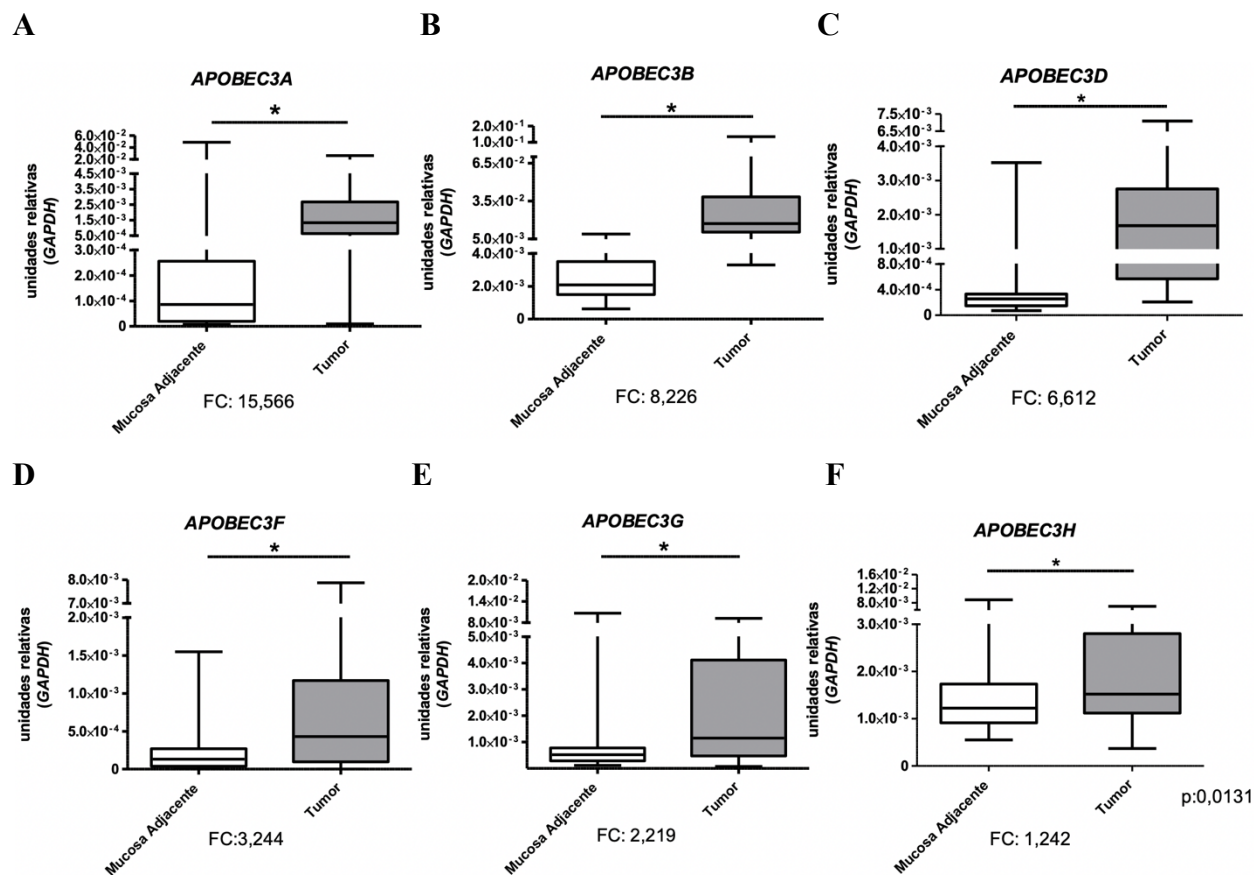


Figura 4.18: Expressão das *APOBEC3s* em pacientes com CEE do INCA. Boxplot mostrando a expressão do RNAm da *APOBEC3A* (A), *APOBEC3B* (B), *APOBEC3D* (C), *APOBEC3F* (D), *APOBEC3G* (E) e *APOBEC3H* (F) em mucosas tumoral e adjacente não tumoral de pacientes com CEE do INCA (49 amostras pareadas para *APOBEC3A*, *3B*, *3D* e *3F*, 47 amostras pareadas para *APOBEC3G* e 43 amostras pareadas para *APOBEC3H*). Teste *Wilcoxon matched-pairs signed rank*, $p < 0,0001$. FC – *Fold Change* de expressão (razão entre a mediana de expressão no tumor e a mediana de expressão na mucosa adjacente não tumoral).

Os pacientes com elevada expressão de *APOBEC3D* apresentaram uma pior sobrevida global quando comparados àqueles com menor expressão desse gene (HR: 4,67 - 95% IC: 1,47-14,88) (**Figura 4.19C**). Enquanto os pacientes com elevada expressão de *APOBEC3G* apresentaram um melhor prognóstico quando comparados àqueles com menor expressão desse gene (HR: 0,4 – 95% IC: 0,16-0,97) (**Figura 4.19E**). Para as demais *APOBECs*, não foram detectadas associações significativas com a sobrevida global dos pacientes (**Figura 4.19**).

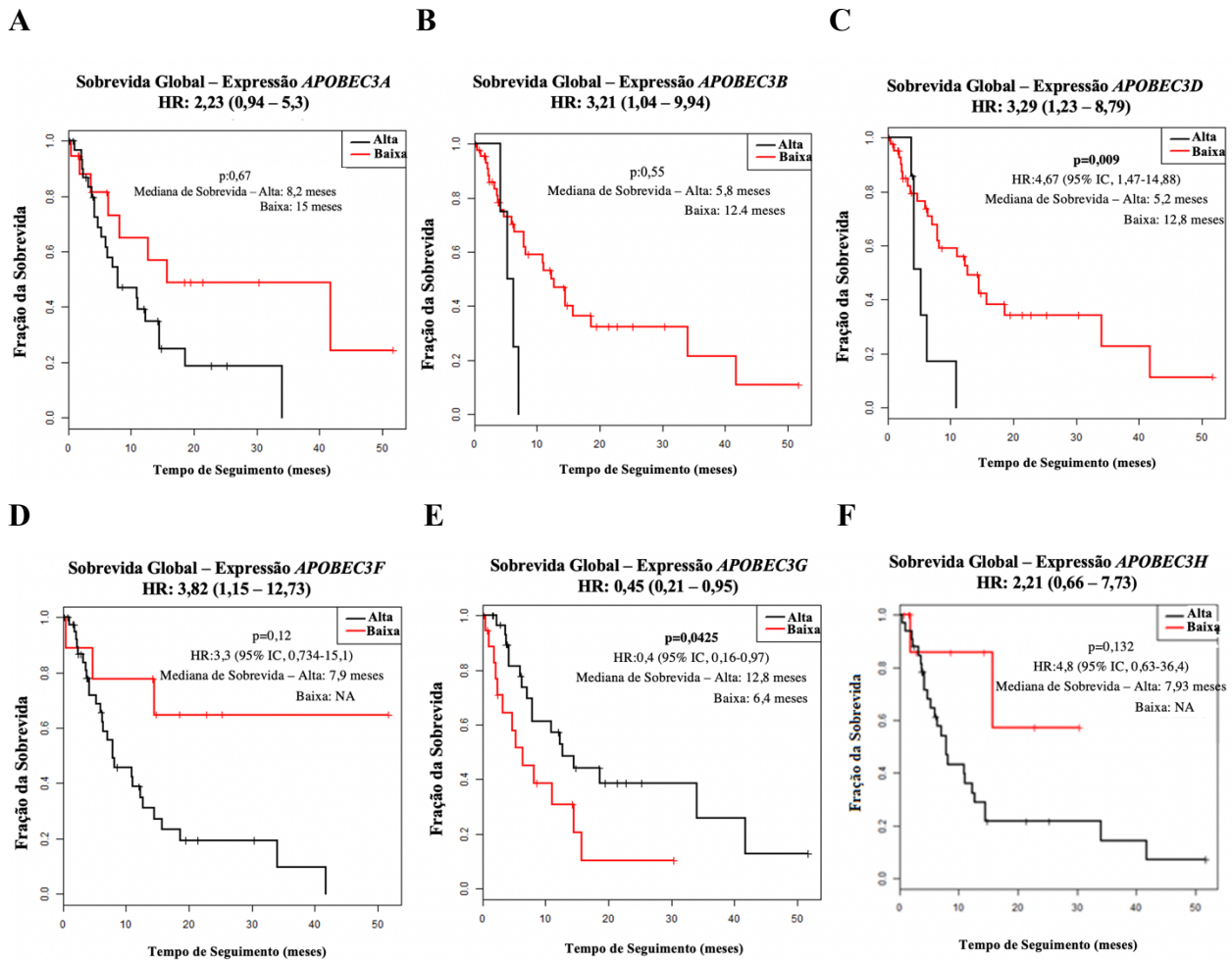


Figura 4.19: O impacto da expressão das *APOBEC3s* na sobrevida global de pacientes com carcinoma epidermoide de esôfago do INCA. Curva de Kaplan Meier mostrando o impacto da expressão do RNAm da *APOBEC3A* (A), *APOBEC3B* (B), *APOBEC3D* (C), *APOBEC3F* (D), *APOBEC3G* (E) e *APOBEC3H* (F) na sobrevida global de pacientes com CEE. A expressão das *APOBECs* foi dividida em octis e o octil superior definiu o grupo com alta expressão. Valores censurados (+) indicam o último tempo de seguimento conhecido dos indivíduos correspondentes. Os dados foram ajustados para o estadiamento do tumor e idade. p em negrito: estatisticamente significativo; IC: intervalo de confiança; HR: razão de chance, do inglês *hazard ratio*, NA, não aplicado.

Além disso, avaliamos a associação da expressão de *APOBEC3s* com os dados clínico-patológicos dos pacientes com CEE (**Tabela 4.5** e **Tabela 4.6**). As variáveis analisadas foram: sexo (masculino e feminino); tabagismo (não tabagista, ex-tabagista e tabagista ativo); etilismo (não etilista, ex-etilista e etilista ativo); localização do CEE (superior, médio e inferior); grau de diferenciação (bem, moderadamente e pouco diferenciado); e estadiamento (estádios I+II e III+IV). Já nas análises referentes à idade, variável contínua, foi utilizada a mediana de expressão desses genes ($>$ ou \leq que a mediana). Foi observada uma maior expressão das *APOBEC3A* (**Tabela 4.5**) e *APOBEC3G* (**Tabela 4.6**) no terço médio ($2,5 \times 10^{-3}$ e $2,1 \times 10^{-3}$, respectivamente) quando comparado ao terço inferior ($1,2 \times 10^{-4}$ e $6,1 \times 10^{-4}$, respectivamente). Algumas análises de associação com os dados clínico-patológicos não puderam ser realizadas devido ao reduzido tamanho amostral em algumas variáveis (não tabagistas e tumores bem diferenciados).

Tabela 4.5: Associação da expressão das *APOBEC3A*, *APOBEC3B* e *APOBEC3D* no tumor e os dados clínico-patológicos dos pacientes com CEE.

Dados Clínico-patológicos	<i>APOBEC3A</i>		<i>APOBEC3B</i>		<i>APOBEC3D</i>	
	Expressão Mediana (min-max)	Valor de p	Expressão Mediana (min-max)	Valor de p	Expressão Mediana (min-max)	Valor de p
Idade	59 anos (39-78)		59 anos (39-78)		59 anos (39-78)	
≤ Mediana	1,6x10 ⁻³ (9,9x10 ⁻⁶ -2,6x10 ⁻²)		2,1x10 ⁻² (5,8x10 ⁻³ -1,3x10 ⁻¹)		1,2x10 ⁻³ (2,8x10 ⁻⁴ -6,2x10 ⁻³)	
> Mediana	1,3x10 ⁻³ (1,7x10 ⁻⁴ -1,9x10 ⁻²)	0,98	1,5x10 ⁻² (3,2x10 ⁻³ – 6,5x10 ⁻²)	0,17	1,8x10 ⁻³ (2,1x10 ⁻⁴ -7,1x10 ⁻³)	0,9
Gênero						
Masculino	1,9x10 ⁻³ (9,9x10 ⁻⁶ -1,9x10 ⁻²)	0,39	2x10 ⁻² (3,2x10 ⁻³ – 1,3x10 ⁻¹)	0,21	1,8x10 ⁻³ (2,1x10 ⁻⁴ -7,1x10 ⁻³)	0,1
Feminino	7,4x10 ⁻⁴ (1,6x10 ⁻⁴ -2,6x10 ⁻²)		1,4x10 ⁻² (3,6x10 ⁻³ – 6x10 ⁻²)		8,5x10 ⁻⁴ (2,5x10 ⁻⁴ - 6,6x10 ⁻³)	
Tabagismo						
Não tabagista	NA	NA	NA	NA	NA	NA
Ex/Ativo	8,8x10 ⁻⁴ (9,9x10 ⁻⁶ – 1,7x10 ⁻²)		1,4x10 ⁻² (3,2x10 ⁻³ -1,3x10 ⁻¹)		8,9x10 ⁻⁴ (2x10 ⁻⁴ -7x10 ⁻³)	
Etilismo						
Não etilista	2,5x10 ⁻³ (7,5x10 ⁻⁴ - 8,4x10 ⁻³)	0,15	1,5x10 ⁻² (1x10 ⁻² -6,2x10 ⁻²)	0,66	6x10 ⁻⁴ (3x10 ⁻⁴ -6x10 ⁻³)	0,63
Ex/Ativo	9,4x10 ⁻⁴ (9,8x10 ⁻⁶ -1,7x10 ⁻²)		1,4x10 ⁻² (3,2x10 ⁻³ -1,3x10 ⁻¹)		1x10 ⁻³ (2x10 ⁻⁴ -7x10 ⁻³)	

Dados Clínico-patológicos	<i>APOBEC3A</i>		<i>APOBEC3B</i>		<i>APOBEC3D</i>	
	Expressão Mediana (min-max)	Valor de p	Expressão Mediana (min-max)	Valor de p	Expressão Mediana (min-max)	Valor de p
Localização do tumor						
Terço Superior	4,2x10 ⁻⁴ (2,7x10 ⁻⁴ - 1,3x10 ⁻³)		9x10 ⁻³ (3,4x10 ⁻³ -1,4x10 ⁻²)		7,9x10 ⁻⁴ (2,6x10 ⁻⁴ -1,8x10 ⁻³)	
Terço Médio	2,5x10⁻³ (3,3x10 ⁻⁴ – 1,9x10 ⁻²)	0,0093	2,5x10 ⁻² (6,4x10 ⁻³ -1x10 ⁻¹)	0,1	1,8x10 ⁻³ (5,4x10 ⁻⁴ -5,7x10 ⁻³)	0,25
Terço Inferior	1,2x10⁻⁴ (9,8x10 ⁻⁶ - 8,4x10 ⁻⁴)		1,7x10 ⁻² (1x10 ⁻² -2,8x10 ⁻²)		1,1x10 ⁻³ (3,7x10 ⁻⁴ -2,8x10 ⁻³)	
Grau de Diferenciação do tumor						
Bem	NA		NA		NA	
Moderadamente	1,6x10 ⁻³ (9,8x10 ⁻⁶ - 2,6x10 ⁻²)	0,98	1,7x10 ⁻² (3,2x10 ⁻³ -1,3x10 ⁻¹)	0,78	1,7x10 ⁻³ (2,1x10 ⁻⁴ -7,1x10 ⁻³)	0,9
Pouco	1,6x10 ⁻³ (2x10 ⁻⁵ – 8,4x10 ⁻³)		1,8x10 ⁻² (4,8x10 ⁻³ -4x10 ⁻²)		1,5x10 ⁻³ (5,1x10 ⁻⁴ -2,5x10 ⁻³)	
Estadiamento do tumor						
I+II	8,2x10 ⁻⁴ (9,8x10 ⁻⁶ - 7,4x10 ⁻³)		2,1x10 ⁻² (3x10 ⁻³ -1x10 ⁻¹)		2,5x10 ⁻³ (2,1x10 ⁻⁴ -6,6x10 ⁻³)	
III+IV	9,1x10 ⁻⁴ (2x10 ⁻⁵ - 1,7x10 ⁻²)	0,58	1,6x10 ⁻² (4,8x10 ⁻³ -1,3x10 ⁻¹)	1,0	1,1x10 ⁻³ (2,8x10 ⁻⁴ -7,1x10 ⁻³)	0,5

NA - não analisado, não foi possível realizar análise devido ao número reduzido de indivíduos em pelo menos um grupo; min – mínimo; max – máximo; em negrito: estatisticamente significativo.

Tabela 4.6: Associação da expressão das *APOBEC3F*, *APOBEC3G* e *APOBEC3H* no tumor e os dados clínico-patológicos dos pacientes com CEE.

Dados Clínico-patológicos	<i>APOBEC3F</i>		<i>APOBEC3G</i>		<i>APOBEC3H</i>	
	Expressão Mediana (min-max)	Valor de p	Expressão Mediana (min-max)	Valor de p	Expressão Mediana (min-max)	Valor de p
Idade	59 anos (39-78)		59 anos (39-78)		59 anos (39-78)	
≤ Mediana	3,5x10 ⁻⁴ (6,4x10 ⁻⁶ -7,8x10 ⁻³)	0,53	1,1x10 ⁻³ (7,4x10 ⁻⁵ -5,9x10 ⁻³)	0,45	1,4x10 ⁻³ (3,7x10 ⁻⁴ -7x10 ⁻³)	0,3
> Mediana	4,4x10 ⁻⁴ (9,9x10 ⁻⁷ -5,2x10 ⁻³)		1,4x10 ⁻³ (1,5x10 ⁻⁴ -9,2x10 ⁻³)		2x10 ⁻³ (7,5x10 ⁻⁴ - 5x10 ⁻³)	
Gênero						
Masculino	4,4x10 ⁻⁴ (6,4x10 ⁻⁶ -7,8x10 ⁻³)	0,3	1,4x10 ⁻³ (8,4x10 ⁻⁵ -9,2x10 ⁻³)	0,1	1,6x10 ⁻³ (3,7x10 ⁻⁴ -7x10 ⁻³)	0,3
Feminino	2,5x10 ⁻⁴ (9,9x10 ⁻⁷ -1,1x10 ⁻³)		9,2x10 ⁻⁴ (7,4x10 ⁻⁵ -8,3x10 ⁻³)		1,3x10 ⁻³ (7,5x10 ⁻⁴ - 4,8x10 ⁻³)	
Tabagismo						
Não tabagista	NA	NA	NA	NA	NA	NA
Ex/Ativo	4,2x10 ⁻⁴ (9,9x10 ⁻⁷ -7,8x10 ⁻³)		1,1x10 ⁻³ (7,4x10 ⁻⁵ -9,2x10 ⁻³)		1,4x10 ⁻³ (3,7x10 ⁻⁴ -7x10 ⁻³)	
Etilismo						
Não etilista	2,3x10 ⁻⁴ (7,8x10 ⁻⁵ -3,5x10 ⁻⁴)	0,19	7,9x10 ⁻⁴ (7,4x10 ⁻⁵ -8,3x10 ⁻³)	0,93	1,6x10 ⁻³ (1x10 ⁻³ -4,8x10 ⁻³)	0,55
Ex/Ativo	4,3x10 ⁻⁴ (9,9x10 ⁻⁷ -7,8x10 ⁻³)		1,1x10 ⁻³ (8,4x10 ⁻⁵ -9,2x10 ⁻³)		1,4x10 ⁻³ (3,7x10 ⁻⁴ -7x10 ⁻³)	

Dados Clínico-patológicos	<i>APOBEC3F</i>		<i>APOBEC3G</i>		<i>APOBEC3H</i>	
	Expressão Mediana (min-max)	Valor de p	Expressão Mediana (min-max)	Valor de p	Expressão Mediana (min-max)	Valor de p
Localização do tumor						
Terço Superior	4,5x10 ⁻⁴ (1,2x10 ⁻⁴ -2,5x10 ⁻³)		9,4x10 ⁻⁴ (9,2x10 ⁻⁴ -1,1x10 ⁻³)		2,4x10 ⁻³ (9,9x10 ⁻⁴ -2,4x10 ⁻³)	
Terço Médio	8,6x10 ⁻⁴ (3,9x10 ⁻⁵ -5,2x10 ⁻³)	0,58	2,1x10⁻³ (2,7x10 ⁻⁴ -9,2x10 ⁻³)	0,042	2x10 ⁻³ (7,5x10 ⁻⁴ - 5x10 ⁻³)	0,32
Terço Inferior	2,5x10 ⁻⁴ (6,5x10 ⁻⁶ -1,2x10 ⁻³)		6,1x10⁻⁴ (8,4x10 ⁻⁵ -1,1x10 ⁻³)		1,3x10 ⁻³ (6x10 ⁻⁴ - 1,4x10 ⁻³)	
Grau de Diferenciação do tumor						
Bem	NA		NA		NA	
Moderadamente	4,2x10 ⁻⁴ (6,5x10 ⁻⁶ -7,8x10 ⁻³)	0,19	1,1x10 ⁻³ (7,4x10 ⁻⁵ -9,2x10 ⁻³)	0,61	1,5x10 ⁻³ (6x10 ⁻⁴ -7x10 ⁻³)	0,2
Pouco	7,4x10 ⁻⁴ (9,9x10 ⁻⁷ -2,8x10 ⁻³)		1,4x10 ⁻³ (2,7x10 ⁻⁴ -4,2x10 ⁻³)		1,3x10 ⁻³ (3,7x10 ⁻⁴ - 2,4x10 ⁻³)	
Estadiamento do tumor						
I+II	1,2x10 ⁻⁴ (6,5x10 ⁻⁶ -2,8x10 ⁻³)		1,8x10 ⁻³ (4x10 ⁻⁴ -8,3x10 ⁻³)		1,8x10 ⁻³ (1x10 ⁻³ - 4,8x10 ⁻³)	
III+IV	4,3x10 ⁻⁴ (9,9x10 ⁻⁷ -7,8x10 ⁻³)	0,3	1,2x10 ⁻³ (7,4x10 ⁻⁵ -9,2x10 ⁻³)	0,28	1,5x10 ⁻³ (3,7x10 ⁻⁴ - 7x10 ⁻³)	0,7

NA - não analisado, não foi possível realizar análise devido ao número reduzido de indivíduos em pelo menos um grupo; min – mínimo; max – máximo; em negrito: estatisticamente significativo.

4.2.2 ENSAIOS *IN VITRO*

Os ensaios *in vitro* tiveram como objetivo avaliar se o mimetismo viral desencadeado pela desmetilação global do DNA e consequente ativação da resposta mediada por IFN desempenhariam um papel na indução da expressão de *APOBECs* em células de CEE. Nesta etapa do estudo, apenas as *APOBECs* cuja expressão foi correlacionada com a presença da assinatura mutacional (*APOBEC3A* e *APOBEC3B*) e aquelas cuja expressão foi associada à sobrevida global dos pacientes com CEE (*APOBEC3D*) foram avaliadas. Para isso, foram utilizadas três abordagens nas linhagens de CEE: tratamento com agente desmetilante a baixas doses, tratamento com RNA de dupla fita e tratamento com interferons.

4.2.2.1 TRATAMENTO COM AGENTE DESMETILANTE A BAIXAS DOSES

O tratamento com o agente desmetilante 5-AZA-CdR, também conhecida como decitabina (Sigma-Aldrich), a 300 nM por seis dias foi realizado em duas linhagens de CEE, TE1 e TE13, com o objetivo de mimetizar a reativação de elementos transponíveis observada em câncer, conforme já descrito na literatura (ROULOIS et al., 2015). Foram realizados três experimentos independentes e em triplicata. Após o tratamento, não foi observado impacto significativo na viabilidade celular na linhagem TE1 (**Figura 4.20A**) ou na linhagem TE13 (**Figura 4.20B**).

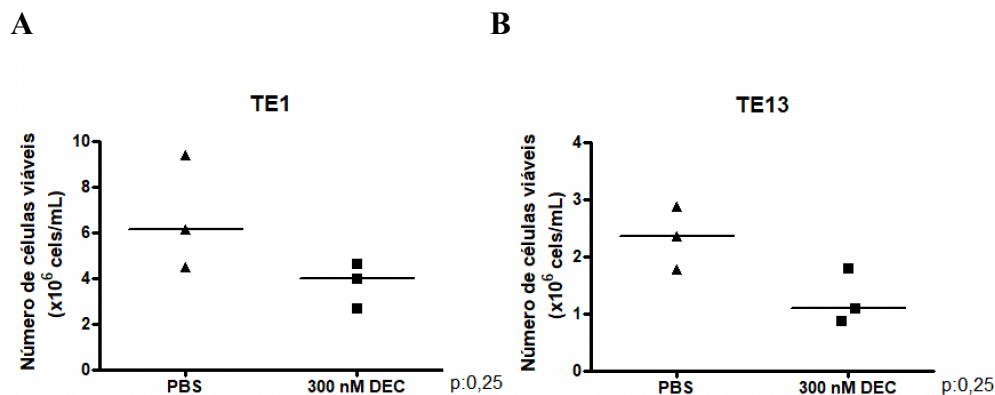


Figura 4.20: Viabilidade celular das linhagens de CEE após o tratamento com decitabina. Gráfico de *Scatter plot* mostrando a viabilidade celular das linhagens TE1 (A) e TE13 (B) após o tratamento com 300 nM de Decitabina por seis dias, avaliada pela técnica de exclusão por azul de tripan. Teste *Wilcoxon signed rank*, $p > 0,05$. PBS, controle do tratamento, DEC, decitabina.

Em seguida, foi avaliada a expressão gênica das *APOBEC3A*, *APOBEC3B* e *APOBEC3D*, bem como dos genes que são induzidos pelo tratamento com decitabina, de acordo com a literatura, *DDX58*, *MDA5* e *IRF7* (**Figura 4.21**). A expressão de nenhum dos genes avaliados diferiu significativamente entre células tratadas e células controle. Porém, o aumento de expressão de *APOBEC3A* foi superior àquele dos genes responsivos ao tratamento em TE-1 após o tratamento, sendo esse aumento significativo quando comparado ao do gene *DDX58* (**Figura 4.21A**). Em TE-13, a indução de *APOBEC3A* e *APOBEC3D* foi semelhante ou superior àquela dos genes responsivos (**Figura 4.21B**). Cabe ressaltar que na maioria das células tratadas com PBS (controle), a expressão de *APOBEC3A* não foi detectada e para viabilizar a comparação entre grupos, consideramos um Ct de 40 para essas amostras.

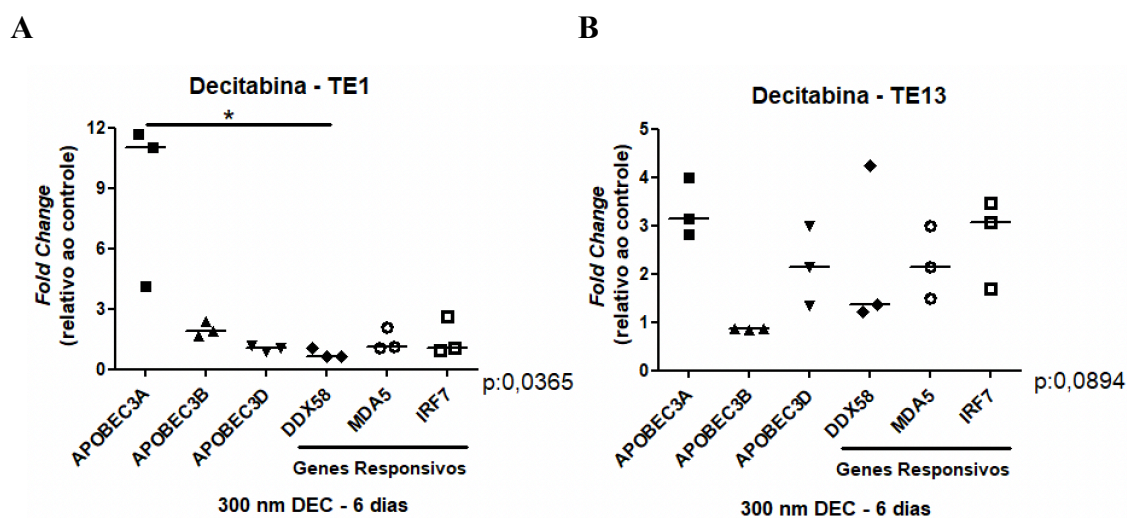


Figura 4.21: Análise da expressão gênica das *APOBECs* e dos genes responsivos após o tratamento com decitabina. Gráfico de *Scatter plot* mostrando a expressão das *APOBEC3A*, *APOBEC3B*, *APOBEC3D*, *DDX58*, *MDA5* e *IRF7* após o tratamento com 300 nM de decitabina por 6 dias nas linhagens TE1 (A) e TE13 (B), utilizando o *Fold Change* de expressão relativo ao controle. Foram feitos três experimentos independentes em triplicata para o controle e o tratado. Teste *Kruskal Wallis* e pós teste de *Dunn*, $p > 0,05$. DEC, decitabina; $*p < 0,05$.

4.2.2.2 TRATAMENTO COM RNA DE DUPLA FITA

O tratamento com RNA de dupla fita foi realizado com POLY I:C, um análogo sintético de RNA de dupla fita, um padrão molecular associado à infecção viral. Esse análogo ativa os receptores de reconhecimento de padrão antiviral, como por exemplo, DDX58-MDA5, levando à ativação de vias inflamatórias, de resposta a interferon, como por exemplo, IRFs (fatores regulatórios de interferon) (CHIAPPINELLI et al., 2015; ONOMOTO et al., 2021; Invivogen, disponível em <https://www.invivogen.com/polyic-hmw>, acesso em 22 de setembro de 2021).

O tratamento com 10 µg/mL de POLY I:C (Invivogen) por 24h, conforme já descrito na literatura (LIM et al., 2009; MATIJEVIĆ, T. & PAVELIĆ, J., 2012), não afetou significativamente a viabilidade das linhagens TE-1 (Figura 4.22A) e TE-13 (Figura 4.22B).

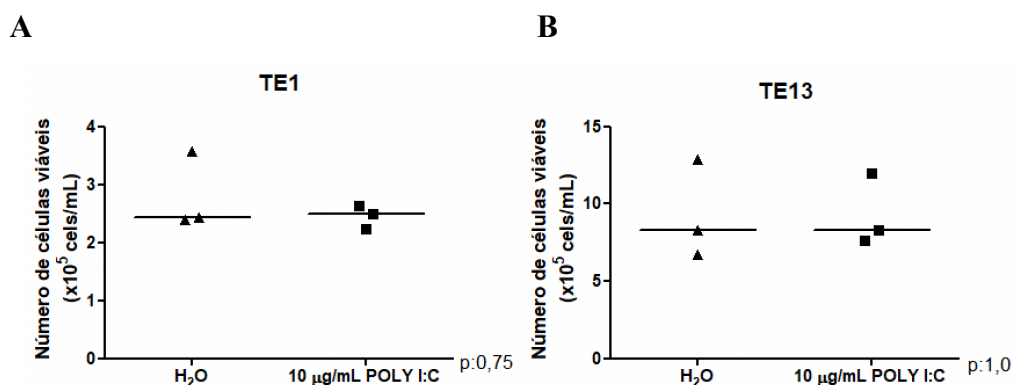


Figura 4.22: Viabilidade celular das linhagens de CEE após o tratamento com POLY I:C. Gráfico de *Scatter plot* mostrando a viabilidade celular das linhagens TE1 (A) e TE13 (B) após o tratamento com 10 µg/mL POLY I:C por 24h, pela técnica de exclusão por azul de tripan. Teste *Wilcoxon signed rank*, $p > 0,05$. H₂O: água ultrapura, controle do tratamento.

Em seguida, foi avaliada a expressão gênica das *APOBEC3A*, *APOBEC3B* e *APOBEC3D*, bem como dos genes responsivos ao tratamento, *DDX58*, *MDA5* e *IRF7* após o tratamento (Figura 4.23). Novamente, nenhuma diferença estatisticamente significativa entre células tratadas e células controle foi observada para nenhum dos genes analisados. Porém, em ambas as linhagens, *MDA5* parece ser o gene mais induzido pelo tratamento, conforme esperado por seu papel no reconhecimento do RNA de dupla-fita. Além disso, em TE1, a *APOBEC3D* apresentou uma média de indução de pelo menos cinco vezes em células tratadas em relação ao controle (Figura 4.23A),

enquanto em TE13, a *APOBEC3A* parece ser a APOBEC potencialmente induzida (**Figura 4.23B**). Novamente, as células controle em geral não apresentaram detecção do RNAm de *APOBEC3A*, tendo sido considerado o Ct de 40 nesses casos para viabilizar as comparações.

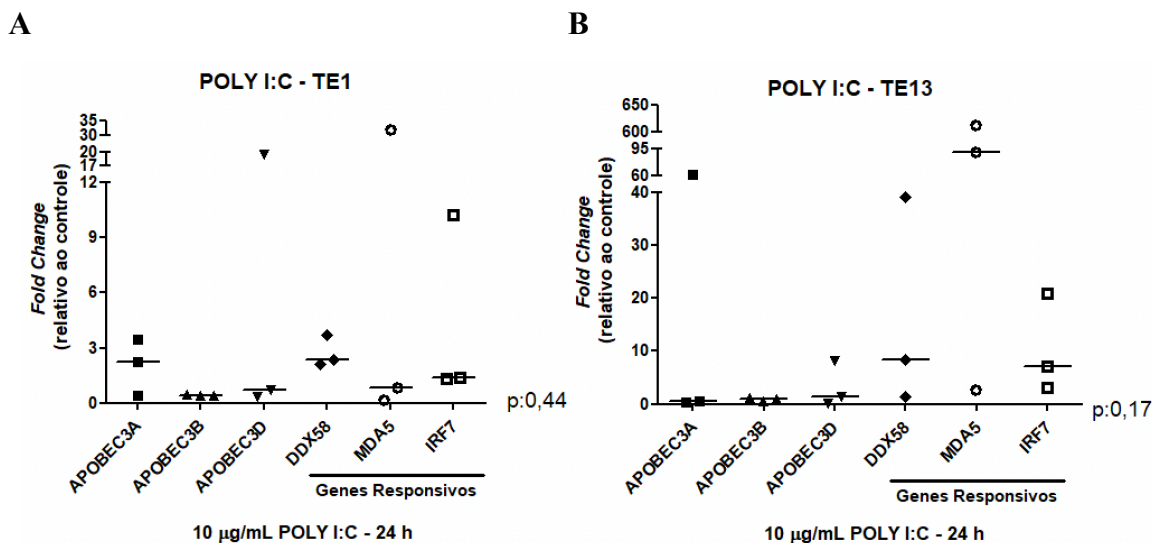


Figura 4.23: Análise da expressão gênica das *APOBECs* e dos genes responsivos após o tratamento com POLY I:C. Gráfico de *Scatter plot* mostrando a expressão das *APOBEC3A*, *APOBEC3B*, *APOBEC3D*, *DDX58*, *MDA5* e *IRF7* após o tratamento com 10 µg/mL POLY I:C por 24h nas linhagens TE1 (A) e TE13 (B), utilizando o *Fold Change* de expressão relativo ao controle. Foram feitos três experimentos independentes em triplicata para o controle e o tratado. Teste *Kruskal Wallis* e pós teste de *Dunn*, $p > 0,05$.

4.2.2.3 TRATAMENTO COM INTERFERONS

O tratamento com interferons foi realizado para mimetizar a resposta à infecção viral e por ser um dos mecanismos de regulação das APOBECs já descrito (BONVIN et al., 2006; CHEN et al., 2006; PENG et al., 2006; KONING et al., 2009; LUCIFORA et al., 2014; WANG et al., 2014). Para realizar este experimento, as células foram tratadas com 100 U/mL de IFN α (R&D Systems) ou 100 ng/mL de IFN β (Peprotech) ou 100 ng/mL de IFN γ (Peprotech) por 24h, conforme já descrito na literatura (GAN et al., 2016).

4.2.2.3.1 TRATAMENTO COM IFN α

Não foi observado impacto do tratamento com IFN α na viabilidade celular na TE1 (**Figura 4.24A**) ou na TE13 (**Figura 4.24B**).

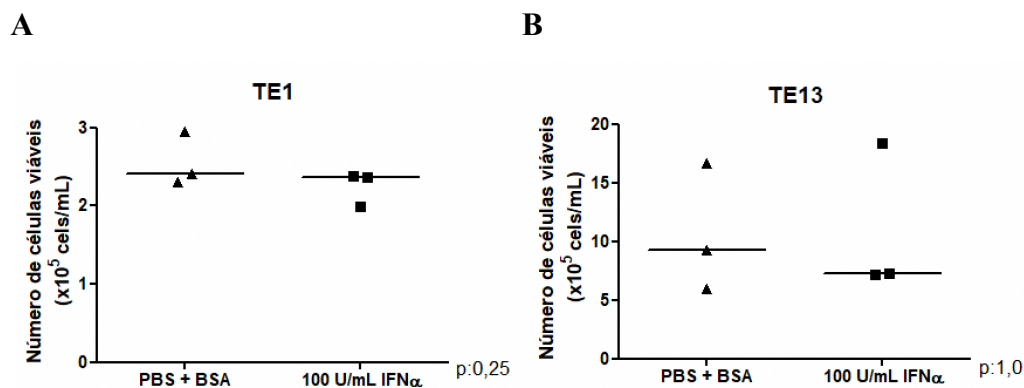


Figura 4.24: Viabilidade celular das linhagens de CEE após o tratamento com IFN α . Gráfico de *Scatter plot* mostrando a viabilidade celular das linhagens TE1 (A) e TE13 (B) após o tratamento com 100 U/mL de IFN α por 24h, pela técnica de exclusão por azul de tripan. Teste *Wilcoxon signed rank*, $p > 0,05$. PBS + BSA a 0,1%, controle do tratamento.

Em seguida, foi avaliada a expressão gênica das *APOBEC3A*, *APOBEC3B* e *APOBEC3D*, bem como do gene responsivo ao tratamento *IRF7* após o tratamento (**Figura 4.25**). Não foram detectadas diferenças significativas para nenhum dos genes avaliados em nenhuma das linhagens. Porém, a indução de *APOBEC3A* frente ao tratamento com IFN α parece ser superior àquela observada para o gene conhecidamente induzido ao tratamento, *IRF7*. Nesse sentido, é importante ressaltar que a expressão de *APOBEC3A* raramente foi detectada nas células controle, reforçando essa possível indução.

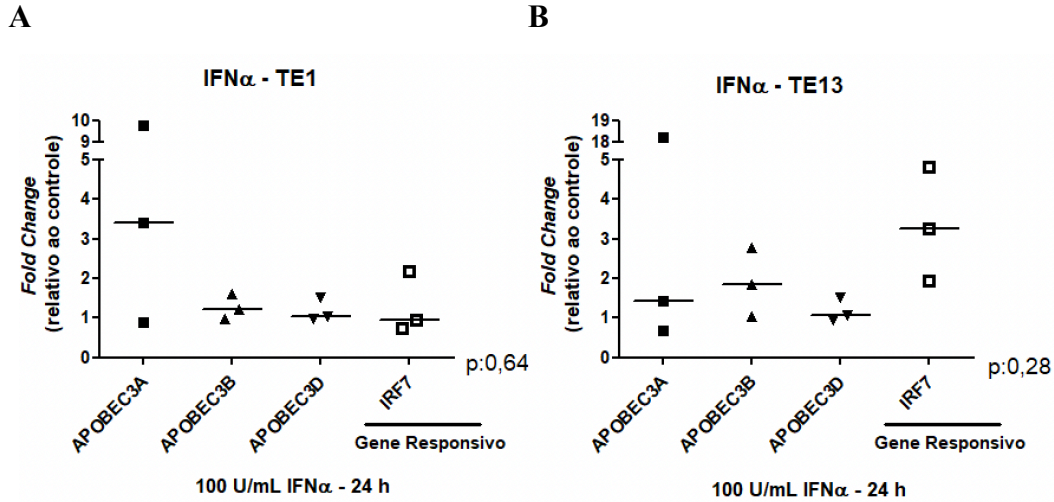


Figura 4.25: Análise da expressão gênica das *APOBECs* e do gene responsivo após o tratamento com IFN α . Gráfico de *Scatter plot* mostrando a expressão das *APOBEC3A*, *APOBEC3B*, *APOBEC3D* e *IRF7* após o tratamento com 100 U/mL IFN α por 24h nas linhagens TE1 (A) e TE13 (B), utilizando o *Fold Change* de expressão relativo ao controle. Foram feitos três experimentos independentes em triplicata para o controle e o tratado. Teste *Kruskal Wallis* e pós teste de *Dunn*, $p > 0,05$.

4.2.2.3.2 TRATAMENTO COM IFN β

O tratamento com IFN β não afetou significativamente a viabilidade de células TE1 (**Figura 4.26A**) ou TE13 (**Figura 4.26B**).

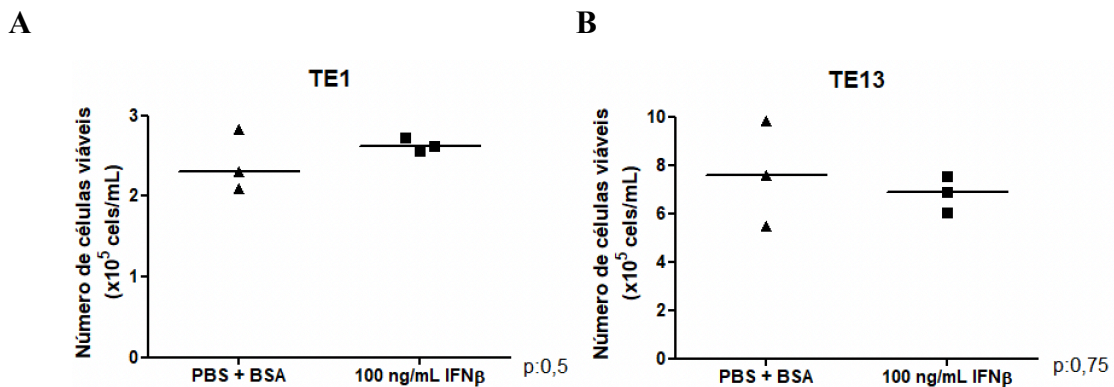


Figura 4.26: Viabilidade celular das linhagens de CEE após o tratamento com IFN β . Gráfico de *Scatter plot* mostrando a viabilidade celular das linhagens TE1 (A) e TE13 (B) após o tratamento

com 100 ng/mL de IFN β por 24h, pela técnica de exclusão por azul de tripan. Teste *Wilcoxon signed rank*, $p > 0,05$. PBS + BSA a 0,1%, controle do tratamento.

A análise de expressão gênica novamente não mostrou diferenças significativas entre células tratadas e células controle (**Figura 4.27**). Porém, em ambas as linhagens, a indução de *APOBEC3A* foi semelhante ou superior àquela observada para *IRF7*.

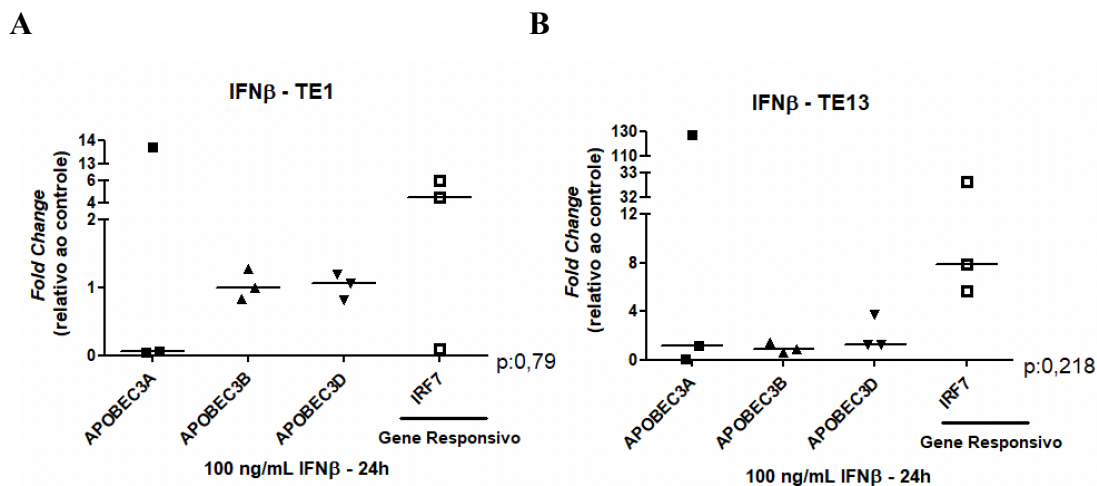


Figura 4.27: Análise da expressão gênica das *APOBECs* e do gene responsivo após o tratamento com IFN β . Gráfico de *Scatter plot* mostrando a expressão das *APOBEC3A*, *APOBEC3B*, *APOBEC3D* e *IRF7* após o tratamento com 100 ng/mL IFN β por 24h nas linhagens TE1 (A) e TE13 (B), utilizando o *Fold Change* de expressão relativo ao controle. Foram feitos três experimentos independentes em triplicata para o controle e o tratado. Teste *Kruskal Wallis* e pós teste de *Dunn*, $p > 0,05$.

4.2.2.3.3 TRATAMENTO COM IFN γ

A viabilidade das células TE1 (**Figura 4.28A**) ou TE13 (**Figura 4.28B**) não foi afetada significativamente pelo tratamento com IFN γ (**Figura 4.28**).

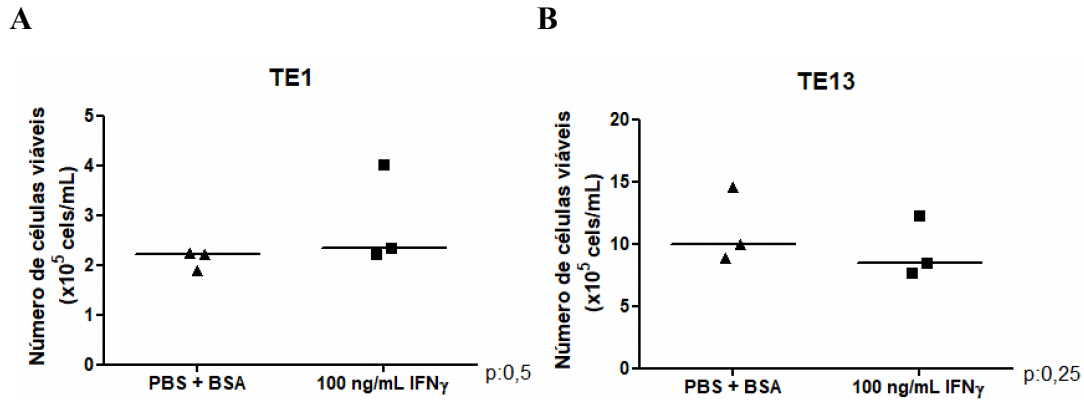


Figura 4.28: Viabilidade celular das linhagens de CEE após o tratamento com IFN γ . Gráfico de *Scatter plot* mostrando a viabilidade celular das linhagens TE1 (A) e TE13 (B) após o tratamento com 100 ng/mL de IFN γ por 24h, pela técnica de exclusão por azul de tripan. Teste *Wilcoxon signed rank*, $p > 0,05$. PBS + BSA a 0,1%, controle do tratamento.

A avaliação de expressão gênica não mostrou diferenças estatisticamente significativas entre células tratadas com IFN γ em comparação a células controle (**Figura 4.29**). Porém, novamente, a indução de *APOBEC3A* parece ser superior àquela de *IRF7* em TE1 (**Figura 4.29A**), enquanto em TE13 a indução de *APOBEC3A* e *3D* é semelhante à do gene responsivo ao tratamento (**Figura 4.29B**).

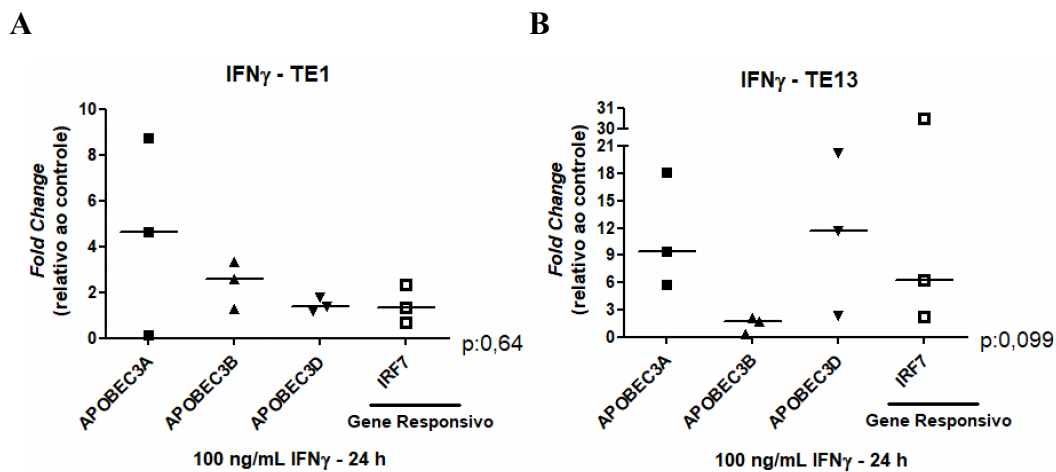


Figura 4.29: Análise da expressão gênica das *APOBECs* e do gene responsivo após o tratamento com IFN γ . Gráfico de *Scatter plot* mostrando a expressão das *APOBEC3A*, *APOBEC3B*, *APOBEC3D* e *IRF7* após o tratamento com 100 ng/mL IFN γ por 24h nas linhagens TE1 (A) e

TE13 (B), utilizando o *Fold Change* de expressão relativo ao controle. Foram feitos três experimentos independentes em triplicata para o controle e o tratado. Teste *Kruskal Wallis* e pós teste de *Dunn*, $p > 0,05$.

Em todos os experimentos *in vitro*, a detecção de *APOBEC3A* entre células controle foi rara, sugerindo a ausência ou baixos níveis de RNAm desta enzima nas linhagens avaliadas. Frente aos dados aqui apresentados, o número de replicatas experimentais será aumentado para tentar consolidar nossos achados.

5 DISCUSSÃO

Essa tese teve como principais objetivos dissecar a contribuição da assinatura mutacional induzida pela atividade de APOBECs em tumores de cabeça e pescoço e esôfago, assim como avaliar quais enzimas estão desreguladas nesses tumores, quais delas mais provavelmente estão associadas ao estabelecimento da assinatura e por quais mecanismos a expressão de *APOBECs* pode ser induzida.

Todos os tumores avaliados apresentaram as assinaturas 2/13, relacionadas à atividade das APOBECs, sendo que em alguns pacientes esta assinatura representou quase a totalidade (Figura 4.3). De forma surpreendente, CEOF HPV⁻ apresentou maior fração de contribuição da assinatura 2/13 em comparação a todos os outros subsítios tumorais. Na literatura, entre os tumores CECP, a maior atividade das APOBECs é geralmente reportada em CECP HPV⁺, uma vez que essas enzimas são de resposta antiviral (RIVA et al., 2021; CANNATARO et al., 2019; HENDERSON et al., 2014). Porém, cabe ressaltar que estes trabalhos não estratificam os subsítios da cabeça e pescoço, o que foi feito no presente estudo. Além disso, comparam a proporção de indivíduos em que a assinatura de APOBEC é detectada e não sua contribuição em cada caso, o que também foi feito nesta tese. Sendo assim, essa maior contribuição da assinatura de APOBEC em CEOF HPV⁻ pode ter sido mascarada em estudos anteriores. Além disso, CECP HPV⁻ também exibiu uma assinatura mutacional associada ao tabagismo (assinatura 5), porém, a assinatura mutacional associada ao tabaco (assinatura 4) não foi detectada. Cabe ressaltar que a assinatura 4 é associada principalmente a um dos carcinógenos do tabaco, o hidrocarboneto policíclico aromático benzopireno (PFEIFER et al 2002; ALEXANDROV et al 2013). Uma vez que o tabaco contém uma mistura de mais de 70 carcinógenos (HOFFMANN & HECHT, 1990; HOFFMANN et al., 2001; PFEIFER et al 2002), podemos supor que outros compostos poderiam atuar por outros mecanismos mutacionais ou até mesmo não genotóxicos. Essa suposição vai de encontro com o grande número de assinaturas mutacionais descritas em câncer cujas causas ainda não foram determinadas, como a própria assinatura 5, associada à idade, tabagismo e defeitos de reparo de DNA (ALEXANDROV et al., 2013; ALEXANDROV et al., 2016; SOUTH et al., 2019).

A principal assinatura mutacional encontrada em CEL foi a assinatura 4, relacionada à exposição ao benzopireno (tabagismo). Neste ponto, nossos dados se aproximam daqueles

reportados na literatura tanto em termos de relatos de perfis mutacionais quanto da associação do desenvolvimento de CEL com o hábito de fumar. Das mutações em *TP53*, por exemplo, 39% já foram relatadas como sendo transversões G>T (equivalentes às mutações de C>A da assinatura 4) em CEL, em comparação a 20% nos demais sítios de CECP (ZHANG et al., 1994). Além disso, um outro trabalho também utilizando dados do TCGA observou resultados similares aos nossos, um maior número de mutações correspondentes à assinatura 4 já foram descritas em CEL em comparação a CECO e CEOF (SOUTH et al., 2019). Nossos achados também reforçam dados epidemiológicos que mostram uma maior razão de risco associada ao tabaco para CEL (6,84) em comparação a CECO (1,35) e CEOF (2,02) (HASHIBE et al., 2007).

Apenas em CEE foi detectada a assinatura 16, sem etiologia conhecida, porém já associada à exposição ao álcool em CEE (LI et al., 2018; MOODY et al., 2021). O etilismo é um dos principais fatores de risco para o desenvolvimento deste tumor, apresentando uma OR de 4,4 em homens e 2,2 em mulheres. Um trabalho do grupo publicado recentemente também identificou a presença da assinatura 16 em CEE a partir de sequenciamento de genoma completo (MOODY et al., 2021). Porém, tal achado foi limitado a populações específicas (brasileira e japonesa). No caso da população brasileira, isto provavelmente está associado aos elevados níveis de consumo de bebidas alcóolicas (volume e frequência), enquanto na população japonesa está associado à presença de uma variante germinativa em *ALDH2* que leva a uma eliminação mais lenta de acetaldeído, o principal metabólito do etanol. Ou seja, o álcool parece ser capaz de contribuir para a carcinogênese esofágica por mecanismos genotóxicos, porém apenas em contextos específicos. No caso dos pacientes do consórcio TCGA, a nacionalidade foi variável (sendo em sua maioria vietnamita com 43,75% dos casos, seguido por 17,7% de brasileiros, 12,5% de russos, 6,25% de norte-americanos – EUA, 5,2% de ucranianos e 14,6% sem identificação). Como em países ocidentais o elevado consumo de bebidas alcóolicas é associado ao desenvolvimento de CEE, isto poderia explicar a detecção da assinatura 16 nos casos avaliados nesta tese.

Outro aspecto importante é a presença de assinaturas associadas à idade em todos os tumores. Essas assinaturas são observadas em diversos tipos de câncer, inclusive em tumores de cabeça pescoço e de esôfago, de acordo com a literatura (ALEXANDROV et al., 2013; ALEXANDROV et al., 2020). Porém, em CEOF, o tipo de assinatura relacionada à idade foi diferente entre tumores HPV⁻ e HPV⁺. Tumores HPV⁻ apresentaram a assinatura 5 enquanto HPV⁺

a assinatura 1. A assinatura 1 está associada à desaminação espontânea de citosinas em sítios CpG, enquanto a assinatura 5 apresenta como etiologia proposta o envelhecimento, a exposição ao tabaco e defeitos na maquinaria de reparo por excisão de nucleotídeo (NER, do inglês *nucleotide excision repair*) (COSMIC; KIM et al., 2020). De acordo com a literatura, a assinatura 1 é resultado de um processo mutacional endógeno em que cada divisão celular é uma possibilidade para essas mutações ocorrerem. Isso também pode explicar o porquê de muitos tumores com elevada contribuição da assinatura 1 serem derivados de epitélio normal com uma elevada taxa de renovação celular (ALEXANDROV et al., 2015; KIM et al., 2020), como acontece com os demais tumores desse estudo, CEE, CEL e CECO. Além disso, já foi descrita uma correlação desta assinatura com a expressão de genes do ciclo celular em câncer de mama (KIM et al., 2020). A assinatura 5, no entanto, não é correlacionada com o ciclo celular, mas sim com o processo de oxidação. O acúmulo de bases lesionadas por oxidação está relacionado à idade (HAMILTON et al., 2001), bem como ao tabagismo. Esses dados também corroboram a associação da assinatura 5 com defeito em reparo por excisão de nucleotídeos (NER), uma via de neutralização de danos oxidativos no DNA (MELIS et al., 2013). Com isso, essa assinatura associada à idade poderia estar relacionada à exposição contínua a um agente mutagênico ambiental/metabólico (KIM et al., 2020). Indivíduos com CEOF HPV⁻ são em geral tabagistas, o que poderia estar contribuindo para a presença da assinatura 5, assim como em CECO. Já CEOF HPV⁺ apresenta um processo de carcinogênese altamente associado a distúrbios do ciclo celular pelo efeito das oncoproteínas E6 e E7 sobre p53 e Rb, respectivamente (GRAHAM et al., 2017), sendo associado com a assinatura 1.

Não houve diferença de contribuição das assinaturas associadas à atividade das APOBECs de acordo com o estadiamento dos tumores e nem por status de HPV em CECO e CEL. A detecção da assinatura já em estágios precoces do desenvolvimento tumoral e a ausência de alteração de sua contribuição em estágios mais tardios sugerem que essas assinaturas são estabelecidas no início do processo de carcinogênese. Já foi demonstrado na literatura que alguns casos de CECF HPV⁻ com atividade precoce de APOBEC apresentaram um score de assinatura de IFN γ maior quando comparado com os casos com ativação tardia, mostrando que esses tumores apresentam uma maior inflamação (MESSERSCHMIDT et al., 2020). Sob inflamação crônica, espécies reativas de oxigênio são produzidas não apenas por células inflamatórias, mas também por células epiteliais, causando danos ao DNA, levando ao desenvolvimento de tumores (PINLAOR et al., 2004). Assim,

o processo inflamatório poderia induzir as APOBECs e a produção de espécies reativas de oxigênio, culminando no processo de carcinogênese. Já em esôfago, as assinaturas mediadas por APOBEC são pouco detectadas em clones derivados de epitélio esofágico normal (MARTINCORENA et al., 2018), porém estão presentes nos tumores (ALEXANDROV et al 2013; ALEXANDROV et al 2020), sugerindo que esses tumores são mais propensos a evoluir de clones com mutagênese de APOBEC ou que essa mutagênese é adquirida durante a evolução do tumor (MARTINCORENA et al., 2018). Em tumores de mama, por exemplo, não foi observada diferença significativa no número de mutações atribuído à assinatura de APOBEC entre tumores primários, tumores com recorrência local e em linfonodos auxiliares metastáticos. Porém, amostras metastáticas apresentaram um aumento da atividade de processos mutacionais mediados por APOBEC quando comparadas a tumores primários. Esse dado sugere que as assinaturas mutacionais mediadas por APOBEC são alterações que já aparecem nos tumores primários, porém a evolução clonal é acompanhada pelo aumento dessa atividade mutacional (ULLAH et al., 2018). Em tumores de bexiga, não há diferença significativa das assinaturas mediadas pelas APOBECs entre tumores de bexiga de alto grau não invasivo ao músculo e tumores de bexiga invasivos ao músculo, sendo predominantes nesses tumores (responsável por 53% de todas as mutações). Isto sugere que essa mutagênese endógena seja a principal causa de mutações no câncer de bexiga em todos os estágios (NASSAR et al., 2019). Já com relação ao status de HPV, a infecção por este vírus não é reconhecida como fator de risco para CECO e CEL. Ou seja, é mais provável que as infecções detectadas não participem do processo de carcinogênese e, portanto, não teriam impacto sobre o estabelecimento de assinaturas mutacionais.

Também não foi observado impacto na sobrevida global em cinco anos de pacientes com CEE, CEL, CECO, CEOF HPV⁻ e CEOF HPV⁺ de acordo com a fração de contribuição dessas assinaturas. Em outros tumores, como carcinoma urotelial, pacientes com elevada mutagênese mediada por APOBECs apresentam melhor prognóstico (75% em 5 anos), o que parece estar associado a uma elevada carga de neo-antígenos que poderão ser reconhecidos pelo sistema imune, por exemplo (ROBERTSON et al., 2017). Pacientes com câncer de pulmão com assinaturas mutacionais associadas à atividade das APOBECs também apresentam melhor prognóstico, sendo estas assinaturas também associadas à resposta à imunoterapia (CHEN et al., 2019). Os nossos dados sugerem que as assinaturas mutacionais sozinhas não impactam na sobrevida dos pacientes

com CEE e CECP. Porém, a combinação com outros fatores como a interação com o sistema imune, assinaturas relacionadas ao microambiente tumoral (assinaturas de interferon resultam em maior inflamação, por exemplo), e a carga mutacional possam ser capazes de impactar na sobrevida e na predição de resposta à inibidores de *checkpoint* imune, conforme mostrado por Cao e colaboradores para tumores de bexiga (CAO et al., 2021).

Uma vez tendo observado diferenças da contribuição da assinatura mutacional de APOBECs nos tumores avaliados, nos perguntamos se a APOBEC envolvida também poderia variar entre eles. Em CEE, nossos resultados mostraram que pacientes que expressam maiores níveis de *APOBEC3A* e *APOBEC3B* apresentam maiores contribuições da assinatura. Apesar das correlações identificadas serem fracas, estudos anteriores já sugeriram um envolvimento da APOBEC3B no desenvolvimento de CEE, devido à presença de transições C:G>T:A em sequências TpC e à superexpressão dessa enzima nesse tumor (LIN et al., 2014; MATSUMOTO et al., 2015). Em tumores de mama, foi demonstrada a superexpressão da *APOBEC3A*, bem como correlação direta significativa com a assinaturas mutacionais 2 e 13 (ρ : 0,29, $p < 0,0001$). Neste estudo também foi demonstrado que a APOBEC3A tem 100 vezes mais atividade de desaminação de citidina do que a APOBEC3B na presença de RNA celular, isso pode explicar o porquê dos níveis mais elevados de expressão de APOBEC3B contribuírem menos para a mutagênese em câncer de mama. Sendo assim, APOBEC3A parece ser a principal fonte de atividade da citidina desaminase em células de carcinoma mamário e um contribuinte importante para a assinatura mutacional mediada por APOBEC (CORTEZ et al., 2019). A expressão da APOBEC3A humana em tecido hepático e cólon murinos promove mutação e leva à carcinogênese *in vivo*. Animais que expressam APOBEC3A exibem assinaturas mutacionais mediadas por APOBEC marcantes (LAW et al., 2020). Em nosso estudo, identificamos a superexpressão de *APOBEC3B*, mas também de *APOBEC3A*, *3D*, *3F*, *3G* e *3H* em CEE quando comparado à mucosa adjacente não tumoral. A expressão das APOBEC3s também foi diretamente correlacionada à contribuição da assinatura mutacional, com algumas especificidades, em CECO (*APOBEC3A*, *APOBEC3B*, *APOBEC3C* e *APOBEC3D*), em CEL (*APOBEC3B*, *APOBEC3C*, *APOBEC3D*, *APOBEC3F* e *APOBEC3G*) e em CEOF HPV⁺ (*APOBEC3A*). Já foi descrito na literatura que as mutações mediadas por APOBECs são bastante prevalentes em CECP (58% dos casos), tendo correlação (ρ :0,25) com a expressão de *APOBEC3A* quando estratificado por status de HPV (FADEN et al., 2017).

Também já foi encontrada correlação significativa da expressão de *APOBEC3B* e a assinatura mutacional em CECP ($\rho:0,144$), porém tumores HPV⁺ apresentam maior expressão desse gene e maior frequência de mutação (HENDERSON et al., 2014). Quase todas as *APOBEC3s* foram encontradas superexpressas nos tumores avaliados quando comparados à mucosa adjacente não tumoral. Mas temos que levar em consideração que as impressões finais das mutações relacionadas com as APOBECs no genoma dependem de outros fatores além do nível de expressão dessas enzimas, além de não sabermos se existe um nível mínimo de expressão para que haja a geração da mutação. Entre estes outros fatores, podemos citar a quantidade de DNA de fita simples acessível (o substrato da enzima) e o mecanismo de correção dessa lesão (enzimas de reparo funcionais ou não). Todos eles contribuem para uma maior ou menor geração de assinaturas mutacionais associadas à atividade das APOBECs (KIM et al., 2020). Assim, nossos dados apontam para diferenças importantes da contribuição da assinatura mutacional de APOBECs entre os sítios de CECP e CEE, além de sugerir a participação de uma ou mais enzimas no processo, principalmente APOBEC3A e APOBEC3B. Porém, como correlações não inferem relações causais, mais estudos são necessários para se tentar determinar as enzimas e os demais mecanismos envolvidos. Além disso, a análise descrita aqui se baseou apenas em um momento específico do processo de carcinogênese, em geral em estágios mais avançados do desenvolvimento tumoral. Nesse sentido, estudos que possibilitem um acompanhamento de todo o processo, como em modelos animais, podem ser bastante benéficos para a compreensão do estabelecimento da assinatura mutacional de APOBECs.

De uma maneira em geral, a expressão das *APOBECs* está desregulada nesses tumores e está correlacionada com as assinaturas mutacionais 2/13. Assim, nos perguntamos quais mecanismos poderiam estar envolvidos na desregulação dessas enzimas. Primeiramente, avaliamos alterações genéticas, porém a frequência de mutações e alterações do número de cópias dos genes que codificam AID/APOBECs é baixa nesses tumores, sendo de no máximo 5% em *AICDA* e para as *APOBECs* chega a no máximo 3% em CEL, por exemplo. Uma outra possibilidade seria avaliar a presença de fusão gênica, como já observada em tumores cervicais. Um polimorfismo de deleção germinativa comum no cluster de *APOBEC3s* resulta em um gene híbrido *APOBEC3A/APOBEC3B*. Esse híbrido pode transcrever uma fusão gênica entre a região codificante de *APOBEC3A* com a região 3'UTR da *APOBEC3B* resultando na perda completa da

sequência codificante da *APOBEC3B*. Essa fusão codifica proteína com sequência de aminoácidos idêntica da *APOBEC3A* e é mais estável, resultando em níveis mais elevados de *APOBEC3A* intracelular (CAVAL et al., 2014; NIK-ZAINAL et al., 2014; REVATHIDEVI et al., 2021). Além disso, já foi mostrado que p53 reprime a expressão de *APOBEC3B* via p21 (PERIYASAMY et al., 2017), o que poderia justificar a maior expressão desse gene. Porém, como os tumores estudados aqui apresentam uma elevada frequência de perda de p53, seja por mutações ou via a oncoproteína viral E6 do HPV, não foi possível avaliar esta associação.

Além de mecanismos genéticos, as alterações epigenéticas também representam um importante mecanismo de regulação da expressão gênica em câncer (DAWSON & KOUZARIDES, 2012). Apesar de poucos estudos até então terem avaliado o papel da metilação do DNA e de miRNAs na regulação da expressão de *APOBECs*, a presença de ilhas CpG em suas regiões promotoras e vias regulatórias sobrepostas nos levaram a avaliar a associação desses mecanismos com a desregulação dessas enzimas nos tumores estudados. Assim como já foi descrito para AID/*APOBECs*, IFN também é capaz de induzir a expressão de miRNAs (SEDDER, 2013). Mir-29a, miR-155 e miR-548, por exemplo, são induzidos por IFN (SEDDER, 2013), o que poderia explicar a correlação direta entre a expressão desses miRNAs e a expressão de *AICDA* observada em nosso estudo. Por outro lado, uma regulação pós-transcricional de *APOBEC3A* e *APOBEC3G* por miR-2909 já foi descrita (KAUL et al., 2015; CAO & WU, 2018). Curiosamente, esse miRNA não é expresso nos tumores avaliados no banco de dados do TCGA. Já em relação à metilação do DNA, os dados são ainda mais escassos na literatura. Alterações pontuais em *APOBECs* específicas e tumores específicos foram identificadas em nosso estudo, como para *APOBEC3D* em CEE. Nesse caso, para a sonda cg16066354 (localizada na região promotora do gene), uma menor metilação foi observada nos tumores em comparação à mucosa adjacente, além de uma correlação inversa entre metilação e expressão, o que poderia explicar, ao menos em parte, a superexpressão de *APOBEC3D* em CEE. Entretanto, já que essa *APOBEC* não foi associada à presença da assinatura mutacional neste tipo tumoral, não prosseguimos com essa avaliação. Porém, não podemos descartar que a metilação do DNA possa participar da regulação da expressão dessas enzimas em situações específicas.

Nossos dados sugerem que alterações genéticas e epigenéticas não são os principais mecanismos de regulação das *APOBECs* nos tumores avaliados. Uma vez que o mecanismo mais

bem descrito na literatura de regulação dessas enzimas é a infecção viral (BONVIN et al., 2006; ENDO et al., 2007; JIANG et al., 2007; WATASHI et al., 2013; LUCIFORA et al., 2014; WANG et al., 2014) e como CEECP (com exceção de CEOF HPV⁺) e CEE não são reconhecidos como tumores causados por vírus (PARKIN, 2006; HERBSTER et al., 2012; ABNET et al., 2018; LEEMANS et al., 2018; JOHNSON et al., 2020), nos perguntamos se mecanismos que mimetizam uma infecção viral poderiam estar envolvidos na regulação dessas enzimas. Já foi observado na literatura que tumores de uma forma geral apresentam um perfil de hipometilação global do DNA. Isto pode levar à ativação de elementos transponíveis do genoma, cuja expressão pode desencadear uma resposta imune antiviral mediada por interferon em células tumorais (CHIAPPINELLI et al., 2015; ROULOIS et al., 2015). Assim, avaliamos a correlação entre os níveis de metilação desses elementos e a expressão de *APOBECs*. Foram encontradas diversas correlações significativas entre os níveis de metilação de ALU, LINE e LTRs e a expressão das *APOBEC3s* nos tumores avaliados, sendo que podemos destacar o elevado número de correlações inversas significativas da metilação de elementos ALU com a expressão de *APOBEC3A* em CEE. Na literatura, a hipometilação global após tratamento com agente desmetilante já foi associada à indução de outra enzima de restrição viral, ADAR1 (desaminase de adenosina específica de RNA, do inglês *Adenosine Deaminase RNA Specific*). Esta indução foi mediada pela reativação de elementos ALU e ativação do receptor MDA5. Assim, a terapia epigenética desencadeia o mimetismo viral induzido por ALU, porém, a ativação de ADAR1 minimiza estes efeitos em tumores colorretais (MEHDIPOUR et al., 2020, com correção em MEHDIPOUR et al., 2021). No nosso estudo, uma hipometilação global desencadeada por ALU, gerando um mimetismo viral, poderia levar a um aumento de expressão de *APOBEC3A*.

Um outro aspecto que pode ser relevante na regulação de *APOBECs* e, conseqüentemente, no estabelecimento de assinaturas mutacionais relacionadas à atividade dessas enzimas é a composição do microambiente tumoral. Conforme mencionado anteriormente, a expressão de *APOBECs* pode ser regulada por uma sinalização mediada por citocinas (BONVIN et al., 2006; CHEN et al., 2006; PENG et al., 2006; KONING et al., 2009; LUCIFORA et al., 2014; WANG et al., 2014), produzidas principalmente por células do sistema imune. Baseado nisso, realizamos uma análise de composição do microambiente tumoral para cada paciente do consórcio TCGA analisado neste estudo e avaliamos sua associação com a contribuição das assinaturas 2/13 e com

a expressão das *APOBECs* correlacionadas com essa assinatura em cada tipo tumoral. Em geral, nossos dados mostraram correlações da assinatura e da expressão de *APOBECs* com a presença de macrófagos, exceto em CEOF HPV⁺. Além disso, em CECO, correlações diretas foram observadas entre a presença de células TCD4 de memória ativadas e células TCD8 e a contribuição das assinaturas 2/13, assim como mostrado anteriormente em tumores de pulmão (CHEN et al., 2019). Em CECP, já foi mostrada a associação da *APOBEC3H* e o infiltrado de células TCD8⁺ (LIU et al., 2020). Além disso, em CECP, tumores HPV⁻ com enriquecimento de assinatura de APOBEC apresentam aumento de células supressoras mieloides, células TCD8⁺ de memória efetoras e células T *helper* do tipo 17 (MESSERSCHMIDT et al., 2020). Assim, a hiperativação de APOBECs e a consequente geração de neo-antígenos poderia estar levando à ativação de uma resposta imune antitumoral, que poderia ser potencializada pelo tratamento com imunoterapia em pacientes com CECO. Por outro lado, as análises de RNAseq a partir de biópsias tumorais não nos permitem determinar quais tipos celulares de fato expressam as *APOBECs*, ou seja, as correlações observadas podem ser um reflexo tanto da sinalização entre células do sistema imune e células tumorais, quanto da expressão de *APOBECs* pelas próprias células do infiltrado inflamatório. Por isso, nosso próximo passo foi avaliar a expressão gênica dessas enzimas em células isoladas a partir de RNAseq de experimentos de *single cell*. Porém, apenas dados de CEE estavam disponíveis nos bancos de dados avaliados.

As análises de *single cell* de CEE mostraram que todas as *APOBECs* são expressas em células epiteliais. A *APOBEC3A*, correlacionada com a presença das assinaturas mutacionais 2/13 neste tipo tumoral, apresentou maior detecção em células epiteliais (com 7,0% de células com detecção de expressão) e mieloides (com 15,8% de células com detecção de expressão). Uma vez que macrófagos são células de origem mielóide, esta observação pode explicar, ao menos em parte, a observação anterior. Já a *APOBEC3B*, também correlacionada à presença da assinatura de APOBECs em CEE, foi detectada principalmente em células epiteliais (11,4%). Dados esses semelhantes ao que foi observado em CECP HPV⁻ por outros autores, com maior detecção de *APOBEC3B* e *APOBEC3C* em células malignas (MESSERSCHMIDT et al., 2020). Por outro lado, as *APOBECs* (*APOBEC3C*, *APOBEC3D*, *APOBEC3F* e *APOBEC3G*) cuja expressão não foi correlacionada à presença da assinatura mutacional foram detectadas principalmente em células T e células B, o que pode indicar que a maior expressão desses genes observada em CEE em relação

à mucosa adjacente pode ser uma consequência da presença de um infiltrado inflamatório composto por esses tipos celulares. De fato, um trabalho anterior do nosso grupo mostrou a presença de muitos clones de células B infiltrando o tumor e a abundância dessas células em estruturas linfoides terciárias em CEE (BARROS et al., 2020). Sendo assim, a caracterização do microambiente tumoral, dissecando a interação com as células do sistema imune, é essencial para a melhor compreensão da carcinogênese esofágica e o papel funcional das APOBECs nesse tumor.

A partir das análises *in silico*, reunimos informações que sugerem uma participação de APOBEC3A e APOBEC3B, expressas pelas células tumorais, no estabelecimento de assinaturas mutacionais em CEE. Além disso, a indução de *APOBEC3A* nesses tumores pode estar relacionada à ativação de elementos ALU, mediada pela hipometilação desses elementos. Devido à consistência desses dados e à disponibilidade de linhagens celulares de CEE em nosso laboratório, este tumor foi escolhido como modelo. Primeiramente, validamos a expressão gênica das *APOBECs* em pacientes com CEE do INCA. As *APOBEC3s* estão mais expressas no tumor quando comparado à mucosa adjacente não tumoral, validando os achados do RNAseq de CEE (Figura 4.9A) e corroborando a literatura (LIN et al., 2014; MATSUMOTO et al., 2015). Pacientes com CEE com alta expressão da *APOBEC3D* apresentaram um pior prognóstico quando comparados aos pacientes com menor expressão desse gene. Porém, pacientes com alta expressão da *APOBEC3G* apresentaram uma melhor sobrevida global quando comparados aos pacientes com menor expressão desse gene. Esses dados sugerem que a *APOBEC3D* poderia desempenhar um papel na progressão do tumor, enquanto a *APOBEC3G* exerceria um papel protetor. Ambas são enzimas de restrição viral e de edição de DNA (OLSON et al., 2018). Já foi demonstrada na literatura uma correlação positiva entre a expressão da *APOBEC3G* e genes de células T, inclusive há uma colocalização da APOBEC3G e de linfócitos T infiltrados em carcinoma de ovário seroso de alto grau (LEONARD et al., 2016). Além disso, a elevada expressão desse gene já foi associada a um melhor prognóstico neste tumor, indicando ser um novo biomarcador para linfócitos T infiltrados na massa tumoral e de prognóstico favorável para o carcinoma de ovário (LEONARD et al., 2016). Em nossas análises de CEE, foi observada também uma correlação direta significativa entre a expressão de *APOBEC3G* e a presença de células TCD8 ($\rho:0,27$, $p:0,01$ dados não mostrados) e células T CD4 de memória em repouso ($\rho:0,27$, $p:0,009$, dados não mostrados).

Com relação às associações da expressão das *APOBECs* com os dados clínicos patológicos de pacientes com CEE, as *APOBEC3A* e *APOBEC3G* estão mais expressas no terço médio quando comparado ao terço inferior. Isso é bastante interessante porque o ADE acomete principalmente o terço inferior do esôfago, enquanto o CEE ocorre mais frequentemente em células escamosas que revestem os terços superior e médio (ENZINGER & MAYER, 2003).

Mediante a validação da expressão das *APOBECs* nos pacientes com CEE do INCA, foram realizados ensaios *in vitro* em linhagens de CEE para investigar os mecanismos de regulação dessas enzimas. Foram utilizadas três abordagens, com foco em mecanismos que poderiam mimetizar a infecção viral: tratamento com agente desmetilante a baixas doses, com RNA de dupla fita e com interferons ($IFN\alpha$, $IFN\beta$ e $IFN\gamma$).

O tratamento com agente desmetilante a baixas doses não impactou significativamente a viabilidade celular das duas linhagens estudadas, apesar de ser uma droga tóxica. Na literatura, o tratamento com baixas doses de decitabina (100nM em 24h e 48h) em linhagens de leucemia mieloide não impactou na proliferação dessas células. Em doses mais altas, esse tratamento apresentou um efeito citotóxico, reduzindo a proliferação celular de maneira dose-dependente (1 μ M em 48h, 10 μ M e 25 μ M em 24h e 48h) (ALVES et al., 2019). Em leucemia mieloide aguda, baixas doses de decitabina (100 nM e 400 nM) em 24h aumentaram significativamente a proporção de células apoptóticas quando comparadas às células não tratadas (MPAKOU et al., 2021). Assim, os dados sugerem que a toxicidade pode variar dependendo da dose e do tipo celular estudado. Apesar da falta de significância estatística, os controles positivos, assim como *APOBEC3A* e *APOBEC3D*, parecem ter sido induzidos pelo tratamento, principalmente na linhagem TE13. Já na TE1, a indução da *APOBEC3A* foi significativamente maior do que a indução do controle positivo *DDX58*. Ou seja, essa hipometilação global gerada pelo agente desmetilante poderia estar levando à ativação de elementos transponíveis, mimetizando uma infecção viral e levando à ativação dos receptores *DDX58* e *MDA5*, desencadeando assim uma resposta inflamatória via interferon (possível indução de *IRF7*) (CHIAPPINELLI et al., 2015). Como *APOBECs* também são considerados genes responsivos a IFN (SEDGGER et al., 2013), este mecanismo também poderia estar levando à indução das *APOBECs*, mais especificamente, *APOBEC3A*. Na literatura, já foi demonstrado que a *APOBEC3A* é induzida por $IFN\alpha$ em diversas células hematopoiéticas, como macrófagos e linfócitos T (PENG et al., 2006; KONING et al., 2009) e em hepatócitos

humanos primários ou hepatócitos infectados por HBV, por exemplo (BONVIN et al., 2006; LUCIFORA et al., 2014).

O próximo passo foi o tratamento com POLY I:C (Invivogen), um análogo sintético de RNA de dupla fita, padrão molecular associado com infecção viral (ZHAO et al., 2012). Esse tratamento não impactou na viabilidade celular em nenhuma das linhagens estudadas. Com relação à expressão gênica, houve indução não significativa de expressão dos controles positivos, mais especificamente *MDA5*, o sensor viral, em ambas as linhagens. Conforme mostrado na literatura, tanto *MDA5* quanto *DDX58* são induzidos por esse tratamento em células de câncer colorretal e mama, bem como *IRF3/7*, levando à produção de interferon do tipo 1, incluindo $IFN\alpha$ e $IFN\beta$ (ZHAO et al., 2012). Já em relação às *APOBECs*, na TE13, observamos um aumento de expressão de *APOBEC3A* e, na TE1, de *APOBEC3D*, porém em níveis menores do que foi observado para os controles positivos.

Finalmente, realizamos o tratamento com interferons, do tipo 1 ($IFN\alpha$ e $IFN\beta$) e do tipo 2 ($IFN\gamma$), mimetizando a etapa final da infecção viral, produção de citocinas. Esses tratamentos não impactaram na viabilidade celular das linhagens. Com relação à expressão gênica, maiores níveis de RNAm de *IRF7*, nosso controle positivo, foram observados após o tratamento. Para os interferons do tipo 1, parece haver indução basicamente da *APOBEC3A* em ambas as linhagens, em níveis superiores ao que foi observado para *IRF7*. Já após o tratamento com $IFN\gamma$, o aumento de expressão de *APOBEC3A* na linhagem TE1 e de *APOBEC3A* e *3D* na linhagem TE13 foi inferior ao aumento de *IRF7*. Já foi demonstrado que as *APOBEC3s* são induzidas por IFN de resposta inata ($IFN\alpha$) e de resposta imune adaptativa ($IFN\gamma$) (SEDGER, 2013). De acordo com a literatura, os interferons são produzidos pela maioria das células, com exceção do interferon do tipo II ($IFN\gamma$), produzidos pelas células do sistema imune, mais especificamente, as células hematopoiéticas, tendo como célula-alvo a maioria das células. Os interferons tem papel importante durante a infecção viral. Os $IFN\alpha/\beta$ são produzidos rapidamente quando fatores virais, como glicoproteínas de envelope viral e RNA de dupla fita, interagem com receptores de reconhecimento de padrão como os receptores toll-like e receptores citosólicos. Essa interação vírus-hospedeiro ativa vias de sinalização celular (os fatores regulatórios de interferon, como por exemplo o *IRF7*, $NF-\kappa B$, via de *JAK-STAT*) que levam à expressão de $IFN\alpha/\beta$. Já o $IFN\gamma$ (interferon do tipo 2) é induzido por estimulação mediada por receptores, em resposta a citocinas

produzidas precocemente (interleucinas 2 e 18 e ao próprio IFN α/β) ou por estimulação através de receptores de células T ou receptores de células *natural killer*. A sinalização de IFN se dá por meio de receptores transmembrana, ativando principalmente a via de JAK-STAT. A expressão de IFN γ induzida por citocinas e receptores de células T ativa vias de transdução de sinal envolvendo fatores de transcrição como NFAT, STATs e NF- κ B, resultando em indução e ativação de fatores antivirais intrínsecos, como a proteína quinase ativada por RNA (PKR, do inglês *RNA-activated protein kinase*) (MALMGAARD, 2004; FENSTERL & SEN, 2009; CHOW & GALE, 2015). Além disso, já foi mostrado que o tratamento com agente desmetilante (100 nM de decitabina por 72h) ativa sensores de RNA de dupla fita (DDX58-MDA5), causando uma resposta de interferon do tipo 1 (indução de IFN β) e apoptose em linhagens de câncer de ovário (CHIAPPINELLI et al., 2015).

Outro ponto importante é que doenças como a síndrome de Sjogren, uma doença sistêmica autoimune crônica caracterizada por infiltração linfocítica nas glândulas salivares e lacrimais resultando em olhos e bocas secas em mulheres de meia idade, apresenta um aumento de interferon do tipo 1 circulante (IFN α) desencadeado pelo aumento de expressão de LINE1 endógeno. Os indivíduos acometidos pela doença apresentam superexpressão de APOBEC3A e APOBEC3B associadas com o aumento de transcrito de LINE1. Os autores sugerem que essas enzimas possam contribuir para a desregulação de LINE1 endógeno nessa doença (MAVRAGANI & CROW, 2010), porém não podemos descartar que a maior produção de IFN α possa levar à indução de *APOBECs*.

Além da reativação de retroelementos do genoma, mais recentemente foi mostrado que DNA circular extracromossomal (DNAecc), formado a partir da ligação de DNA genômico fragmentado liberado por células apoptóticas, pode promover respostas imunológicas. Interferons do tipo 1 (IFN α e IFN β) são induzidos significativamente em uma ampla faixa de concentração de DNAecc (10-240ng/mL) em células dendríticas derivadas de medula óssea transfectadas com DNAecc. Esses elementos também são capazes de apresentar uma atividade imunoestimulante maior (indução de IL-6 e TNF- α) em macrófagos derivados de medula óssea, particularmente em baixas concentrações (10 e 30 ng/mL). A natureza circular desses elementos é fundamental para sua atividade imunoestimulante forte através da via de sinalização de Sting (estimulador de genes interferon) (WANG et al 2021). Esses elementos podem estar livres na circulação, sugerindo sua

potente utilidade como um novo tipo de biomarcador de detecção precoce, de avaliação de tratamento e de acompanhamento de doença (ZHU et al., 2018). Isto poderia ocorrer também em CEE, em que já foi sugerido que DNAAecc são comuns e funcionais já que são diferencialmente expressos entre a mucosa tumoral e epitélio adjacente pareado (SUN et al., 2021).

Como conclusão geral, a hipometilação global (representada pelo tratamento com decitabina) poderia levar à ativação de sensores de RNA de dupla fita (representada pelo tratamento com RNA de dupla fita e pela possível indução dos receptores de RNA de dupla fita, que parece ser o MDA5, no nosso caso), causando uma resposta por interferons (tratamentos com interferons). Tratamento com agente desmetilante ativando sensores de RNA de dupla fita que culmina em uma resposta imune por interferon já foi demonstrado na literatura para câncer de ovário (CHIAPPINELLI et al., 2015) e câncer colorretal (ROULOIS et al., 2015). Os nossos dados sugerem que essa via poderia estar induzindo a expressão da *APOBEC3A* que, por sua vez, poderia estar contribuindo para o estabelecimento da assinatura mutacional em CEE. Isso porque a expressão desse gene foi encontrada aumentada nos tumores, apresentou correlação significativa com a contribuição das assinaturas 2/13 e foi detectada em células de origem epitelial (ou seja, células tumorais) nos experimentos de *single cell*. Porém, não podemos deixar de ressaltar a importância da interação com as células do sistema imune, mais especificamente com os macrófagos.

Esse estudo apresentou algumas limitações, entre elas as análises baseadas em correlações. Essas são evidências fracas, em que não há determinação de causalidade, sendo necessárias mais análises para confirmar tal achado, principalmente envolvendo ensaios *in vitro* e *in vivo*. Os coeficientes de correlação foram significativos, porém fracos (com rho em torno 0,2 a 0,3) para a maioria das análises, sugerindo que outros mecanismos possam desempenhar um papel no estabelecimento das assinaturas mutacionais associadas às APOBECs em câncer. Além disso, pode refletir outra limitação do trabalho, que diz respeito ao momento da análise. As mutações podem ser geradas pelas APOBECs ao longo de todo o processo de carcinogênese, enquanto a expressão dessas enzimas foi determinada no ponto de aquisição da amostra e não necessariamente no momento da mutagênese ativa (PETLJAK et al., 2019; PETLJAK & MACIEJOWSKI, 2020). Ainda, conforme discutido anteriormente, a expressão das *APOBECs* nos tumores pode ser oriunda das células do sistema imune (NG et al., 2019). Outra limitação importante foram os ensaios *in*

vitro. Foi observada grande heterogeneidade entre replicatas em pelo menos um experimento de cada tratamento, apresentando desvios-padrão consideráveis. A baixa expressão dos genes analisados, principalmente em células não tratadas pode explicar tal fenômeno. Sendo assim, são necessários mais experimentos para confirmar a indução sugerida aqui.

Para comprovar nossas hipóteses, além de aumentar o número de replicatas, é necessário realizar a análise de expressão proteica pós-tratamentos. Ainda, a realização de experimentos de modulação de vias de sinalização, por exemplo utilizando inibidores específicos de IFN α , IFN β e IFN γ e de vias ativadas por eles, poderiam nos ajudar a entender os mecanismos mais provavelmente envolvidos. Ainda, a modulação da expressão das *APOBECs*, principalmente de APOBEC3A, faz parte dos nossos planos futuros. Finalmente, a assinatura mutacional deve ser avaliada como desfecho em ensaios *in vitro* a fim de comprovar se APOBEC3A é capaz de gerar a assinatura mutacional em CEE.

6 CONCLUSÕES

- CEE e CEL apresentaram um maior número de mutações quando comparados aos outros tumores do CECP;
- As principais assinaturas mutacionais encontradas nos tumores foram as assinaturas 1/6 (relacionada à idade e ao defeito por mau-pareamento), 2/13 (associadas à atividade de APOBECs), 4 (associada à exposição ao benzopireno), 5 (relacionada à idade, ao tabagismo e ao defeito no reparo por excisão de nucleotídeos), 6/15/21 (relacionadas ao defeito no reparo por mau-pareamento) e a assinatura 16 (sem etiologia conhecida, porém artigos demonstram associação à exposição ao álcool);
- Em todos os tumores, foram encontradas as assinaturas 2/13, associadas à atividade das APOBECs;
- Quando comparada a fração de contribuição das assinaturas 2/13 entre os tumores, tumores CEOF HPV⁻ apresentaram maior contribuição dessas assinaturas;
- Não houve diferença da fração de contribuição das assinaturas 2/13 de acordo com o estadiamento do tumor (tumores precoces *versus* avançados) em CEE, CEL, CECO, CEOF HPV⁻ e o CEOF HPV⁺ e por status de HPV em CEL e CECO;
- Houve correlação direta significativa entre as assinaturas 2/13 com a expressão das *APOBEC3A* e *APOBEC3B* em CEE; com a expressão das *APOBEC3B*, *APOBEC3C*, *APOBEC3D*, *APOBEC3F* e *APOBEC3G* em CEL; com a expressão das *APOBEC3A*, *APOBEC3B*, *APOBEC3C* e *APOBEC3D* em CECO; com a expressão da *APOBEC3A* em CEOF HPV⁺. Além de correlação inversa significativa com a expressão de *AICDA* e *APOBEC2* em CECO;
- Não foi observada uma associação entre a contribuição das assinaturas 2/13 e a sobrevida global dos pacientes com CEE, CEL, CECO, CEOF HPV⁻ e CEOF HPV⁺;
- As *APOBEC3s* estão superexpressas nos tumores quando comparados à mucosa adjacente não tumoral em CEE, CEL e CECO. Em CEOF, essas enzimas estão superexpressas nos tumores HPV⁺ quando comparados a tumores HPV⁻;
- Mutações e alterações do número de cópias nos genes que codificam as APOBECs não são frequentes nesses tumores, bem como as alterações epigenéticas (miRNAs e metilação do

DNA) parecem não ser o principal mecanismo de regulação dessas enzimas em CEE, CEL, CECO, CEOF HPV⁻ e CEOF HPV⁺;

- Foram observadas diversas correlações indiretas significativas entre os níveis de metilação de elementos transponíveis com a expressão das *APOBECs* em CEE, CEL, CECO, CEOF HPV⁻ e CEOF HPV⁺, em que se pode destacar a correlação inversa dos níveis de metilação de ALU com a expressão da *APOBEC3A* em CEE;
- Foram encontradas correlações significativas entre a presença de células do sistema imune com as assinaturas 2/13 e a expressão das *APOBECs*, como por exemplo, com a presença de macrófagos em CEE, CEL e CECO e com mastócitos em CEOF HPV⁺;
- Em modelo de *single cell* de CEE, todas as *APOBECs* foram detectadas em células epiteliais com níveis variáveis entre pacientes;
- As *APOBEC3A* e *APOBEC3B* são mais detectadas em células epiteliais e a *APOBEC3A* também é mais detectada em células mieloides em CEE;
- As *APOBEC3s* estão mais expressas no tumor quando comparado à mucosa adjacente não tumoral em pacientes com CEE do INCA;
- Foi observado um aumento de expressão das *APOBEC3A* e *APOBEC3G* no terço médio quando comparado ao terço inferior em pacientes com CEE do INCA;
- Pacientes com CEE com alta expressão da *APOBEC3D* apresentam um pior prognóstico quando comparados a pacientes com menor expressão desse gene. Porém, pacientes com alta expressão da *APOBEC3G* apresentaram uma melhor sobrevida global quando comparados a pacientes com menor expressão desse gene;
- Nenhum dos tratamentos *in vitro* impactou na viabilidade celular das linhagens de CEE;
- Os tratamentos com decitabina, POLY I:C e com interferons (IFN α , IFN β e IFN γ) parecem induzir a expressão da *APOBEC3A*, bem como dos controles positivos dos tratamentos, sendo mais evidente na linhagem TE13. Porém, são necessários mais experimentos para confirmar a indução de maneira significativa.

Como conclusão final, *APOBEC3A* parece ter um papel importante no estabelecimento da assinatura mutacional em CEE, além de sua interação com células do sistema imune presentes no microambiente tumoral, os macrófagos, por exemplo. Essa enzima parece ser regulada por

interferons produzidos pela resposta à ativação de elementos móveis devido à hipometilação global encontrada em tumores, como em CEE.

7 PERSPECTIVAS

A perspectiva desse trabalho é aumentar o número experimental dos tratamentos realizados para confirmar a indução de expressão de *APOBECs* de maneira significativa nas linhagens de CEE avaliadas nesta tese (TE1 e TE13), bem como em outra linhagem de CEE (OE21), além de avaliar a metilação e a expressão dos elementos móveis nessas linhagens. Após a identificação dos mecanismos e vias de sinalização (vias de interferons) mais provavelmente envolvidos na indução das *APOBECs* em linhagens de CEE, modular tais vias com uso de inibidores específicos para cada tipo de interferon ($IFN\alpha$, $IFN\beta$ e $IFN\gamma$) e das vias de sinalização ativadas por eles. Ainda, modular a expressão das *APOBECs* e avaliar a presença das assinaturas mutacionais 2/13 no genoma dessas células.

8 REFERÊNCIA BIBLIOGRÁFICA

ABNET, C.C.; ARNOLD, M.; WEI, W.Q. Epidemiology of Esophageal Squamous Cell Carcinoma. **Gastroenterology**, 154(2):360-373, 2018.

ADLI, M. The CRISPR tool kit for genome editing and beyond. **Nat Commun**, 9(1):1911, 2018.

ALEXANDROV, L.B. et al. Signatures of mutational processes in human cancer. **Nature**, 500(7463):415-21, 2013.

ALEXANDROV, L.B.; JONES, P.H.; WEDGE, D.C.; SALE, J.E.; CAMPBELL, P.J., et al. Clock-like mutational processes in human somatic cells. **Nat Genet**, 47(12):1402–7, 2015.

ALEXANDROV, L.B.; JU, Y.S.; HAASE, K.; LOO, P.; MARTINCORENA, I., et al. Mutational signatures associated with tobacco smoking in human cancer. **Sci (New York, N.Y.)**, 354(6312):618–22, 2016.

ALEXANDROV, L.B. et al. The repertoire of mutational signatures in human cancer. **Nature**, 578(7793):94-101, 2020.

ALLEN, J.W.; RICHARDSON, J.D.; EDWARDS, M.J. Squamous cell carcinoma of the esophagus: a review and update. **Surg Oncol**, 6(4):193-200, 1997.

ALVES, J.; DEXHEIMER, G.M.; RECKZIGEL, L.; GOETTERT, M.; BIOLCHI, V.; ABUJAMRA, A.L. Changes in IDH2, TET2 and KDM2B Gene Expression After Treatment With Classic Chemotherapeutic Agents and Decitabine in Myelogenous Leukemia Cell Lines. **J Hematol**, 8(3): 89–101, 2019.

AN, J.; RAO, A.; KO, M. TET family dioxygenases and DNA demethylation in stem cells and cancers. **Experimental & Molecular Medicine**, Seoul, v. 49, n. 4, p. e323, 2017.

ANDERS, S.; PYL, P.T.; HUBER, W. HTSeq--a Python framework to work with high-throughput sequencing data. **Bioinformatics**, 31(2):166-9, 2015.

ASHMAN, J.N.;PATMORE, H.S.; CONDON, L.T.; CAWKWELL, L.; STAFFORD, N.D.;GREENMAN, J. Prognostic value of genomic alterations in head and neck squamous cell carcinoma detected by comparative genomic hybridisation. **Br J Cancer**, 89(5):864-9, 2003.

AXELL, T. Occurrence of leukoplakia and some other oral white lesions among 20,333 adult Swedish people. **Community Dent Oral Epidemiol**,15:46-51, 1987.

BABOCI, L. et al. Evidence of the causal role of human papillomavirus type 58 in an oropharyngeal carcinoma. **Virology**, 10(1): 334, 2013.

BARNES, L. et al. World health organization classification of tumours. **Pathology & Genetics - Head and neck Tumours**. Lyon: IARC Press, 168p, 2005.

BARROS, L.R.C.; SOUZA-SANTOS, P.T.; PRETTI, M.A.M.; VIEIRA, G.F.; BRAGATTE, M.A.S.; MENDES, M.F.A.; DE FREITAS, M.V.; SCHERER, N.M.; DE OLIVEIRA, I.M.; RAPOZO, D.C.M.; FERNANDES, P.V.; SIMÃO, T.A.; SOARES-LIMA, S.C.; BORONI, M.; RIBEIRO PINTO, L.F.; BONAMINO, M.H. High infiltration of B cells in tertiary lymphoid structures, TCR oligoclonality, and neoantigens are part of esophageal squamous cell carcinoma microenvironment. **J Leukoc Biol**, 108(4):1307-1318, 2020.

BAYLIN, S.B.; OHM, J.E. Epigenetic gene silencing in cancer - a mechanism for early oncogenic pathway addiction? **Nat Rev Cancer**, v. 6, n. 2, p. 107-116, 2006.

BAYLIN, S.B.; JONES, P.A. A decade of exploring the cancer epigenome - biological and translational implications. **Nat Rev Cancer**, v. 11, n. 10, p. 726-734, 2011.

BERGAMO, N.A et al. Classic and molecular cytogenetic analyses reveal chromosomal gains and losses correlated with survival in head and neck cancer patients. **Clin Cancer Res**, 11: 621–631, 2005.

BIRD, A. DNA methylation patterns and epigenetic memory. **Genes Dev**, v. 16, n. 1, p. 6-21, 2002.

BLOT, W.J. et al. Smoking and drinking in relation to oral and pharyngeal cancer. **Cancer Res** 48, 3282–3287, 1988.

BOCKMÜHL, U.; ISHWAD, C.S.; FERRELL, R.E.; GOLLIN, S.M. Association of 8p23 deletions with poor survival in head and neck cancer. **Otolaryngol Head Neck Surg**, 124(4):451-5, 2001.

BOGERD, H.P.; WIEGAND, H.L.; HULME, A.E.; GARCIA-PEREZ, J.L.; O'SHEA, K.S.; MORAN, J.V.; CULLEN, B.R. Cellular inhibitors of long interspersed element 1 and Alu retrotransposition. **Proc Natl Acad Sci U S A**, 103(23):8780-5, 2006.

BONVIN, M.; ACHERMANN, F.; GREEVE, I et al. Interferon-inducible expression of APOBEC3 editing enzymes in human hepatocytes and inhibition of hepatitis B virus replication. **Hepatology (Baltimore, Md)**, 43:1364, 2006.

BOSMAN, F.T.; CARNEIRO, F.; HRUBAN, R.H.; THEISE, N.D. WHO classification of tumours of the digestive system. 4th ed. Lyon: **IARC Press**; 2010.

BRAY, F.; FERLAY, J.; SOERJOMATARAM, I.; SIEGEL, R.L.; TORRE, L.A.; JEMAL, A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. **CA Cancer J Clin**, 68(6):394-424, Nov 2018.

CANCER GENOME ATLAS NETWORK et al. Comprehensive genomic characterization of head and neck squamous cell carcinomas. *Nature*, 517(7536): 576–582, 2015.

CANCER GENOME ATLAS RESEARCH NETWORK et al. Integrated genomic characterization of oesophageal carcinoma. **Nature**, 541(7636):169-175, 2017.

CANNATARO, V.L.; GAFFNEY, S.G.; SASAKI, T.; ISSAEVA, N.; GREWAL, N.K.S.; GRANDIS, J.R.; YARBROUGH, W.G.; BURTNES, B.; ANDERSON, K.S.; TOWNSEND, J.P. APOBEC-induced mutations and their cancer effect size in head and neck squamous cell carcinoma. **Oncogene**, 38, 3475–3487, 2019.

CAO, R.; YUAN, L.; MA, B.; WANG, G.; TIAN, Y. Tumour microenvironment (TME) characterization identified prognosis and immunotherapy response in muscle-invasive bladder cancer (MIBC). **Cancer Immunol Immunother**, 70(1):1-18, 2021.

CAO, W.; WU, W. Apolipoprotein B mRNA Editing Enzyme, Catalytic Polypeptide-Like Gene Expression, RNA Editing, and MicroRNAs Regulation. **Methods Mol Biol**, 1699:75-81, 2018.

CASARTELLI, N. et al. The antiviral factor APOBEC3G improves CTL recognition of cultured HIV-infected T cells. **J. Exp. Med**, 207, 39–49, 2010.

CASTELLSAGUÉ, X.; MUÑOZ, N.; DE STEFANI, E.; VICTORA, C.G.; CASTELLETTO, R.; ROLÓN, P.A.; QUINTANA, M.J. Independent and joint effects of tobacco smoking and alcohol drinking on the risk of esophageal cancer in men and women. **Int J Cancer**, 82(5):657-64, 1999.

CASTELLSAGUÉ, X.; MUÑOZ, N.; DE STEFANI, E.; VICTORA, C.G.; CASTELLETTO, R.; ROLÓN, P.A. Influence of mate drinking, hot beverages and diet on esophageal cancer risk in South America. **Int J Cancer**, 88(4):658-64, 2000.

CASTELLSAGUE, X.; ALEMANY, L.; QUER, M. et al. HPV involvement in head and neck cancers: comprehensive assessment of biomarkers in 3680 patients. **J Natl Cancer Inst**, 108(6):1-12, 2016.

CAVAL, V.; SUSPÈNE, R.; SHAPIRA, M.; VARTANIAN, J.-P.; WAIN-HOBSON, S. A prevalent cancer susceptibility APOBEC3A hybrid allele bearing APOBEC3B 3'UTR enhances chromosomal DNA damage. **Nat. Commun**, 5:5129, 2014.

CHAI, E.Z.; SIVEEN, K.S.; SHANMUGAM, M.K.; ARFUSO, F.; SETHI, G. Analysis of the intricate relationship between chronic inflammation and cancer. **The Biochemical journal**, 468:1, 2015.

CHEN, L.Q.; HU, C.Y.; GHADIRIAN, P.; DURANCEAU, A. Early detection of esophageal squamous cell carcinoma and its effects on therapy: an overview. **Dis Esophagus**, 12(3):161-7, 1999.

CHEN, K.; HUANG, J.; ZHANG, C.; HUANG, S et al. Alpha interferon potently enhances the anti-human immunodeficiency virus type 1 activity of APOBEC3G in resting primary CD4 T cells. **Journal of virology**, 80:7645, 2006.

CHEN, Y.A.; LEMIRE, M.; CHOUFANI, S.; BUTCHER, D.T.; GRAFODATSKAYA, D.; ZANKE, B.W.; GALLINGER, S.; HUDSON, T.J.; WEKSBERG, R. Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. **Epigenetics**, 8(2):203-9, 2013.

CHEN, H.; CHONG, W.; TENG, C.; YAO, Y.; WANG, X.; LI, X. The immune response-related mutational signatures and driver genes in non-small-cell lung cancer. **Cancer Sci**, 110(8):2348-2356, 2019.

CHIAPPINELLI, K.B.; STRISSEL, P.L.; DESRICHARD, A.; LI, H.; HENKE, C. et al. Inhibiting DNA Methylation Causes an Interferon Response in Cancer via dsRNA Including Endogenous Retroviruses. **Cell**, 162(5):974-86, 2015.

CHOW, K.T.; GALE, M. JR. SnapShot: Interferon Signaling. **Cell**, 163(7):1808-1808.e1, 2015.

CHUANG, S.C.; HASHIBE, M.; SCELO, G. et al. Risk of second primary cancer among esophageal cancer patients: a pooled analysis of 13 cancer registries. **Cancer Epidemiol Biomarkers Prev**, 17(6):1543-9, 2008.

CHUNG, C.S. et al. Risk factors for second primary neoplasia of esophagus in newly diagnosed head and neck cancer patients: a case-control study. **BMC Gastroenterology**, 13:154, 2013.

CIESIELSKI P, JÓŹWIAK P, KRZEŚLAK A. TET proteins and epigenetic modifications in cancers. **Postepy Hig Med Dosw (Online)**, 69:1371-83, 2015.

COCA-PELAZ, A.; RODRIGO, J.P.; TAKES, R.P.; SILVER, C.E.; PACCAGNELLA, D.; RINALDO, A.; HINNI, M.L.; FERLITO, A. Relationship between reflux and laryngeal cancer. **Head Neck**, 35(12):1814-8, 2013.

COHEN, D.J.; AJANI, J. An expert opinion on esophageal cancer therapy. **Expert Opin Pharmacother**, 12(2):225-39, 2011.

COLAPRICO, A.; SILVA, T.C.; OLSEN, C.; GAROFANO, L.; CAVA, C.; GAROLINI, D.; SABEDOT, T.; MALTA, T.; PAGNOTTA, S.M.; CASTIGLIONI, I.; CECCARELLI, M.;

NOUSHMEHR, G.B.H. TCGAbiolinks: An R/Bioconductor package for integrative analysis of TCGA data. **Nucleic Acids Research**, 44 (8):e71, 2016.

CONTICELLO, S. G. The AID/APOBEC family of nucleic acid mutators. **Genome Biology**, London, v. 9, n. 6, p. 229, 2008.

CORLEY, D.A.; BUFFLER, P.A. Oesophageal and gastric cardia adenocarcinomas: analysis of regional variation using the Cancer Incidence in Five Continents database. **Int J Epidemiol**, 30(6):1415-25, 2001.

CORTEZ, L.M.; BROWN, A.L.; DENNIS, M.A. et al. APOBEC3A is a prominent cytidine deaminase in breast cancer. **PLoS Genet**, 15(12):e1008545, 2019.

DA COSTA, N.M.; SOARES LIMA, S.C.; DE ALMEIDA SIMÃO, T.; RIBEIRO PINTO, L.F. The potential of molecular markers to improve interventions through the natural history of oesophageal squamous cell carcinoma. **Biosci Rep**, 33(4), 2013.

DAHER, G.S.A. et al. Características epidemiológicas de casos de câncer de boca registrados em hospital de Uberaba no período 1999-2003: um alerta para a necessidade de diagnóstico precoce. **Revista Brasileira de Epidemiologia**, São Paulo, v. 11, n. 4, p. 584-596, 2008.

DAHL, C.; GRØNBÆK, K.; GULDBERG, P. Advances in DNA methylation: 5-hydroxymethylcytosine revisited. **Clinica Chimica Acta: International Journal of Clinical Chemistry**, Amsterdam, v. 412, n. 11-12, p. 831-836, 2011.

DANIELS, T.F. et al. Lipoproteins, cholesterol homeostasis and cardiac health. **Int. J.Biol.Sci.** 5, 474-488, 2009.

DAWSON, M.A.; KOUZARIDES, T. Cancer epigenetics: from mechanism to therapy. **Cell**, 150(1):12-27, 2012.

DE A SIMÃO, T.; SOUZA-SANTOS, P.T.; DE OLIVEIRA, D.S.; BERNARDO, V.; LIMA, S.C.; RAPOZO, D.C.; KRUEL, C.D.; FARIA, P.A.; RIBEIRO PINTO, L.F.; ALBANO, R.M. Quantitative evaluation of SPRR3 expression in esophageal squamous cell carcinoma by qPCR and its potential use as a biomarker. **Exp Mol Pathol**, 91(2):584-9, 2011.

DE CUBAS, A.A.; DUNKER, W.; ZANINOVICH, A.; ONGO, R.A.; BHATIA, A.; PANDA, A.; BECKERMANN, K.E.; BHANOT, G.; GANESAN, S.; KARIJOLICH, J.; RATHMELL, W.K. DNA hypomethylation promotes transposable element expression and activation of immune signaling in renal cell cancer. **JCI Insight**, 5(11):e137569, 2020.

DE SUMMA, S.; MALERBA, G.; PINTO, R.; MORI, A.; MIJATOVIC, V.; TOMMASI, S. GATK hard filtering: tunable parameters to improve variant calling for next generation sequencing targeted gene panel data. **BMC Bioinformatics**, 18(Suppl 5):119, 2017.

DEDEOGLU, F.; HORWITZ, B.; CHAUDHURI, J.; ALT, F.W.; GEHA, R.S. Induction of activation-induced cytidine deaminase gene expression by IL-4 and CD40 ligation is dependent on STAT6 and NFkappaB. **International immunology**, 16:395, 2004.

DEDIVITIS, R.A. et al. Características clínico-epidemiológicas no carcinoma espinocelular de boca e orofaringe. **Rev Bras Otorrinolaringol**, 70:35-40, 2004.

DEJEUX, E. et al. Identification and quantification of differentially methylated loci by the pyrosequencing technology. **Methods in Molecular Biology**, Clifton, v. 507, p. 189-205, 2009.

DENIS, H.; NDLOVU, M.N.; FUKS, F. Regulation of mammalian DNA methyltransferases: a route to new mechanisms. **EMBO Rep**, v. 12, n. 7, p. 647-656, 2011.

DIKSHIT, R.P.; GILLIO-TOS, A.; BRENNAN, P.; DE MARCO, L.; FIANO, V.; MARTINEZ-PEÑUELA, J.M.; BOFFETTA, P.; MERLETTI, F. Hypermethylation, risk factors, clinical

characteristics, and survival in 235 patients with laryngeal and hypopharyngeal cancers. **Cancer**, 110(8):1745-51, 2007.

DOBROSSY, L. Epidemiology of head and neck cancer: magnitude of the problem. **Cancer and Metastasis Rev**, 24(1):9-17, Jan 2005.

EL-NAGGAR, A.K.; CHAN, J. K. C.; GRANDIS, J. R. et al. World health organization classification of tumours. **WHO Classification of Head and Neck Tumors**. Lyon: IARC Press. 133p, 2017.

EL-SERAG, H.B.; ERGUN, G.A.; PANDOLFINO, J.; FITZGERALD, S.; TRAN, T.; KRAMER, J.R. Obesity increases oesophageal acid exposure. **Gut**, 56(6):749-55, 2007.

ENDO, Y.; MARUSAWA, H.; KINOSHITA, K.; MORISAWA, T.; SAKURAI, T.; OKAZAKI, I.M.; WATASHI, K.; SHIMOTOHNO, K.; HONJO, T.; CHIBA, T. Expression of activation-induced cytidine deaminase in human hepatocytes via NF-kappaB signaling. **Oncogene**, 26:5587, 2007.

ENDO, Y.; MARUSAWA, H.; KOU, T.; NAKASE, H.; FUJII, S.; FUJIMORI, T.; KINOSHITA, K.; HONJO, T.; CHIBA, T. Activation-induced cytidine deaminase links between inflammation and the development of colitis-associated colorectal cancers. **Gastroenterology**, 135:889, 2008

ENZINGER, P.C.; MAYER, R.J. Esophageal cancer. **N Engl J Med**, 349(23):2241-52, 2003.

ESLICK, G.D. Epidemiology of esophageal cancer. **Gastroenterol Clin North Am**, 38(1):17-25, vii, 2009.

ETARD, C. et al. Lack of Apobec2-related proteins causes a dystrophic muscle phenotype in zebrafish embryos. **J. Cell Biol**, 189, 527–539, 2010.

EVANS, M.; NEWCOMBE, R.; FIANDER, A. et al. Human Papillomavirus-associated oropharyngeal cancer: an observational study of diagnosis, prevalence and prognosis in a UK population. **BMC Cancer**, 13(220):1-10, 2013.

FADEN, D.L.; THOMAS, S.; CANTALUPO, P.G.; AGRAWAL, N.; MYERS, J.; DERISI, J. Multi-modality analysis supports APOBEC as a major source of mutations in head and neck squamous cell carcinoma. **Oral Oncol**, 74:8-14, 2017.

FALK, G.W. Risk factors for esophageal cancer development. **Surg Oncol Clin N Am**, 18(3):469-85, 2009.

FANJUL-FERNÁNDEZ, M.; QUESADA, V.; CABANILLAS, R.; CADIÑANOS, J.; FONTANIL, T.; OBAYA, A.; RAMSAY, A.J.; LLORENTE, J.L.; ASTUDILLO, A.; CAL, S.; LÓPEZ-OTÍN, C. Cell-cell adhesion genes CTNNA2 and CTNNA3 are tumour suppressors frequently mutated in laryngeal carcinomas. **Nat Commun**, 4:2531, 2013.

FEINBERG, A.P.; TYCKO, B. The history of cancer epigenetics. **Nat Rev Cancer**, v. 4, n. 2, p. 143-153, 2004.

FENSTERL, V. & SEN, G.C. Interferons and viral infections. **Biofactors**, 35(1):14-20, 2009.

FERLAY, J.; SOERJOMATARAM, I.; DIKSHIT, R. et al. Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. **Int J Cancer**, 136(5):E359-E386, 2015.

FERREIRA, C.G.; CASALI, J.C. **Oncologia Molecular**. Rio de Janeiro: Atheneu, 664p, 2010.

FERREIRA, L.A.M & TUNISSIOLLI, N.M. CÂNCER DE CABEÇA E PESCOÇO: Fatores de Risco, Sinais e Sintomas. **Jornal eletrônico do Instituto NANOCELL**, v.5, n. 02, 2017.

FRANCHINI, D.M.; PETERSEN-MAHRT, S.K. AID and APOBEC deaminases: balancing DNA damage in epigenetics and immunity. **Epigenomics**, 6(4):427-43, 2014.

FRAUER, C. et al. Recognition of 5-hydroxymethylcytosine by the Uhrf1 SRA domain. **PLoS One**, San Francisco, v. 6, n. 6, p. e21306, 2011.

FURLAN, C.; POESEL, J.; BARZAN, L. et al. Prognostic significance of LINE-1 hypomethylation in oropharyngeal squamous cell carcinoma. **Clin Epigenetics**, 9(58):1-11, 2017.

GAN, H. et al. Interferon- γ promotes double-stranded RNA-induced TLR3-dependent apoptosis via upregulation of transcription factor Runx3 in airway epithelial cells. **Am J Physiol Lung Cell Mol Physiol**, 311: L1101-L1112, 2016.

GEHRING, J.S.; FISCHER, B.; LAWRENCE, M.; HUBER, W. SomaticSignatures: inferring mutational signatures from single-nucleotide variants. **Bioinformatics**, 31(22):3673-5, 2015.

GEHRING, J. SomaticCancerAlterations: Somatic Cancer Alterations. **R package version 1.14.0**, 2017. Available on <http://bioconductor.org/packages/release/data/experiment/html/SomaticCancerAlterations.html>

GILLISON, M.L. et al. Evidence for a casual association between human papillomavirus and a subset of head and neck cancers. **J Natl Cancer Inst**, 92(9): 709-720, 2000.

GLOBOCAN 2020. (<http://gco.iarc.fr/today>) World Health Organization. International Agency for Research on Cancer 2021. Último acesso em 17 de setembro de 2021.

GRAHAM, S.V. The human papillomavirus replication cycle, and its links to cancer progression: a comprehensive review. **Clin Sci (Lond)**, 131(17):2201-2221, 2017.

GRANT, W.B. Solar ultraviolet irradiance and cancer incidence and mortality. **Adv Exp Med Biol**, 810:52-68, 2014.

GUO, J. U. et al. Emerging roles of TET proteins and 5-hydroxymethylcytosines in active DNA demethylation and beyond. *Cell Cycle*, Georgetown, v. 10, n. 16, p. 2662-2668, 2011.

HAMILTON, M.L.; REMMEN, H.V.; DRAKE, J.A.; YANG, H.; GUO, Z.M.; KEWITT, K.; WALTER, C.A.; RICHARDSON, A. Does oxidative damage to DNA increase with age? **Proc Nat Acad Sci**, 98(18):10469–74, 2001.

HARRIS, R. S.; LIDDAMENT, M. T. Retroviral restriction by APOBEC proteins. **Nature Review Immunology**, London, v. 4, n. 11, p. 868-8877, 2004.

HASHIBE, M.; BRENNAN, P.; BENHAMOU, S.; CASTELLSAGUE, X.; CHEN, C.; CURADO, M.P., et al. Alcohol drinking in never users of tobacco, cigarette smoking in never drinkers, and the risk of head and neck cancer: pooled analysis in the International Head and Neck Cancer Epidemiology Consortium. **J Natl Cancer Inst**, 99(10):777–789, 2007.

HENDERSON, S.; CHAKRAVARTHY, A.; SU, X.; BOSHOFF, C.; FENTON, T.R. APOBEC-mediated cytosine deamination links PIK3CA helical domain mutations to human papillomavirus-driven tumor development. **Cell Rep**, 7(6):1833-41, 2014.

HERBSTER, S.; FERRARO, C.T.; KOFF, N.K.; ROSSINI, A.; KRUEL, C.D.; ANDREOLLO, N.A.; RAPOZO, D.C.; BLANCO, T.C.; FARIA, P.A.; SANTOS, P.T.; ALBANO, R.M.; SIMÃO, T.DE A.; PINTO, L.F. HPV infection in Brazilian patients with esophageal squamous cell carcinoma: interpopulational differences, lack of correlation with surrogate markers and clinicopathological parameters. **Cancer Lett**, 326(1):52-8, 2012.

HERCEG, Z. Epigenetics and cancer: towards an evaluation of the impact of environmental and dietary factors. **Mutagenesis**, v. 22, n. 2, p. 91-103, 2007.

HOFFMANN, D.; HECHT, S. Handbook of Experimental Pharmacology. 94/1 ed. Cooper CS and Grover PL (eds) Heidelberg: Springer-Verlag pp. 63–102, 1990.

HOFFMANN, D.; HOFFMANN, I.; EL-BAYOUMY, K. The less harmful cigarette: a controversial issue. a tribute to Ernst L. Wynder. **Chem Res Toxicol**, 14(7):767-90, 2001.

HOLLIDAY, R. Epigenetics: a historical overview. **Epigenetics**, v. 1, n. 2, p. 76-80, 2006.

HONGO, M.; NAGASAKI, Y.; SHOJI, T. Epidemiology of esophageal cancer: Orient to Occident. Effects of chronology, geography and ethnicity. **J Gastroenterol Hepatol**, 24(5):729-35, 2009.

HUBER, M. A.; TANTIWONGKOSI, B. Oral and oropharyngeal cancer. **Med Clin North Am**, 98(6):1299-1321, 2014.

HWANG, J.K. ET AL. Related mechanisms of antibody somatic hypermutation and class switch recombination. **Microbiol. Spectr.** 3, MDNA3-0037-2014, 2015.

IARC. Tobacco smoking. IARC monographs on the evaluation of the carcinogenic risk of chemicals to humans. Lyon, France: **IARC**, 38:35-394, 1986.

INCA. Monitoramento das Ações de Controle dos Cânceres de Lábio e Cavidade Oral. **Informativo Detecção Precoce**, ano 7, n. 1, 2016. Disponível em: <http://www1.inca.gov.br/inca/Arquivos/informativodeteccaoprecoce_janabr2016.pdf>.

INCA. Estimativa 2020: incidência de câncer no Brasil. **Instituto Nacional de Câncer José Alencar Gomes da Silva**, Ministério da Saúde, 2019.

ISHWARAN, H.; BLACKSTONE, E.H.; APPERSON-HANSEN, C.;RICE, T.W. A novel approach to cancer staging: application to esophageal cancer. **Biostatistics**, 10(4):603-20, 2009.

ITO, F.; FU, Y.; KAO, S.A.; YANG, H.; CHEN, X.S. Family-Wide Comparative Analysis of Cytidine and Methylcytidine Deamination by Eleven Human APOBEC Proteins. **J Mol Biol**, 429(12): 1787–1799, 2017.

JEMAL, A.; SIEGEL, R.; WARD, E.; MURRAY, T.; XU, J.; SMIGAL, C.; THUN, M.J. Cancer statistics, 2006. **CA Cancer J Clin**, 56(2):106-30, 2006.

JEMAL, A.; BRAY, F.; CENTER, M.M.; FERLAY, J.; WARD, E.; FORMAN, D. Global Cancer Statistics. **CA Cancer J Clin**, 61(2):69-90, 2011.

JIANG, C.; FOLEY, J.; CLAYTON, N.; KISSLING, G.; JOKINEN, M.; HERBERT, R.; DIAZ, M. Abrogation of lupus nephritis in activation-induced deaminase-deficient MRL/lpr mice. **Journal of immunology**, 178:7422, 2007.

JIN, C.; QIN, T.; BARTON, M.C.; JELINEK, J.; ISSA, J-P.J. Minimal role of base excision repair in TET-induced global DNA demethylation in HEK293T cells. **Epigenetics**, 10(11): 1006–1013, 2015.

JOENSUU, H.; JYRKKIÖ, S.; KELLOKUMPU-LEHTINEN, P-L.; KOURI, M.; ROBERTS, P.J. & TEPPU, L. (TOIM.) Syöpätaudit. Helsinki: Kustannus Oy Duodecim, 2013.

JOHNSON, D.E.; BURTNES, B.; LEEMANS, C.R.; LUI, V.W.Y.; BAUMAN, J.E.; GRANDIS, J.R. Head and neck squamous cell carcinoma. **Nat Rev Dis Primers**, 6(1):92, 2020.

JOHNSON, N.W. et al. Squamous cell carcinoma and precursor lesions of the oral cavity: epidemiology and aetiology. **Periodontology**, 57:19-37, 2011.

KAUL, D.; ARORA, M.; GARG, A.; SHARMA, S. MALT1 induced immune response is governed by miR-2909 RNomics. **Mol Immunol**, 64 (1):210–217, 2015.

KIM, D.; PERTEA, G.; TRAPNELL, C.; PIMENTEL, H.; KELLEY, R.; SALZBERG, S.L. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. **Genome Biol**, 14(4):R36, 2013.

KIM, Y.A.; WOJTOWICZ, D.; SARTO BASSO, R.; SASON, I.; ROBINSON, W.; HOCHBAUM, D.S.; LEISERSON, M.D.M.; SHARAN, R.; VADIN, F.; PRZYTYCKA, T.M. Network-based approaches elucidate differences within APOBEC and clock-like signatures in breast cancer. **Genome Med**, 12(1):52, 2020.

KNISBACHER, B. A.; GERBER, D.; LEVANON, E. Y. DNA Editing by APOBECs: a genomic preserver and transformer. **Trends in Genetics**, Amsterdam, v. 32, n. 1, p. 16-28, 2016.

KOH, K. P. et al. Tet1 and Tet2 regulate 5-hydroxymethylcytosine production and cell lineage specification in mouse embryonic stem cells. **Cell Stem Cell**, Cambridge, v. 8, n. 2, p. 200-213, 2011.

KONING, F.A.; NEWMAN, E.N.; KIM, E.Y.; KUNSTMAN, K.J.; WOLINSKY, S.M.; MALIM, M.H. Defining APOBEC3 expression patterns in human tissues and hematopoietic cell subsets. **Journal of virology**, 83:9474, 2009.

KUDO, Y. et al. Loss of 5-hydroxymethylcytosine is accompanied with malignant cellular transformation. **Cancer Science**, Tokyo, v. 103, n. 4, p. 670-676, 2012.

KUMAR, M.; NANAVATI, R.; MODI, T.G.; DOBARIYA, C. Oral cancer: Etiology and risk factors: A review. **J Cancer Res Ther**, 12(2):458-63, 2016.

LAW, E.K.; LEVIN-KLEIN, R.; JARVIS, M.C.; KIM, H. APOBEC3A catalyzes mutation and drives carcinogenesis in vivo. **J Exp Med**, 217(12):e20200261, 2020.

LAWRENCE, M.S. et al. Mutational heterogeneity in cancer and the search for new cancer-associated genes. **Nature**, 499:214–218, 2013.

LAWRENCE, M.S. et al. Discovery and saturation analysis of cancer genes across 21 tumour types. **Nature**, 505:495–501, 2014.

LEE, Y.C. et al. Revisit of field cancerization in squamous cell carcinoma of upper aerodigestive tract: better risk assessment with epigenetic markers. **Cancer Prev Res**, 4:1982-1992, 2011.

LEEMANS, C.R.; BRAAKHUIS, B.J.; BRAKENHOFF, R.H. The molecular biology of head and neck cancer. **Nat Rev Cancer**, 11(1):9-22, 2011.

LEEMANS, C.R.; SNIJDERS, P.J.F.; BRAKENHOFF, R.H. The molecular landscape of head and neck cancer. **Nat Rev Cancer**, 18(5):269-282, 2018.

LEONARD, B. et al. APOBEC3G Expression Correlates with T-Cell Infiltration and Improved Clinical Outcomes in High-grade Serous Ovarian Carcinoma. **Clin Cancer Res**, 22(18):4746-55, 2016.

LI, X.C.; WANG, M.Y.; YANG, M.; DAI, H.J.; ZHANG, B.F.; WANG, W.; CHU, X.L., WANG, X.; ZHENG, H.; NIU, R.F.; ZHANG, W.; CHEN, K.X. A mutational signature associated with alcohol consumption and prognostically significantly mutated driver genes in esophageal squamous cell carcinoma. **Ann Oncol**, 29(4):938-944, 2018.

LIAN, C. G. et al. Loss of 5-hydroxymethylcytosine is an epigenetic hallmark of melanoma. **Cell**, Cambridge, v. 150, n. 6, p. 1135-1146, 2012.

LIAO, W. et al. APOBEC-2, a cardiac-and skeletal muscle-specific member of the cytidine deaminase super gene family. **Biochem. Biophys. Res. Commun.** 260, 398–404, 1999.

LIJINSKY, W. Chemistry and biology of N-nitroso compounds. **Cambridge University Press**, 1992.

LIM, D.M. et al. TLR3-mediated NF- κ B signaling in human esophageal epithelial cells. *Am J Physiol Gastrointest Liver Physiol* 297: G1172–G1180, 2009.

LIMA, S.C.; HERNÁNDEZ-VARGAS, H.; SIMÃO, T.; DURAND, G.; KRUEL, C.D.; LE CALVEZ-KELM, F.; RIBEIRO PINTO, L.F.; HERCEG, Z. Identification of a DNA methylome signature of esophageal squamous cell carcinoma and potential epigenetic biomarkers. **Epigenetics**, 6(10):1217-27, 2011.

LIMA, S.C.S. **Alterações epigenéticas e suas consequências funcionais em carcinoma epidermoide de esôfago**. Tese (Doutorado em Oncologia) – Instituto Nacional de Câncer, Rio de Janeiro, 2012.

LIN, D.C. et al. Genomic and molecular characterization of esophageal squamous cell carcinoma. **Nat Genet**, 46(5):467-73, 2014.

LIU, C. & XU, D. Inhibition of histone deacetylases. In: *epigenetics protocols*. New Jersey: **Humana Press**, p. 87-97. (Methods in Molecular Biology, 287), 2004.

LIU, Q., LUO, Y-W. et al. Association between APOBEC3H-Mediated Demethylation and Immune Landscape in Head and Neck Squamous Carcinoma. **Biomed Res Int**, 2020:4612375, 2020.

LOPEZ, J. et al. The context and potential of epigenetics in oncology. **Br J Cancer**, v. 100, n. 4, p. 571-577, 2009.

LOVE, M.I.; HUBER, W.; ANDERS, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. **Genome Biol**, 15(12):550, 2014.

LUCIFORA, J.; XIA, Y.; REISINGER, F.; ZHANG, K et al. Specific and nonhepatotoxic degradation of nuclear hepatitis B virus cccDNA. **Science (New York, NY)**, 343:1221, 2014.

MA, L.J.; LI, W.; ZHANG, X.; HUANG, D.H.; ZHANG, H.; XIAO, J.Y.; TIAN, Y.Q. Differential gene expression profiling of laryngeal squamous cell carcinoma by laser capture microdissection and complementary DNA microarrays. **Arch Med Res**, 40(2):114-23, 2009.

MALMGAARD, L. Induction and regulation of IFNs during viral infections. **J Interferon Cytokine Res**, 24(8):439-54, 2004.

MARSIT, C.J.; POSNER, M.R.; MCCLEAN, M.D.; KELSEY, K.T. Hypermethylation of E-cadherin is an independent predictor of improved survival in head and neck squamous cell carcinoma. **Cancer**, 113(7):1566-71, 2008.

MARTINCORENA, I.; FOWLER, J.C.; WABIK, A. et al. Somatic mutant clones colonize the human esophagus with age. **Science**, 362(6417): 911–917, 2018.

MARUCCIA, M.; ONESTI, M.G.; PARISI, P.; CIGNA, E.; TROCCOLA, A.; SCUDERI, N. Lip cancer: a 10-year retrospective epidemiological study. **Anticancer Res**, 32(4):1543-6, 2012.

MATIJEVIĆ, T. & PAVELIĆ, J. Poly(I:C) treatment influences the expression of calreticulin and profilin-1 in a human HNSCC cell line: a proteomic study. **Tumor Biol**, 33:1201–1208, 2012.

MATSUMOTO, T.; SHIMIZU, T.; TAKAI, A.; MARUSAWA, H. Exploring the Mechanisms of Gastrointestinal Cancer Development Using Deep Sequencing Analysis. **Cancers (Basel)**, 7(2):1037-51, 2015.

MAVRAGANI, C.P.; CROW, M.K. Activation of the type I interferon pathway in primary Sjogren's syndrome. **J Autoimmun**, 35(3):225-31, 2010.

MCCABE & DLAMINI Z. The molecular mechanisms of oesophageal cancer. **International Immunopharmacology**, 5:1113–1130, 2005.

MCLARTY, A.J.; DESCHAMPS, C.; TRASTEK, V.F.; ALLEN, M.S.; PAIROLERO, P.C.; HARMSSEN, W.S. Esophageal resection for cancer of the esophagus: long-term function and quality of life. **Ann Thorac Surg**, 63(6):1568-72, 1997.

MEHDIPOUR, P.; MARHON, S.A.; ETTAYEBI, I.; CHAKRAVARTHY, A.; HOSSEINI, A.; WANG, Y.; DE CASTRO, F.A.; LOO YAU, H.; ISHAK, C.; ABELSON, S.; O'BRIEN, C.A.; DE CARVALHO, D.D. Epigenetic therapy induces transcription of inverted SINEs and ADAR1 dependency. **Nature**, 588(7836):169-173, 2020.

MEHDIPOUR, P.; MARHON, S.A.; ETTAYEBI, I. et al. Publisher Correction: Epigenetic therapy induces transcription of inverted SINEs and ADAR1 dependency. **Nature** 591, E20 (2021).

MELIS, J.P.; VAN STEEG, H.; LUIJTEN, M. Oxidative DNA damage and nucleotide excision repair. **Antioxid Redox Signal**, 18(18):2409–19, 2013.

MERLEVEDE, J. et al. Mutation allele burden remains unchanged in chronic myelomonocytic leukaemia responding to hypomethylating agents. **Nat. Commun.**, 7, 10767, 2016.

MESSERSCHMIDT, C. et al. Distinct immune evasion in APOBEC-enriched, HPV-negative HNSCC. **Int J Cancer**, 147(8):2293-2302, 2020.

MIKL, M.C. et al. Mice deficient in APOBEC2 and APOBEC3. **Mol. Cell Biol**, 25, 7270–7277, 2005.

MOLLOY, P.L.; WATT, F. DNA methylation and specific protein-DNA interactions. **Philos Trans R Soc Lond B Biol Sci**, v. 326, n. 1235, p. 267-275, 1990.

MOODY, S. et al. Mutational signatures in esophageal squamous cell carcinoma from eight countries with varying incidence. **Nature Genetics**, 53(11):1553-1563, 2021.

MORGANELLA S et al. The topography of mutational processes in breast cancer genomes. **Nat Commun**, 7:11383, 2016.

MPAKOU, V.; SPATHIS, A.; BOUHLA, A. et al. Synergistic inhibitory effects of low-dose decitabine in combination with bortezomib in the AML cell line Kasumi-1. **Exp Ther Med**, 21(3):195, 2021.

MURAMATSU, M. et al. Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. **Cell** 102, 553–563, 2000.

MURATA, A. et al. TET family proteins and 5-hydroxymethylcytosine in esophageal squamous cell carcinoma. **Oncotarget**, Albany, v. 6, n. 27, p. 23372-23382, 2015.

NABEL, C.S.; MANNING, S.A.; KOHLI, R.M. The curious chemical biology of cytosine: deamination, methylation, and oxidation as modulators of genomic potential. **ACS Chem Biol**, 7(1):20-30, 2012.

NAN, X. et al. Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. **Nature**, v. 393, n. 6683, p. 386-389, 1998.

NASSAR, A.; UMETON, R.; KIM, J. et al. Mutational Analysis of 472 Urothelial Carcinoma Across Grades and Anatomic Sites. **Clin Cancer Res**, 25(8):2458-2470, 2019.

NEWMAN, A.M.; LIU, C.L.; GREEN, M.R. et al. Robust enumeration of cell subsets from tissue expression profiles. **Nat Methods**, 12:453-457, 2015.

NG, J.C.F.; QUIST, J.; GRIGORIADIS, A.; MALIM, M.H.; FRATERNALI, F. Pan-cancer transcriptomic analysis dissects immune and proliferative functions of APOBEC3 cytidine deaminases. **Nucleic Acids Res**, 47 1178–1194. 10.1093/nar/gky1316, 2019.

NICHOLSON, A.G.; TSAO, M.S.; TRAVIS, W.D.; PATIL, D.T.; GALATEAU-SALLE, F.; MARINO, M.; DACIC, S.; BEASLEY, M.B.; BUTNOR, K.J.; YATABE, Y.; PASS, H.I.; RUSCH, V.W.; DETTERBECK, F.C.; ASAMURA, H.; RICE, T.W.; RAMI-PORTA, R. Eighth Edition Staging of Thoracic Malignancies: Implications for the Reporting Pathologist. **Arch Pathol Lab Med**, 142(5):645-661, 2018.

NICIURA, S.C.M. & SARAIVA, N.Z. Epigenética: bases moleculares, efeitos na fisiologia e na patologia, e implicações para a produção animal e vegetal. Brasília: Embrapa, 2014.

NIK-ZAINAL, S.; ALEXANDROV, L.B.; WEDGE, D.C.; VAN LOO, P.; GREENMAN, C.D.; RAINE, K.; JONES, D.; HINTON, J., et al. Mutational processes molding the genomes of 21 breast cancers. **Cell**, 149(5):979-93, 2012.

NIK-ZAINAL, S.; WEDGE, D.C.; ALEXANDROV, L.B.; PETLJAK, M.; BUTLER, A.P.; BOLLI, N.; STRATTON, M.R. Association of a germline copy number polymorphism of APOBEC3A and APOBEC3B with burden of putative APOBEC-dependent mutations in breast cancer. **Nat. Genet**, 46:487–491, 2014.

NOONE, A.M.; HOWLADER, N.; KRAPCHO, M.; MILLER, D.; BREST, A.; YU, M.; RUHL, J.; TATALOVICH, Z.; MARIOTTO, A.; LEWIS, D.R.; CHEN, H.S.; FEUER, E.J.; CRONIN, K.A. (EDS). SEER Cancer Statistics Review, 1975-2015, **National Cancer Institute**. Bethesda, MD, https://seer.cancer.gov/csr/1975_2015/, based on November 2017 SEER data submission, posted to the SEER web site, April 2018.

NORMAN, J.M. et al. The antiviral factor APOBEC3G enhances the recognition of HIV-infected primary T cells by natural killer cells. **Nat. Immunol**, 12, 975–983, 2011.

OLIVIER M, HOLLSTEIN M, HAINAUT P. TP53 mutations in human cancers: origins, consequences, and clinical use. **Cold Spring Harb Perspect Biol**, 2:a001008, 2010.

OLSON, M.E.; HARRIS, R.S.; HARKI, D.A. APOBEC Enzymes as Targets for Virus and Cancer Therapy. **Cell Chem Biol**, 25(1):36-49, 2018.

ONOMOTO, K.; ONOGUCHI, K.; YONEYAMA, M. Regulation of RIG-I-like receptor-mediated signaling: interaction between host and viral factors. **Cell Mol Immunol**, 18(3):539-555, 2021.

PARKIN, D.M.; BRAY, F.; FERLAY, J.; PISANI, P. Global cancer statistics, 2002. **CA Cancer J Clin**, 55(2):74-108, 2005.

PARKIN, D.M. The global health burden of infection-associated cancers in the year 2002. **International journal of cancer**, 118:3030, 2006.

PATEL, K.; BHAT, F.A.; PATIL, S.; ROUSTRAY, S.; MOHANTY, N.; NAIR, B.; SIDRANSKY, D.; GANESH, M.S.; RAY, J.G.; GOWDA, H.; CHATTERJEE, A. Whole-Exome Sequencing Analysis of Oral Squamous Cell Carcinoma Delineated by Tobacco Usage Habits. **Front Oncol**, 11:660696, 2021.

PATIL, V.; WARD, R.L. & HESSON, L.B. The evidence for functional non-CpG methylation in mammalian cells. **Epigenetics**, 9(6): 823–828, 2014.

PATTERSON, K. et al. DNA methylation: bisulphite modification and analysis. **Journal of Visualized experiments**, Cambridge, v. 56, p. 3170, 2011.

PENA-DIAZ, J. et al. Noncanonical mismatch repair as a source of genomic instability in human cells. **Molecular cell**, 47:669–680, 2012.

PENG, G.; LEI, K.J.; JIN, W.; GREENWELL-WILD, T.; WAHL, S.M. Induction of APOBEC3 family proteins, a defensive maneuver underlying interferon-induced anti-HIV-1 activity. **The Journal of experimental medicine**, 203:41, 2006.

PERIYASAMY, M.; SINGH, A.K.; GEMMA, C. et al. p53 controls expression of the DNA deaminase APOBEC3B to limit its potential mutagenic activity in cancer cells. **Nucleic Acids Res**, 45(19):11056-11069, 2017.

PETLJAK, M.; ALEXANDROV, L.B.; PRICE, S.; BRAMMELD, J.S. et al. Characterizing Mutational Signatures in Human Cancer Cell Lines Reveals Episodic APOBEC Mutagenesis. **Cell**, 176(6):1282-1294.e20, 2019.

PETLJAK, M. & MACIEJOWSKI, J. Molecular Origins of APOBEC-Associated Mutations in Cancer. **DNA Repair (Amst)**, 94:102905, 2020.

PFEIFER G.P.; DENISSENKO, M.F.; OLIVIER, M.; TRETAKOVA, N.; HECHT, S.S.; HAINAUT, P. Tobacco smoke carcinogens, DNA damage and p53 mutations in smoking-associated cancers. **Oncogene**, 21(48):7435-51, 2002.

PFEIFER, G.P. Environmental exposures and mutational patterns of cancer genomes. **Genome medicine**, 2:54, 2010.

PICKENS, A.; ORRINGER, M.B. Geographical distribution and racial disparity in esophageal cancer. **Ann Thorac Surg**, 76(4):S1367-9, 2003.

PINLAOR, S.; MA, N.; HIRAKU, Y.; YONGVANIT, P.; SEMBA, R.; OIKAWA, S. et al. Repeated infection with *Opisthorchis viverrini* induces accumulation of 8-nitroguanine and 8-oxo-7,8-dihydro-2'-deoxyguanine in the bile duct of hamsters via inducible nitric oxide synthase. **Carcinogenesis**, 25:1535-42, 2004.

QUEIROGA, R.C.; PERNAMBUCO, A.P. Câncer de esôfago: epidemiologia, diagnóstico e tratamento. **Revista Brasileira de Cancerologia**, 52(2): 173-178, 2006.

RAKYAN, V.K.; DOWN, T.A.; BALDING, D.J.; BECK, S. Epigenome-wide association studies for common human diseases. **Nat Rev Genet**, 12:529-41, 2011.

READ, S.A.; DOUGLAS, M.W. Virus induced inflammation and cancer development. **Cancer letters**, 345:174, 2014.

REBHANDL, S.; HUEMER, M.; GREIL, R.; GEISBERGER, R. AID/APOBEC deaminases and cancer. **Oncoscience**, 2(4):320-33, 2015.

REFSLAND, E.W.; STENGLEIN, M.D.; SHINDO, K.; ALBIN, J.S.; BROWN, W.L.; HARRIS, R.S. Quantitative profiling of the full APOBEC3 mRNA repertoire in lymphocytes and tissues: implications for HIV-1 restriction. **Nucleic Acids Res**, 38(13):4274-84, 2010.

REICHENBACH, Z.W. et al. Clinical and translational advances in esophageal squamous cell carcinoma. **Adv Cancer Res**, 144:95-135, 2019.

REVATHIDEVI, S.; MURUGAN, A.K.; NAKAOKA, H.; INOUE, I.; MUNIRAJAN, A.K. APOBEC: A molecular driver in cervical cancer pathogenesis. **Cancer Lett**, 496:104-116, 2021.

REZENDE, C.P. et al. Mudança do paradigma da saúde bucal em pacientes com câncer de boca e orofaringe. **Revista Brasileira de Cirurgia de Cabeça e Pescoço**, São Paulo, v. 36, n. 1, p. 19-22, 2007.

RIBEIRO PINTO, L.F.; TEIXEIRA ROSSINI, A.M.; ALBANO, R.M.; FELZENSZWALB, I.; DE MOURA GALLO, C.V.; NUNES, R.A.; ANDREOLLO, N.A. Mechanisms of esophageal cancer development in Brazilians. **Mutat Res**, 544(2-3):365-73, 2003.

RIVA, G.; ALBANO, C.; GUGLIESI, F.; PASQUERO, S.; PACHECO, S.F.C.; PECORARI, G.; LANDOLFO, S.; BIOLATTI, M.; DELL'OSTE, V. HPV Meets APOBEC: New Players in Head and Neck Cancer. **Int J Mol Sci**, 22(3):1402, 2021.

ROBERTSON, A.G. et al. Comprehensive Molecular Characterization of Muscle-Invasive Bladder Cancer. **Cell**, 171(3):540-556.e25, 2017.

ROGOZIN, I.B. et al. APOBEC4, a new member of the AID/APOBEC family of polynucleotide (deoxy) cytidine deaminases predicted by computational analysis. **Cell Cycle** 4, 1281–1285, 2005.

ROSSINI, A. et al. TP53 mutation profile of esophageal squamous cell carcinomas of patients from Southeastern Brazil. **Mutat Res**, 696(1):10-5, 2010.

ROULOIS, D.; LOO YAU, H.; SINGHANIA, R.; WANG, Y.; DANESH, A.; SHEN, S.Y.; HAN, H.; LIANG, G.; JONES, P.A.; PUGH, T.J.; O'BRIEN, C.; DE CARVALHO, D.D. DNA-Demethylating Agents Target Colorectal Cancer Cells by Inducing Viral Mimicry by Endogenous Transcripts. **Cell**, 162(5):961-73, 2015.

SALTER, J.D.; BENNETT, R.P.; SMITH, H.C. The APOBEC Protein Family: United by Structure, Divergent in Function. **Trends Biochem Sci**, 41(7):578-594, 2016.

SATO, Y. et al. Deficiency in APOBEC2 leads to a shift in muscle fiber type, diminished body mass, and myopathy. **J. Biol. Chem**, 285, 7111–7118, 2010.

SAXONOV, S.; BERG, P.; BRUTLAG, D.L. A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. **Proc Natl Acad Sci U S A**, 103:1412–7, 2006.

SAWADA, G. et al. Genomic Landscape of Esophageal Squamous Cell Carcinoma in a Japanese Population. **Gastroenterology**, 150(5):1171-1182, 2016.

SCHÄFER, A. Gadd45 proteins: key players of repair-mediated DNA demethylation. In: LIEBERMANN, D.; HOFFMAN, B. (eds). **Gadd45 Stress Sensor Genes**. New York: Springer, p. 35-50, 2013. (Advances in Experimental Medicine and Biology, v. 793).

SCHMITTGEN, T.D. & LIVAK, K.J. Analyzing real-time PCR data by the comparative C(T) method. **Nat Protoc**, 3: 1101-1108, 2008.

SEDDER, L.M. microRNA control of interferons and interferon induced anti-viral activity. **Mol Immunol**, 56(4):781-93, 2013.

SEPEHR, A.; TANIÈRE, P.; MARTEL-PLANCHE, G.; ZIA'EE, A.A.; RASTGAR-JAZII, F.; YAZDANBOD, M.; ETEMAD-MOGHADAM, G.; KAMANGAR, F.; SAIDI, F.; HAINAUT, P. Distinct pattern of TP53 mutations in squamous cell carcinoma of the esophagus in Iran. **Oncogene**, 20(50):7368-74, 2001.

SEVERI, F. et al. Analysis of reptilian APOBEC1 suggests that RNA editing may not be its ancestral function. **Molecular Biology and Evolution**, Chicago, v. 28, 1125–1129, 2011.

SHARMA, S. et al. APOBEC3A cytidine deaminase induces RNA editing in monocytes and macrophages. **Nat. Commun**, 6, 6881, 2015.

SHI, X. et al. Loss of 5-Hydroxymethylcytosine is an independent unfavorable prognostic factor for esophageal squamous cell carcinoma. **PLoS One**, San Francisco, v. 11, n. 4, e0153100, 2016.

SIMÃO, T.DE.A.; SIMÕES, G.L.; RIBEIRO, F.S.; CIDADE, D.A.; ANDREOLLO, N.A.; LOPES, L.R.; MACEDO, J.M.; ACATAUASSU, R.; TEIXEIRA, A.M.; FELZENSZWALB, I.; PINTO, L.F.; ALBANO, R.M. Lower expression of p14ARF and p16INK4a correlates with higher DNMT3B expression in human oesophageal squamous cell carcinomas. **Hum Exp Toxicol**, 25(9):515-22, 2006.

SIRIWARDENA, S.; CHEN, K.; BHAGWAT, A.S. The Functions and Malfunctions of AID/APOBEC Family Deaminases: the known knowns and the known unknowns. **Chem Rev**, 116(20): 12688–12710, 2016.

SMEDS, J.; BERGGREN, P.; MA, X.; XU, Z.; HEMMINKI, K.; KUMAR, R. Genetic status of cell cycle regulators in squamous cell carcinoma of the oesophagus: the CDKN2A (p16(INK4a) and p14(ARF)) and p53 genes are major targets for inactivation. **Carcinogenesis**, 23(4):645-55, 2002.

SOARES-LIMA, S.C. et al. Upper Aerodigestive Tract Squamous Cell Carcinomas Show Distinct Overall DNA Methylation Profiles and Different Molecular Mechanisms behind WNT Signaling Disruption. **Cancers (Basel)**, 13(12):3014, 2021.

SOBIN, L.H.; GOSPODAROWICZ, M.K.; WITTEKIND, CH. TNM Classification of Malignant Tumours. **UICC International Union Against Cancer**, 7th ed., 2009.

SOUTH, A.P.; DEN BREEMS, N.Y.; RICHA, T.; NWAGU, U.; ZHAN, T.; POOJAN, S.; MARTINEZ-OUTSCHOORN, U.; JOHNSON, J.M.; LUGINBUHL, A.J.; CURRY, J.M. Mutation signature analysis identifies increased mutation caused by tobacco smoke associated DNA adducts in larynx squamous cell carcinoma compared with oral cavity and oropharynx. **Sci Rep**, 9(1):19256, 2019.

SOUZA-SANTOS, P.T.; SOARES LIMA, S.C.; NICOLAU-NETO, P.; BORONI, M.; MEIRELES DA COSTA, N.; BREWER, L.; MENEZES, A.N.; FURTADO, C.; MOREIRA, M.A.M.; SEUANEZ, H.N.; DE ALMEIDA SIMÃO, T.; RIBEIRO PINTO, L.F. Mutations, Differential Gene Expression, and Chimeric Transcripts in Esophageal Squamous Cell Carcinoma Show High Heterogeneity. **Transl Oncol**, 11(6):1283-1291, 2018.

SPECHLER, S.J.; ZEROOGIAN, J.M.; ANTONIOLI, D.A.; WANG, H.H.; GOYAL, R.K. Prevalence of metaplasia at the gastro-oesophageal junction. **Lancet**, 3;344(8936):1533-6, 1994.
STEIN, H.J.; BRÜCHER, B.L.; SENDLER, A.; SIEWERT, J.R. Esophageal cancer: patient evaluation and pre-treatment staging. **Surg Oncol**, 10(3):103-11, 2001.

STRANSKY, N.; EGLOFF, A.M.; TWARD, A.D.; KOSTIC, A.D.; CIBULSKIS, K et al. The mutational landscape of head and neck squamous cell carcinoma. **Science**, 333(6046):1157-60, 2011.

STRATTON, M.R.; CAMPBELL, P.J.; FUTREAL, P.A. The cancer genome. **Nature**, 458:719–724, 2009.

SUN, Z.; JI, N.; ZHAO, R.; LIANG, J.; JIANG, J.; TIAN, H. Extrachromosomal circular DNAs are common and functional in esophageal squamous cell carcinoma. **Ann Transl Med**, 9(18):1464, 2021.

TA HILIANI, M. et al. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. **Science**, New York, v. 324, n. 5929, p. 930-935, 2009.

TANAKA, Y.; MARUSAWA, H.; SENO, H et al. Anti-viral protein APOBEC3G is induced by interferonalpha stimulation in human hepatocytes. **Biochemical and biophysical research communications**, 341:314, 2006.

THERNEAU, T. A Package for Survival Analysis in S. version 2.38 [Internet], 2015. Accessed on 2019 May 30. Available from: Available from: URL: <http://CRAN.R-project.org/package=survival>.

THOMAS, R.M.; SOBIN, L.H. Gastrointestinal cancer. **Cancer**, 75(1 Suppl):154-70, 1995.

TOH, Y.; OKI, E.; OHGAKI, K.; SAKAMOTO, Y.; ITO, S.; EGASHIRA, A.; SAEKI, H.; KAKEJI, Y.; MORITA, M.; SAKAGUCHI, Y.; OKAMURA, T.; MAEHARA, Y. Alcohol drinking, cigarette smoking, and the development of squamous cell carcinoma of the esophagus: molecular mechanisms of carcinogenesis. **Int J Clin Oncol**, 15(2):135-44, 2010.

TUYNS, A.J. Epidemiology of alcohol and cancer. **Cancer Res**, 39(7 Pt 2):2840-3, 1979.

ULLAH, I.; KARTHIK, G-M.; ALKODSI, A. et al. Evolutionary history of metastatic breast cancer reveals minimal seeding from axillary lymph nodes. **J Clin Invest**, 128(4): 1355–1370, 2018.

VAISSIÈRE, T. et al. Quantitative analysis of DNA methylation profiles in lung cancer identifies aberrant DNA methylation of specific genes and its association with gender and cancer risk factors. **Cancer Research**, Chicago, v. 69, n. 1, p. 243-252, 2009.

VAN MONSJOU, H.S. et al. Oropharyngeal squamous cell carcinoma: a unique disease on the rise? **Oral Oncol**, 46(11): 780-785, 2010.

VIEL, A.; BRUSELLES, A.; MECCIA, E.; FORNASARIG, M.; QUAIA, M.; CANZONIERI, V.; POLICICCHIO, E. et al. A Specific Mutational Signature Associated with DNA 8-Oxoguanine Persistence in MUTYH-defective Colorectal Cancer. **EBioMedicine**, 20:39-49, 2017.

WANG, Z.; WAKAE, K.; KITAMURA, K.; AOYAMA, S.; LIU, G.; KOURA, M.; MONJURUL, A.M.; KUKIMOTO, I.; MURAMATSU, M. APOBEC3 deaminases induce hypermutation in human papillomavirus 16 DNA upon beta interferon stimulation. **Journal of virology**, 88:1308, 2014.

WANG, Y.; WANG, M.; DJEKIDEL, M.N.; CHEN, H.; LIU, D.; ALT, F.W.; ZHANG, Y. eccDNAs are apoptotic products with high innate immunostimulatory activity. **Nature**, <https://doi.org/10.1038/s41586-021-04009-w>, 2021.

WATASHI, K.; LIANG, G.; IWAMOTO, M et al. Interleukin-1 and tumor necrosis factor-alpha trigger restriction of hepatitis B virus infection via a cytidine deaminase activation-induced cytidine deaminase (AID). **The Journal of biological chemistry**, 288:31715, 2013.

WEBER, M.; SCHÜBELER, D. Genomic patterns of DNA methylation: targets and function of an epigenetic mark. **Curr Opin Cell Biol**, v. 19, n. 3, p. 273-280, 2007.

WRIGHT, G.; MORGAN, M. Y. Alcohol and tobacco misuse: reducing aerodigestive cancer risk. **World J Hepatol**, 5(8):452-457, 2013.

YANG, W.F.; QIN, N.; SONG, X.; JIANG, C.; LI, T.; JI, P.; LI, Y.; DING, D.; WANG, C.; DAI, J.; JIN, G.; CHEN, T.W.; CHANG, Y.S.; OUYANG, D.Q.; LIAO, G.Q.; HU, Z.; CHANG, K.P.; SU, Y.X.; MA, H. Genomic Signature of Mismatch Repair Deficiency in Areca Nut-Related Oral Cancer. **J Dent Res**, 99(11):1252-1261, 2020.

ZHANG, L.F.; HEMMINKI, K.; SZYFTER, K.; SZYFTER, W.; SÖDERKVIST, P. p53 mutations in larynx cancer. **Carcinogenesis**, 15(12):2949-51, 1994.

ZHANG, L. et al. Genomic analyses reveal mutational signatures and frequently altered genes in esophageal squamous cell carcinoma. **Am J Hum Genet**,96(4):597-611, 2015.

ZHANG, X. et al. Dissecting esophageal squamous-cell carcinoma ecosystem by single-cell transcriptomic analysis. **Nat Commun**, 12(1):5291, 2021.

ZHAO, X et al. Poly I:C-Induced Tumor Cell Apoptosis Mediated by Pattern-Recognition Receptors. **Cancer Biother Radiopharm**, 27(9):530-4, 2012.

ZHENG, Y.; JOYCE, B.T.; LIU, L.; ZHANG, Z.; KIBBE, W.A.; ZHANG, W.; HOU, L. Prediction of genome-wide DNA methylation in repetitive elements. **Nucleic Acids Res**, 45(15):8697-8711, 2017.

ZHU, J.; CHEN, S.; ZHANG, F.; WANG, L. Cell-Free eccDNAs: A New Type of Nucleic Acid Component for Liquid Biopsy? **Mol Diagn Ther**, 22(5):515-522, 2018.

APÊNDICE A - Termo de consentimento informado - INCA



MINISTÉRIO DA SAÚDE

Instituto Nacional de Câncer

Comitê de Ética em Pesquisa-INCA

PROT ##### – Aprovado __/__/2011

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO 1

Projeto de Pesquisa “Marcadores Moleculares de Tumores Secundários de Esôfago em Pacientes com Carcinoma de Células Escamosas em Cabeça e Pescoço: Uso Potencial no Diagnóstico Precoce e na Terapêutica de Rotina” (CEE)

Nome do Voluntário: _____

Você é um paciente matriculado no INCA com diagnóstico confirmado de câncer de esôfago e, em breve, será submetido a um tratamento que poderá ser cirúrgico ou não (Quimioterapia/Radioterapia). Para seu estudo e planejamento terapêutico será necessária a realização de uma Endoscopia Digestiva Alta (exame para avaliar características do seu tumor como tamanho, localização, etc.). Por isso está sendo convidado a participar de um projeto de pesquisa que envolve a coleta de uma biópsia do seu tecido tumoral, de tecido normal próximo ao tumor (quando possível) e uma amostra do seu sangue. Esta biópsia poderá ser realizada no momento em que for feita sua EDA e/ou no momento da cirurgia (caso seja indicado o tratamento cirúrgico). Para que você possa decidir se quer participar ou não deste projeto, precisa conhecer seus benefícios, riscos e consequências. Após receber todas as informações abaixo você poderá fornecer seu consentimento por escrito, caso queira participar.



MINISTÉRIO DA SAÚDE

Instituto Nacional de Câncer

Comitê de Ética em Pesquisa-INCA

PROT #### – Aprovado __/__/2011

PROPÓSITO DO PROJETO DE PESQUISA

Este projeto de pesquisa é um esforço conjunto entre pesquisadores para utilizar os recentes avanços na medicina na tentativa de diagnosticar precocemente casos de câncer de esôfago que surgem como segundos tumores primários (um segundo tumor que ocorre em alguns pacientes que já tiveram câncer de cabeça e pescoço). Tanto os tumores de esôfago quanto os de cabeça e pescoço são muito comuns em nosso país e frequentemente causam a morte desses indivíduos. Para que você possa decidir se quer participar ou não deste projeto de pesquisa, precisa conhecer seus benefícios, riscos e implicações.

OBJETIVOS DO PROJETO DE PESQUISA

Os objetivos deste projeto são a comparação das alterações moleculares entre câncer de cabeça e pescoço e câncer de esôfago buscando descobrir marcadores moleculares com maior eficiência para o diagnóstico precoce do câncer de esôfago ou identificar os indivíduos que possuem um alto risco de desenvolver um segundo tumor.

PROCEDIMENTOS DO PROJETO DE PESQUISA

Se concordar em participar deste projeto será necessário coletar amostras do tumor, e quando possível, do tecido normal próximo à lesão. Esta coleta poderá ser realizada na endoscopia digestiva alta e/ou durante o procedimento cirúrgico. Este procedimento não comprometerá o diagnóstico e não irá interferir no tratamento. Além disto, será coletada uma pequena amostra de sangue, o equivalente a duas colheres de sopa, para futuras pesquisas. Para



MINISTÉRIO DA SAÚDE
Instituto Nacional de Câncer
Comitê de Ética em Pesquisa-INCA
PROT ##### – Aprovado __/__/2011

realizar a endoscopia, todo paciente precisa que seja puncionada uma veia do braço com uma agulha para a aplicação de soro e medicamentos. Nesse momento aproveitaremos a punção para a coleta do sangue.

Ao concordar em participar deste projeto, os pesquisadores participantes consultarão seus registros do prontuário e poderão utilizar para pesquisas parte do material coletado que foi armazenado.

RISCOS

MÉTODOS ALTERNATIVOS

Caso você não deseje participar deste projeto de pesquisa, basta que você não assine este Termo de Consentimento, e nenhuma amostra será utilizada nem os seus registros médicos serão consultados.

RISCOS

O seu tratamento será exatamente o mesmo caso você participe ou não deste projeto de pesquisa. Não haverá risco adicional para a obtenção de amostras para o presente projeto.

BENEFÍCIOS

Não haverá benefícios imediatos a você, porém a sua participação nesse projeto de pesquisa poderá contribuir para o entendimento de casos como o seu no futuro. Este projeto



MINISTÉRIO DA SAÚDE
Instituto Nacional de Câncer
Comitê de Ética em Pesquisa-INCA
PROT ##### – Aprovado __/__/2011

poderá, no entanto, resultar em benefícios para outros pacientes que terão o diagnóstico de câncer de esôfago e que irão fazer o tratamento.

ACOMPANHAMENTO, ASSISTÊNCIA E RESPONSÁVEIS

Durante o projeto você terá o acompanhamento normal junto ao seu médico do INCA que é responsável pela sua assistência. Além disso, a Dr^a Izabela Costa da Cirurgia de Cabeça e Pescoço (tel: 3207-6782), o Dr. Eduardo Pinto, da cirurgia abdominopélvica (tel: 3207-1161) e o pesquisador responsável pelo projeto, Dr. Luis Felipe Ribeiro Pinto (tel: 3207-6560), estarão disponíveis para esclarecer eventuais dúvidas.

CARÁTER CONFIDENCIAL DOS REGISTROS

Além da equipe de saúde que cuidará de você, seus registros médicos poderão ser consultados pelo Comitê de Ética em Pesquisa e equipe de pesquisadores envolvidos. As amostras de tecido serão processadas de tal forma que sua privacidade e identidade sejam preservadas. Seu nome não será revelado ainda que as informações de seu registro médico sejam utilizadas para propósitos educativos ou de publicação, que ocorrerão independentemente dos resultados obtidos.



MINISTÉRIO DA SAÚDE

Instituto Nacional de Câncer

Comitê de Ética em Pesquisa-INCA

PROT #### – Aprovado __/__/2011

TRATAMENTO MÉDICO EM CASO DE DANOS

Caso haja a necessidade de atendimento médico, durante ou após este projeto, este ficará a cargo da instituição. Seu tratamento e acompanhamento médico independem de sua participação neste projeto de pesquisa.

CUSTOS

Se você concordar com o uso das amostras biológicas conforme escrito acima, é necessário esclarecer que você não terá qualquer custo ou forma de pagamento pela participação no projeto de pesquisa. A participação no projeto é voluntária e você não sofrerá nenhuma penalidade caso não autorize a sua participação. Todo o seu tratamento e acompanhamento médico serão os mesmos, independente de sua decisão de autorizar ou não a participação no projeto.

BASES DA PARTICIPAÇÃO

É importante que você saiba que a sua participação neste projeto de pesquisa é completamente voluntária e que você pode recusar-se a participar ou interromper sua participação a qualquer momento sem penalidades ou perda de benefícios aos quais você tem

direito. Em caso de desistência, a equipe assistente deve ser comunicada e a coleta de amostras para os exames relativos ao estudo será imediatamente interrompida.

GARANTIA DE ESCLARECIMENTOS

Nós estimulamos a você ou seus familiares a fazer perguntas a qualquer momento durante o andamento desse projeto de pesquisa. Neste caso, por favor, ligue para o Dr. Luis Felipe Ribeiro Pinto no telefone (21) 3207-6560. Se você tiver perguntas com relação a seus direitos como participante desse projeto de pesquisa, também pode contar com uma terceira pessoa imparcial, o Comitê de Ética em Pesquisa do INCA, Rua André Cavalcanti 37, telefone (21) 3207-6551. Email: CEP@inca.gov.br



MINISTÉRIO DA SAÚDE

Instituto Nacional de Câncer

Comitê de Ética em Pesquisa-INCA

PROT #### – Aprovado __/__/2011

DECLARAÇÃO DE CONSENTIMENTO E ASSINATURA

Li as informações acima e entendi o propósito deste projeto de pesquisa assim como os benefícios e riscos potenciais da participação no mesmo. Tive a oportunidade de fazer perguntas e todas foram respondidas. Eu, por intermédio deste, dou livremente meu consentimento para participar neste projeto de pesquisa e:

- () Expresso que desejo ser reconvocado a oferecer um novo consentimento para cada projeto que venha a utilizar agora, ou no futuro, material biológico coletado pelo presente projeto de pesquisa, indicando as pessoas abaixo relacionadas como representantes em caso de incapacidade minha para prover tal consentimento:

_____ (NOME EM LETRA FORMA)

_____ (GRAU DE PARENTESCO)

_____ (TELEFONE)

_____ (NOME EM LETRA FORMA)

_____ (GRAU DE PARENTESCO)

_____ (TELEFONE)



MINISTÉRIO DA SAÚDE

Instituto Nacional de Câncer

Comitê de Ética em Pesquisa-INCA

PROT ##### – Aprovado __/__/2011

() Não gostaria de ser reconvocato a oferecer um novo consentimento para cada projeto de pesquisa que venha a utilizar agora, ou no futuro, material biológico coletado pelo presente projeto de pesquisa.

Entendo que não receberei compensação monetária por minha participação neste projeto.

Eu recebi uma cópia assinada deste formulário de consentimento.

_____ / ____ / ____
(Assinatura do Paciente) dia mês ano

(Nome do Paciente – letra de forma)

_____ / ____ / ____
(Assinatura de Testemunha, se necessário) dia mês ano

Eu, abaixo assinado, expliquei completamente os detalhes relevantes deste projeto ao paciente indicado acima e/ou pessoa autorizada para consentir pelo paciente.

_____ / ____ / ____
(Assinatura da pessoa que obteve o consentimento) dia mês ano

APÊNDICE B – Prontuários – INCA

REANÁLISE DE PRONTUÁRIOS – Câncer de Esôfago – INCA

CEE

ADE

Preenchedor: _____ Data: ____/____/____ Matrícula do Pcte: _____

Nome: _____		Data Matrícula: ____/____/____		Gênero: Masculino <input type="checkbox"/> Feminino <input type="checkbox"/>	
Nacionalidade: Brasil <input type="checkbox"/> Outro <input type="checkbox"/>		Naturalidade (UF): _____		Residência: Cidade: _____ UF: _____	
Cor da pele: Branca <input type="checkbox"/> Não-branca <input type="checkbox"/>		Data Nascimento: ____/____/____		Idade: _____ Profissão: _____	
Escolaridade: Analfabeto <input type="checkbox"/> 1º Grau Incompleto <input type="checkbox"/> 1º Grau Completo <input type="checkbox"/> 2º Grau Incompleto <input type="checkbox"/> 2º Grau Completo <input type="checkbox"/> Superior Incompleto <input type="checkbox"/> Superior Completo <input type="checkbox"/> Pós-Graduação <input type="checkbox"/> Não Informa <input type="checkbox"/>		Estado Civil: Casado <input type="checkbox"/> União Estável <input type="checkbox"/> Divorciado <input type="checkbox"/> Solteiro <input type="checkbox"/> Hist Familiar de Câncer de Esôfago: Não <input type="checkbox"/> Sim <input type="checkbox"/>			

HÁBITOS DO PACIENTE						
Tabagismo: Não <input type="checkbox"/> Sim <input type="checkbox"/> Ex <input type="checkbox"/> Não Informa <input type="checkbox"/>			Refluxo: Não <input type="checkbox"/> Sim <input type="checkbox"/> Não Informa <input type="checkbox"/>			
Etilismo: Não <input type="checkbox"/> Sim <input type="checkbox"/> Ex <input type="checkbox"/> Não Informa <input type="checkbox"/>			Barret: Não <input type="checkbox"/> Sim <input type="checkbox"/> Não Informa <input type="checkbox"/>			
Status Nutricional: Desnutrição Grave <input type="checkbox"/> Desnutrição Moderada <input type="checkbox"/> Desnutrição Leve <input type="checkbox"/> Eutrófico <input type="checkbox"/> Sobrepeso <input type="checkbox"/> Obeso <input type="checkbox"/> Não Informa <input type="checkbox"/>						

INFORMAÇÕES SOBRE A DOENÇA						
Data do Diagnóstico (Lauda DIPAT): ____/____/____			Esôfago é STP: Não <input type="checkbox"/> Sim <input type="checkbox"/>		CID PTP: _____	
Tto do PTP: Nenhum <input type="checkbox"/> Cirurgia <input type="checkbox"/> Rxt <input type="checkbox"/> QT <input type="checkbox"/> Rxt+QT <input type="checkbox"/> Rxt+BQT <input type="checkbox"/> Rxt+QT+BQT <input type="checkbox"/> Outro <input type="checkbox"/> Outro: _____						
Localização: C15.0 Esôfago Cervica; <input type="checkbox"/> C15.1 Esôfago Torácico <input type="checkbox"/> C15.2 Esôfago Abdominal <input type="checkbox"/> C15.3 Esôfago Torácico 1/3 Sup <input type="checkbox"/> C15.4 Esôfago Torácico 1/3 Med <input type="checkbox"/> C15.5 Esôfago Torácico 1/3 Inf <input type="checkbox"/> C15.8 Lesão Invasiva <input type="checkbox"/> C15.9 Não especificado <input type="checkbox"/>						
Grau de Diferenciação: I (Bem dif) <input type="checkbox"/> II (Mod dif) <input type="checkbox"/> III (Pouco dif) <input type="checkbox"/>			Uso de Sonda NasoEntérica: Não <input type="checkbox"/> Sim <input type="checkbox"/> Não Informa <input type="checkbox"/>			
TNM (Abdome): cT: 1a <input type="checkbox"/> 1b <input type="checkbox"/> 2 <input type="checkbox"/> 3 <input type="checkbox"/> 4a <input type="checkbox"/> 4b <input type="checkbox"/> is <input type="checkbox"/> X <input type="checkbox"/> cN: 0 <input type="checkbox"/> 1 <input type="checkbox"/> 2 <input type="checkbox"/> 3 <input type="checkbox"/> X <input type="checkbox"/> cM: 0 <input type="checkbox"/> 1 <input type="checkbox"/> X <input type="checkbox"/>						
Estadiamento (Abdome): IA <input type="checkbox"/> IB <input type="checkbox"/> IIA <input type="checkbox"/> IIB <input type="checkbox"/> IIIA <input type="checkbox"/> IIIB <input type="checkbox"/> IIIC <input type="checkbox"/> IV <input type="checkbox"/>						

TRATAMENTO						
Tto Inicial: Nenhum <input type="checkbox"/> Cirurgia <input type="checkbox"/> Rxt <input type="checkbox"/> Rxt+QT <input type="checkbox"/> Rxt+BQT <input type="checkbox"/> Rxt+QT+BQT <input type="checkbox"/> Suporte <input type="checkbox"/>						
Local Tto inicial: For a do INCA <input type="checkbox"/> No INCA <input type="checkbox"/> Início Tto inicial: ____/____/____ Fim do Tto inicial: ____/____/____						
Doença Residual após Tto inicial: Não <input type="checkbox"/> Sim <input type="checkbox"/>						
Tto Adjuvante: Nenhum <input type="checkbox"/> Radioterapia <input type="checkbox"/> Rxt+QT <input type="checkbox"/> Tto Resgate: Cirurgia <input type="checkbox"/> Suporte <input type="checkbox"/> Nenhum <input type="checkbox"/> Tto Resgate: ____/____/____						

RXT: Curativa <input type="checkbox"/> Neoadjuvante <input type="checkbox"/> Paliativa <input type="checkbox"/>	Completo? Não <input type="checkbox"/> Sim <input type="checkbox"/>
QT: Curativa <input type="checkbox"/> Neoadjuvante <input type="checkbox"/> Paliativa <input type="checkbox"/>	Completo? Não <input type="checkbox"/> Sim <input type="checkbox"/>

INFORMAÇÕES SOBRE CIRURGIA (ESOFAGECTOMIA)										
Cirurgia:	VLP <input type="checkbox"/>	Transhiatal <input type="checkbox"/>	Transtorácica <input type="checkbox"/>	Linfadenectomia:	2 Campos <input type="checkbox"/>	3 Campos <input type="checkbox"/>	Não Realizada <input type="checkbox"/>	Não Informa <input type="checkbox"/>		
Reconstrução Trato Alimentar:	Cólon Direito <input type="checkbox"/>		Cólon Esquerdo <input type="checkbox"/>		Cólon Transverso <input type="checkbox"/>		Estômago <input type="checkbox"/>			
Margens Livres:	Não <input type="checkbox"/>	Sim <input type="checkbox"/>	Sim, após ampliação <input type="checkbox"/>		Não avaliado <input type="checkbox"/>					
TNM (Abdome):	cT:	1a <input type="checkbox"/>	1b <input type="checkbox"/>	2 <input type="checkbox"/>	3 <input type="checkbox"/>	4a <input type="checkbox"/>	4b <input type="checkbox"/>	is <input type="checkbox"/>	X <input type="checkbox"/>	
	cN:	0 <input type="checkbox"/>	1 <input type="checkbox"/>	2 <input type="checkbox"/>	3 <input type="checkbox"/>	X <input type="checkbox"/>	cM:	0 <input type="checkbox"/>	1 <input type="checkbox"/>	X <input type="checkbox"/>
Estadiamento (Abdome):	IA <input type="checkbox"/>	IB <input type="checkbox"/>	IIA <input type="checkbox"/>	IIB <input type="checkbox"/>	IIIA <input type="checkbox"/>	IIIB <input type="checkbox"/>	IIIC <input type="checkbox"/>	IV <input type="checkbox"/>		

ACOMPANHAMENTO										
Remissão Completa:	Não <input type="checkbox"/>	Sim <input type="checkbox"/>	Recidiva Presente:	Não <input type="checkbox"/>	Sim <input type="checkbox"/>	Data Recidiva:	____/____/____			
Local Recidiva:	Local <input type="checkbox"/>	LNFD <input type="checkbox"/>	Distância <input type="checkbox"/>	Última Consulta:	____/____/____		C/ Evid. de doença <input type="checkbox"/>	S/ Evid. de doença <input type="checkbox"/>		
Data HCIV:	____/____/____		Data Óbito:	____/____/____		Causa Óbito:	_____			
Apresentou STP:	Não <input type="checkbox"/>	Sim <input type="checkbox"/>	Deteção STP:	____/____/____		Sincrônico:	Não <input type="checkbox"/>	Sim <input type="checkbox"/>	CID STP:	_____
Metástase a distância:	Não <input type="checkbox"/>	Sim <input type="checkbox"/>	Sítio de Metástase:	_____						

ANEXO – Carta de aprovação do projeto sob o número 116/11 pelo Comitê de Ética em Pesquisa do INCA



Memo 440/11-CEP-INCA

Rio de Janeiro, 5 de dezembro de 2011.

A(o): Dr. Luis Felipe Ribeiro Pinto
Pesquisador(a) Principal

Registro CEP nº 116/11 (Este nº. deve ser citado nas correspondências referentes a este estudo)
CAAE – 0086.0.007.000-11
Título do Estudo – Marcadores Moleculares de Tumores Secundários de Esôfago em Pacientes com Carcinoma de Células Escamosas em Cabeça e Pescoço: Uso Potencial no Diagnóstico Precoce e na Terapia de Rotina

Prezado(a) Pesquisador(a),

Informo que o Comitê de Ética em Pesquisa do INCA **aprovou** após reanálise o estudo intitulado: **Marcadores Moleculares de Tumores Secundários de Esôfago em Pacientes com Carcinoma de Células Escamosas em Cabeça e Pescoço: Uso Potencial no Diagnóstico Precoce e na Terapia de Rotina versão 2**, bem como seu TCLE para estudo prospectivo **CECP versão 3** e TCLE para estudo retrospectivo **CECP e CEE versão 3**, em 30 de novembro de 2011.

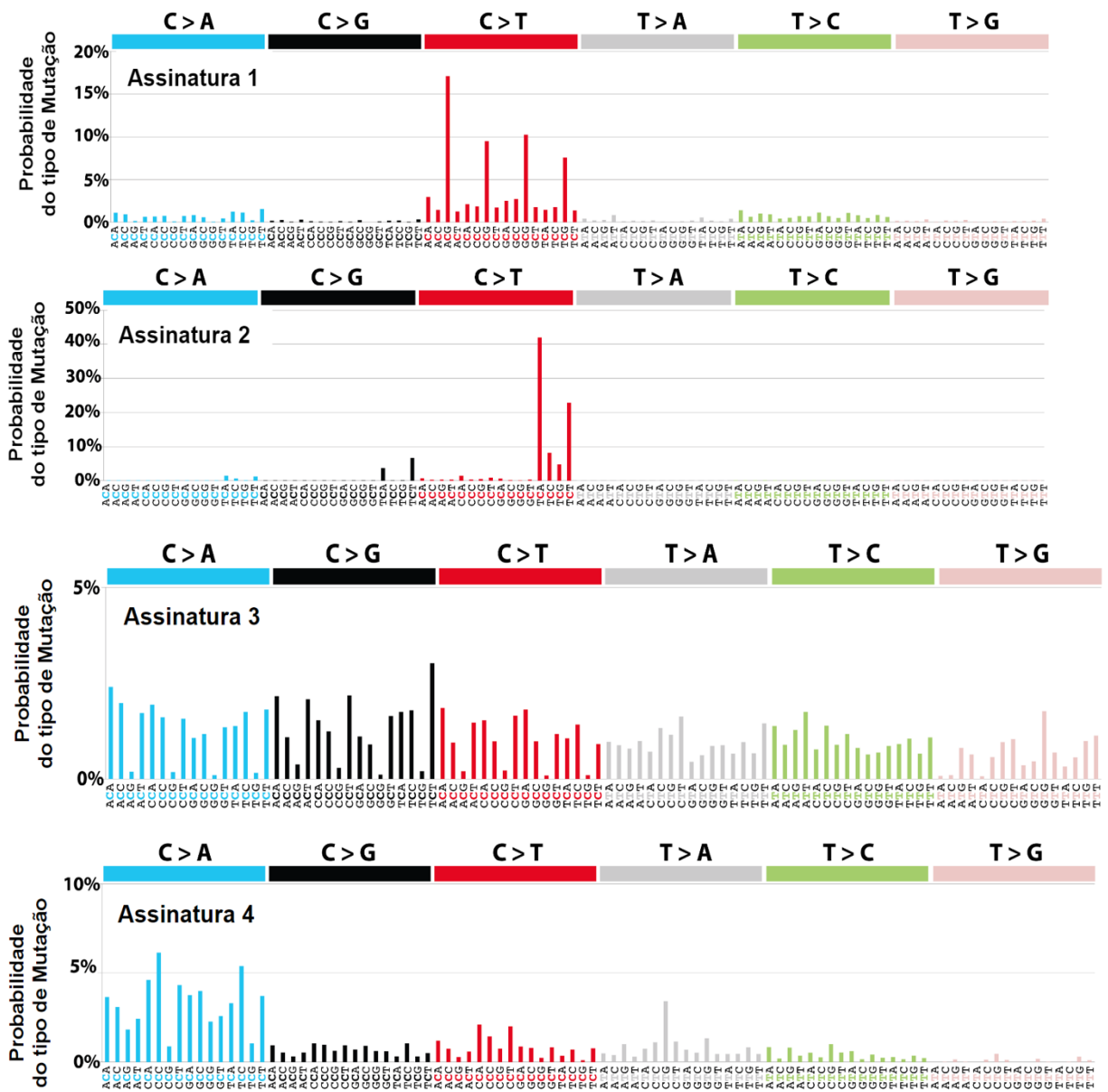
Ressalto o(a) pesquisador(a) responsável deverá apresentar relatórios semestrais a respeito do seu estudo.

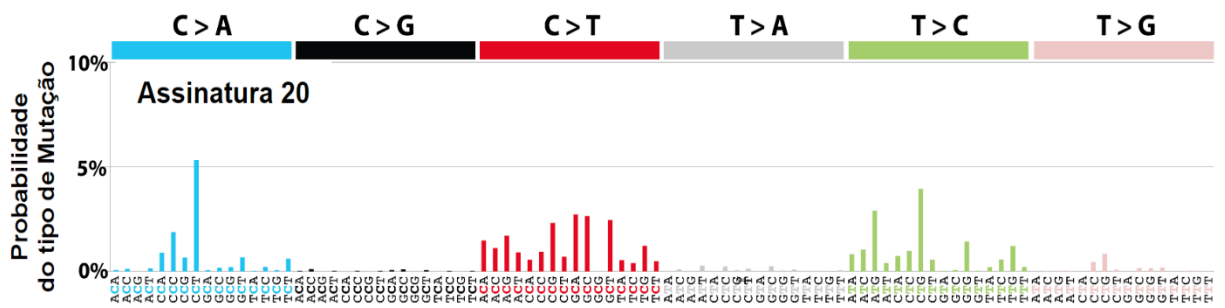
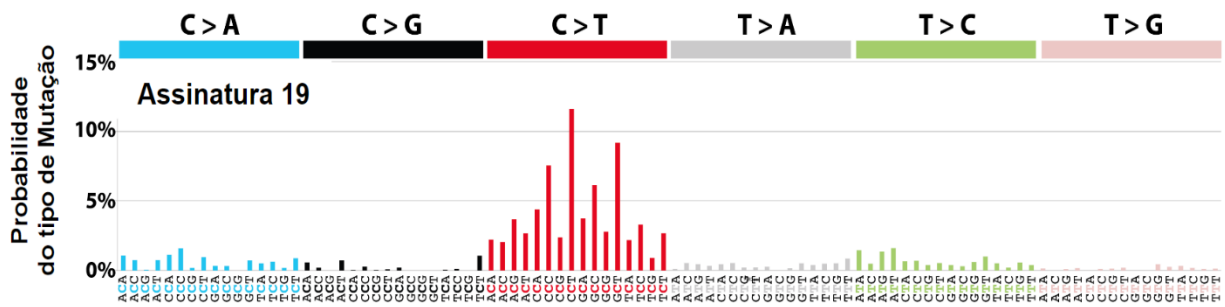
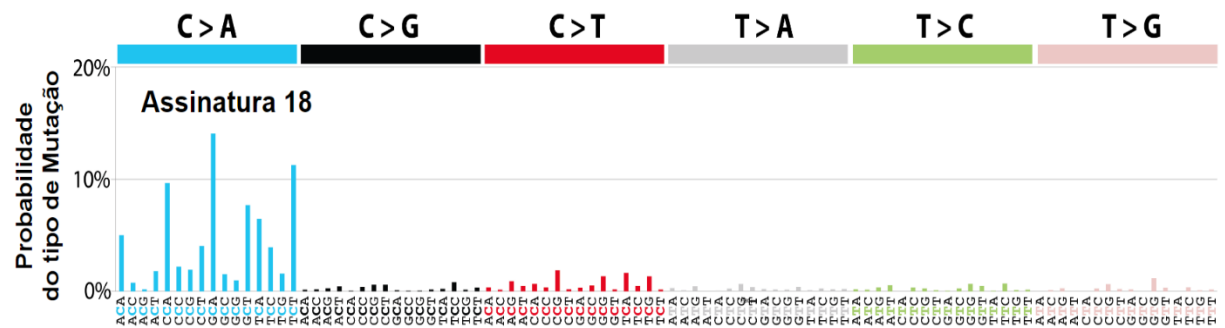
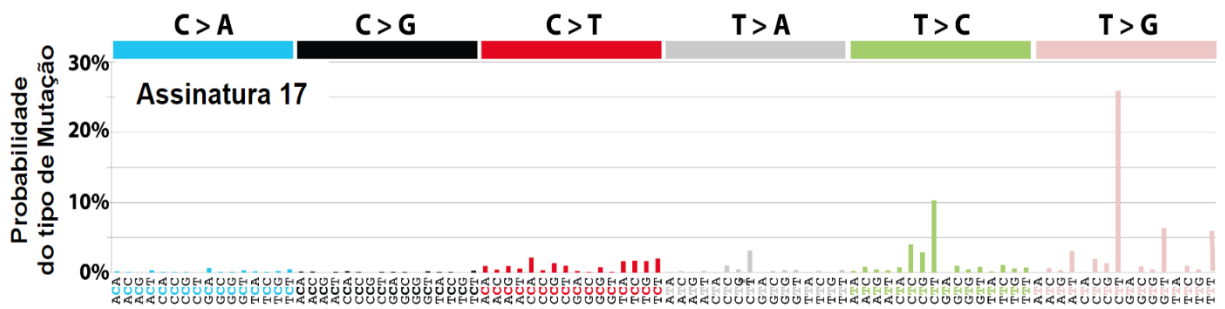
Atenciosamente,

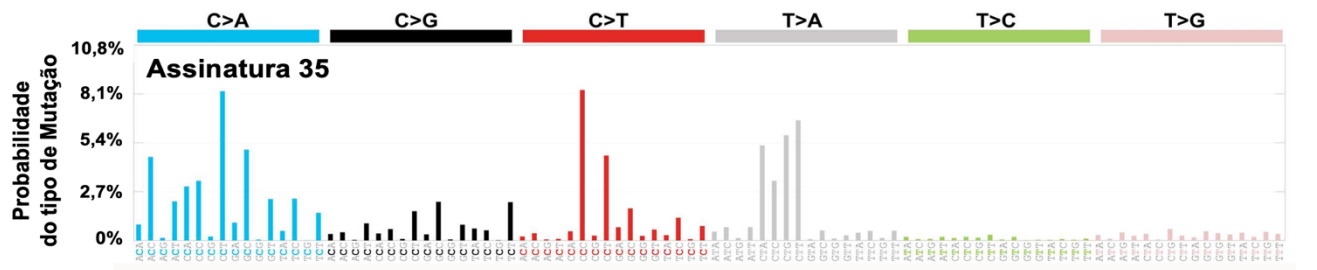
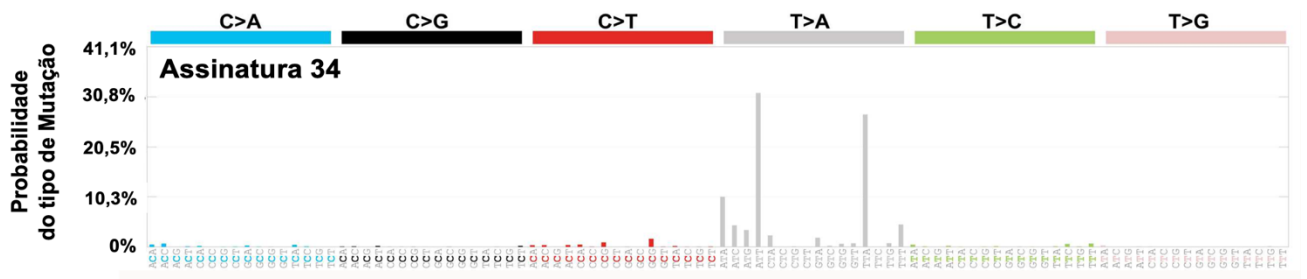
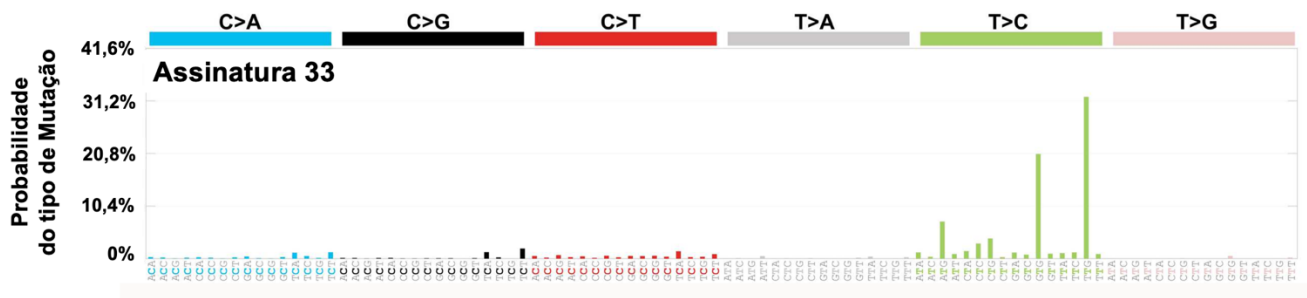
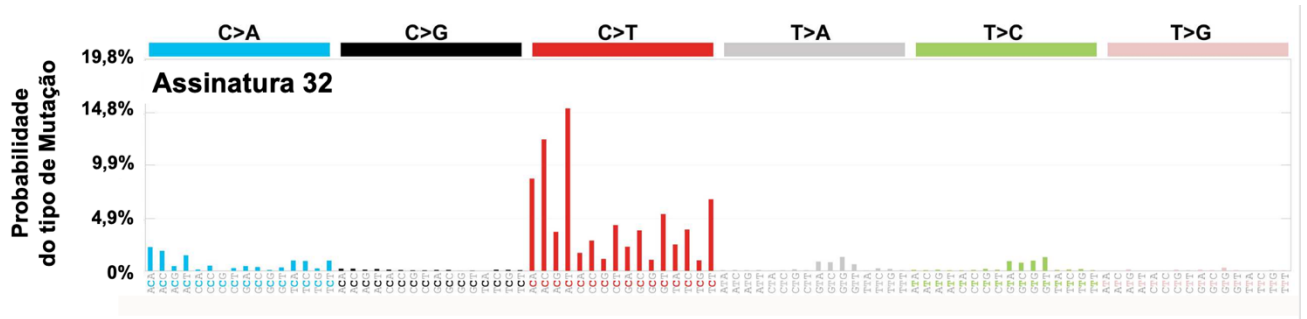

Dr. Carlos Henrique D. Silva
Coordenador do Comitê de Ética em Pesquisa
CEP-INCA

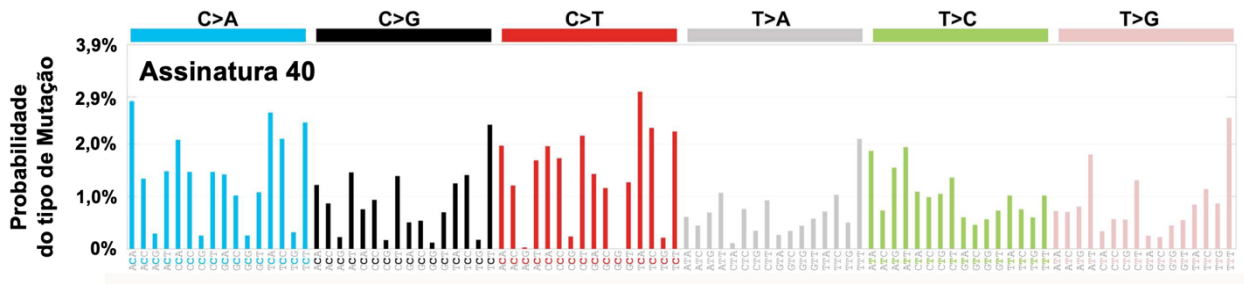
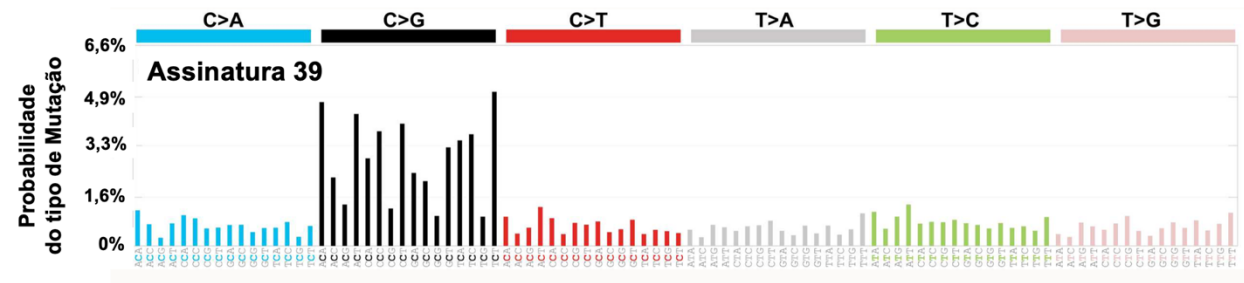
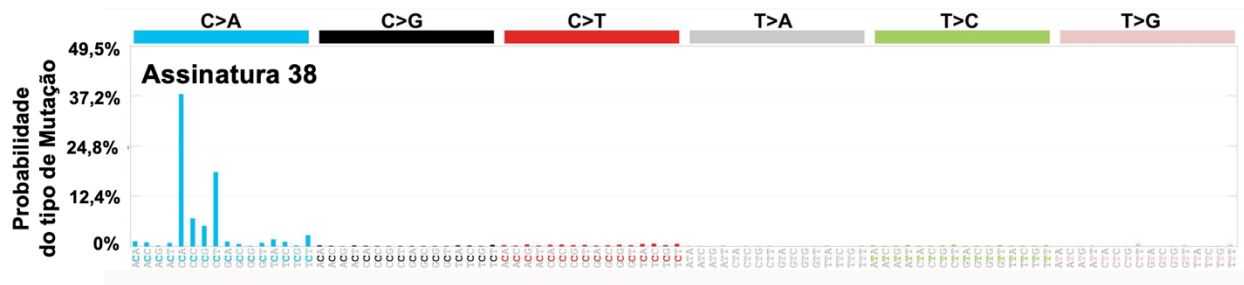
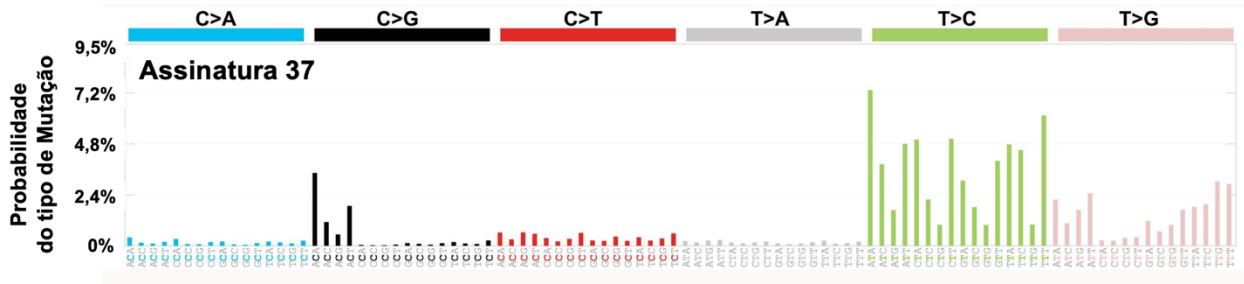
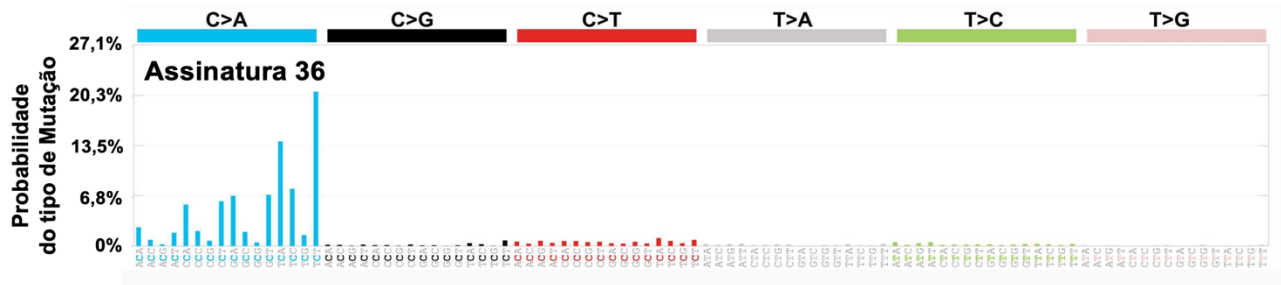
C/c – Dr. Luiz Augusto Maltoni – Direção Geral

FIGURA SUPLEMENTAR 1 - As diferentes assinaturas mutacionais presentes em tumores. Neste gráfico, o eixo X indica o tipo de mutação observada com sua frequência sendo representada em barras coloridas respectivas, de acordo com o contexto. Já o eixo Y representa a probabilidade do tipo de mutação observada. Fonte: COSMIC; Adaptado de ALEXANDROV et al., 2013; Adaptado de ALEXANDROV et al., 2020.









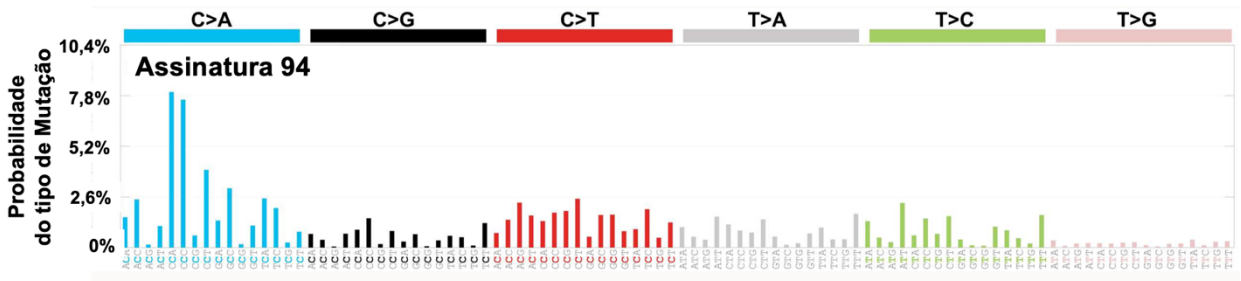


Tabela Suplementar 1: *Nível de metilação das APOBECs entre as mucosas adjacente e tumoral de pacientes com CEE, CEL, CECO, CEOF HPV⁻ e CEOF HPV⁺ do INCA e correlação desse nível com a expressão dessas enzimas nesses tumores no banco de dados do TCGA.*

Sondas	Posição	TIPO DE CÂNCER				
		CEE	CEL	CECO	CEOF HPV ⁻	CEOF HPV ⁺
cg27562311	TSS1500	$\Delta\beta$:-0,004 Rho:NA	$\Delta\beta$:- 0,033 Rho:NA	$\Delta\beta$:0,012 Rho:NA	$\Delta\beta$:- 0,0947 Rho:NA	$\Delta\beta$:- 0,2164 Rho:NA
cg10465202	TSS1500	$\Delta\beta$:-0,03 Rho:NA	$\Delta\beta$:- 0,069 Rho:NA	$\Delta\beta$:-0,03 Rho:NA	$\Delta\beta$:0,016 Rho:NA	$\Delta\beta$:- 0,265 Rho:NA
cg02926127	Corpo	$\Delta\beta$:- 0,02 Rho:0,0397	$\Delta\beta$:-0,017 Rho:0,145	$\Delta\beta$:-0,0037 Rho: 0,04	$\Delta\beta$:0,0058 Rho:-0,35	$\Delta\beta$:-0,001 Rho:-0,20
cg06193723	Corpo	$\Delta\beta$:- 0,0297 Rho:NA	$\Delta\beta$:NA Rho:NA	$\Delta\beta$:-0,017 Rho: NA	$\Delta\beta$: 0,039 Rho:NA	$\Delta\beta$:-0,0028 Rho:NA
cg07999668	Corpo	$\Delta\beta$:-0,037 Rho:0,0703	$\Delta\beta$:- 0,084 Rho:0,00148	$\Delta\beta$:- 0,123 Rho:-0,037	$\Delta\beta$:0,06 Rho:-0,095	$\Delta\beta$:-0,272 Rho:0,036
AICDA						
cg10989381	Corpo	$\Delta\beta$:NA Rho: 0,117	$\Delta\beta$:-0,047 Rho: 0,133	$\Delta\beta$:NA Rho: 0,1958	$\Delta\beta$:- 0,213 Rho:0,03	$\Delta\beta$:-0,026 Rho:-0,06
cg15494122	Corpo	$\Delta\beta$:-0,081 Rho:0,0964	$\Delta\beta$:- 0,0824 Rho:0,0964	$\Delta\beta$:-0,04 Rho:0,08	$\Delta\beta$:- 0,1059 Rho:0,01	$\Delta\beta$:- 0,2146 Rho: 0,0008
cg17070108	Corpo	$\Delta\beta$:-0,006 Rho:0,143	$\Delta\beta$:-0,01 Rho: 0,072	$\Delta\beta$:-0,003 Rho:-0,02	$\Delta\beta$:- 0,0165 Rho:-0,059	$\Delta\beta$:- 0,0276 Rho:0,23
cg18775247	Corpo	$\Delta\beta$:-0,052 Rho:NA	$\Delta\beta$:- 0,108 Rho:NA	$\Delta\beta$:-0,033 Rho:NA	$\Delta\beta$:- 0,1362 Rho:NA	$\Delta\beta$:- 0,1419 Rho:NA
cg24412650	Corpo	$\Delta\beta$:0,0008 Rho:0,002	$\Delta\beta$:- 0,02 Rho:0,1694	$\Delta\beta$:0,0029 Rho:-0,1	$\Delta\beta$:- 0,0375 Rho:-0,29	$\Delta\beta$:-0,0227 Rho:0,16
cg00245465	Enhancer 3'UTR	$\Delta\beta$:0,023 Rho:-0,0464	$\Delta\beta$:-0,067 Rho:0,0305	$\Delta\beta$:-0,029 Rho:0,0934	$\Delta\beta$:-0,047 Rho:-0,075	$\Delta\beta$:- 0,29 Rho: 0,306
cg08709276	Enhancer TSS1500	$\Delta\beta$:0,007 Rho:-0,099	$\Delta\beta$:0,004 Rho:-0,007	$\Delta\beta$:0,003 Rho:-0,07	$\Delta\beta$: 0,0457 Rho:0,13	$\Delta\beta$:0,0006 Rho:-0,14
cg10567498	Enhancer TSS1500	$\Delta\beta$: 0,14 Rho: -0,06	$\Delta\beta$: 0,063 Rho: -0,22	$\Delta\beta$:0,035 Rho:-0,09	$\Delta\beta$:-0,082 Rho:-0,21	$\Delta\beta$:-0,0658 Rho:-0,08
APOBEC2						
cg24654907	Enhancer TSS1500	$\Delta\beta$:-0,003 Rho:NA	$\Delta\beta$:-0,01 Rho:NA	$\Delta\beta$:-0,0009 Rho:NA	$\Delta\beta$:- 0,025 Rho:NA	$\Delta\beta$:-0,0096 Rho:NA
cg22682201	Enhancer TSS200	$\Delta\beta$: 0,38 Rho:0,178	$\Delta\beta$: 0,14 Rho: -0,106	$\Delta\beta$:0,08 Rho:-0,055	$\Delta\beta$:0,087 Rho:-0,21	$\Delta\beta$:-0,0571 Rho:-0,21

APOBEC2

cg17548735	<i>Enhancer</i> 5'UTR/ 1 éxon	$\Delta\beta$: 0,42 Rho: -0,096	$\Delta\beta$: 0,217 Rho: -0,131	$\Delta\beta$:0,009 Rho:-0,06	$\Delta\beta$: -0,09 Rho:-0,29	$\Delta\beta$: -0,33 Rho: -0,57
cg22375610	<i>Enhancer</i> 1 éxon	$\Delta\beta$: 0,24 Rho: -0,084	$\Delta\beta$: 0,10 Rho: -0,09	$\Delta\beta$:0,04 Rho:-0,006	$\Delta\beta$: -0,107 Rho:-0,20	$\Delta\beta$: -0,27 Rho: -0,47
cg20172737	<i>Enhancer</i> Corpo	$\Delta\beta$: 0,38 Rho: 0,0155	$\Delta\beta$:0,06 Rho: -0,096	$\Delta\beta$: -0,002 Rho:0,099	$\Delta\beta$: -0,13 Rho:-0,3	$\Delta\beta$: -0,4618 Rho:-0,126
cg12253732	3'UTR	$\Delta\beta$:0,01 Rho: 0,193	$\Delta\beta$: -0,02 Rho: -0,066	$\Delta\beta$:0,01 Rho: -0,2	$\Delta\beta$: -0,022 Rho: -0,45	$\Delta\beta$:0,0031 Rho:0,09

TIPO DE CÂNCER

Sondas	Posição	TIPO DE CÂNCER		TIPO DE CÂNCER			
		CEE	CEL	CECO	CEOF	HPV ⁻	HPV ⁺
<i>APOBEC3A</i>	cg18746774	TSS1500	$\Delta\beta$:-0,01 Rho:-0,0762	$\Delta\beta$:- 0,03 Rho:-0,0058	$\Delta\beta$:-0,004 Rho:0,107	$\Delta\beta$:0,044 Rho:0,39	$\Delta\beta$:-0,0175 Rho:0,035
	cg02869653	TSS1500	$\Delta\beta$:-0,006 Rho:0,08	$\Delta\beta$:-0,01 Rho:-0,028	$\Delta\beta$:0,001 Rho: 0,19	$\Delta\beta$:- 0,018 Rho:0,2	$\Delta\beta$:-0,0059 Rho:0,16
	cg01397001	TSS1500	$\Delta\beta$:-0,028 Rho:-0,0926	$\Delta\beta$:- 0,01 Rho:-0,08	$\Delta\beta$:-0,005 Rho:0,12	$\Delta\beta$:- 0,06 Rho:0,14	$\Delta\beta$:- 0,036 Rho:-0,07
	cg02587894	TSS1500	$\Delta\beta$:-0,009 Rho:NA	$\Delta\beta$:-0,034 Rho:NA	$\Delta\beta$:-0,008 Rho:NA	$\Delta\beta$:-0,0004 Rho:NA	$\Delta\beta$:- 0,05 Rho:NA
	cg24360617	TSS200	$\Delta\beta$: 0,16 Rho: -0,198	$\Delta\beta$:-0,06 Rho:- 0,26	$\Delta\beta$:- 0,12 Rho:- 0,14	$\Delta\beta$:0,027 Rho:-0,39	$\Delta\beta$:- 0,09 Rho:- 0,28
	cg27504369	5'UTR/ 1éxon	$\Delta\beta$: 0,14 Rho: -0,1960	$\Delta\beta$:- 0,07 Rho:- 0,19	$\Delta\beta$:- 0,27 Rho:-0,1	$\Delta\beta$:-0,027 Rho:-0,27	$\Delta\beta$:- 0,15 Rho:-0,24
	cg21345826	5'UTR/ 1éxon	$\Delta\beta$: 0,27 Rho: -0,129	$\Delta\beta$:-0,1 Rho:-0,06	$\Delta\beta$:-0,2 Rho:0,1208	$\Delta\beta$:- 0,257 Rho:0,23	$\Delta\beta$:- 0,528 Rho:-0,09
	cg22954818	Corpo	$\Delta\beta$:- 0,1 Rho:-0,1480	$\Delta\beta$:- 0,07 Rho:-0,048	$\Delta\beta$:-0,08 Rho:0,12	$\Delta\beta$:0,068 Rho:0,027	$\Delta\beta$:0,03 Rho:0,025
cg06932384	3'UTR	$\Delta\beta$:-0,01 Rho: -0,00317	$\Delta\beta$:-0,014 Rho:-0,17	$\Delta\beta$:-0,02 Rho: 0,186	$\Delta\beta$:- 0,035 Rho:0,31	$\Delta\beta$:-0,011 Rho:0,12	
<i>APOBEC3B</i>	cg11816043	Enhancer TSS1500	$\Delta\beta$: 0,0864 Rho:-0,183	$\Delta\beta$: 0,07 Rho:- 0,21	$\Delta\beta$:0,019 Rho:0,0046	$\Delta\beta$:-0,068 Rho:0,016	$\Delta\beta$:- 0,12 Rho:- 0,38
	cg16045423	Enhancer TSS1500	$\Delta\beta$: 0,132 Rho:- 0,35	$\Delta\beta$:0,038 Rho:- 0,37	$\Delta\beta$:-0,0638 Rho:-0,095	$\Delta\beta$:-0,02 Rho:0,18	$\Delta\beta$:- 0,15 Rho:- 0,41
	cg14387414	Enhancer TSS1500	$\Delta\beta$:NA Rho:- 0,44	$\Delta\beta$:NA Rho:- 0,29	$\Delta\beta$:NA Rho:-0,02	$\Delta\beta$:NA Rho:-0,08	$\Delta\beta$:NA Rho:-0,19
	cg14194956	Enhancer TSS200	$\Delta\beta$:-0,0037 Rho: -0,16	$\Delta\beta$:-0,0039 Rho:0,12	$\Delta\beta$:-0,0115 Rho:0,04	$\Delta\beta$:0,02 Rho:0,31	$\Delta\beta$:-0,0145 Rho:-0,048
	cg25787886	Enhancer TSS200	$\Delta\beta$:0,0027 Rho:-0,125	$\Delta\beta$:0,0027 Rho:-0,067	$\Delta\beta$:-0,0017 Rho:0,09	$\Delta\beta$: 0,009 Rho:0,11	$\Delta\beta$:-0,011 Rho:0,10
	cg26000393	Enhancer TSS200	$\Delta\beta$:0,00547 Rho:-0,057	$\Delta\beta$:0,004 Rho:0,036	$\Delta\beta$:-0,0034 Rho:-0,006	$\Delta\beta$:-0,002 Rho:0,22	$\Delta\beta$:-0,0033 Rho:-0,0016
	cg21707131	Enhancer TSS200	$\Delta\beta$:NA Rho: -0,5	$\Delta\beta$:NA Rho:- 0,20	$\Delta\beta$:NA Rho:-0,08	$\Delta\beta$:NA Rho:-0,31	$\Delta\beta$:NA Rho:-0,07
	cg01089751	Enhancer 5'UTR/ 1 éxon	$\Delta\beta$:NA Rho:0,0625	$\Delta\beta$:NA Rho:0,038	$\Delta\beta$:NA Rho:0,07	$\Delta\beta$:NA Rho:0,2	$\Delta\beta$:NA Rho:-0,16
cg06837067	Corpo	$\Delta\beta$:NA Rho: -0,392	$\Delta\beta$:NA Rho:-0,1	$\Delta\beta$:NA Rho:- 0,17	$\Delta\beta$:NA Rho:-0,16	$\Delta\beta$:NA Rho:- 0,47	

TIPO DE CÂNCER							
Sondas	Posição	CEE	CEL	CECO	CEOF		
					HPV ⁻	HPV ⁺	
<i>APOBEC3C</i>	cg11413071	TSS1500	$\Delta\beta$:0,0249 Rho:NA	$\Delta\beta$:0,0136 Rho:NA	$\Delta\beta$:-0,0317 Rho:NA	$\Delta\beta$: -0,0874 Rho:NA	$\Delta\beta$:-0,018 Rho:NA
	cg21310636	TSS1500	$\Delta\beta$:0,007 Rho:NA	$\Delta\beta$:0,003 Rho:NA	$\Delta\beta$:-0,0048 Rho:NA	$\Delta\beta$: -0,0269 Rho:NA	$\Delta\beta$:-0,0037 Rho:NA
	cg06523744	TSS200	$\Delta\beta$:0,0034 Rho:-0,032	$\Delta\beta$:-0,007 Rho: 0,26	$\Delta\beta$:-0,018 Rho: 0,17	$\Delta\beta$: 0,033 Rho:0,2	$\Delta\beta$:0,006 Rho:0,054
	cg10316474	TSS200	$\Delta\beta$:0,0119 Rho: -0,32	$\Delta\beta$: -0,059 Rho:-0,09	$\Delta\beta$: -0,046 Rho:-0,02	$\Delta\beta$: 0,0158 Rho:-0,16	$\Delta\beta$:0,0025 Rho:-0,062
	cg04835613	5'UTR/ lêxon	$\Delta\beta$:-0,003 Rho:-0,07	$\Delta\beta$:-0,0004 Rho: 0,198	$\Delta\beta$:-0,0059 Rho:0,088	$\Delta\beta$:0,005 Rho:0,016	$\Delta\beta$:-0,0058 Rho: -0,34
	cg07186138	Corpo	$\Delta\beta$:0,010 Rho: -0,268	$\Delta\beta$: -0,042 Rho:-0,018	$\Delta\beta$: -0,045 Rho:0,048	$\Delta\beta$: 0,0107 Rho:-0,05	$\Delta\beta$:-0,013 Rho:-0,17
	cg27451450	3'UTR	$\Delta\beta$:NA Rho:-0,059	$\Delta\beta$:NA Rho:0,097	$\Delta\beta$:NA Rho:0,049	$\Delta\beta$:NA Rho:0,02	$\Delta\beta$:NA Rho:0,013
<i>APOBEC3D</i>	cg26530502	TSS1500	$\Delta\beta$: 0,151 Rho:NA	$\Delta\beta$: 0,05 Rho:NA	$\Delta\beta$:0,039 Rho:NA	$\Delta\beta$:0,0104 Rho:NA	$\Delta\beta$: -0,115 Rho:NA
	cg02038910	TSS1500	$\Delta\beta$: NA Rho: -0,256	$\Delta\beta$:NA Rho:0,012	$\Delta\beta$:NA Rho:-0,067	$\Delta\beta$:NA Rho:-0,3	$\Delta\beta$:NA Rho:0,03
	cg16066354	<i>Enhancer</i> 5'UTR/ lêxon	$\Delta\beta$: -0,27 Rho: -0,68	$\Delta\beta$:0,235 Rho: -0,2	$\Delta\beta$:-0,1194 Rho: -0,54	$\Delta\beta$: 0,11 Rho: -0,63	$\Delta\beta$: 0,05 Rho: -0,55
	cg10555000	Corpo	$\Delta\beta$:0,0054 Rho:NA	$\Delta\beta$: -0,106 Rho:NA	$\Delta\beta$:-0,047 Rho:NA	$\Delta\beta$: -0,083 Rho:NA	$\Delta\beta$: -0,1519 Rho:NA
	cg08665930	<i>Enhancer</i> Corpo	$\Delta\beta$: -0,098 Rho: -0,51	$\Delta\beta$:0,0549 Rho: -0,195	$\Delta\beta$:-0,076 Rho: -0,55	$\Delta\beta$:0,017 Rho:-0,40	$\Delta\beta$:-0,143 Rho: -0,53
	cg18707706	3'UTR	$\Delta\beta$:NA Rho: 0,342	$\Delta\beta$:NA Rho: 0,498	$\Delta\beta$:NA Rho: 0,46	$\Delta\beta$:NA Rho: 0,53	$\Delta\beta$:NA Rho: 0,43

		TIPO DE CÂNCER					
		CEE	CEL	CECO	CEOF		
Sondas	Posição				HPV ⁻	HPV ⁺	
<i>APOBEC3F</i>	cg10191799	TSS1500	$\Delta\beta$:-0,05 Rho:0,08	$\Delta\beta$:-0,134 Rho:0,27	$\Delta\beta$:0,09 Rho:0,069	$\Delta\beta$:-0,0152 Rho:0,26	$\Delta\beta$:0,0704 Rho:0,37
	cg17385662	TSS1500	$\Delta\beta$:-0,069 Rho:0,062	$\Delta\beta$:-0,09 Rho:0,29	$\Delta\beta$:0,043 Rho:0,07	$\Delta\beta$:0,044 Rho:0,07	$\Delta\beta$:0,0736 Rho:0,48
	cg10016348	TSS1500	$\Delta\beta$:NA Rho:0,24	$\Delta\beta$:NA Rho:0,39	$\Delta\beta$:NA Rho:0,068	$\Delta\beta$:NA Rho:0,10	$\Delta\beta$:NA Rho:0,43
	cg07963345	5'UTR/ 1 éxon	$\Delta\beta$:0,011 Rho:-0,055	$\Delta\beta$:0,005 Rho:-0,00018	$\Delta\beta$:0,007 Rho:0,01	$\Delta\beta$:0,019 Rho:0,25	$\Delta\beta$:0,0014 Rho:0,002
	cg17035591	Corpo	$\Delta\beta$:NA Rho:-0,21	$\Delta\beta$:NA Rho:0,106	$\Delta\beta$:NA Rho:-0,13	$\Delta\beta$:NA Rho:0,099	$\Delta\beta$:NA Rho:-0,39
	cg09582545	3'UTR/ Corpo	$\Delta\beta$:0,0785 Rho:0,27	$\Delta\beta$:0,06 Rho:0,29	$\Delta\beta$:0,097 Rho:0,087	$\Delta\beta$:-0,01 Rho:0,42	$\Delta\beta$:-0,0259 Rho:-0,02
	cg19333614	3'UTR	$\Delta\beta$:0,032 Rho:0,23	$\Delta\beta$:-0,295 Rho:0,61	$\Delta\beta$:-0,235 Rho:0,41	$\Delta\beta$:-0,19 Rho:0,62	$\Delta\beta$:-0,21 Rho:0,60
<i>APOBEC3G</i>	cg13440882	TSS1500	$\Delta\beta$:0,014 Rho:NA	$\Delta\beta$:-0,009 Rho:NA	$\Delta\beta$:-0,026 Rho:NA	$\Delta\beta$:-0,0197 Rho:NA	$\Delta\beta$:0,0046 Rho:NA
	cg26082280	TSS1500	$\Delta\beta$:0,028 Rho:NA	$\Delta\beta$:-0,032 Rho:NA	$\Delta\beta$:0,019 Rho:NA	$\Delta\beta$:-0,07 Rho:NA	$\Delta\beta$:-0,029 Rho:NA
	cg13256546	TSS1500	$\Delta\beta$:NA Rho:-0,08	$\Delta\beta$:NA Rho:0,19	$\Delta\beta$:NA Rho:0,059	$\Delta\beta$:NA Rho:0,11	$\Delta\beta$:NA Rho:0,15
	cg22902400	5'UTR/ 1 éxon	$\Delta\beta$:NA Rho:-0,43	$\Delta\beta$:NA Rho:-0,17	$\Delta\beta$:NA Rho:-0,21	$\Delta\beta$:NA Rho:-0,54	$\Delta\beta$:NA Rho:-0,3
	cg23896816	5'UTR/ 1 éxon	$\Delta\beta$:0,005 Rho:-0,37	$\Delta\beta$:0,0022 Rho:-0,08	$\Delta\beta$:-0,0139 Rho:-0,07	$\Delta\beta$:0,0178 Rho:-0,16	$\Delta\beta$:0,0019 Rho:0,005
	cg06398754	Corpo	$\Delta\beta$:-0,062 Rho:0,17	$\Delta\beta$:-0,018 Rho:0,26	$\Delta\beta$:0,01 Rho:0,01	$\Delta\beta$:-0,056 Rho:0,17	$\Delta\beta$:-0,109 Rho:-0,29
	cg24519526	3'UTR	$\Delta\beta$:0,03 Rho:0,15	$\Delta\beta$:-0,021 Rho:0,4	$\Delta\beta$:-0,034 Rho:0,32	$\Delta\beta$:-0,053 Rho:0,35	$\Delta\beta$:-0,0176 Rho:0,01

		TIPO DE CÂNCER				
		CEE	CEL	CECO	CEOF	
Sondas	Posição				HPV-	HPV+
cg06229674	TSS1500	$\Delta\beta$:0,0013 Rho:NA	$\Delta\beta$:-0,01 Rho:NA	$\Delta\beta$:-0,019 Rho:NA	$\Delta\beta$:-0,0085 Rho:NA	$\Delta\beta$:0,026 Rho:NA
cg18696237	TSS1500	$\Delta\beta$:-0,013 Rho:-0,055	$\Delta\beta$:0,007 Rho:-0,15	$\Delta\beta$:0,023 Rho:0,0316	$\Delta\beta$: 0,074 Rho:0,07	$\Delta\beta$: 0,053 Rho:-0,03
cg24515575	TSS1500	$\Delta\beta$:-0,007 Rho:0,0407	$\Delta\beta$:-0,007 Rho: -0,203	$\Delta\beta$:0,0002 Rho:0,04	$\Delta\beta$: -0,035 Rho:0,15	$\Delta\beta$:-0,01 Rho:0,01
<i>APOBEC3H</i>						
cg09135399	TSS200	$\Delta\beta$: 0,07 Rho:-0,09	$\Delta\beta$: -0,074 Rho:-0,02	$\Delta\beta$: -0,096 Rho:0,0027	$\Delta\beta$:0,0041 Rho:-0,18	$\Delta\beta$: -0,037 Rho:-0,24
cg24424889	<i>Enhancer</i> 5'UTR	$\Delta\beta$: 0,063 Rho: -0,42	$\Delta\beta$:-0,004 Rho: -0,405	$\Delta\beta$:-0,06 Rho: -0,29	$\Delta\beta$:0,12 Rho:-0,32	$\Delta\beta$:-0,17 Rho: -0,42
cg17972162	<i>Enhancer</i> Corpo	$\Delta\beta$: -0,084 Rho: -0,25	$\Delta\beta$:-0,049 Rho: -0,35	$\Delta\beta$:-0,065 Rho:-0,0767	$\Delta\beta$:0,087 Rho:-0,17	$\Delta\beta$: -0,17 Rho: -0,43
cg03582215	3'UTR	$\Delta\beta$:-0,048 Rho:0,13	$\Delta\beta$: -0,13 Rho:-0,018	$\Delta\beta$:-0,103 Rho: 0,42	$\Delta\beta$: -0,11 Rho: 0,565	$\Delta\beta$:-0,0229 Rho:0,23

NA: Não aplicado, Em negrito: estatisticamente significativo, UTR: região não traduzida; TSS: sítio de início da transcrição; $\Delta\beta$: mediana da diferença de valor de beta de metilação entre tumor e mucosa adjacente não tumoral.