

**RASTREAMENTO GLOBAL DE ALTERAÇÕES
GENÔMICAS EM CARCINOMAS DE CABEÇA E
PESCOÇO DE PACIENTES JOVENS POR
SEQUENCIAMENTO TOTAL DO EXOMA**

PRISCILA MAYRINK DE MIRANDA

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

Área de Concentração: Oncologia

Orientadora: Profa. Dra Silvia Regina Rogatto

São Paulo

2017

FICHA CATALOGRÁFICA

Preparada pela Biblioteca da Fundação Antônio Prudente

Miranda, Priscila Mayrink

Rastreamento global de alterações genômicas em carcinomas de cabeça e pescoço de pacientes jovens por sequenciamento total do exoma / Priscila Mayrink de Miranda – São Paulo; 2017.

107p.

Tese (Doutorado)-Fundação Antônio Prudente.

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

Orientadora: Silvia Regina Rogatto

Descritores: 1. Carcinoma de Células Escamosas/Carcinoma, Squamous Cell. 2. Exoma/Exome. 3. Neoplasias de Cabeça e Pescoço/Head and Neck Neoplasms 4. Mapeamento Cromossômico/Chromosome Mapping.

*“Conheça todas as teorias, domine todas as técnicas,
mas ao tocar uma alma humana,
seja apenas outra alma humana”*

Carl Gustav Jung

DEDICATÓRIA

*Aos meus pais, José Zaidan e Maria Claret,
Às minhas irmãs, Aline e Danielly,
E a Júlia, a minha fonte de infinita inspiração e alegria.
A vocês, que com muito amor, carinho, estímulo constante e fé,
deram-me tudo que necessitei para chegar até aqui
e continuar sempre em frente na busca da
realização dos meus sonhos!*

AGRADECIMENTOS

A **Deus** pela vida e por tudo que a complementa!

A minha orientadora, **Dra Silvia Regina Rogatto**, pelas oportunidades oferecidas, que proporcionaram muito mais que um crescimento profissional. Será sempre uma referência. Muito obrigada por tudo.

Ao **Dr. Fabio A. Marchi**, pela contribuição e responsabilidade com a análise de Bioinformática.

Ao **Prof. Dr. Luiz Paulo Kowalski**, por sempre estar disposto a compartilhar da sua experiência e conhecimento em Tumores de Cabeça e Pescoço.

Aos **membros da banca examinadora**, por aceitarem o convite a fim de contribuírem para a melhoria desse trabalho, com seus conhecimentos científicos e experiências profissionais.

A equipe do laboratório NeoGene: **André Guollo, Caroline, Deborah, Fábio, Flávia, Graziela, Hellen, Isabella, Jennice, Juan, Julia, Luisa, Lídia, Maísa, Márcia, Marco, Mariana, Mateus, Rolando, Rogerio, Sara, Dra Sandra, Tatiane, Vanessa**. *“Na vida você vai perceber que há um propósito para todos que encontrar pelo caminho. E, o mais importante, é que todos vão te ensinar algo”*.

A equipe do laboratório de Oncogenética: **Fernanda, Karina, Kelvin, Renata, Dr. Alexandre, Dra Nirvana, Amanda e Lisley**, por todo apoio, inclusive médico e pessoal.

Aos **funcionários** da POS-GRADUAÇÃO e BIBLIOTECA, por todo aprendizado e apoio envolvidos. Em especial à Ana Kuninari, Karla Cristina, Cássia, Luciana, Suely e Vanuza.

À **FAPESP** (Fundação de Amparo a Pesquisa do Estado de São Paulo. Processo: 2013/06897-2), pelo suporte financeiro.

As amigas do “*Priscila Dinah*” (**Fernanda, Graziela, Iara e Karina**) pelos momentos de risadas, pelo incentivo, amizade e companheirismo.

Aos meus amigos, parceiros da vida: **Juliana, Lívia, Luana, Roberta,**

Tetê, Rafael, Stela Mars, Bianca, Mari Rosendo, Nayara, Nara, Thati Mello e Carol Sabino. Obrigada por estarem *SEMPRE* presentes e me darem suporte em todos os momentos da minha vida. "*A amizade é um amor que nunca morre*".

A **Hellen** e **Luisa**, meu agradecimento e admiração por fazerem o melhor para ajudar ao próximo e por me auxiliarem incondicionalmente.

À **Lucélia Landini, Lais e Luciana**, por toda força ao longo deste período e por me acolherem como filha e irmã. Minha família paulistana, obrigada pelo imenso carinho.

Ao **NABEM**, pelo acolhimento, apoio e sustentação. Em especial, **Adriana, Graça, Felipe, Maurice, Nelson e Xavier**.

Ao **João**, obrigada por renovar e fortalecer em mim a capacidade de enxergar o que é essencial e a não ter medo de assumir o que nos impulsiona a seguir em frente.

A Tia **Conceição Telles**, minha mentora! Minha professora por anos no primário e que nunca deixou de me acompanhar e incentivar. Obrigada principalmente por toda força emocional na fase de conclusão desse trabalho.

Aos mais importantes da minha vida: meus pais (**José** e **Claret**), minhas irmãs (**Aline** e **Danielly**) e à minha sobrinha (**Júlia**), a mais doce de todas as inspirações. Vocês são minha base, fonte dos melhores ensinamentos, amor e respeito. Em vocês, eu me inspiro todos os dias. De vocês, eu recebo a força para encarar os desafios e para me reerguer quando necessário. Com vocês, eu recarrego as energias para seguir adiante e, acima de tudo, com os reais valores para manter viva a essência de ser mais Humana. Amo vocês.

E com imenso carinho e respeito, agradeço aos **Pacientes**, sem os quais, esse trabalho seria impossível de ser realizado. "*O sofrimento físico tem o dom de igualar estudiosos e iletrados*".

RESUMO

Miranda PM. **Rastreamento global de alterações genômicas em carcinomas de cabeça e pescoço de pacientes jovens por sequenciamento total do exoma.** São Paulo; 2017. [Tese de Doutorado-Fundação Antônio Prudente].

O carcinoma de células escamosas (CCE) de cabeça e pescoço (CP) é um dos mais comuns do mundo. Os fatores de risco mais importantes descritos para CCECP são o consumo de álcool e o uso de tabaco, além da infecção pelo vírus do papiloma humano. A maioria dos casos é diagnóstica do entre a quinta e oitava décadas de vida, entretanto, nos últimos anos tem sido observado um aumento da incidência da doença em pacientes com idade inferior a 45 anos. Nestes casos, são relatadas características mais graves da doença e nem sempre há relação com os fatores de riscos descritos para esses tumores. Não há consenso em literatura sobre os aspectos etiológicos, curso clínico e prognóstico do CCECP de pacientes jovens. Apesar de sua importância, estudos com este grupo de tumores são escassos em literatura. Métodos recentes de sequenciamento de alto desempenho em exons (exoma) permitem identificar mutações e variações no genoma. O exoma de 35 pacientes com idade inferior a 49 anos e dos seus tumores (CEC de cavidade oral e orofaringe) foram investigados com o intuito de identificar alterações somáticas associadas com características clínico patológicas. O DNA do tecido tumoral e sangue periférico de cada paciente foram utilizados no preparo das bibliotecas (*Nextera Exome Enrichment* Illumina kit) para o sequenciamento do exoma total (SET) (HiSeq2500, Illumina). A cobertura de mapeamento alcançou uma média de 48,7 vezes por amostra, sendo uma amostra excluída por não ter atingido os critérios de qualidade. Após processamento dos dados e identificação das variantes foi realizada uma comparação entre amostras tumorais e normais, selecionando apenas variantes exclusivas do tumor de cada paciente, não

sinônimas e patogênicas (segundo os bancos de dados: SIFT, POLYPHEN2_HDIV, POLYPHEN2_HVAR, LRT, MutationTaster, FATHMM, Radial SVM, LR, GERP++, PhyloP, SiPhy). Foram identificadas 3963 variantes em exons e 28.206 SNVs não sinônimos mapeados em 10.092 genes, exclusivos dos tumores. Apenas cinco variantes patogênicas foram identificadas no gene *TP53*, em cinco pacientes (14,7%). Duas delas (rs28934578 e rs121912656) foram relacionadas à síndrome de Li-Fraumeni e uma ao carcinoma hepatocelular (rs28934571). A análise das variantes quanto a infecção pelo HPV revelou que a rs28934578 e rs121912656 estavam presentes apenas nos casos HPV positivos enquanto rs28934571 foi identificada apenas em casos HPV negativos. A variante rs28934578 foi identificada apenas entre os não tabagistas enquanto os tabagistas apresentaram as variantes rs121912656 e rs28934571. Foram identificadas variantes no genes *FANCA* (rs17233141 e rs17227403) e *FANCD2* (rs3864017). Mutações patogênicas nestes genes estão associadas com a Anemia de Fanconi e pacientes jovens acometidos por esta síndrome tem risco aumentado de desenvolvimento de CCECP. Quatro pacientes do nosso estudo apresentaram a variante rs3864017 do gene *FANCD2* no sangue periférico e no tumor, a qual é considerada como benigna no banco de dados ClinVar enquanto a variante rs17233141 do gene *FANCA* é considerada patogênica por esse mesmo banco de dados. Os casos HPV negativos apresentaram uma maior frequência de alterações. Apenas os casos HPV positivos apresentaram a variante no oncogene *PIK3CA* (rs2230461). Foi verificada uma forte presença de componentes imunológicos alterados nas vias enriquecidas quanto ao status do HPV. Nos casos HPV positivos essa associação foi mais proeminente, por exemplo, o incluindo o componente imune adaptativo, a via sinalização do *PD1* e *TCR*, os quais são potenciais alvos para imunoterapia (considerando as variantes com frequência mínima de 20%). Foram identificados 10 SNV (*single nucleotide variants*) comuns em pelo menos três dos sete pacientes (48,7%) que apresentaram metástase pulmonar. Essas alterações foram consideradas patogênicas em pelo menos cinco bancos de dados. O único

fator clinico comum entre esses pacientes foi o estadiamento N2. Os resultados deste estudo foram comparados com dados publicados em literatura e revelaram um perfil de mutação característico de pacientes jovens. As alterações identificadas neste trabalho tem potencial para contribuir para melhor entendimento da biologia de tumores dos pacientes jovens e para a identificação de marcadores diagnósticos, prognósticos e de tratamento.

SUMMARY

Miranda PM. **[Global screening of genetic alterations in head and neck carcinomas of young patients by whole exome sequencing]**. São Paulo; 2017, [Tese de Doutorado-Fundação Antônio Prudente].

The head and neck squamous cell carcinoma (HNSCC) is one of the most common cancers worldwide. The main risk factors described for HNSCC are consumption of alcohol and tobacco, and also infection by the Human Papilloma Virus (HPV). Most cases are diagnosed between the fifth and eighth decades of age. However, in the last years, an increased incidence of the disease has been described in young patients (< 45 years old). In these cases, more severe aspects of the disease are described, which is not always related to the risk factors frequently associated with these tumors. There is no consensus in literature about the etiological aspects, clinical course and prognosis of HNSCC from young patients. Despite its relevance, studies involving this group of tumors are scarce in literature. Recent high-throughput sequencing methods in exons (exoma) allow the identification of mutations and genome variations. In this study, the exome of 35 young patients (≤ 49 years) with oral and or pharynx carcinomas were sequenced in order to identify somatic alterations that might be associated with clinic-pathological characteristics. The DNA of the tumoral tissue and peripheral blood samples in each patient were used in the preparation of libraries (Nextera Exome Enrichment Illumina kit) for the total exon sequencing (SET) (HiSeq2500, Illumina). The median sequencing coverage has reached 48.7x; and one case was excluded for not attaining the quality of the parameters. After processing data and identifying variants, a comparison was made between tumoral and normal samples, selecting only exclusive tumor variants from each patient, non-synonymous and pathogenic (in accordance to the following data banks: SIFT, POLYPHEN2_HDIV, POLYPHEN2_HVAR, LRT, MutationTaster, FATHMM, Radial SVM, LR, GERP++, PhyloP, SiPhy).

We identified 3,963 exonic variants. A total of 28,206 non-synonymous SNVs mapped in 10,092 genes, tumor-exclusives was reported. Only five pathogenic variants of the *TP53* gene, in five patients, were identified (14,7%). Considering the identified variants in the *TP53* gene, two were related to the Li-Fraumeni syndrome (rs28934578 e rs121912656) and one to the hepatocellular carcinoma (rs28934571). The rs28934578 and rs121912656 variants were present only on HPV positive cases while rs28934571 was identified only in HPV negative cases. The rs28934578 variant was identified only among non smokers while smokers patients presented the rs121912656 and rs28934571 variants. FANCA (rs17233141 e rs17227403) and FANCD2 (rs3864017) variants were identified, both related to Fanconi Anemia, which has been associated with an increasing risk of young patients developing HNSCC. Four patients in our study presented the rs3864017 variant of the FANCD2 gene in both, t, peripheral blood and in the tumor sample, which is considered benign at the ClinVar database while the rs17233141 variant of the FANCA gene is considered pathogenic by the same database. The HPV negative cases presented a higher frequency of alterations. Only HPV positive cases presented variants in oncogene, PIK3CA (rs2230461). A strong presence of immunological components altered in the enriched pathways concerning the HPV status was verified. HPV positive cases presented this association more prominent, for example, including the immune adaptive component, the signaling pathway of PD1 and TCR, which are potential targets for immunotherapy (variants with minimum 20% of frequency). Ten common SNVs were identified in at least three of seven patients (48,7%) that presented pulmonary metastasis. These alterations were considered pathogenic in at least five databases. The only common clinical factor between these patients was N2 staging. The results of this study were compared with published data in literature and revealed a characteristic mutation profile in young patients. Our data have potential of contributing on better understanding tumor biology of young patients and the identification of diagnostic, prognostic and treatment markers.

LISTA DE FIGURAS

Figura 1	Interpretação dos resultados provenientes da genotipagem do HPV.....	32
Figura 2	Fluxograma das etapas realizadas no grupo de 35 pacientes com CCECP.....	46
Figura 3	Características clínico patológicas por paciente incluso no estudo.....	52
Figura 4	Distribuição das médias de variantes por subgrupos de pacientes com CCECP.....	58
Figura 5	Eletroferograma das alterações no gene <i>TP53</i> , identificadas nos pacientes.....	75

LISTA DE TABELAS

Tabela 1	Descrição das características clínicas e histopatológicas dos pacientes incluídos no estudo.....	29
Tabela 2	Resumo dos resultados de limpeza, mapeamento dos <i>reads</i> e cobertura, após o sequenciamento do exoma de amostras de sangue periférico (N) e tecido tumoral (T) dos pacientes com CCECP.....	54
Tabela 3	Resumo dos genes que apresentaram variantes (SNV) em pelo menos 20% dos casos.....	60
Tabela 4	Genes em destaque nos principais estudos de rastreamento de alterações em CCECP identificados por SET.....	65
Tabela 5	Variantes exclusivas nos casos estudados de acordo com o sítio anatômico.....	67
Tabela 6	Principais vias canônicas envolvendo as alterações detectadas nos casos HPV positivos.....	69
Tabela 7	Principais vias canônicas envolvendo os genes alterados nos casos HPV negativos.....	71

LISTA DE QUADROS

Quadro 1	Estadiamento clínico da AJCC para os cânceres da cavidade oral e orofaringe.....	4
Quadro 2	Resumo dos principais estudos utilizando SET para rastreamento de alterações em casos de CCECP, principalmente envolvendo tumores da cavidade oral e orofaringe.....	11
Quadro 3	Sequência de iniciadores do gene <i>TP53</i> avaliados por sequenciamento direto.....	47
Quadro 4	Alterações identificadas no gene <i>TP53</i>	61
Quadro 5	Variantes identificadas nos genes relacionados a Anemia de Fanconi.....	63
Quadro 6	Variantes (SNVs) mais frequentes e patogênicos entre os pacientes que desenvolveram metástase	73

LISTA DE SIGLAS E ABREVIATURAS

AJCC*	The American Joint Committee on Cancer
CCECP	Carcinomas de Células Escamosas de Cabeça e Pescoço
CCEOF	Carcinomas de Células Escamosas da Orofaringe
CEP	Comitê de Ética em Pesquisa
COSMIC*	Catalog of Somatic Mutations in Cancer
DNA*	Deoxyribonucleic Acid
dNTPs	Deoxinucleotídeos trifosfatados
dUTP	Trifosfato deoxiuridina
EEO*	Expanded Exome Oligos
EWS*	Enrichment Wash Solution
EEB*	Enrichment Elution Buffer
EHB*	Enrichment Hybridization Buffer
ETB*	Elute Target Buffer
EtOH	Etanol
EXO1	Exonuclease I
FAP	Fundação Antônio Prudente
FFPE*	Formalin-Fixed Paraffin-Embedded
GPCR*	G-protein coupled receptors
GST	Genes supressores tumorais
HPV*	Human Papilloma Virus
HR-HPV	High Riskk Human Papilloma Virus
IGV*	Integrated Genomic Viewer
INCA	Instituto Nacional do Câncer
NGS*	Next generation sequencing
PCR*	Polymerase Chain Reaction
qPCR*	Quantitative Real Time – PCR
QT	Quimioterapia
RSB*	Ressuspension Buffer
RTx	Radioterapia

UNG*	AmpErase uracil-N-glicosilase
SAP*	Shrimp Alkaline Phosphatase
SEER*	The Surveillance, Epidemiology, and end Results
SET	Sequenciamento do exoma total
SPB*	Sample Purification Beads
SMB*	Streptavidin Magnetic Beads
SNP*	Single Nucleotide Polymorphisms
SNV*	Single Nucleotide Variants

*siglas ou abreviaturas derivadas do inglês

ÍNDICE

1	INTRODUÇÃO	1
1.1	Carcinomas de Células Escamosas de Cabeça e Pescoço (CCECP)	1
1.2	Fatores de Risco Associados ao Desenvolvimento de CCECP	6
1.3	Sequenciamento de Alto Desempenho e Alterações Genômicas em CCECP	9
1.4	Pacientes Jovens e CCECP	22
2	OBJETIVOS	26
2.1	Objetivo Geral.....	26
2.2	Objetivos Específicos	26
3	MATERIAL E MÉTODOS.....	27
3.1	Caracterização das Amostras.....	27
3.2	Pacientes.....	28
3.3	Extração do DNA Genômico.....	30
3.4	Genotipagem do Papilomavirus Humano (HPV)	30
3.5	Sequenciamento Total do Exoma.....	32
3.5.1	Construção das Bibliotecas de Exoma	33
3.5.2	Sequenciamento do Exoma.....	41
3.5.3	Análise do Sequenciamento do Exoma	43
3.6	Sequenciamento Direto do Gene TP53	47
4	RESULTADOS.....	50
4.1	Avaliação das Características Clínico-Patológicas.....	50
4.2	Análise dos Dados do Sequenciamento de Exoma Total (SET).....	53
4.2.1	Análise dos Dados Brutos Gerados pelo SET	53
4.2.2	Alterações Genômicas nos Carcinomas Orais e de Orofaringe de Pacientes Jovens	56
4.2.3	Perfil Geral de Alterações Somáticas e Não Sinônimas	59

4.2.3.1	Mutação no gene TP53	61
4.2.3.2	Genes Relacionados a Anemia de Fanconi (AF)	61
4.2.3.3	Comparação com os Principais Genes Descritos em Literatura Relacionados ao CCECP	64
4.2.4	Comparação Entre as Alterações Segundo o Sítio Anatômico (Cavidade Oral e Orofaringe)	66
4.2.5	Comparação entre os Casos Segundo o Status do HPV (Positivo e Negativo)	68
4.2.6	Comparação entre as Alterações Genômicas e os Principais Fatores de Risco	71
4.2.7	Mutações Exclusivas em Pacientes que Apresentaram Metástases .	71
4.3	Confirmação das Alterações Identificadas no Gene TP53	74
5	DISCUSSÃO	76
6	CONCLUSÃO	92
7	REFERENCIAS BIBLIOGRAFICAS	93

ANEXO

Anexo 1 Carta de aprovação do Comitê de Ética em Pesquisa-CEP

1 INTRODUÇÃO

1.1 CARCINOMAS DE CÉLULAS ESCAMOSAS DE CABEÇA E PESCOÇO (CCECP)

O câncer de cabeça e pescoço se refere a um grupo heterogêneo de tumores, resultante de várias estruturas anatômicas, incluindo os ossos craniofaciais, tecidos moles, glândulas salivares, pele e mucosas do trato aerodigestivo superior. O carcinoma de células escamosas de cabeça e pescoço (CCECP) é a doença mais comum dessa região, responsável por mais de 90% dos casos (PAI e WESTRA 2009). Frequentemente, o termo câncer de cabeça e pescoço é usado para descrever todos os carcinomas que surgem a partir do epitélio que reveste o trato sinonasal, cavidade oral, faringe e laringe, mostrando evidência de diferenciação escamosa microscópica (PAI e WESTRA 2009; LEEMANS et al. 2011).

A incidência mundial e a taxa de mortalidade desses tumores ainda são altas. O câncer de lábio e cavidade oral, por exemplo, de acordo com última estimativa realizada pelo GLOBOCAN (2012), representaram cerca de 300.000 casos em 2012 (2,1% do total de todos os tipos de câncer, com exceção de câncer de pele não melanoma). Foram observadas 145.000 mortes no mundo (1,8%), das quais 77% ocorreram em regiões menos desenvolvidas sócio economicamente (FERLAY et al. 2015). A doença ainda é mais incidente em homens. Porém, tem sido observado o aumento da

incidência mundial deste tumor entre as mulheres e o declínio entre os homens (CURADO e HASHIBE 2009; UDEABOR et al. 2012).

Para o ano de 2016, o Instituto Nacional do Câncer (INCA) do Brasil, estimou 11,27 novos casos de câncer de cavidade oral a cada 100 mil homens e 4,21 a cada 100 mil mulheres. De acordo com a distribuição dos tumores mais incidentes por regiões do país, sem considerar os tumores de pele não melanoma, o câncer da cavidade oral em homens é mais frequente na Região Sudeste (14,58/100 mil), ocupando a quarta posição. Para as mulheres, o câncer da cavidade oral é mais frequente na Região Nordeste (4,11/100 mil) seguido da região Sudeste (5,29/100 mil) (Ministério Da Saúde 2016).

Os CCECP acometem mais frequentemente indivíduos com idade mais avançada, entre a sexta e oitava décadas de vida, sendo 64 anos a idade média de apresentação da doença (National Cancer Institute-NIH 2016). Embora os CCECP ocorram frequentemente em pacientes com idade superior a 50 anos e estejam associados a fatores de riscos bem estabelecidos, estudos epidemiológicos passaram a apontar um aumento na incidência da doença em crianças e adultos jovens (18-45 anos), especialmente para os tumores da orofaringe e cavidade oral. Aproximadamente 5,5% desses pacientes tem idade inferior a 40 anos (TOROSSIAN et al. 2000; JEMAL et al. 2006; TONER e O'REGAN 2009; KAMINAGAKURA et al. 2012; HILLY et al. 2013; BODNER et al. 2014). Entre os pacientes jovens, a distribuição entre homens e mulheres é distinta daquela observada entre os pacientes que desenvolve a doença acima dos

50 anos. Enquanto que, entre os pacientes em idade mais avançada, a taxa de pacientes do sexo masculino acometidos pela doença está em torno de 70% dos casos, entre os jovens correspondem a 50-60% dos casos (VAN MONSJOU et al. 2013). Estimativas mais recentes não evidenciam a distribuição detalhada da doença por faixa etária e, embora o interesse seja crescente, ainda poucos estudos abordam o tema.

O prognóstico para os pacientes com CCECP é determinado principalmente pelo estadiamento clínico ao diagnóstico, o qual considera a extensão do tumor, presença de metástase nos linfonodos e metástase à distância para classificação (LEEMANS et al. 2011). Para os cânceres da cavidade oral e orofaringe, o sistema de estadiamento clínico recomendado é o da *The American Joint Committee on Cancer-AJCC*. Doença em estágio precoce é definida por doença localmente limitada ou ausência de linfonodos metastáticos (T1-2, N0). Os estádios III e IV implicam grupos variados, que podem ser caracterizados por doença limitada localmente com metástases nodais extensivas, localmente avançados e sem metástases linfonodais; ou localmente avançados com metástases nodais (Quadro 1) (EDGE e COMPTON 2010).

Quadro 1 - Estadiamento clínico da AJCC para os cânceres da cavidade oral e orofaringe.

Estadiamento Clínico	TNM
0	Tis N0 M0
I	T1 N0 M0
II	T2 N0 M0
III	T3 N0 M0 T1-3 N1 M0
IVA	T4a N0/1 M0 T1-4a N2 M0
IVB	T4b qqN M0 qqT N3 M0
IVC	qqT qqN M1

Fonte: EDGE e COMPTON (2010). T: tumor primário; N: Linfonodos; M: Metástase. Tis: carcinoma *in situ*; qq: *qualquer*. T0: sem evidência de tumor primário; T1-T4: o tamanho do tumor e/ou disseminação da doença nas proximidades. Quanto maior o número de T, maior o tumor e/ou sofreu maior disseminação para os tecidos próximos; N0 e M0: linfonodos não foram comprometidos pelo câncer e não foi identificada metástase à distância, respectivamente; M1: metástase à distância. N1-3: os números correspondem ao tamanho, localização e/ou o número de linfonodos com a doença.

As modalidades de tratamento para o CCECP incluem cirurgia, radioterapia, quimioterapia e imunoterapia. Para os casos em que a doença é diagnosticada precocemente (estádio I e II), o tratamento indicado é a cirurgia ou a radioterapia, tendo em geral, um prognóstico favorável. Em caso de tumores em estágios mais avançados (estádio III e IVA/B), o tratamento é a cirurgia combinada com a radioterapia pós-operatória, seguida ou não por quimioterapia (BELCHER et al. 2014). Em casos de tumores não ressecáveis, o tratamento com cisplatina continua sendo a melhor opção. Para pacientes com metástase, o tratamento vai depender das condições do paciente, sendo que os pacientes debilitados são submetidos normalmente ao tratamento paliativo (MARUR e FORASTIERE

2016).

AMIT et al. (2015) relataram que os pacientes T3N1 apresentavam alto risco de falha de resposta ao tratamento, semelhante ao observado em pacientes estadio IVa. Tal observação é sugestiva da necessidade de re-estratificação desses pacientes T3N1 para IVa, de modo a auxiliar, potencialmente, a identificação de pacientes para os quais o tratamento adjuvante concomitante seja necessário.

Embora novas estratégias terapêuticas venham sendo desenvolvidas, a taxa de mortalidade destes carcinomas continua alta, devido ao seu potencial infiltrativo, alta taxa de recorrência e/ou ao diagnóstico tardio, muito comum em países em desenvolvimento, influenciado pela situação socioeconômica (TSANTOULIS et al.2007; GOURIN et al. 2009; BAYKUL et al. 2010; OLSEN et al. 2015; GIEFING et al. 2016; WANG et al. 2016). A combinação de mais de uma técnica terapêutica é uma alternativa para os casos com prognóstico ruim em que um tratamento único não seria o suficiente para melhoria do paciente (MARTA et al. 2015a; MARTA et al. 2015b RODRIGO et al. 2015) .

Em alguns casos de pacientes com a doença em estadio mais avançado, é comum a presença de resistência inata ou adquirida aos agentes quimioterápicos. É crescente o número de estudos em CCECP visando esclarecer os mecanismos moleculares de quimio-resistência (e fluxo de drogas, apoptose, mecanismos de reparo de DNA, transição epitélio mesênquima, autofagia e miRNA) com o objetivo de melhorar as opções de tratamento (WANG et al. 2016).

Imunoterapia é outra abordagem que vem sendo fortemente considerada para os casos de CCECP. Desequilíbrio no funcionamento do sistema imunológico humano pode estar ligado ao início e ao desenvolvimento da carcinogênese. Elevado número de macrófagos, células T, contagem absoluta de linfócitos, redução da expressão do antígeno de leucócito tumoral, células *natural killer* (NK) e deficiência na apresentação de antígenos foram identificados em pacientes com CCECP (LI Q et al. 2015). Futuros estudos podem envolver estratégias imunoterapêuticas em combinação a outros tratamentos clássicos para obter melhores respostas ao tratamento.

1.2 FATORES DE RISCO ASSOCIADOS AO DESENVOLVIMENTO DE CCECP

Inicialmente, os principais fatores de risco associados ao aumento da incidência dos CCECP foram o hábito tabagista e o consumo de álcool, responsáveis por 80-90% dos casos (RODRIGUEZ et al. 2004; HASHIBE et al. 2007). O fumo está associado com um aumento de mais de 10 vezes no risco de desenvolvimento de todos os cânceres de cabeça e pescoço e é responsável por 42% das mortes pela doença. Já o etilismo pesado corresponde a, aproximadamente 16% dos óbitos (CASSIDY et al. 2010; Ministério da Saúde 2016). Estudos apontam um sinergismo entre esses dois fatores, o que leva a um aumento no risco de até 30 vezes de o paciente desenvolver estes tumores (HASHIBE et al. 2007; JABER 2010;

MAJCHRZAK et al. 2014).

A diminuição global do uso do tabaco e seus produtos derivados levou a redução da incidência de CCECP, com exceção dos carcinomas de orofaringe, que teve um aumento da incidência, em parte atribuível a infecção pelo vírus do papiloma humano (HPV) (LI R et al. 2015).

O HPV é composto por DNA de dupla fita, de aproximadamente 6.800 a 8.400 pares de bases, não envelopados, de simetria icosaédrica, com 72 capsômeros. Seu ciclo de vida difere das demais famílias de vírus, pois a infecção requer a disponibilidade de células epiteliais da mucosa ou epiderme, que sejam hábeis à proliferação (células da camada basal) (ZUR HAUSEN 1996, 2002). Ao entrar em contato com uma célula epitelial, o vírus pode permanecer em sua forma episomal ou integrar-se ao genoma do hospedeiro, ambas as formas podem induzir o processo carcinogênico (MOODY e LAIMINS 2010).

O genoma do HPV apresenta três domínios: uma região não codificadora, a URR (do inglês, *upstream regulatory region*), uma região de função precoce (*early*, em inglês) com seis ORFs (do inglês, *Open Reading Frames*) - *E1*, *E2*, *E4*, *E5*, *E6* e *E7*, e uma região de função tardia (do inglês, *late*) contendo os genes *L1* e *L2*, responsáveis por codificar proteínas do capsídeo viral (ZUR HAUSEN 2002). As oncoproteínas virais *E6* e *E7* são as relacionadas ao potencial oncogênico do HPV. Cada uma dessas proteínas, expressas no núcleo, interfere em uma variedade de mecanismos celulares regulatórios predominantemente anti-proliferativos, como os que envolvem as proteínas p53 e pRb (MCLAUGHLIN-DRUBIN e MÜNGER 2009). O gene

TP53, o qual é inativado em nível proteico pela E6, de um modo geral, não se apresenta mutado nos tumores HPV positivos, enquanto nos tumores negativos o gene apresenta-se frequentemente mutado (SHI et al. 2009; VAN MONSJOU et al. 2013). Os casos HPV positivos, apresentam inativação da proteína pRb pela oncoproteína viral E7 e expressam a proteína p16, embora sua atividade supressora de tumor seja perdida por ser mediada pela pRb (inativa) (LOWY e MUNGER 2010; VAN MONSJOU et al. 2013).

Este aumento na incidência de tumores HPV-positivos pode ser explicado por mudanças no comportamento sexual incluindo o aumento do número de parceiros sexuais e início de atividade sexual em idade mais jovem. A presença de verrugas ano-genitais em homens e mulheres podem contribuir para colonização viral da mucosa oral. Uma parcela significativa dos pacientes com CCECP são HPV-positivos com história de elevado número de parceiros sexuais (SHIBOSKI et al. 2005; D'SOUZA et al. 2007 e 2014; MAJCHRZAK et al. 2014). CHATURVEDI et al. (2011) estimaram um aumento de 225% entre os anos de 1988 e 2004 na incidência dos carcinomas de células escamosas de orofaringe (CCEOF) associados à infecção pelo HPV. Os autores relatam que a incidência do vírus em CCECP deverá superar a incidência de câncer de colo de útero até 2020, o qual foi o primeiro tumor associado a infecção pelo HPV.

A base biológica e os mecanismos moleculares envolvidos nos CCECP HPV positivos e negativos estão sob intensa investigação nos últimos anos. Quando positivo para o HPV, o CCEOF é considerado como

uma entidade clínico-patológica distinta e associada a um melhor prognóstico e resposta ao tratamento (quimioterapia e radioterapia) (DAHLSTRAND et al. 2008; VIDAL e GILLISON 2008; DESCHLER et al. 2014; MAJCHRZAK et al. 2014; SEPIASHVILI et al. 2015; AMBULOS et al. 2016; MARUR E FORASTIERE et al. 2016; OKAMI 2016; TINHOFER et al. 2016). A sobrevida dos pacientes com CCECP HPV positivos é relatada como maior do que a observada em casos HPV negativos (sobrevida em três anos de 84% vs. 57%, respectivamente) (ANG et al. 2010). Embora estudos em literatura associem uma melhor resposta ao tratamento para pacientes HPV positivos, ainda existem questionamentos que necessitam estudos mais elaborados para elucidar o papel do HPV nesses tumores, principalmente nos casos mais avançados da doença (MAJCHRZAK et al. 2014; SEPIASHVILI et al. 2015; AMBULOS et al. 2016; MARUR e FORASTIERE et al. 2016; OKAMI 2016; TINHOFER et al. 2016).

1.3 SEQUENCIAMENTO DE ALTO DESEMPENHO E ALTERAÇÕES GENÔMICAS EM CCECP

Com o rápido avanço das técnicas de sequenciamento genômico foi possível, nos últimos anos, a identificação de novos genes e mutações genéticas em uma variedade de tumores (BRENNER et al. 2011, ROBINSON et al. 2013, RYAN et al. 2013, ASANGANI et al. 2014). Estudos utilizando a tecnologia de *next-generation sequencing* (NGS) levaram a descobertas importantes e surpreendentes para a investigação da base do

câncer. O desafio atual com a enorme quantidade de dados gerados, é extrair informações clínicas relevantes que possam ser utilizadas no diagnóstico, prognóstico e para orientar o desenvolvimento e aplicação de terapias específicas (MAJEWSKI et al. 2011; KU et al. 2012; TABATABAEIFAR et al. 2014).

Dentre as diferentes abordagens do *NGS*, o sequenciamento do exoma total (SET), o qual consiste na captura de sequências codificadoras de proteínas (éxons) por meio de hibridização, permite identificar mutações e variações no genoma tumoral. Esta estratégia pode confirmar algumas alterações bem estabelecidas e desvendar outras desconhecidas. Mais barata e rápida do que o mapeamento tradicional do genoma, esta técnica permite a identificação de genes responsáveis por doenças raras ainda não diagnosticadas ou com diagnósticos imprecisos (BAMSHAD et al. 2011; MAJEWSKI et al. 2011; STRANNEHEIM e WEDELL et al. 2016).

Esforços tem sido concentrados para a caracterização genômica de CCECP com o objetivo de aumentar nosso conhecimento da biologia tumoral e alavancar os resultados para otimizar a gestão de CCECP. O Quadro 2 apresenta o resumo dos primeiros estudos utilizando SET para rastreamento de alterações em casos de CCECP.

Quadro 2 - Resumo dos principais estudos utilizando SET para rastreamento de alterações em casos de CCECP, principalmente envolvendo tumores da cavidade oral e orofaringe.

Referência	Amostras	Biblioteca	Plataforma	Média cobertura	Idade (anos)	Sítio anatômico	Principais genes	Software**
Agrawal et al. 2011	32 TU / 32 NL	SureSelect (Agilent)/ CCDS (NimbleGen)	Illumina GAI/HiSeq e ABI SOLiD	77x/44x	N/A	Cavidade oral (25), orofaringe (4), hipofaringe (1), laringe (2)	<i>TP53, CDKN2A, FAT1, PTEN, PIK3CA, HRAS, EGFR, FBXW7, NOTCH1</i>	Picard, SAMtools,
Stransky et al. 2011	74 TU / 74 NL	CCDS (NimbleGen)	Illumina GAI/HiSeq	149x/152x	N/A	Cavidade oral, Orofaringe, Hipofaringe, Laringe e Cavidade sinusal	<i>TP53, CDKN2A, PTEN, PIK3CA, HRAS, NOTCH1, IRF6, TP63</i>	Picard, BWA, IGV
Nichols et al. 2012a	2 TU / 2 NL	SureSelect (Agilent)	ABISOLiD	28,1x	1:49 e 51	Tonsila e base de língua	<i>ZNF3 10 229, 470, 543, 616, 664, 638, 716,e799 MUC4, 6 12, e16</i>	GATK ANNOVAR
Nichols et al. 2012b	6 linhagens CCECP	NimbleGen	Illumina HiSeq 2000	129x	-	Cell lines Cal27, Detroit562, FaDu, SCC4, SCC15, SCC25	<i>NOTCH, TP53, CDKN2A,PIK3CA, NRAS,MET, SMAD4</i>	GATK COSMIC
IPTofICGC 2013	50TU / 50 NL	SeqCap EZ System (NimbleGen)	Illumina HiSeq 2000 / Roche GS-FLX	36,67x/ 37,58x (HiSeq2000) e 35.43x/24.2x (Roche)	1:26 a 70	Gingivo-bucal COCE	<i>USP9X, MLL4, ARID2, UNC13C e TRPM3</i>	FASTQC, BWA, SAMtools, GATK, Genome MuSiC, MutSigCV v1.3

Cont/ Quadro 2

Referência	Amostras	Biblioteca	Plataforma	Média cobertura	Idade (anos)	Sítio anatômico	Principais genes	Software**
Lui et al. 2013	45 TU / 45 NL	-	Illumina	N/A	I:19 a 90 M:61	Cavidade oral (24), orofaringe (3), hipofaringe (3), laringe (16), cavidade nasal (1)	<i>PIK3CA, PIK3CG, PTEN, PIK3R5, PIK3AP1, PIK3R1, HRAS, KRAS, JAK3, JAK1, STAT1, IL6ST</i>	-
Lechner et al. 2013	20 TU (HPV+) 20 TU (HPV-) (FFPE)	-	Illumina HiSeq 2000	50%	HPV+: I: 42-81 HPV-: I:45-77	Orofaringe (11- Tonsila e 9- base de língua)	Exclusivas HPV: <i>TP53, CCND1, CDK2A/B, PIK33CA, PTEN, FBXW7, SOX2</i> (predominante em HPV+)	BWA; Picard; SAMtools; GATK; dbSNP; COSMIC
Biswas et al. 2014	17 TU/ 17NL	Illumina/ NimbleGen	Illumina HiSeq 2000/ Ion Torrent	44,94±1.83	I: 26 a 70	Gingivo-bucal COCE	<i>TP53, CASP8, FAT1, NOTCH1, USP9X, ARID2, MLL4, PIK3CA, HRAS</i>	Picard, SAMTools, GATK, Annovar, Oncotator, GenomeMuSiC, COSMIC
Pickering et al. 2014a ***	16 jovens (TU + NAd) 28 adultos (TU + NAd)	NimbleGen	SOLiD / Illumina HiSeq2000	N/A	Jovens I:<45 Adultos I:>45 anos	Língua (Cavidade Oral)	<i>FAT1, TP53, PIK3CA, CDKN2A, CASP8, AJUBA, NOTCH1, MLL2, NSD1, HLA-A, TGFB2</i>	Partek, ASCAT , GISTIC, R software, MultSig

Cont/ Quadro 2

Referência	Amostras	Biblioteca	Plataforma	Média cobertura	Idade (anos)	Sítio anatômico	Principais genes	Software**
Li R et al. 2015	6 TU língua 5 controles de estudo prévio (Agrawal et al. 2011)	SureSelect (Agilent)	Illumina HiSeq 2500	116x/ 77x	Não fumantes: M: 50,4 Fumante: M: 61,9	Língua	<i>Não fumantes:</i> <i>CTNNA3, EIF3A, EP300, FXR1, NEK8, NOTCH1 (NOTCH2, NOTCH3), PIK3CA, PKHD1L1, PTCHD2, RALGAPB, SPEN, UBR4</i> <i>Fumantes:</i> <i>CD101, LAMA1, NCAPD3, RIMBP2, SI, SYNE, TP53, TERT</i>	CASAVA, dbSNP
Samman et al. 2015	10 TU/10 NAd (FFPE) Carcinoma verrucoso oral 20 TU 20 NAd (FFPE) COCE	Agilent exome capture kit (Agilent)	Illumina HiSeq 2500	90x	I: 43 a 96	Língua, Assoalho Bucal	<i>TP53, CDKN2A, NOTCH1, NOTCH2, FAT1</i>	Picard, GATK, Varscan2, DAVID.
TCGA 2015	279 TU	SureSelect (Agilent)	Illumina HiSeq2000	95x	M: 61	Cavidade Oral, Orofaringe, laringe, hipofaringe	<i>FAT1, TP53, PIK3CA, CDKN2A, CASP8, AJUBA. NOTCH1, KMT2D, NSD1, HLA-A</i>	Picard, Firehose, MutSigCV, dRanger

Cont/ Quadro 2

Referência	Amostras	Biblioteca	Plataforma	Média cobertura	Idade (anos)	Sítio anatômico	Principais genes	Software**
Vettore et al. 2015	18TU / 18NL	SureSelect (Agilent)	Illumina HiSeq 2000	93x	I: 21 a 89	Língua	<i>TP53, DST, RNF213</i>	BWA, SAMTools, GATK, Partek, Genomics Suite, GATK, COSMIC
Upadhyay et al. 2016	23 TU (HPV-)	Illumina / NimbleGen	Illumina HiSeq 1500	N/A	I: 23 a 76	Língua	<i>DLL4, NOTCH1, NOTCH4, MAML3, DTX1, RBPJ</i>	BWA, Picard, GATK, COSMIC, dbSNP, Oncotator v1.1.6.0.
Liu et al. 2016	3 TU / 3NL	SureSelect (Agilent)	Illumina HiSeq 2000	N/A	N/A	N/A	<i>NOTCH1</i>	BWA, Bowtie, VCFtools, MuTect30, SomaticSniper, ANNOVAR, COSMIC

TU: amostra de tecido tumoral; **NL:** amostra pareada de sangue periférico; **NAd:** amostra pareada de tecido normal adjacente; **HPV-:** negativo para papilomavirus humano; **HPV+:** positivo para papilomavirus humano; **FFPE:** *formalin-fixed paraffin-embedded* tissue (tecido tumoral em parafina). **IPTofICGC:** *India Project Team of International Cancer Genome Consortium*; **COCE:** Carcinoma oral de células escamosas. ***estudo envolvendo pacientes jovens. **softwares citados pelos artigos referentes a etapa de análise dos dados de sequenciamento do exoma. I: intervalo de faixa etária, em anos. M: média de idade, em anos.

Os dados obtidos desses estudos confirmam a natureza heterogênea e complexa dos CCECP, com elevado número de mutações somáticas diferentes e em diversos genes. Além de confirmar o padrão de interação alterado em certas vias de sinalização, juntamente com anormalidades cromossômicas e mutações bem estabelecidas em genes previamente conhecidos em CCECP como *TP53*, *CDKN2A*, *PIK3CA*, *PTEN* e *HRAS*, a técnica de SET permitiu a identificação de novos genes como *FBXW7*, *NOTCH1*, *IRF6* e *TP63* e seu envolvimento fundamental na patogênese de CCECP (AGRAWAL et al. 2011; LEEMANS et al. 2011; STRANKY et al. 2011; PICKERING et al. 2013, TABATABAEIFAR et al. 2014; YAP et al. 2015).

AGRAWAL et al. (2011) utilizaram a metodologia de SET em 32 amostras CCECP, incluindo os subsítios da cavidade oral (n=25), orofaringe (n=4), hipofaringe (n=1) e laringe (n=2). Além dos genes mutados previamente descritos em CCECP como *TP53*, *CDKN2A*, *PIK3CA* e *HRAS*, os autores também relataram mutações em *FBXW7* e *NOTCH1*. O gene *NOTCH1* está envolvido na regulação da diferenciação de células escamosa (NGUYEN et al. 2006) e foi identificado entre 11% a 19,3% dos tumores, em artigos publicados posteriormente a AGRAWAL et al. (2011). Aproximadamente 40% das 28 mutações identificadas no gene *NOTCH1* produziam produtos gênicos truncados, sugerindo o papel deste gene como um supressor tumoral. Outra observação importante foi que apenas 18% dos 28 tumores não associados à infecção pelo HPV tiveram mutações ativadoras em oncogenes, contrastando com os 89% que possuíam

mutações inativadoras em genes supressores tumorais. Esta distinção é fundamental, porque a maior parte da nova geração de terapias é direcionada para oncogenes ativados (AGRAWAL et al. 2011). Mutações no gene *FBXW7* foram identificadas em 22% dos tumores HPV-positivos por LECHNER et al. (2013).

O segundo estudo descrito em CCECP foi relatado por STRANSKY et al. (2011). Os autores utilizaram a técnica SET em 74 amostras pareadas (tumor e sangue periférico) englobando a maioria dos sítios anatômicos de CCECP (cavidade oral, orofaringe, hipofaringe, laringe e cavidade sinusal). Foram identificados genes previamente descritos em CCECP (*TP53*, *CDKN2A*, *PTEN*, *PIK3CA* e *HRAS*) e pelo menos 30% das mutações encontradas ocorriam em genes que regulam a diferenciação celular (por exemplo, *NOTCH1*, *IRF6* e *TP63*). Estes achados sugeriram que a desregulação do processo de diferenciação é um dos principais iniciadores da carcinogênese em cabeça e pescoço. Posteriormente, outros estudos foram publicados confirmando as alterações já relatadas em CCECP e evidenciando o potencial da metodologia para identificar alvos terapêuticos.

Um estudo piloto foi publicado em dois carcinomas de cabeça e pescoço (um paciente do sexo masculino, idade de 49 anos, não fumante, tumor de tonsila T2N0, HPV16+ *versus* um paciente do sexo feminino, idade de 81 anos, fumante, carcinoma oral T2N0 e negativo para HPV 16) (NICHOLS et al.2012a). Foram realizadas as análises de SET pareadas entre tumor e sangue destes pacientes e os autores relataram que o caso HPV+ apresentava 58 mutações e o caso negativo, 73. Foram verificadas

múltiplas mutações em genes *zinc fingers* e genes da mucina (*MUC 4, 6 12, e 16*). Posteriormente, o mesmo grupo relatou a análise por sequenciamento de 535 genes associados com câncer em linhagens celulares de tumores cabeça e pescoço HPV negativas (*CAL27, DETROIT562, FaDu, SCC4, SCC15 e SCC25*) (NICHOLS et al. 2012b). Os autores relataram uma média de 38 mutações para cada linhagem celular. Todas as linhagens apresentavam mutações em *TP53* e *CDKN2A* (exceto *SCC15*) (NICHOLS et al. 2012b).

A análise dos dados de SET em 151 amostras de CCECP, realizada por LUI et al. 2013, revelou que a via *PI3K (phosphoinositide 3-kinase)* como a mais frequentemente alterada (30,5%) nestes tumores, com uma taxa de mutações significativamente maior. Todos os tumores com mutação simultânea de vários genes da via *PI3K* eram de estádios avançados (IV), sugerindo que alterações nesta via estão relacionadas à agressividade dos CCECP. Mutações em genes desta via podem também revelar biomarcadores preditivos para a seleção do tratamento (LUI et al.2013).

Em 2014, LI et al. compararam os dados genômicos de 39 linhagens celulares de carcinoma epidermóide com 106 CCECP. Amplificações em oito genes (*PIK3CA, EGFR, CCND2, KDM5A, ERBB2, PMS1, FGFR1 e WHSCIL1*) e deleções em cinco genes (*CDKN2A, SMAD4, NOTCH2, ARN e TRIM33*) foram encontrados tanto em linhagens celulares derivadas de CCECP quanto nos tumores analisados. Dezesete genes com mutação foram observados apenas em linhagens celulares de carcinomas epidermóides (>10%), sugerindo que estas mutações poderiam surgir pela

imortalização das células para o estabelecimento das linhagens. Por outro lado 11 genes com mutação foram observados em >10 % dos CCECP. Foram relatados vários genes alterados da via EGFR nas linhagens celulares e nos tumores. Foi observado o perfil farmacológico de oito agentes anti-tumorais em seis linhagens de tumores de cabeça e pescoço sugerindo que a mutação em *PIK3CA* pode servir como um biomarcador preditivo para tratamento que visam a via *EGFR/PI3K*. Estes resultados sugeriram uma correlação entre as mutações genéticas em linhagens celulares e tumores humanos dando uma nova visão sobre a seleção de modelos pré-clínicos para CCECP (LI et al. 2014).

LI R et al. (2015) avaliaram os dados clínico-patológicos de 89 pacientes com CCECP e selecionaram seis tumores de língua de pacientes não fumantes para sequenciamento do exoma (*Agilent's SureSelect Paired-EndVersion 2.0 Human Exome Kit*). Os dados do SET desses pacientes foram comparados com os obtidos em cinco amostras de pacientes fumantes de um estudo prévio de AGRAWAL et al. (2011). Entre os não fumantes foram observadas mutações recorrentes nos genes *CTNNA3*, *EIF3A*, *EP300*, *FXR1*, *NEK8*, *NOTCH1* (*NOTCH2* e *NOTCH3*), *PIK3CA*, *PKHD1L1*, *PTCHD2*, *RALGAPB*, *SPEN* e *UBR4*. As mutações recorrentes detectadas em fumantes foram nos genes *CD101*, *LAMA1*, *NCAPD3*, *RIMBP2*, *SI*, *SYNE* e *TP53*. Todos os tumores de pacientes fumantes (n=5) apresentaram mutações em *TP53* enquanto que entre os não fumantes apenas um de seis casos apresentou mutações em *TP53*. Não houve diferença estatística significativa entre os resultados obtidos em fumantes e

não fumantes, porém o número amostral desse estudo foi muito pequeno.

O *India Project Team of the International Cancer Genome Consortium* (2013) avaliaram o exoma de 50 carcinomas de células escamosas de gengiva bucal (alveolares) (CEB-GB) pareados com o sangue periférico do mesmo paciente (IlluminaHiSeq 2500 e Roche GS-FLX). Foram identificados genes significativamente alterados, incluindo *USP9X*, *MLL4*, *ARID2*, *UNC13C* e *TRPM3*. Para a confirmação desses dados, foi realizado o sequenciamento destes genes em 60 em amostras pareadas N e T independentes. Os autores relataram frequências similares das mutações envolvendo estes cinco genes. Na análise de SET foram relatados genes já descritos como alterados em CCECP de outros sítios anatômicos (*TP53*, *FAT1*, *CASP8*, *HRAS* e *NOTCH1*). Também foram encontrados genes com amplificações recorrentes (por exemplo, *Drosha*, *YAP1*) ou deleções (por exemplo, *DDX3X*) em CEB-GB. Foi encontrada uma alta proporção de transversões C>G e elevada taxa de mutações entre os usuários de tabaco. Muitas vias enriquecidas para as alterações genômicas foram específicas para CEB-GB. Este estudo revelou subtipos moleculares com perfis distintos de mutações, como pacientes que apresentam mutações em *CASP8* com ou sem mutações em *FAT1* (INDIA PROJECT TEAM OF THE INTERNATIONAL CANCER GENOME CONSORTIUM 2013).

Os dados do The Cancer Genome Atlas (TCGA) em 2015 avaliaram 279 amostras de pacientes com CCECP e revelam que, em média, ocorrem 140 mutações por tumor. Nesse estudo, as mutações ocorreram em quatro vias reguladoras, incluindo: controle de ciclo celular (96%), crescimento e

proliferação (62%), apoptose (44%) e diferenciação escamosa (64%) (PICKERING et al. 2013; TCGA 2015). É importante ressaltar que para cada uma dessas vias regulatórias, há opções terapêuticas em desenvolvimento ou em fase de ensaio clínico (BIRKELAND e BRENNER 2015).

Um dos aspectos de grande interesse em ser elucidado por SET é o perfil mutacional em tumores HPV positivos comparado com HPV negativos. Além do *status* de infecção pelo HPV, dados recentes de SET em CCECP, identificaram múltiplos genes comumente alterados e vias envolvidas neste processo. De forma geral, tumores HPV positivos apresentam uma menor taxa de mutações por tumor, o gene *TP53* raramente é mutado assim como é observado a perda de função em p16INK4a quando comparados a tumores HPV negativos, o que reflete diferenças biológicas nesses tumores (AGRAWAL et al. 2011; STRANSKY et al. 2011; KANDOTH et al. 2013).

Em um estudo mais recente, CHUNG et al. (2015) avaliaram 617 genes associados ao câncer, descritos em trabalhos previamente publicados, em tumores fixados em formalina e em blocos de parafina. O propósito do estudo foi comparar se os dados obtidos eram compatíveis com os resultados obtidos a partir de SET de amostras de tecido congelado, uma vez que na prática clínica, muitos laboratórios tem acesso apenas a esse tipo de material. Foi observada positividade para o HPV em uma das amostras. Os dados corroboraram com os resultados disponíveis nos bancos de dados do TCGA, FoundationOne/Foundation Medicine, (Cambridge, MA) e University of Chicago dataset (Chicago cohort). O perfil genômico de alteração é diferente entre pacientes HPV positivos e HPV

negativos, porém em ambos houve predomínio de alterações em genes supressores tumorais em relação aos oncogenes. A via PI3K foi a principal via alterada em amostras HPV positivas, incluindo alterações em *PTEN* (15%), *AKT1* (5%), *RICTOR* (4%), *mTOR* (2%), *AKT2* (2%), *PIK3R1* (2%) e *PIK3CA*. Em tumores HPV negativos, as principais vias alteradas foram reparo de DNA- TP53, ciclo celular e PIK3. Os resultados obtidos mostraram que o perfil genômico de tumores em parafina obtidos a partir da rotina clínica, assemelham-se ao perfil genômico de alterações de tumores de tecido congelado.

Embora muitas alterações tenham sido descritas nestes tumores, é necessária a melhor compreensão da biologia tumoral desta doença tão distinta em nível molecular, para potencializar o desenvolvimento de terapias individualizadas e alvo dirigidas. Outro aspecto a ser considerado é que as alterações mais frequentes encontradas por NGS em CCECP são em genes supressores tumorais (GST) e não em oncogenes. Isso representa uma dificuldade do ponto de vista terapêutico devido ao fato de que, em geral, é mais difícil interferir na perda de função de GST comparado a inibição do aumento da atividade em oncogenes. Entre os 15 genes frequentemente alterados em CCECP (*TP53*, *NOTCH1*, *CDKN2A*, *PIK3CA*, *FBXW7*, *HRAS*, *SYNE1*, *FAT1*, *MLL2*, *CASP8*, *PTEN*, *NSD1*, *MLL3*, *EP300* e *AJUBA*), dois são identificados como oncogenes (*PIK3CA* e *HRAS*) e apenas *PIK3CA* é atualmente considerado como candidato a alvo terapêutico (KANG et al. 2015). Além disso, a implementação de terapia personalizada para pacientes com CCECP é um processo lento quando comparado a alguns tumores

devido a alta complexidade genômica e alta frequência de mutações como demonstrado anteriormente (The Cancer Genome Atlas Network 2015). Entretanto, à medida que são desenvolvidas técnicas mais robustas que possibilitam o rastreamento de alterações e contribuem para identificar mutações em CCECP, aumentam as chances de entendimento da doença. Infelizmente a complexidade e variedade de alterações, dificultam o desenvolvimento de um único agente terapêutico para ser utilizado no tratamento de CCECP, exigindo em um futuro próximo a incorporação de vários agentes e terapia multimodal (LI et al. 2014; SHALEM et al. 2014, WANG et al. 2014; CHUNG et al. 2015; SUN e CALIFANO 2014; BIRKELAND e BRENNER 2015; KANG et al. 2015; TINHOVER et al. 2016).

1.4 PACIENTES JOVENS E CCECP

Não há consenso em literatura a respeito da etiologia, curso clínico e prognóstico do CCECP de pacientes jovens. Estudos em pacientes jovens com CCECP são escassos, principalmente aqueles enfocando os aspectos moleculares.

Além do aumento da incidência de CCECP em pacientes com idades inferiores a 45 anos, observa-se que em pacientes mais jovens, a doença ocorre na ausência de fatores de riscos típicos para este tipo de tumor, como o consumo de álcool, tabaco e infecção pelo HPV (PATEL et al. 2011; VAN MONSJOU et al. 2013; MAJCHRZAK et al. 2014; ADEEL et al. 2016). Além disso, a maioria dos jovens não tabagistas e não etilistas são mulheres (VAN

MONSJOU et al. 2013).

A exposição aos fatores de risco pode ser similar em pacientes jovens e adultos, mas a questão é se essa exposição é longa o suficiente para tornar esses agentes etiológicos os principais nos casos de pacientes jovens. Alterações germinativas e doenças que predisõem ao desenvolvimento de câncer (como a Anemia de Fanconi, Disqueratose Congênita, Síndrome de Bloom, Xeroderma Pigmentoso) e casos de transplantes de medula óssea foram identificados como fatores causais em um subgrupo de pacientes jovens. O verdadeiro impacto da idade e o comportamento clínico ainda tem que ser melhor elucidado (VAN MONSJOU et al. 2013; MAJCHRZAK et al. 2014).

São muitas as discordâncias em relação ao prognóstico dos pacientes jovens. KURIAKOSE et al. (1992), SOUDRY et al. (2010), HILLY et al. (2013), MAJCHRZAK et al.(2014) entre outros estudos sugeriram que a doença em jovens tem o diagnóstico realizado em estágios mais avançados e, portanto, o curso clínico da doença é mais agressivo e há maior taxa de recorrência em pacientes jovens quando comparados aos casos de CCECP em pacientes com idade mais avançada. Outros estudos reportaram que o prognóstico é similar entre os dois grupos (MYERS et al. 2000; PITMAN et al. 2000; MANUEL et al. 2003). HYAM et al. (2003) relataram que a sobrevida global foi menor em pacientes com idade superior a 60 anos, embora a maioria dos jovens tenham apresentado tumores de alto grau. Este achado em pacientes mais velhos pode estar associado com a presença de comorbidades tais como doenças pulmonares ou cardíacas.

Os resultados do estudo realizado por UDEABOR et al. (2012)

mostraram que os pacientes jovens com CCECP podem apresentar um prognóstico melhor, especialmente em termos de sobrevida global. Em contrapartida, MYERS et al. (2000) relataram que quando o carcinoma espinocelular de boca se desenvolve sem os fatores de risco tradicionalmente reconhecidos, a taxa de sobrevida é menor, independentemente da idade.

KAMINAGAKURA et al. (2016) avaliaram a expressão, por imunohistoquímica, de enzimas envolvidas na biotransformação do álcool e tabaco (CYP1A1, CYP1B1, ALDH1A1 e ALDH2) entre 59 pacientes com idade superior a 59 anos e 40 pacientes mais jovens (inferior a 40 anos de idade). A taxa de sobrevida livre de recidiva em pacientes jovens foi relatada como associada com o aumento da expressão de ALDH1A1. Os pacientes com mais de 50 anos apresentaram expressão aumentada de CYP1B1 comparados aos pacientes jovens, fato que parece interferir de forma positiva na melhoria de sobrevida livre de recidiva desses pacientes. Estes resultados sugerem que em células expostas ao tabaco, ocorre um aumento da expressão dessa enzima resultando em maior desintoxicação de substâncias cancerígenas. O tabaco e álcool são metabolizados por meio de uma série de vias, de forma complexa, dinâmica e integrativa com outras vias, tais como vias da proliferação celular, diferenciação celular, apoptose e biotransformação de outras substâncias. Tais processos metabólicos envolvem um grande número de proteínas que podem desempenhar outras funções além de metabolização. Novos estudos envolvendo outros genes e proteínas são necessárias a fim de complementar os resultados desta pesquisa.

Até o momento e para o nosso conhecimento, apenas PICKERING et al. (2014a) utilizaram a metodologia de SET com o foco de detalhar a diferença molecular entre os pacientes com CCECP que desenvolveram a doença em idade mais jovem. Os autores realizaram o sequenciamento de alto desempenho em 16 pacientes jovens com tumores de língua (pareados com amostras de tecido normal) assim como de 28 pacientes com idade superior a 45 anos. Foi detectada uma média de 28 e 83 mutações nos tumores de pacientes jovens e mais velhos, respectivamente. A variação do número de mutações não foi surpreendente, pois é dependente da idade e tabagismo. Embora considerado um grupo pequeno e sem significância estatística, entre os pacientes jovens foi observada maior frequência de mutações em *TP53* (93,8%) enquanto no grupo de pacientes mais velhos esta frequência foi de 53%. Os genes *FAT1*, *TP53* e *PIK3CA* mostraram tendência à significância estatística comparando as amostras de pacientes jovens e mais velhos, embora a frequência de mutações tenha sido maior em pacientes mais velhos. Os genes com maior frequência de mutações encontrados nos pacientes jovens foram *TP53*, *NOTCH1*, *CDKN2A*, *CASP8* e *FAT1* (PICKERING et al. 2014a). Entre os resultados apresentados, as alterações genômicas observadas não foram capazes de explicar o aumento da incidência de tumor de língua em pacientes jovens.

Os resultados obtidos até o momento utilizando dados de SET indicam a necessidade de estudos adicionais que explorem os aspectos envolvidos no desenvolvimento de CCECP em pacientes jovens.

2 OBJETIVOS

2.1 OBJETIVO GERAL

Investigar o perfil de alterações somáticas em carcinomas orais e de orofaringe de pacientes jovens usando sequenciamento total do exoma.

2.2 OBJETIVOS ESPECÍFICOS

- Avaliar o perfil de alteração genética em amostras pareadas de tecido normal (leucócitos) e tumoral de pacientes com carcinomas orais e de orofaringe e comparar os dados gerados com aqueles depositados em bancos de dados públicos;
- Identificar e comparar as alterações genéticas detectadas segundo a localização anatômica (oral e orofaringe), o consumo de álcool, uso de tabaco, infecção pelo vírus do papiloma humano (HPV) e outros dados clínicos e histopatológicos;
- Comparar os resultados obtidos com aqueles identificados em pacientes com câncer de mesma localização com idade inferior a 45 anos e superior a 60 anos disponíveis no TCGA;
- Definir uma assinatura genética mutacional para pacientes com câncer oral e de orofaringe em idade mais jovem do que a frequentemente observada na prática clínica (maior ou igual a 60 anos).

3 MATERIAL E MÉTODOS

3.1 CARACTERIZAÇÃO DAS AMOSTRAS

Ao delineamento deste estudo, o Banco de Tumores do A.C.Camargo Cancer Center, SP (AC Project) contava com 2300 casos com tecido congelado de tumores primários de cabeça e pescoço, registrados pelo Departamento de Cirurgia de Cabeça e Pescoço. Entre estes casos, 523 eram de pacientes com idade inferior a 49 anos englobando todas as topografias (incluindo casos de tumores em tireóide e outras topografias da cabeça e pescoço). Após uma criteriosa análise foram selecionados amostras de tecido normal e tumoral de 35 pacientes atendidos na Instituição entre 1995 a 2014. O estudo foi aprovado pelo Comitê de Ética em Pesquisa (CEP) da Fundação Antônio Prudente (FAP), sob o parecer N° 1727/12 (Anexo 1).

Todos os pacientes foram tratados e acompanhados pela equipe do Departamento de Cirurgia de Cabeça e Pescoço da Instituição. Os registros médicos de todos os pacientes foram examinados para a obtenção de informações clínicas e histopatológicas segundo a ferramenta de visualização integrada de prontuários do hospital, H²TC - A.C.Camargo Cancer Center (<http://intranet.cipe.accamargo.org.br/h2tc/>). Estas informações foram confirmadas pela equipe do Departamento de Cirurgia de Cabeça e Pescoço da Instituição.

3.2 PACIENTES

Foram selecionados para este estudo os casos com disponibilidade de material (DNA tumoral e de sangue periférico) no banco de macromoléculas da Instituição e com qualidade necessária para os experimentos. Foram excluídos os pacientes que apresentavam história prévia de câncer, presença de outro câncer sincrônico ou metacrônico e ou eram portadores de síndromes que aumentam o risco de desenvolvimento de CCECP como a Anemia de Fanconi, Disqueratose Congênita, Xeroderma Pigmentoso, Ataxia Telangiectasia, Síndrome de Bloom e Síndrome de Li-Fraumeni.

No total, foram incluídos no estudo 35 pacientes com informações clínicas e patológicas completas sendo 24 casos de carcinomas de cavidade oral (mucosa bucal, triângulo retromolar, palato duro 2/3 anteriores da língua, soalho de boca e região jugal) e 11 de orofaringe (base de língua, tonsila, palato mole e loja amigdalina). Os pacientes selecionados tinham idade igual ou inferior a 49 anos ao diagnóstico (variação de 20 a 49 anos, média de 40,14 anos) e não foram submetidos a tratamento quimioterápico ou radioterápico prévios à obtenção das amostras. A maioria dos pacientes é do sexo masculino (N=25; 71,4%), consumidor de álcool (N=21; 60%) e tabaco (N=19; 54,2%).

Os dados clínicos e histopatológicos dos casos incluídos no estudo estão resumidos na Tabela 1.

Tabela 1 - Descrição das características clínicas e histopatológicas dos pacientes incluídos no estudo.

Variável	Carcinomas Cavidade Oral N= 24 (%)	Carcinomas Orofaringe N=11 (%)	Total Pacientes N=35 (%)
Sexo			
Feminino	8 (33,3)	2 (18,2)	10 (28,6)
Masculino	16 (66,7)	9 (81,8)	25 (71,4)
Tabagismo			
Sim	11 (44)	8 (73)	19(54,2)
Não	13 (56)	2 (18)	15 (42,8)
NI	-	1 (9)	1 (3)
Etilismo			
Sim	15 (62,5)	6 (54,5)	21 (60)
Não	9 (37,5)	4 (36)	13 (37)
NI	-	1 (9,5)	1 (3)
HPV			
Positivo	7 (29,1)	5 (45,5)	12 (34,3)
Negativo	17 (70,9)	6 (54,5)	23 (65,7)
Evolução clínica			
Vivo	13 (54)	6 (54,5)	19(54)
Óbito pela doença	9 (37,5)	4 (36)	13 (37)
Perda de seguimento	2 (8,5)	1 (9,5)	3 (9)
Tamanho Tumoral (T)			
1-2	14 (58,3)	3 (27)	17 (48,5)
3-4	10 (41,7)	8 (73)	18 (51,5)
Comprometimento de Linfonodos (N)			
0	12 (50)	4 (36)	16 (45,7)
1-2	12 (50)	5 (46)	17 (48,6)
3	-	2 (18)	2 (5,7)
Invasão perineural			
Presente	11 (45,8)	3 (27)	14 (40)
Não detectada	13 (54,2)	8 (73)	21 (60)
Invasão angiolinfática			
Presente	4 (16,7)	-	4(12)
Não detectada	20 (83,3)	11 (100)	31(88)
Recidiva			
Sim	11 (45,8)	3 (27)	14 (40)
Não	13 (54,2)	8 (73)	21 (60)
Segundo tumor primário			
Sim	1 (4)	2 (18)	3(8,5)
Não	23 (96)	9 (82)	32 (91,5)
Metástase à distância			
Sim (pulmonar e cervical)	6 (25)	1 (9)	7(20)
Não	18 (75)	10 (91%)	28 (80)

3.3 EXTRAÇÃO DO DNA GENÔMICO

O procedimento de extração do DNA foi realizado pela equipe técnica do Banco de Macromoléculas do A.C.Camargo Cancer Center, SP. O DNA do tecido tumoral e de linfócitos de sangue periférico foi extraído utilizando o *Gentra Puregene Tissue Kit* (Qiagen) e o *QIACUBE DNA Blood kit* (Qiagen), respectivamente, segundo as recomendações do fabricante. Posteriormente o DNA foi armazenado a -20°C, quantificado no sistema *NanoDrop ND 1000* e a qualidade verificada utilizando o *Bioanalyzer 2100* (Agilent Technologies).

3.4 GENOTIPAGEM DO PAPILOMAVIRUS HUMANO (HPV)

A genotipagem do HPV foi realizada utilizando o kit *LINEAR ARRAY HPV Genotyping Test* (Roche Molecular Systems, Inc., Branchburg, NJ, USA). Este método é baseado na amplificação do DNA alvo pela PCR utilizando iniciadores para mais de 37 subtipos de HPV, seguido de hibridação dos produtos amplificados com sondas de oligonucleotídeos e detecção dos produtos amplificados e ligados às sondas, por determinação colorimétrica. A detecção e a determinação do genótipo foram efetuadas usando o DNA amplificado desnaturado e uma matriz de sondas de oligonucleotídeos, a qual permite a identificação dos genótipos do HPV. A amplificação de um grande espectro de genótipos do HPV, incluindo 13 de alto risco (16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59 e 68), utiliza

iniciadores biotinilados em uma região conservada do HPV. Para proporcionar um controle adequado da qualidade do DNA, um par de iniciadores adicional para o gene da β -globina humana é adicionado à mistura de iniciadores. Todos os procedimentos seguiram as recomendações do fabricante.

A amplificação das amostras tumorais foi realizada em uma reação contendo 50 μ L da amostra tumoral (300ng) 1U de DNA polimerase AmpliTaq Gold, Mg^{2+} , deoxinucleosídeo 5'-trifosfatos (dNTPs) e iniciadores biotinilados para a amplificação do gene da beta-globina (268pb) e das regiões específicas de HPV (450pb). Foram também incluídos na reação o trifosfato deoxiuridina (dUTP) e a enzima AmpErase uracil-N-glicosilase (UNG). As amostras foram mantidas em termociclador por 2 min a 55°C para ação da enzima UNG com a finalidade de remover as bases uracil de qualquer produto de amplificação contaminante existente na mistura de reação. Após esse período foi realizada a desnaturação da fita de DNA a 95°C por 9 min e a reação foi submetida a 40 ciclos de 95°C por 30 min, 55°C por 1min e 72°C por 1min. Foi realizado o aquecimento a 72°C por 5 min e após a amplificação foi acrescido ao produto uma solução de desnaturação, levando a formação de fitas simples. O produto amplificado foi adicionado a bandejas contendo tiras de genotipagem revestidas com sondas dos diferentes tipos de HPV. Os procedimentos de hibridação, lavagem e detecção foram realizados de acordo com as especificações do fabricante. O resultado foi interpretado com o auxílio do Guia de Referência do Teste de Genotipagem *LINEAR ARRAY HPV*. A Figura 1 ilustra o método

de análise dos resultados.

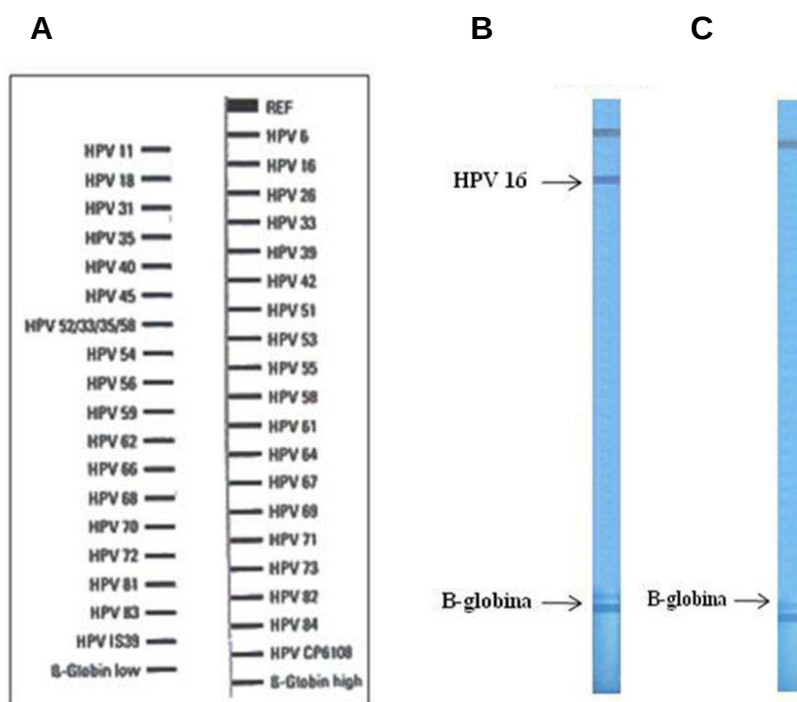


Figura 1 - Interpretação dos resultados provenientes da genotipagem do HPV. **A.** Mapa utilizado para identificar os diferentes genótipos. **B.** Resultado de um caso HPV 16 e positividade para a B-globina; **C.** caso HPV negativo e positividade para a B-globina.

3.5 SEQUENCIAMENTO TOTAL DO EXOMA

O sequenciamento de alto desempenho foi realizado no equipamento HiSeq 2500 (Illumina), Centro de Genômica Funcional da ESALQ – USP, Piracicaba, São Paulo.

O sequenciador HiSeq2500 é capaz de gerar fragmentos de 200 pb e cerca de 300 Gb de dados, por flow-cell (lâmina de sequenciamento), em uma única corrida (http://www.illumina.com/systems/hiseq_2000_1000.ilmn). Cada *flow-cell* é dividida em oito *lanes* (canaletas), cuja capacidade de

geração de dados é de aproximadamente 37Gb cada. Com estas informações é possível calcular e estimar a cobertura das sequências. Para o preparo das bibliotecas e realização do sequenciamento foi utilizado o *Nextera Exome Enrichment* (Illumina, San Diego, California, USA), o qual permite a captura 62Mb do genoma. O procedimento inclui a fragmentação do DNA e a ligação dos adaptadores usando enzimas transposases (denominado tagmentação), um procedimento que diminui o tempo de processamento e pode ser finalizado em dois dias e meio.

3.5.1 Construção das Bibliotecas de Exoma

As amostras de DNA de tecido tumoral (n=35) e sangue periférico (n=35) foram diluídas para uma concentração final de 5ng/μL em um volume de 10 μL, de acordo com as recomendações do fabricante (*Nextera Rapid Capture Guide, Part#15037436 Rev.A*). Para essa diluição as amostras foram quantificadas por método fluorimétrico específico e sensível para quantificar a dupla fita de DNA (*QubitdsDNA BR Assay kit*). Para cada amostra foi utilizado 1μL de *dsDNA BR Reagent Reagent* diluído em 199μL de tampão fornecido pelo kit e 2μL da amostra. A concentração foi aferida usando o *Qubit® Fluorometer 2.0* (Thermo Fisher Scientific, Waltham, MA, USA) (de acordo como as instruções do fabricante disponível em https://tools.thermofisher.com/content/sfs/manuals/Qubit_dsDNA_BR_Assay_UG.pdf).

As etapas de construção das bibliotecas de exoma foram: 1) fragmentação do DNA genômico pelo uso das transposases e ligação dos

adaptadores; 2) limpeza do produto da fragmentação; 3) amplificação pela PCR convencional; 4) limpeza do produto da PCR; 5) duas hibridações de 16 horas e posterior lavagem das bibliotecas; 6) amplificação das bibliotecas pela PCR convencional; 7) validação das bibliotecas por meio de PCR quantitativa (qPCR). As etapas de limpeza do produto das etapas de fragmentação, PCR e bibliotecas foram feitas utilizando *beads* magnéticas. A quantificação da biblioteca foi realizada pela PCR quantitativa em tempo real (qPCR) para detectar pequenas concentrações (pM–picomol) das amostras que foram sequenciadas posteriormente. As etapas da construção da biblioteca estão descritas em detalhes abaixo:

- 1 *Fragmentação do DNA, ligação e tagmentação de adaptadores:* Esse processo consiste na fragmentação do DNA genômico ao mesmo tempo em que são adicionados adaptadores nas extremidades, pelas enzimas transposases, permitindo a amplificação pela PCR nos processos subsequentes. A fragmentação do DNA e a ligação dos adaptadores ocorreram a partir da adição de 25µL de *Tagment DNA Buffer* e 15µL de *Tagment DNA Enzyme*, 50ng de DNA genômico (5ng/µL). As amostras foram homogeneizadas a 1800 rpm por um minuto e centrifugadas a 280xg por mais um minuto. Após centrifugação, as amostras foram incubadas a 58°C por 10 minutos e posteriormente foram adicionados 15µL de *Stop Tagment Buffer*, que inibe a ligação dos adaptadores. Em seguida, procedeu-se a homogeneização e centrifugação das amostras a 1800rpm e 280xg, respectivamente, por um minuto e posterior incubação à temperatura

ambiente por quatro minutos.

- 2 *Limpeza do produto tagmentado:* Este processo purifica o DNA tagmentado do transpossomos (Nextera). É uma etapa crítica porque quando não removidos, os transpossomos (Nextera) podem se ligar fortemente às extremidades de DNA e interferir nos processos subsequentes. Para a sua adequada remoção, 65µL de *Sample Purification Beads* (SPB) foram adicionados ao produto tagmentado, sendo então homogeneizada a reação. Em seguida, o produto foi incubado por oito minutos à temperatura ambiente e por dois minutos (ou até o sobrenadante ficar claro) em raque magnética. Foram descartados 130µL do sobrenadante e, sem perturbar as *beads*, foram adicionados 200µL de etanol 80% (EtOH 80%). Após incubação por 30 segundos, o etanol foi removido. Foi realizada mais uma lavagem com EtOH 80% nas mesmas condições. Após a remoção do sobrenadante, as amostras foram incubadas por 10 minutos à temperatura ambiente na raque magnética para secagem das *beads*. O DNA tagmentado purificado foi ressuspendido em 22,5µL de *Resuspension Buffer* (RSB). Em seguida, a reação foi incubada por dois minutos à temperatura ambiente e dois minutos (ou até o sobrenadante ficar claro) na raque magnética. Foram transferidos 20µL do sobrenadante para novos tubos de PCR.
- 3 *Primeira amplificação pela PCR convencional:* Este processo amplifica o DNA tagmentado e purificado pela PCR sendo adicionados os adaptadores “*Index 1*” e “*Index 2*” necessários para a

identificação das amostras na etapa de sequenciamento. Nesta etapa foram adicionados 5µL de *Index 1* em todas amostras, 5µL de *Index 2* específico para cada amostra e 20µL de *Nextera Library Amplification Mix*. O produto foi homogeneizado a 1200 rpm por um minuto e centrifugado a 280xg por um minuto. As condições de ciclagem foram três minutos a 72°C, 30 segundos a 98°C seguidos por 10 ciclos de 10 segundos a 98°C, 30 segundos a 60°C e 30 segundos a 72°C e extensão final por cinco minutos a 72°C. A reação foi realizada no termociclador Veriti 96-Well Thermal Cycler (Thermo Fisher Scientific).

- 4 *Limpeza do produto amplificado pela PCR:* A biblioteca de DNA foi purificada com *Sample Purification Beads (SPB)* para remoção de produtos indesejáveis. As amostras foram removidas do termociclador e centrifugadas por um minuto a 280xg. Foram adicionados 90µL de SPB aos 50µL de amostra e foram homogeneizados por um minuto a 1800 rpm. As amostras foram incubadas a temperatura ambiente por 10 minutos e em seguida por dois minutos (ou até o sobrenadante ficar claro) em estante magnética. Após a remoção do sobrenadante, foram adicionados 200 µL de EtOH 80% (sem perturbar as *beads*), incubados por 30 segundos e descartado o sobrenadante, com os tubos ainda na estante magnética. A etapa de adição de EtOH 80% foi repetida. Após a retirada do etanol, os tubos foram mantidos a temperatura ambiente em estante magnética por 10 minutos para secagem das *beads*. O DNA amplificado e purificado foi

ressuspendido em 27 μ L de RSB e incubados a temperatura ambiente por dois minutos e em estante magnética por dois minutos (ou até o sobrenadante ficar claro). Foram transferidos 25 μ L do sobrenadante para um novo tubo para a quantificação da biblioteca pelo método fluorimétrico Qubit.

- 5 *Primeira hibridação:* Nesta etapa a biblioteca de DNA foi hibridada com sondas biotiniladas responsáveis por capturar as regiões alvos de interesse (exons). As amostras na concentração de 500ng foram combinadas em “*pools*”. Cada *pool* contém oito amostras em um volume final de 40 μ L. Foram adicionados 50 μ L de *Enrichment Hybridization Buffer* (EHB) e 10 μ L de *Expanded Exome Oligos* (EEX) em cada *pool* de amostras. Após homogeneização (1200rpm por um minuto) e centrifugação (280xg por um minuto), o *pool* de amostras foi incubado em termociclador (Veriti 96-Well Thermal Cycler, Thermo Fisher Scientific) a 95°C por 10 minutos, seguido de 18 ciclos de um minuto de incubação por temperatura, começando a 94°C e diminuindo 2°C por ciclo até chegar aos 58°C, e permaneceu nesta temperatura entre 16 e 18 horas. Este tempo de hibridação recomendado garante que as regiões alvo se liguem completamente às sondas de captura.
- 6 *Primeira captura:* Este processo usa *beads* marcadas com estreptavidina para capturar as sondas biotiniladas hibridadas para as regiões alvo de interesse. Dois procedimentos de lavagem aquecida foram utilizados para que ligações não específicas fossem removidas

e a biblioteca enriquecida fosse eluída e preparada para uma segunda etapa de hibridação. Para isso, após 16-18h, o *pool* de amostras foi retirado do termociclador, centrifugado por um minuto a 280xg e adicionados 250µL de *Streptavidin Magnetic cBeads* (SMB). As amostras foram então homogeneizadas por cinco minutos a 1200rpm, incubadas a temperatura ambiente por 25 minutos e centrifugadas a 280xg por um minuto. O *pool* de amostras foi incubado em estante magnética por dois minutos (ou até o sobrenadante ficar claro) e o sobrenadante descartado. Foram então adicionados 200µL de *Enrichment Wash Solution* (EWS) aos *pools*, que foram homogeneizados a 1800 rpm por quatro minutos. As amostras foram incubadas a 50°C por 30 minutos, e por mais dois minutos (ou até o sobrenadante ficar claro) em estante magnética. Em seguida, o sobrenadante foi descartado. A lavagem com EWS seguido de duas incubações e descarte do sobrenadante foram repetidos mais uma vez. Foram adicionados e homogeneizados ao *pool* de amostras 23 µL de uma solução de pré-eluição (28,5µL de *Enrichment Elution Buffer 1-EE1* e 1,5µL de 2N NaOH) por dois minutos a 1800rpm, seguindo de incubação à temperatura ambiente por dois minutos e por mais dois minutos (ou até o sobrenadante ficar claro) em estante magnética. Vinte e um µl do sobrenadante foram transferidos para um novo tubo de *ependorf* com 4µL de *Elute Target Buffer 2* (ET2), homogeneizados a 1200rpm por um minuto e centrifugados por 280xg por um minuto.

- 7 *Segunda hibridação e captura:* Esta etapa consiste na combinação da biblioteca de DNA eluída na etapa anterior com adição de sondas de captura para regiões marcadas de interesse (éxons). A repetição deste processo assegura alta especificidade das regiões capturadas. Os passos da segunda hibridação e segunda captura são os mesmos descritos nas etapas 5 e 6.
- 8 *Limpeza das bibliotecas enriquecidas:* Foram adicionados 45µL de SPB ao *pool* de amostras, homogeneizados a 1800 rpm por um minuto e centrifugados a 280xg por um minuto. Em seguida, as amostras foram incubadas por 10 minutos em temperatura ambiente e em estante magnética por dois minutos (ou até o sobrenadante ficar claro) sendo em seguida descartado o sobrenadante. Com os tubos ainda na estante magnética, foram adicionados 200 µL de EtOH 80%, incubados por 30 segundos e descartado o sobrenadante. Foi repetida esta etapa de lavagem com EtOH 80%. As bibliotecas foram incubadas por 10 minutos, à temperatura ambiente, para secarem completamente. Em seguida, foram retiradas da estante magnética, ressuspendidas em 27,5µL de RSB e homogeneizadas a 1800 rpm por um minuto. As bibliotecas foram incubadas em temperatura ambiente por dois minutos e em raque magnética por dois minutos (ou até o sobrenadante ficar claro), sendo então transferidos 25µL do sobrenadante para um novo tubo.
- 9 *9) Segunda amplificação pela PCR convencional:* A biblioteca de DNA purificada foi amplificada pela PCR convencional. A cada *pool* de

amostras foram adicionados 20µL de *Nextera Enrichment Amplification Mix* (NEM) e 5µL de *PCR Primer Cocktail* (PPC). As condições de ciclagem, foram 30 segundos a 98°C seguidos de 12 ciclos de 10 segundos a 98°C 30 segundos a 60°C e 30 segundos a 72°C com um tempo de extensão final de cinco minutos a 72°C. A reação foi realizada no termociclador (Veriti 96-Well Thermal Cycler, Thermo Fisher Scientific),

- 10 *Limpeza do produto amplificado pela PCR:* A biblioteca de DNA foi purificada para a remoção de produtos indesejáveis. Nesta etapa foram adicionados 90µL de SPB aos 50µL de bibliotecas e homogeneizados por um minuto a 1800 rpm. As bibliotecas foram incubadas à temperatura ambiente por 10 minutos. Em seguida, os tubos foram transferidos para estante magnética e incubados por dois minutos (ou até o sobrenadante ficar claro). O sobrenadante foi descartado, e ainda em estante magnética, foram adicionados 200µL de EtOH 80% às bibliotecas, incubados por 30 segundos e descartado o sobrenadante novamente. Essa etapa de lavagem foi repetida uma vez e no final, os tubos foram retirados da estante magnética. O DNA amplificado e purificado foi ressuspendido a partir da homogeneização com 32µL de RSB e incubados à temperatura ambiente por dois minutos. Na sequência, as bibliotecas foram transferidas para estante magnética e incubadas por dois minutos (ou até o sobrenadante ficar claro). Após o sobrenadante ficar límpido, foram transferidos 30µL do sobrenadante para novos tubos para

proceder à etapa de sequenciamento do exoma.

3.5.2 Sequenciamento do Exoma

Foi realizada a validação das bibliotecas pela PCR quantitativa (qPCR), utilizando o *KAPA Illumina GA primers-SYBR ABI Prism kit* (Kapa Biosystems, Wilmington, MA, USA). A quantificação das bibliotecas pós-enriquecimento é importante para otimizar a qualidade dos dados obtidos pelo sequenciamento pois permite criar uma densidade ótima de *cluster* em cada *lane*.

Após a validação das bibliotecas pela qPCR, os *pools* de amostras foram incluídos na *flowcell* seguidos de amplificação clonal em fase sólida no equipamento cBot (Illumina, Inc, San Diego, CA, USA). Cada *pool* continha oito amostras, que por sua vez foram colocadas em *lanes* separadamente. Cada *flow-cell* é dividida em oito *lanes* (canaletas).

A capacidade de geração de dados é de aproximadamente 37,5Gb. Com estas informações e com os dados do tamanho da sequência do exoma (62Mb = 0,062Gb baseado no protocolo utilizado neste trabalho), foi estimado uma média de cobertura por amostra de 75x.

Na *flowcell* estão plotadas as sondas correspondentes às sequências contidas nos adaptadores utilizados na construção das bibliotecas. Estas sondas funcionam como iniciadores para que a amplificação em ponte ocorra na etapa de clusterização. Assim, as bibliotecas são agrupadas em *clusters*, ou seja, são produzidas múltiplas cópias de fragmentos específicos por meio da desnaturação da dupla fita de DNA e incorporação de

nucleotídeos na nova cadeia de DNA crescente. Esse princípio da clusterização que ocorre na *flowcell*, é semelhante ao observado em uma PCR convencional. Conforme cada base é incorporada à cadeia crescente de DNA, o *cluster* emite uma fluorescência correspondente à base e assim, sucessivamente até o completo sequenciamento dos fragmentos (200pb em corrida *pair-end*). Após o término da clusterização, os fragmentos são desnaturados de modo que Das de fita única fiquem ligados à *flowcell* para posterior sequenciamento. Os sinais fluorescentes são lidos e interpretados por *software* específico, possibilitando a construção da sequência. Todos os procedimentos seguiram as recomendações do fabricante.

Após o sequenciamento, as bases foram nomeadas(*base-calling*) e os dados extraídos utilizando os programas SAV (*Software AnalysisViewer*; https://support.illumina.com/sequencing/sequencing_software/sequencing_analysis_viewer_sav.html) e CASAVA (https://support.illumina.com/sequencing/sewquencing_software/casava.htm). Após essa etapa, cada caso foi identificado e separado do *pools* de amostras sequenciadas em cada *lane*, correspondendo a etapa denominada de demultiplexação. Os dados foram compactados em 440Gb e foram transferidos e armazenados via *download* para o servidor do Laboratório NeoGene, Faculdade de Medicina, UNESP, Botucatu, para a análise de bioinformática.

3.5.3 Análise do Sequenciamento do Exoma

Considerando que este estudo foi delineado para a análise das alterações somáticas, foi realizada a comparação entre amostras tumorais (T) e normais (N) pareadas, selecionando apenas variantes exclusivas do tumor de cada paciente. A comparação entre N e T é uma estratégia extremamente válida para a identificação de variantes genéticas e candidatos a *drivers* (JONES et al. 2015)

As primeiras etapas do estudo consistiram na análise do controle de qualidade, alinhamento e mapeamento das sequências obtidas. No processo de sequenciamento, as bases foram nomeadas (*base-calling*) e os dados extraídos, compactados e transferidos para o servidor. A limpeza dos dados brutos e a filtragem de contaminantes (como adaptadores e vetores) foram realizadas utilizando o programa *Seqclean* v.1.9.10 (<https://github.com/ibest/seqclean>) e o banco de dados *UniVec* (<ftp://ftp.ncbi.nlm.nih.gov/pub/UniVec/>). A limpeza foi realizada de acordo com a qualidade das bases sequenciadas, comprimento dos *reads* e possíveis moléculas contaminantes. Foram descartados os *reads* com *QScore* abaixo de 26 e/ou 24 e de comprimento inferior a 60 pb. Posteriormente, os *reads* foram mapeados utilizando o algoritmo de alinhamento local do programa *Bowtie2* v.2.1.0 (<http://bowtie-bio.sourceforge.net/bowtie2/>). Esta estratégia permite a obtenção de um maior número de alinhamento dos *reads*, pois as sequências não são descartadas quando não há o pareamento exato em ambas as extremidades. O arquivo gerado foi convertido para o formato “sam”

(*SequenceAlignment/Map*) em sua forma binária “.bam”, o que é necessário para a continuidade do *pipeline*.

Após o alinhamento, as sequências foram ordenadas para a retirada dos *reads* duplicados, que são compostos por uma série de fragmentos que se iniciam e terminam exatamente na mesma base. Os *reads* duplicados podem ocorrer em consequência da amplificação pela PCR realizada durante o preparo das bibliotecas, fragmentação do DNA no mesmo local, artefatos devido a regiões repetitivas e *background* na detecção de *clusters* na fase do sequenciamento. Esta etapa é necessária para que a cobertura das regiões genômicas não seja superestimada e não haja impacto na eficiência de detecção das variantes. Para esta etapa foi utilizado o programa PICARD v.1.101 (<http://broadinstitute.github.io/picard/>).

Em seguida à obtenção dos *reads*, o conjunto de sequências foi nomeado de acordo com sua origem, etapa necessária para entrada no programa GATK (v.3.2.2.gec30cee/<http://gatkforums.broadinstitute.org/gatk/discussion/2806/howtoapply-hard-filters-to-a-call-set>) (MCKENNA et al. 2010). Foram identificadas as variantes de base única e pequenas inserções e deleções (*Indels*) utilizando as ferramentas *mpileup* do programa *Samtools* v.1.2.1 (<http://samtools.sourceforge.net/samtools.shtml>) (LI et al. 2009). Estas variantes foram selecionadas de acordo com a qualidade de base acima de 30 (*Qscore* > 30) e cobertura mínima de 10x por sítio de alteração.

Para gerar as variantes, foram utilizados o *Unified Genotyper* dos programas *GATK* v.3.2.2 (MCKENNA et al. 2010) e *Samtools* v.1.2.1 (LI et al.

2009). A seguir, utilizou-se os comandos de geração das variantes com *Samtools* e filtragens das variantes com o conjunto de ferramentas *VCFtools* v.0.1.12 (<https://vcftools.github.io/index.html>) (DANECEK et al. 2011). Posteriormente, foi realizada a comparação entre as sequências geradas considerando a qualidade de mapeamento e o genoma de referência.

A união entre as variantes provenientes do programa *Samtools* e *GATK* foi realizada após a obtenção dos arquivos VCFs. Para aumentar a confiabilidade na detecção das mutações somáticas foi utilizado o algoritmo MuTect (<http://www.broadinstitute.org/cancer/cga/MuTect>). Esta ferramenta aplica um classificador Bayesiano seguido de seis filtros internos para a detecção de mutações com frações de alelos muito baixas garantindo alta especificidade. As variantes foram anotadas utilizando o programa *Annovar* (versão 14, Dec 2015) (WANG et al. 2010).

A nominação das variantes foi realizada de acordo com os bancos de dados *known Gene* (UCSC), *ensemble Gene* (Ensemble Database), RefSeq, GWAS Catalog, clinvar, cosmic70 1000 Genomas, 6500 Exomas 1000 Genomas, snp138 e dbfnsp (ljb26). O banco de dados dbfnsp, contém as informações de patogenicidade dos seguintes bancos de dados: SIFT, POLYPHEN2_HDIV, POLYPHEN2_HVAR, LRT, MutationTaster, FATHMM, Radial SVM, LR, GERP++, PhyloP, SiPhy. Todos os bancos de dados estão disponíveis para *download* no site do programa *Annovar* (<http://www.openbioinformatics.org/annovar>).

Para a seleção de variantes e para identificar alterações candidatas a *drivers*, foram consideradas apenas aquelas não sinônimas, patogênicas e

presente em mais de 20% dos casos . Essa frequência mínima é um critério utilizado para a identificação de genes *drivers*, como descrito por VOGELSTEIN et al. (2013).

Após a seleção das variantes exclusivas em amostras tumorais, foram feitas comparações entre grupos de pacientes de acordo com os dados clínicos, incluindo o consumo de álcool, hábito tabagista, infecção pelo HPV e o sítio anatômico (cavidade oral e orofaringe). O fluxograma descrevendo as comparações realizadas no estudo está representado na Figura 2.

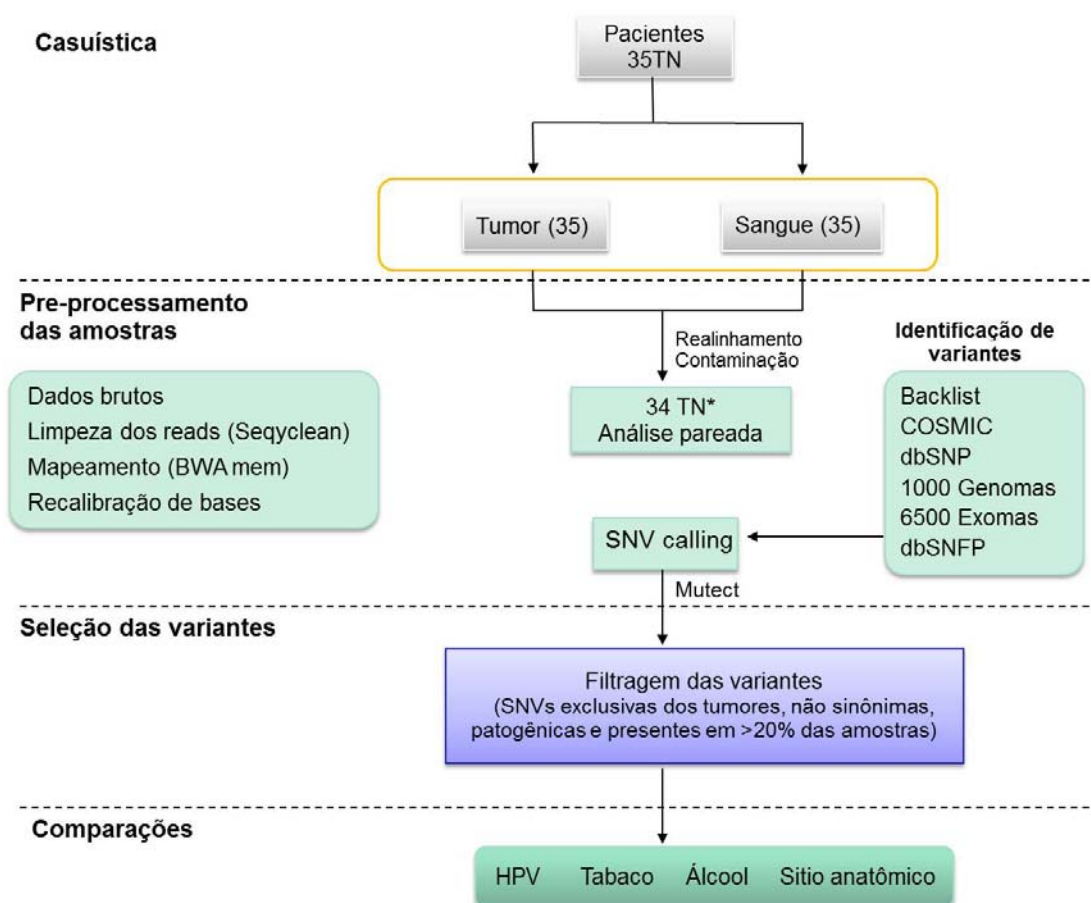


Figura 2 - Fluxograma das etapas realizadas no grupo de 35 pacientes com CCECP. *Após o sequenciamento, uma amostra foi excluída por baixa cobertura.

3.6 SEQUENCIAMENTO DIRETO DO GENE TP53

Amostras pareadas de DNA de sangue periférico e do tumor de cinco pacientes foram selecionadas para serem confirmadas como apresentando alterações no gene *TP53* (detectado na análise de sequenciamento do exoma). O DNA foi amplificado pela PCR foi analisado pelo sequenciamento direto para a confirmação das mutações identificadas nos exons 5/6, 7 e 8 do gene *TP53*. O método consiste na adição de dideoxiribonucleotídeos (nucleotídeos modificados que não possuem um grupamento hidroxila OH no carbono 3' da pentose) marcados com fluoróforos que atuam inviabilizando a elongação da cadeia de nucleotídeos pela DNA polimerase após serem inseridos na cadeia recém-sintetizada.

Para este procedimento, foram utilizados iniciadores flanqueando as regiões contendo as alterações detectadas no sequenciamento do exoma (Quadro 3).

Quadro 3 - Sequência de iniciadores do gene *TP53* avaliados por sequenciamento direto.

Gene	Exon	Sequência (5' - 3'):
<i>TP53-Fb</i>	5 -6	GAAAAGCTCCTGAGGTGTAGACG
<i>TP53-Rb</i>	5-6	CTCATGGGGTTATAGGGAGGTCA
<i>TP53-Fb</i>	7	CTGCTTGCCACAGGTCTCC
<i>TP53-Rb</i>	7	CAGTAAGGAGATTCCCCGC
<i>TP53-Fa</i>	8-9	GGAATCTCCTTACTGCTCCAC
<i>TP53-Ra</i>	8-9	ACCAGGAGCCATTGTCTTTGAG

Todos os iniciadores foram diluídos a uma concentração estoque de 100 μ M e diluídos a uma concentração de 10 μ M. Para a amplificação pela PCR foi utilizado o kit *Platinum Taq DNA Polymerase High Fidelity* (Invitrogen /Life Technologies, Carlsbad, CA, USA), que apresenta alta fidelidade durante o processo de extensão. A reação foi realizada com 30ng de DNA em um volume final de 10 μ L 10 μ M de cada um dos iniciadores (F e R) (Prodinol-IDT, Belo Horizonte, MG, Brazil), 0,2 mM de dNTP Mix (incluindo dATP, dCTP, dGTP e dTTP) (Invitrogen /Life Technologies, Carlsbad, CA, USA), tampão *High Fidelity PCR Buffer* a 1X e 0,5 unidade da enzima *Platinum Taq DNA Polymerase High Fidelity*. As condições da reação foram: 94°C por 2 minutos; 35 ciclos a 94°C por 30 segundos, 55°C por 30 segundos, 68°C por 45 segundos e uma etapa de extensão final a 68°C por 10 minutos. Após a reação de amplificação, a quantidade e qualidade dos produtos amplificados foram analisadas em gel de agarose a 1,5%, corados com *GelRed* (Biotium, VWR, Leuven, Belgium).

Os produtos da PCR foram purificados, por metodologia enzimática, utilizando as enzimas *Exonuclease I* (ExoI - que degrada DNA de fita simples e excesso de iniciadores) e *Shrimp Alkaline Phosphatase* (SAP – que desfosforila os desoxinucleotídeos excedentes) (USB Corporation, Cleveland, OH, USA). A reação de purificação foi realizada em um volume de 10 μ L, contendo 8 μ L do produto da PCR 1x SAP Dilution Buffer 1U da enzima SAP e 3U da enzima ExoI, submetida às condições de incubação a 37°C por 30 minutos e posterior inativação das enzimas a 80°C por 15 segundos. O produto final da purificação foi utilizado como molde para as

reações de sequenciamento. A reação de sequenciamento foi realizada com o *Big Dye Terminator v3.1 Cycle Sequencing Kit* (Applied Biosystems/Life Technologies, Foster City, CA, USA) em um volume de 10 µL, contendo 1 µL do *BigDye Terminator v3.1* (que contém os dideoxynucleotídeos trifosfatados marcados com fluorescência), 500 mM de um dos iniciadores e 5x *BDT v3.1 Sequencing Buffer*. A reação de precipitação foi realizada como recomendado, utilizando EDTA 125mM, pH 8,0, acetato de sódio 3M, pH 5,2 e etanol 100% e 70%. O produto precipitado foi seco a 95°C por 20 minutos e ressuspenso em 13 µL de *formamida Hi-Di* (Applied Biosystems, Woolston, Warrington, UK) seguido por desnaturação a 95°C por 3 minutos. A solução obtida foi então utilizada para eletroforese no analisador automático de DNA, o sequenciador automático *ABI Prism modelo 3130XL* (Applied Biosystems Foster City, CA, USA).

Os dados do sequenciamento foram avaliados quanto a sua qualidade utilizando o software *CLC Main Workbench 6.7* (CLC bio, Aarhus, DK). Foram consideradas na análise somente as sequências que apresentaram similaridade com a região gênica esperada, quando analisadas pelo BLAST (alinhamento pelo próprio programa, *CLC Main Workbench 6.7*).

4 RESULTADOS

4.1 AVALIAÇÃO DAS CARACTERÍSTICAS CLÍNICO-PATOLÓGICAS

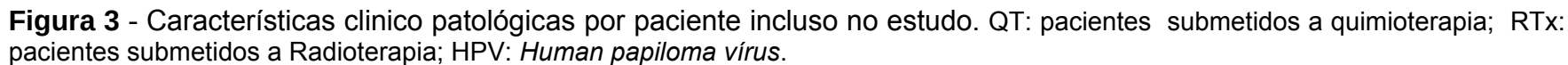
O grupo de 35 pacientes avaliados por sequenciamento do exoma apresentou idade variando de 20 a 49 anos, com média de 40,14 anos. Quatro pacientes (11%) apresentavam história de câncer em parentes de primeiro grau (carcinoma de mama, próstata, pulmão e melanoma). Treze pacientes (37%) foram a óbito em decorrência da doença, sete (53,8%) deles faleceram em até um ano após o diagnóstico. Foi perdido o seguimento clínico de três casos inclusos no estudo. Dezesete pacientes (48,5%) apresentaram tumores em estádios T1-T2 e 18 (51,5%) estádios T3-T4. Invasão perineural e invasão angiolinfática foram identificadas em 14 (40%) e 4 (11,5%) casos, respectivamente. Apenas dois pacientes apresentaram metástases em linfonodos (N3). Sete pacientes apresentaram metástase em pulmão; um caso também apresentou metástase cervical (Tabela 1).

Foram observadas margens cirúrgicas livres em 24 (68,5%) dos casos. A maioria dos pacientes não apresentou recidiva (n=21 / 60%) e três pacientes (8,5%) tiveram um segundo tumor primário (localizado em mucosa jugal, esôfago e base de língua). Foram realizados tratamentos radioterápico e quimioterápico em 45,7% dos pacientes (após cirurgia) (Tabela 1).

Entre os 35 pacientes avaliados, 12 (34,3%) tumores foram positivos

para a presença do HPV sendo predominante o HPV16 (6 CEC de cavidade oral e 5 de orofaringe). Um CEC de cavidade oral apresentou o genótipo HPV 55.

A Figura 3 ilustra as características clínicas e histopatológicas dos pacientes incluídos neste estudo.



4.2 ANÁLISE DOS DADOS DO SEQUENCIAMENTO DE EXOMA TOTAL (SET)

4.2.1 Análise dos Dados Brutos Gerados pelo SET

Na Tabela 2 está disposto um resumo dos dados brutos gerados após a análise de limpeza, mapeamento e de-duplicação das amostras sequenciadas. A amostra CPJ33 apresentou uma baixa cobertura e foi excluída das análises subsequentes. A cobertura de mapeamento alcançou uma média de 48,7 vezes por amostra. Como o sequenciamento realizado foi do tipo *paired-end*, haviam dois *reads* de no máximo 100 pb para cada fragmento. Para prosseguir com as análises, os *reads* tiveram de ser pareados com o intuito de formar fragmentos maiores (aproximadamente 200 pb). Isto é fundamental para o adequado alinhamento das sequências na fase do mapeamento. A porcentagem de *reads* pareados foi superior a 99,13% para todas as amostras, confirmando que houve similaridade quase que total entre os dados de exoma e o genoma de referência, conforme descrito na Tabela 2. Foram detectados *reads* que não puderam ser pareados, denominados *single-end*, isto ocorre possivelmente por artefatos do sequenciamento. Em conjunto, os *reads paired-end* e *single-end* compõem o total de sequências mantidas após a etapa de limpeza dos dados. Após a limpeza e o mapeamento, foi realizada a avaliação de SNVs (*Single Nucleotide Variants*) sendo identificadas, em média, 37.454 SNVs (variando de 2497 a 64202) as amostras tumorais e normais.

Tabela 2 - Resumo dos resultados de limpeza, mapeamento dos *reads* e cobertura, após o sequenciamento do exoma de amostras de sangue periférico (N) e tecido tumoral (T) dos pacientes com CCECP.

Amostra	Nº reads mapeados	% de reads alinhados	Cobertura média	Fração da cobertura, com pelo menos:			
				10X	20X	40X	50X
CPJ1N	69637530	99.59	50,1	81,13%	65,21%	41,27%	33,25%
CPJ1T	77727186	99.53	93,8	80,67%	66,02%	44,19%	36,49%
CPJ2N	87755202	99.67	115,8	87,52%	75,75%	52,74%	43,51%
CPJ2T	60276444	99.69	81,7	79,92%	63,81%	40,03%	31,70%
CPJ3N	63425796	99.72	77,1	81,48%	64,52%	39,06%	30,87%
CPJ3T	79505000	99.58	106,9	84,40%	69,64%	45,98%	38,09%
CPJ4N	23154702	99.82	28,7	34,79%	9,13%	0,68%	0,26%
CPJ4T	28957549	99.79	36,0	43,49%	18,59%	3,57%	1,55%
CPJ5N	22978141	99.80	28,6	29,32%	5,75%	0,37%	0,16%
CPJ5T	58006419	99.77	73,4	75,48%	44,84%	13,41%	7,50%
CPJ6N	32944018	99.81	41,3	50,72%	16,83%	1,95%	0,76%
CPJ6T	11680814	99.84	14,9	9,11%	0,58%	0,06%	0,04%
CPJ7N	31185459	99.77	38,8	46,11%	15,51%	2,10%	0,87%
CPJ7T	70870141	99.66	88,0	78,28%	51,64%	16,21%	8,82%
CPJ8N	29977571	99.79	47,5	72,70%	41,78%	10,69%	5,39%
CPJ8T	37708181	99.83	54,3	78,09%	52,91%	20,25%	13,04%
CPJ9N	21080252	99.99	52,1	61,30%	34,81%	11,83%	6,45%
CPJ9T	12356309	99.36	31,0	59,97%	27,14%	5,62%	2,61%
CPJ10N	27464782	99.89	36,6	65,38%	44,44%	18,87%	12,64%
CPJ10T	22136814	99.13	54,8	77,93%	52,95%	20,62%	13,39%
CPJ11N	91921507	99.74	62,2	89,98%	81,07%	61,28%	52,65%
CPJ11T	61199029	99.72	37,8	88,36%	74,20%	40,22%	27,14%
CPJ12N	92755819	99.79	61,7	88,15%	77,74%	58,28%	50,46%
CPJ12T	59397009	99.76	42,1	83,23%	69,02%	46,61%	37,10%
CPJ13N	82375497	99.76	54,5	90,00%	80,98%	58,79%	48,43%
CPJ13T	34272323	99.61	19,7	76,35%	44,41%	7,71%	2,78%
CPJ14N	80764789	99.65	99,8	83,06%	69,13%	46,67%	38,48%
CPJ14T	80277186	99.65	102,4	82,99%	68,89%	46,40%	38,26%
CPJ15N	31535889	99.82	44,5	73,77%	44,72%	13,56%	7,70%
CPJ15T	20072372	99.65	27,8	55,58%	23,06%	4,14%	1,99%
CPJ16N	29091685	99.79	38,2	45,22%	14,93%	1,89%	0,72%
CPJ16T	29534585	99.82	38,2	53,46%	21,92%	4,82%	2,17%
CPJ17N	61971697	99.65	82,4	83,63%	67,65%	41,45%	32,64%
CPJ17T	69886454	99.52	90,5	81,67%	65,36%	42,13%	34,66%
CPJ18N	61546325	99.67	81,0	82,94%	66,86%	41,08%	32,39%
CPJ18T	47088086	99.61	70,8	81,47%	62,24%	32,99%	24,25%
CPJ19N	68906965	99.58	41,9	89,58%	77,31%	45,80%	32,11%
CPJ19T	78244420	99.78	50,0	84,74%	71,94%	48,76%	39,53%
CPJ20N	3620385	99.79	4,8	0,57%	0,02%	0,00%	0,00%

Cont/ Tabela 2

Amostra	Nº reads mapeados	% de reads alinhados	Cobertura média	Fração da cobertura, com pelo menos:			
				10X	20X	40X	50X
CPJ20T	20119550	99.77	26,2	28,52%	6,06%	0,59%	0,29%
CPJ21N	20119550	99.77	56,3	90,25%	81,54%	60,23%	49,09%
CPJ21T	83580760	99.79	62,3	90,53%	82,49%	62,85%	52,75%
CPJ22N	35122269	99.69	24,0	77,64%	51,05%	15,24%	8,02%
CPJ22T	71610064	99.81	53,6	88,30%	78,27%	55,91%	45,43%
CPJ23N	27516886	99.73	19,7	71,04%	40,29%	9,20%	4,47%
CPJ23T	42833720	99.67	29,2	82,46%	61,82%	24,75%	14,34%
CPJ24N	5900276	99.76	4,3	8,07%	1,10%	0,18%	0,09%
CPJ24T	76429876	99.74	53,6	89,91%	80,97%	58,68%	47,05%
CPJ25N	27515188	99.80	35,4	47,44%	16,42%	2,46%	1,03%
CPJ25T	10961533	99.78	14,7	6,47%	0,35%	0,02%	0,01%
CPJ26N	34761478	99.56	51,9	61,26%	23,97%	2,63%	0,72%
CPJ26T	39405211	99.79	50,7	44,99%	26,55%	12,82%	9,17%
CPJ27N	43368688	99.80	30,5	80,36%	58,00%	25,17%	17,01%
CPJ27T	81477484	99.77	51,1	90,13%	80,23%	55,29%	43,96%
CPJ28N	62210183	99.82	38,4	87,97%	73,56%	41,20%	28,65%
CPJ28T	26661593	99.82	18,6	66,20%	34,79%	9,95%	5,54%
CPJ29N	89720052	99.75	61,4	91,80%	84,72%	66,55%	56,06%
CPJ29T	48833933	99.63	30,4	79,35%	58,07%	26,17%	16,84%
CPJ30N	58325725	99.73	38,2	87,47%	73,00%	39,89%	26,92%
CPJ30T	68932679	99.73	45,3	88,56%	75,31%	46,90%	35,38%
CPJ31N	54964521	99.71	35,7	85,46%	69,16%	35,57%	23,43%
CPJ31T	59697391	99.72	39,6	86,84%	71,75%	40,10%	28,31%
CPJ32N	66575933	99.80	41,5	89,20%	77,16%	46,44%	33,36%
CPJ32T	205221820	99.79	127,3	95,89%	93,27%	87,45%	83,87%
CPJ33N	73389868	99.77	45,7	90,51%	80,03%	51,50%	38,57%
CPJ33T*	49527	99.63	0,2	0,00%	0,00%	0,00%	0,00%
CPJ34N	70500933	99.80	43,7	89,94%	78,46%	48,54%	35,61%
CPJ34T	60973880	99.79	36,0	86,39%	68,87%	35,82%	24,62%
CPJ35N	13216685	99.78	9,0	35,81%	7,78%	0,93%	0,51%
CPJ35T	47145227	99.71	30,8	83,78%	63,99%	27,30%	16,34%

* Este caso foi excluído pela baixa cobertura. Os outros casos com a cobertura baixa foram mantidos, por passaram por análise individual posteriormente, a fim de garantir a veracidade das alterações identificadas.

4.2.2 Alterações Genômicas nos Carcinomas Orais e de Orofaringe de Pacientes Jovens

Foram consideradas apenas as alterações em regiões de exons, exclusivas das amostras dos tumores primários dos pacientes selecionados.

Os reagentes utilizados para o preparo das bibliotecas que foram sequenciadas são capazes de capturar regiões estendidas aos exons, como regiões intrônicas, intergênicas e UTRs. Após a análise do resumo de todas as alterações exclusivas do tecido tumoral por paciente, incluindo número de alterações identificadas nos bancos de dados 1000 genomas (1000G-All), 6500 Exoma (ESP6500si_ALL), dbSNP e COSMIC (*Catalog Of Somatic Mutations In Cancer*), foi possível verificar uma variação de 158 a 20525 alterações entre os 35 pacientes. Foi observado maior número de alterações nos pacientes CPJ32 e CPJ34. O paciente CPJ33 apresentou uma média de cobertura de sequenciamento muito baixa (0,2) e por esta razão não foram identificadas alterações nesta amostra sendo excluído da análise final. As amostras com cobertura abaixo de 20x, foram mantidas no estudo para a comparação com o perfil mutacional dos demais casos.

Em média, foram identificadas 6570 variantes em regiões intergênicas 11754 em intrônicas e 3963 em regiões de exons. Considerando o total de todas as variantes identificadas pelo sequenciamento, em média 29.467 já foram descritas no dbSNP e 180 descritas no COSMIC.

Foram realizadas comparações entre os casos deste estudo incluindo o subsítio anatômico e presença ou ausência dos principais fatores de risco

(consumo de álcool, tabaco e infecção pelo HPV). A Figura 4 mostra a distribuição da média das variantes (SNPs e Indels) em cada grupo de comparação. A média das Indels foi maior entre os pacientes HPV negativos (HPV-), tumores da cavidade oral e também para o grupo de pacientes que não são consumidores de bebidas alcoólicas, usuários de tabaco e HPV+ (A-T-HPV+) (Figura 4A e 4B).

Em todos os grupos houve predomínio de alterações *inframe deletion* (Figura 4A) e sinônimas (Figura 4B). Em relação aos SNVs, houve uma inversão entre as médias dos grupos HPV positivos e negativos. Neste caso, os casos HPV positivos apresentaram uma média maior de alterações comparado ao grupo de pacientes HPV negativos (Figura 4C).

Os resultados a seguir, foram selecionados a partir das variantes exclusivas dos tumores e não sinônimas e segundo a patogenicidade predita de acordo com os bancos de dados *Sift*, *Polyphen2_HDIV*, *Polyphen2_HVAR*, *Mutation taster*, *Mutation Assessor*.

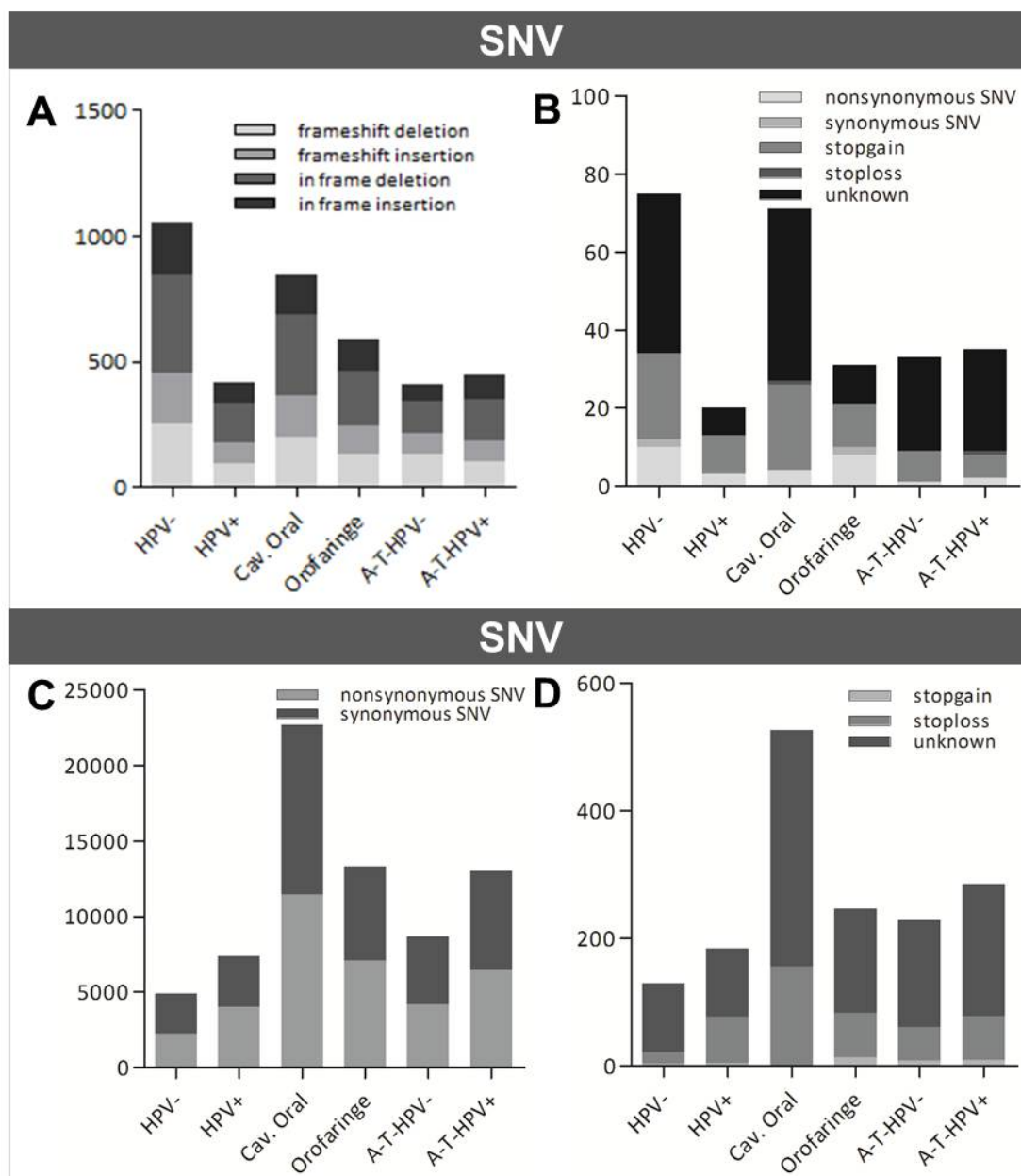


Figura 4 - Distribuição das médias de variantes por subgrupos de pacientes com CCECP. **A** Inserções e deleções e **B-D**. SNVs (*Single Nucleotide Variants*). HPV-: HPV negativo; HPV+: HPV positivo; A-T-HPV-: não consumidores de bebida alcoólica, não tabagista e HPV negativo; A-T-HPV+: não consumidores de bebida alcoólica, não tabagista e HPV positivo.

4.2.3 Perfil Geral de Alterações Somáticas e Não Sinônimas

O perfil mutacional foi heterogêneo entre os tumores avaliados. Foram identificados 28.206 SNVs não sinônimas em 10.092 genes exclusivos dos tumores. A frequência de SNVs entre os pacientes foi bastante diversificada. Considerando as SNVs identificados em pelo menos 20% dos pacientes, foram identificadas 705 SNVs em 506 genes (dados não mostrados). Treze destes genes já foram descritos no COSMIC, sendo 54% relacionados a patologias da tireoide. Dentre os 506 genes identificados neste estudo 47 estão descritos no banco de dados GSEA (Tabela 3), a maioria atua como fatores de transcrição (N=21) e marcadores de diferenciação celular (N=12), seguidos por oncogenes (N=6), genes de reparo a danos de DNA (N=5) e fatores de crescimento (N=3).

As alterações identificadas nos 506 genes (GSEA) foram comparadas às alterações identificadas no TCGA, de acordo com a idade (abaixo de 49 anos e pacientes com idade superior a 49 anos). A maioria das alterações foi predita como patogênica em apenas um banco de dados, não incluindo *Sift/Polyphen*. A Tabela 3 resume as variantes detectadas em pelo menos 20% dos casos deste estudo comparadas com as descritas no banco de dados TCGA de acordo com a idade (<49 anos: 75 casos e >49 anos: 333 casos).

No presente estudo não foi observado um perfil mutacional diferente entre os casos comparando pacientes com idade inferior a 40 anos (n=15), 41 a 45 anos (n=14) e 46 a 49 anos (n=6). Por esta razão, estes casos foram considerados como um único grupo definido como apresentando idade

inferior a 49 anos, o qual ainda assim se mantém como idade mais jovem com relação a média de idade ao diagnóstico dos pacientes com CCECP.

Tabela 3 - Resumo dos genes que apresentaram variantes(SNVs) em pelo menos 20% dos casos.

Gene	Nº casos (%)	TCGA <49 anos	TCGA >49 anos
	Nosso estudo	N=75	N=333
		Nº casos	Nº casos
TLR10	8 (23,5)	0	4
MGA	9 (26,5)	0	8
BCLAF1	7 (20,6)	2	6
RAD17	7 (20,6)	3	1
KLRB1	7 (20,6)	0	0
DCLRE1A	7 (20,6)	0	5
NANOG	7 (20,6)	0	1
USP6	7 (20,6)	0	5
TTLL5	7 (20,6)	0	1
NR1D2	7 (20,6)	0	2
SIRPB1	7 (20,6)	1	7
PROP1	7 (20,6)	0	0
IGHMBP2	7 (20,6)	0	6
CD27	7 (20,6)	1	1
BCL9	7 (20,6)	1	7
CHAF1A	7 (20,6)	0	3
ZSCAN2	7 (20,6)	1	2
ZNF80	7 (20,6)	1	2
ZNF180	7 (20,6)	0	4
TSHR	7 (20,6)	0	1
TNFRSF11B	7 (20,6)	0	2
TNFRSF10B	7 (20,6)	1	0
THRAP3	7 (20,6)	1	2
SPIB	7 (20,6)	0	1
SPEN	7 (20,6)	2	20
SELP	7 (20,6)	1	6
SALL3	7 (20,6)	1	9
PRDM2	7 (20,6)	1	7
NOV	7 (20,6)	1	3
KLRC2	7 (20,6)	0	1
KIR2DL1	7 (20,6)	0	2
INSM2	7 (20,6)	0	1
HOXB3	7 (20,6)	2	3
HOXA7	7 (20,6)	0	2
HOXA1	7 (20,6)	0	2
GYPA	7 (20,6)	0	3
EML4	7 (20,6)	0	2
ELF1	7 (20,6)	0	4
CEBPZ	7 (20,6)	0	2
CD44	7 (20,6)	2	1
CD163	7 (20,6)	1	10
CASC5	7 (20,6)	1	10
BTLA	7 (20,6)	0	3
BMP8B	7 (20,6)	0	3
ATF7IP	7 (20,6)	1	2
POLQ	7 (20,6)	0	14
EXO1	8 (23,5)	1	1

4.2.3.1 Mutação no gene *TP53*

Considerando a alta frequência de mutações no gene *TP53* descrita em carcinomas de cabeça e pescoço (AGRAWAL et al. 2011; STRANSKY et al. 2011; PICKERING et al. 2013; PICKERING et al. 2014a e b), foram investigadas as variantes deste gene. Foram identificadas cinco variantes (uma em cada paciente) preditas como patogênicas no gene *TP53* (Quadro 4). Quatro casos eram de carcinomas orais e um (CPJ07) foi diagnosticado com CEC de orofaringe.

Quadro 4 - Alterações identificadas no gene *TP53*.

ID	Variante <i>TP53</i>	cDNA	Proteína	dbSNP	Exon	Idade	HPV
CPJ19	g.7577534C>A	c.747G>T	p.R249S	rs28934571	7	49	-
CPJ31	g.7577547C>T	c.734G>A	p.G245D	rs12191265 6	7	42	+
CPJ07	g.7578403C>T	c.527G>A	p.C176Y	0	5	35	+
CPJ03	g.7578406C>T	c.524G>A	p.R175H	rs28934578	5	35	+
CPJ27	g.7577036del1	c.902_902delC	P301Qfs*44	0	8	43	-

ID: identificação do paciente; T: amostra de tecido tumoral. +: HPV positivo; -: HPV negativo. IARC TP53 Database - <http://p53.iarc.fr/>.

4.2.3.2 Genes Relacionados a Anemia de Fanconi (AF)

Os pacientes selecionados para este estudo não se enquadram critérios clínicos da Anemia de Fanconi. Entretanto, pacientes com AF tem risco aumentado de desenvolver CCECP em idade mais jovem (ROSENBERG et al. 2003; ALTER et al. 2005) e por esta razão as alterações nos genes envolvidos na AF foram investigadas em maiores detalhes. Entre os genes descritos em literatura como relacionados a

AF(*FANCA*/ *B*/*C*/*D1*/*D2*/*E*/ *F*/ *G*/ *I*/ *J*/ *L*/ *M*/ *N*/ *O*/ *P*) foram identificados SNVs nos genes *FANCD2* e *FANCA* (Quadro 5). Para a análise de variantes nesses genes relacionados ao desenvolvimento da AF, foram analisados as variantes presentes no DNA proveniente do tecido tumoral e sangue periférico de cada paciente. A variante detectada no gene *FANCD2* foi considerada patogênica em um banco de dados (MutationTaster_pred) e estava presente no tecido tumoral de quatro casos (CPJ05, CPJ11, CPJ24 e CPJ32) e no tumor e sangue periférico de quatro pacientes (CPJ02, CPJ22, CPJ27 e CPJ31). A variante rs17233141 (*FANCA*) foi descrita como patogênica em dois bancos de dados (Polyphen2_pred e FATHMM_pred). Todos as SNVs identificados são classificados como *missense*. É interessante notar que apenas a variante rs3864017 do gene *FANCD2* foi detectada tanto no sangue periférico como no tumor de quatro pacientes, embora tenha sido considerada pelo banco de dados ClinVar como benigna. Entre esses quatro pacientes, o acometimento da cavidade oral e orofaringe foi distribuído igualitário (dois pacientes para cada localização) e três desses pacientes apresentaram infecção por HPV (HPV16). A variante rs17233141 do gene *FANCA* é considerada patogênica pelo banco de dados ClinVar.

Quadro 5 - Variantes identificadas nos genes relacionados a Anemia de Fanconi.

<i>Variante</i>	<i>Gene</i>	<i>dbSNP</i>	<i>T</i>				<i>T/N</i>			
			<i>Caso</i>	<i>Sítio Anatômico</i>	<i>Idade</i>	<i>HPV</i>	<i>Caso</i>	<i>Sítio Anatômico</i>	<i>Idade</i>	<i>HPV</i>
g.10106532C>T* c.2141C>T p.P714L	FANCD2	rs3864017	CPJ05	Orofaringe	39	+	CPJ02	Orofaringe	37	+
			CPJ 11	Orofaringe	44	-	CPJ22	Cav.Oral	46	+
			CPJ 24	Cav. Oral	44	-	CPJ27	Cav.Oral	43	-
			CPJ32	Orofaringe	43	-	CPJ31	Orofaringe	42	+
g.89833576G>C c.2574C>G p.S858R	FANCA	rs17233141	CPJ04	Cav. Oral	20	-	--			
g.89805301G>C c.4249C>G p.H1417D	FANCA	rs17227403	CPJ05	Orofaringe	39	+	--			

* Variante patogênica em um banco de dados. T: exclusivo da amostra tumoral; T/N: alteração presente no tumor e sangue periférico. Os números representam a identificação do paciente; Cav: cavidade; HPV:(-) : negativo and (+) positivo. g.:localização genômica da variante; c.: cDNA; p.:proteína.

4.2.3.3 Comparação com os Principais Genes Descritos em Literatura Relacionados ao CCECP

Os genes descritos pelos principais estudos envolvendo SET e CCECP foram investigados e estão listados na Tabela 5. A análise das sequências desses genes revelou uma baixa frequência de alterações em pacientes jovens. O gene *CASP8* foi o único da lista que apresentou uma frequência de 20% de alteração entre os casos aqui estudados (Tabela 4).

Tabela 4 - Genes em destaque nos principais estudos de rastreamento de alterações em CCECP identificados por SET.

Gene	Casos com alterações N (%)	Referência
<i>AKT2</i>	-	Chung et al. 2015
<i>ARID2</i>	-	IPTICG 2013
<i>CASP8</i>	7 (20,6)	IPTICG 2013, Bswas et al. 2014
<i>CDC101</i>	-	Li R et al. 2015
<i>CDKN2A</i>	1 (3)	Agrawal et al. 2011; Stransky et al. 2011; Nichols et al. 2012b; Li et al. 2014, Samman et al. 2015
<i>CCND2</i>	-	Li et al. 2014
<i>CTNNA3</i>	3 (8,8)	Li R et al. 2015; Upadyay et al.2016.
<i>DLL4</i>	-	Bswas et al. 2014;
<i>DTX2</i>	-	Upadyay et al.2016.
<i>DST</i>	-	Vettore et al. 2015.
<i>EGFR</i>	3 (8,8)	Agrawal et al. 2011; Li et al.2014
<i>EIF3A</i>	-	Li R et al. 2015;
<i>EP300</i>	1 (3)	Li R et al. 2015;
<i>FAT1</i>	3 (8,8)	Agrawal et al. 2011; IPTICG 2013; Vettore et al. 2015, Bswas et al. 2014; Samman et al. 2015.
<i>FBXW7</i>	-	Agrawal et al. 2011; Stransky et al. 2011
<i>FXR1</i>	-	Li R et al. 2015
<i>HRAS</i>	1 (3)	Agrawal et al. 2011; IPTICG 2013; Stransky et al. 2011; Bswas et al. 2014.
<i>IRF6</i>	-	Stransky et al. 2011
<i>LAMA1</i>	3 (8,8)	Li R et al. 2015
<i>MLL4</i>	-	IPTICG 2013, Bswas et al. 2014
<i>NCAPD3</i>	2 (5,9)	Li R et al. 2015
<i>NEK8</i>	-	Li R et al. 2015
<i>NOTCH1</i>	1 (3)	Agrawal et al. 2011; Stransky et al. 2011; IPTICG 2013; Li et al. 2014, Bswas et al. 2014; Samman et al. 2015; Upadyay et al. 2016.
<i>PIK3CA</i>	1 (3)	Agrawal et al. 2011; Stransky et al. 2011; Lui et al. 2013; Li et al. 2014; Li R et al. 2015; Chung et al. 2015, Bswas et al. 2014
<i>PKHD1L1</i>	1 (3)	Li R et al. 2015
<i>PTCHD2</i>	1 (3)	Li R et al. 2015
<i>PTEN</i>	-	Agrawal et al. 2011; Stransky et al. 2011; Chung et al. 2015;
<i>RALGAPB</i>	-	Li R et al. 2015
<i>RIMBP2</i>	1 (3)	Li R et al. 2015
<i>SI</i>	-	Li R et al. 2015
<i>RNF213</i>	-	Vettore et al. 2015.
<i>SPEN</i>	6 (17,6)	Li R et al. 2015
<i>SYNE</i>	2 (5,9)	Li R et al. 2015
<i>TP53</i>	5 (14,7)	Agrawal et al. 2011; IPTICG 2013; Li R et al. 2015; Chung et al. 2015; Nichols et al. 2012b; Vettore et al. 2015, Bswas et al. 2014; Samman et al. 2015.
<i>TP63</i>	-	Stransky et al. 2011
<i>TRPM3</i>	6 (17,6)	IPTICG 2013
<i>UBR4</i>	1 (3)	Li R et al. 2015
<i>UNC13C</i>	1 (3)	IPTICG 2013
<i>USP9X</i>	-	IPTICG 2013, Bswas et al. 2014

4.2.4 Comparação Entre as Alterações Segundo o Sítio Anatômico (Cavidade Oral e Orofaringe)

Foram identificadas 11.459 SNVs em 6.087 genes entre os pacientes com diagnóstico de CEC da cavidade oral (N=24) e 7.082 SNVs em 4.469 genes entre os pacientes com diagnóstico de CEC da orofaringe (N=11). A maioria das SNVs se distribuíram de forma heterogênea entre os pacientes. A Tabela 5 apresenta as variantes identificadas em no mínimo 20% dos casos de acordo com a patogenicidade e sua função segundo o GSEA, nos CEC de cavidade oral e orofaringe.

Tabela 5 - Variantes exclusivas nos casos estudados de acordo com o sítio anatômico.

	<i>Gene</i>	<i>dbSNP</i>	<i>cDNA</i>	<i>Proteína</i>	<i>Patogenicidade^a</i>	<i>SIFT POLYPHEN^b</i>	<i>Frequência (%)</i>	<i>Função (GSEA)</i>
Cavidade oral (N=23*)	<i>PDLIM5</i>	rs13107595	c.1475G>A	p.S492N	1	0	26	fator de transcrição
	<i>NCOR1</i>	rs76780359	c.571G>A	p.E191K	5	2	22	fator de transcrição
	<i>NCOR1</i>	rs200020868	c.534G>C	p.K178N	5	2	22	fator de transcrição
	<i>NCOR1</i>	rs150910818	c.515C>T	p.S172L	4	1	22	fator de transcrição
	<i>INSL5</i>	rs549148	c.149A>T	p.Q50L	1	0	26	Citocinas e fatores de crescimento
	<i>FOXD4</i>	rs2492216	c.407A>G	p.D136G	1	0	22	fator de transcrição
CEC Orofaringe (N=11)	<i>SIRPG</i>	rs6043409	c.788T>C	p.V263A	1	0	25	marcadores de diferenciação celular
	<i>SEMA5B</i>	rs2276782	c.2519T>A	p.V840D	1	0	25	citocinas e fatores de crescimento
	<i>POU5F1B</i>	rs6998061	c.527G>A	p.G176E	6	2	25	fator de transcrição
	<i>MICALL1</i>	rs9610875	c.1555G>T	p.A519S	1	0	25	citocinas e fatores de crescimento
	<i>KIR3DL2</i>	rs145326846	c.474G>T	p.E158D	2	1	25	fator de transcrição
	<i>HELQ</i>	rs17006837	c.704T>C	p.L235P	3	1	25	marcadores de diferenciação celular
	<i>GYPA</i>	rs7682260	c.59T>C	p.L20S	1	0	25	reparo de DNA
	<i>CD6</i>	rs11230562	c.650C>T	p.T217M	4	2	25	marcadores de diferenciação celular

*O caso CPJ33 foi excluído das análises devido a qualidade dos dados do sequenciamento. a: Número de banco de dados em que as variantes foram descritas como patogênicas. b: 2) classificadas como patogênicas pelo SIFT/POLYPHEN, 1) patogênicas em pelo menos 1 desses dois bancos, 0) não classificada pelo SIFT/POLYPHEN.

4.2.5 Comparação Entre os Casos Segundo o Status Do HPV (Positivo e Negativo)

Entre os pacientes HPV positivos foram identificadas 4.035 SNVs em 2.904 genes, uma média de 336 SNVs por paciente. Por outro lado, foram identificados 2.252 SNVs em 1.483 genes entre os pacientes HPV negativos. Os casos HPV negativos apresentaram uma maior frequência e média de alterações, aquelas presentes em mais de um caso acometiam genes supressores tumorais. Apenas no grupo de casos HPV positivos foi identificado uma variante em oncogene, o *PIK3CA* (rs2230461).

As Tabelas 6 e 7 apresentam as principais vias alteradas considerando as alterações detectadas nos casos HPV positivos e negativos. Algumas vias foram comuns aos casos HPV positivos e negativos como a envolvida em Diabetes mellitus I, via de sinalização de interferons, sistema imune, entre outras. Foi verificada uma forte presença de componentes imunológicos alterados nas duas vias, porém em casos HPV positivos foi mais proeminente, como por exemplo, o componente imune adaptativo, via sinalização do *PD1* e *TCR* e que são potenciais alvos terapêuticos de imunoterapia.

Tabela 6 - Principais vias canônicas envolvendo as alterações detectadas nos casos HPV positivos.

<i>HPV positivo – Enriquecimento</i>				
Vias alteradas	Banco de dados	Número de genes	<i>p</i> -valor	<i>FDR q</i> -valor
Doença do enxerto contra hospedeiro*	KEGG	10	1.63 e-13	1.76 e-10
Processamento e apresentação de antígeno *	KEGG	12	8.19 e-13	4.41 e-10
Rejeição de alo-enxertos	KEGG	8	1.34 e-10	4.81 e-8
Sinalização por interferon gama *	REACTOME	9	3.62 e-10	9.75 e-8
Diabetes mellitus tipo i*	KEGG	8	4.7 e-10	1.01 e-7
Doença autoimune da tireóide*	KEGG	8	2.23 e-9	4 e-7
Sinalização por interferon *	REACTOME	11	1.04 e-8	1.6 e-6
Translocação do ZAP 70 para sinapse imunológica *	REACTOME	5	2.3 e-8	2.84 e-6
Moléculas de adesão celular CAMS *	KEGG	10	2.37 e-8	2.84 e-6
Miocardite viral*	KEGG	8	3.02 e-8	3.25 e-6
Sinalização de citocina no sistema imune*	REACTOME	13	3.69 e-8	3.61 e-6
Fosforilação de CD3 e das cadeias TC TCR zeta	REACTOME	5	4.96 e-8	4.45 e-6
Sistema imune *	REACTOME	23	8.33 e-8	6.9 e-6
Sinalização PD1	REACTOME	5	9.63 e-8	7.41 e-6
Sinalização TCR a jusante	REACTOME	6	1.53 e-7	1.1 e-5
Rede intestinal imune para produção de IgA	KEGG	6	7.6 e-7	5.12 e-5
Geração de moléculas de mensageiros secundários	REACTOME	5	8.64 e-7	5.47 e-5
Asma *	KEGG	5	1.5 e-6	8.77 e-5
Sinalização TCR	REACTOME	6	1.55 e-6	8.77 e-5
Co-estímulo pela família CD28 FAM	REACTOME	6	3.87 e-6	2 e-4

*vias comuns aos casos HPV negativos.

Tabela 7 - Principais vias canônicas envolvendo os genes alterados nos casos HPV negativos.

<i>HPV negativo – Enriquecimento</i>				
Vias alteradas	Banco de dados	Número de genes	p-valor	FDR q-valor
Via de sinalização olfatória	REACTOME	64	5.44 e-33	4.97 e-30
Transdução olfatória	KEGG	69	9.23 e-33	4.97 e-30
Sinalização a jusante de GPCR	REACTOME	98	6.41 e-32	2.3 e-29
Sinalização por GPCR	REACTOME	103	1.94 e-30	5.23 e-28
Sinalização por interferon gama *	REACTOME	18	2.61 e-13	5.62 e-11
Sistema imune *	REACTOME	69	1.08 e-11	1.95 e-9
Sinalização por interferon *	REACTOME	25	1.47 e-11	2.26 e-9
Diabetes mellitus tipo 1*	KEGG	14	2.29 e-11	3.09 e-9
Doença do enxerto contra hospedeiro *	KEGG	13	1.76 e-10	2.11 e-8
Rejeição de transplante	KEGG	12	6.85 e-10	7.38 e-8
Doença autoimune da tireóide *	KEGG	13	4.28 e-9	4.19 e-7
Sinalização de citocina no sistema imune*	REACTOME	28	1.62 e-8	1.45 e-6
Transporte transmembrânico de pequenas moléculas SM	REACTOME	35	5.13 e-8	4.25 e-6
Processamento e apresentação de antígenos*	KEGG	15	6.55 e-8	5.04 e-6
Moléculas de adesão celular CAMS *	KEGG	18	1.23 e-7	8.85 e-6
Miocardite viral *	KEGG	13	2.51 e-7	1.69 e-5
Interação do receptor ECM	KEGG	13	1.34 e-6	8.5 e-5
Translocação do ZAP 70 para sinapse imunológica *	REACTOME	6	1.79 e-6	1.07 e-4
Asma *	KEGG	8	2.13 e-6	1.21 e-4
SLC mediado pelo transporte de TR transmembrânico	REACTOME	22	4.51 e-6	2.32 e-4

*vias comuns aos casos HPV positivos.

4.2.6 Comparação entre as Alterações Genômicas e os Principais Fatores de Risco

Cinco pacientes (CPJ04, CPJ25, CPJ27, CPJ28, CPJ30) eram não consumidores de bebida alcoólica, não tabagistas e HPV negativos (A-T-HPV-) assim como cinco pacientes (CPJ03, CPJ14, CPJ15, CPJ16, CPJ21) eram não tabagistas, não ingeriam bebidas alcoólicas e eram HPV positivos (A-T-HPV+). Entre esses dois grupos, aqueles positivos para HPV apresentaram maior número de SNVs (6.455) quando comparados aos tumores de pacientes HPV negativos (4.205 SNVs). Variantes patogênicas dos genes *MTCH2*, *PSSR3*, *POU2F3*, *C11orf16*, *KIF6* e *UQCRFS1* foram identificados apenas no grupo A-T-HPV-, com uma frequência de no mínimo 40% dos casos. Foram identificados 704 genes exclusivamente alterados no grupo de tumores A-T-HPV+, todos com alterações presentes em pelo menos três casos (40%), sendo que quatro pacientes apresentavam mutações em 653 genes que não apresentaram variantes no grupo A-T-HPV-, todas patogênicas, entre eles, o gene *BRCA1*.

4.2.7 Mutações Exclusivas em Pacientes que Apresentaram Metástases

Foram identificadas dez SNVs comuns em pelo menos três dos sete pacientes (48,7%) que apresentaram metástase pulmonar. Estas SNVs foram preditos como patogênicos em pelo menos cinco bancos de dados (Quadro 6). Os pacientes CPJ05, CPJ27 e CPJ34 que apresentaram praticamente as mesmas variantes exclusivas, sendo o estadiamento N2 a

única característica em comum entre eles. Estas variantes tem potencial para contribuir para o risco de desenvolvimento de metástases em pacientes jovens com tumores de cabeça e pescoço.

Quadro 6 - Dez SNVs mais frequentes e patogênicos entre os pacientes que desenvolveram metástase.

<i>Gene</i>	<i>Variante</i>	<i>cDNA</i>	<i>Proteína</i>	<i>dbSNP</i>	<i>Nº BD</i>	<i>ID</i>
<i>CTDSP2</i>	g.58220811G>T	c.C322A	p.L108I	-	8	CPJ 05 27, 34
	g.58220831C>G	c.G302C	p.R101T	-	5	CPJ 05 27, 34
<i>KRT18</i>	g.53343084G>C	c.G127C	p.G43R	rs75441140	8	CPJ 05 29, 34
<i>MGA</i>	g.42032383C>G	c.C4567G	p.P1523A	rs17677991	5	CPJ 05 27, 34
<i>MTCH2</i>	g.47647265A>G	c.T683C	p.F228S	rs76666113	7	CPJ 05 27, 34
	g.47647238A>G	c.T710C	p.V237A	rs78071782	5	CPJ 05 27, 34
	g.47660295A>G	c.T235C	p.C79R	-	5	CPJ 05 27, 34
<i>OR12D2</i>	g.29364615G>T	c.T235C	p.C79R	rs9257834	5	CPJ 05 15 27, 34
<i>SORD</i>	g.45365724C>T	c.C1070T	p.P357L	rs76283634	5	CPJ 05 27, 34
<i>TMEM173</i>	g.138856982C>T	c.G878A	p.R293Q	rs7380824	5	CPJ 05 27, 34

° BD: Número de banco de dados onde a variante é descrita como patogênica.

4.3 CONFIRMAÇÃO DAS ALTERAÇÕES IDENTIFICADAS NO GENE TP53

As alterações no gene *TP53* identificadas pelo SET (Quadro 4) foram confirmadas pelo método de sequenciamento de Sanger (Figura 5). Assim como pelo SET, as alterações foram identificadas apenas em amostras dos tumores e estavam ausentes no sangue periférico dos mesmos pacientes, confirmando a reprodutibilidade da metodologia.

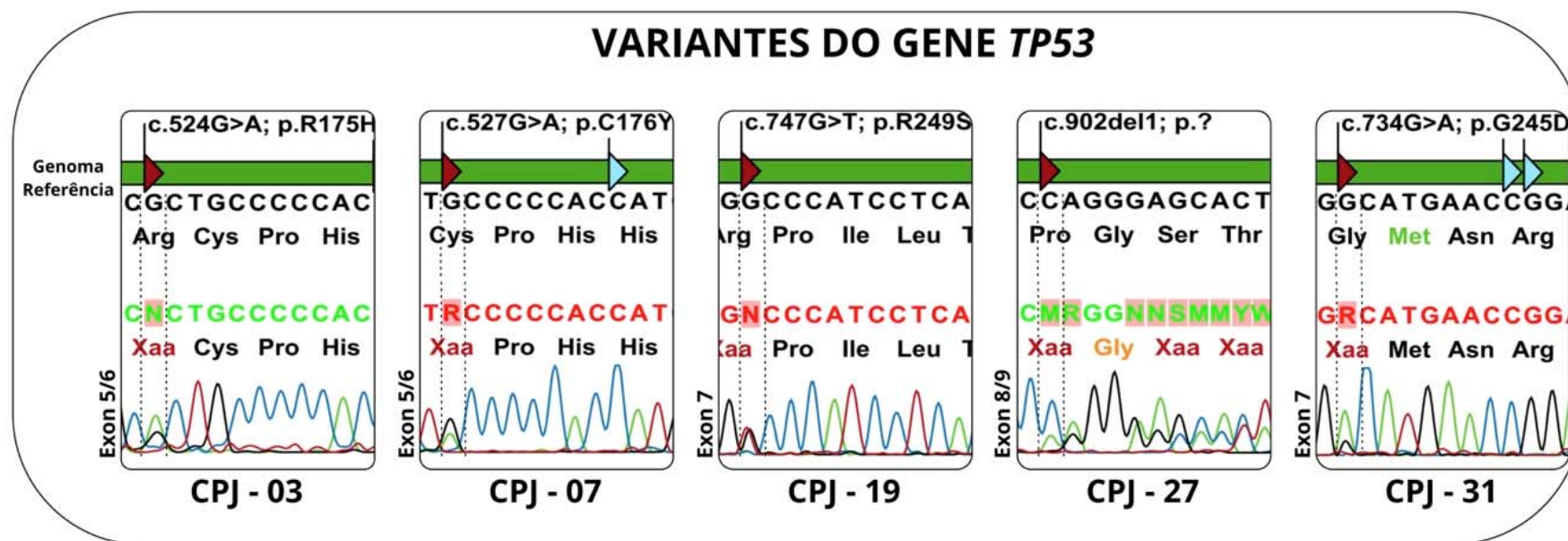


Figura 5 - Eletroferograma das alterações no gene *TP53*, identificadas nos pacientes. CPJ-03 (c.524G>A/p.R175H), CPJ-07 (c.527G>A/p.C176Y), CPJ-19 (c.747G>T/p.R249S), CPJ-27 (c.902del1) e CPJ-31 (c.734G>A/p.G245D).

5 DISCUSSÃO

O carcinoma de células escamosas é a doença mais comum da região da cabeça e pescoço, sendo responsável por >90% dos casos (CASSIDY et al. 2010). O CCECP é um dos tumores mais incidentes da população e é normalmente associado com hábitos tabagista, etilista e infecção pelo papilomavírus humano (HPV). Recentemente, tem-se observado um aumento da incidência de CCECP em pacientes jovens e, nestes casos, o CCECP é considerado uma entidade distinta, com comportamento e prognóstico diferentes dos tumores em pacientes mais velhos.

Nos últimos anos a incidência do carcinoma de células escamosas de cabeça e pescoço vem aumentando entre adultos jovens (<49 anos), embora não haja consenso acerca da etiologia, curso clínico e prognóstico destes casos.

São muitas as discordâncias em relação ao prognóstico dos pacientes com CCECP em idade jovem. Tem sido relato que estes pacientes jovens podem apresentar dois cursos clínicos distintos. Um grupo que se apresenta de forma bem agressiva, com 40% de mortalidade nos dois primeiros anos devido à metástase regional e à distância, e um outro grupo que apresenta um curso indolente nos quais os pacientes permanecem livres da doença por 20 anos (POPOVTZER et al. 2004). Esse dado corrobora com o perfil dos pacientes inclusos nesse estudo, onde 37% foram a óbito em decorrência da doença (período médio de 2 anos do diagnóstico à morte pela doença), sendo que 53,8% desses pacientes faleceram em até um ano após o diagnóstico. Até

o presente momento e a última informação sobre o seguimento clínico (Janeiro 2017) 19 pacientes do estudo estão vivos e em seguimento na instituição, ao passo que quatro deles permanecem em tratamento. Contrário aos jovens, os pacientes mais idosos apresentam um perfil mais homogêneo, visto que a morte pela doença ocorre gradualmente ao longo de 10 anos na grande maioria dos casos (POPOVTZER et al. 2004).

A doença, em geral, acomete mais homens do que mulheres, porém a incidência mundial deste tumor está aumentando entre as mulheres (CURADO e HASHIBE 2009). Ainda assim, o número de homens presentes nessa casuística foi muito maior (10 mulheres e 25 homens). Com relação ao estadiamento clínico, o número de pacientes diagnosticados com tumores primários em estágio T1-T2 foi bem próximo aos pacientes diagnosticados em estágio T3-T4 (48,5% e 51,5%, respectivamente). Para os casos em que a doença é diagnosticada precocemente, o tratamento indicado é a cirurgia ou a radioterapia, tendo em geral um prognóstico favorável. Já para os tumores avançados, o tratamento é a cirurgia combinada com a radioterapia pós-operatória (LEEMANS et al. 2011). Entre os 17 pacientes diagnosticados com estágio T1-T2 29,4% tiveram recidiva, enquanto 50% daqueles com estágio mais avançado (T3-T4 18 casos) recidivaram durante o período do estudo.

O desenvolvimento de CCECP em idade mais jovem apresenta predileção por sítios anatômicos da cavidade oral e orofaringe, comparado a subsítios da laringe e hipofaringe, por exemplo, porém essa predileção não é bem esclarecida (MÜLLER et al. 2008; UDEABOR et al. 2012; BODNER et al. 2014). O levantamento dos dados de registro no Banco de Tumores do Hospital A. C. Camargo-SP revelou a presença de 2300 amostras de tecido

congelado de tumores de cabeça e pescoço, registrados no Departamento de Cirurgia de Cabeça e Pescoço (janeiro de 2013). Dentre esses, foram identificados 523 pacientes com idade inferior a 49 anos englobando todas as topografias. Desconsiderando a topografia com tireoide a maioria desses casos possuíam câncer em região da cavidade oral (65,5%) ou da orofaringe (11,4%). BODNER et al. (2014), avaliaram 156 relatos de casos de pacientes jovens com CCECP observando uma predileção pelas regiões da língua (69%), gengiva (20%) e lábio inferior (5%), totalizando mais 80% dos casos na cavidade oral.

Entre os principais fatores de risco (80-90% dos casos) associados aos CCECP está o consumo de álcool e o de tabaco. Estudos apontam sinergismo entre esses fatores, aumentando em 30 vezes o risco para o desenvolvimento dos tumores de cabeça e pescoço (CASSIDY et al. 2010). Em revisões da literatura acerca de fatores de risco e CCECP em idade jovem, é relatado que muitos desses pacientes diagnosticados antes dos 40 anos, com maior prevalência de não etilistas e não tabagistas (UDEABOR et al. 2012; HILLY et al. 2013; MAJCHRZAK et al. 2013; DESCHLER et al. 2014; BODNER et al. 2014). Os pacientes deste estudo se distribuíram praticamente de forma equivalente entre tabagistas e não tabagistas (N=19 e N=15, respectivamente) e 60% afirmaram ser etilistas. Quinze pacientes eram concomitantemente tabagistas e etilistas, enquanto 10 pacientes (28,6%) declararam-se não tabagistas e não etilistas. A exposição aos fatores de risco pode ser similar em pacientes jovens e adultos, mas a questão é se essa exposição é longa o suficiente para tornar esses agentes etiológicos os principais nos carcinomas de cabeça e pescoço em jovens (LLEWELLYN et al. 2004). Recentemente,

oconsórcio *The International Head and Neck Cancer Epidemiology*-INHANCE, publicou uma análise de dados de 25 estudos caso-controle em câncer de cabeça e pescoço avaliando adultos jovens (<45 anos 2.010 casos e 4.042 controles) e mais velhos (>45 anos 17.700 casos e 22.704 controles). O estudo demonstrou que a contribuição do uso de álcool e tabaco para o desenvolvimento desses tumores em adultos mais jovens é menor que aqueles atribuídos para os mesmos tumores em casos em idade mais avançada (TOPORCOV et al. 2015).

O HPV de alto risco (HR-HPV) tem sido associado a um subgrupo de CCECP (GILLISON et al. 2000; ZUR HAUSEN 2002), principalmente de orofaringe. Estima-se que aproximadamente 30.000 casos de câncer de orofaringe estão associados ao HPV em todo o mundo (D'SOUZA e DEMPSEY, 2011). Entre os 12 casos HPV positivos (34,3%), somente um carcinoma de cavidade oral apresentava positividade para o HPV-55. Em concordância com o descrito em literatura, o HPV-16, de alto risco, foi o tipo predominante sendo encontrado em 91,6% dos casos. Aproximadamente 50% dos tumores de orofaringe são relatados como HPV-16 enquanto apenas 20% dos cânceres de cavidade oral apresentam positividade para o HPV-16 (MAJCHRZAK et al. 2014).

O carcinoma de células escamosas de orofaringe quando positivo para HPV, é considerado como uma entidade clínico-patológica distinta e associada a um melhor prognóstico do paciente (DAHLSTRAND et al. 2008; VIDAL e GILLISON 2008). Neste estudo, a infecção por HPV foi de 29,1% entre os 24 pacientes com carcinomas da cavidade oral e de 45,5% entre os casos de carcinoma de orofaringe (considerando a exclusão do paciente CPJ33). Entre

os pacientes com carcinomas de orofaringe HPV positivos, nenhum foi a óbito, até o presente, e somente um deles teve metástase, enquanto 67% dos pacientes HPV negativos faleceram.

Os primeiros estudos publicados utilizando NGS em CCECP datam de 2011 (Quadro 2). A qualidade dos dados de sequenciamento é determinada por uma série de fatores, como a cobertura, comprimento dos *reads* e plataforma utilizada. A média de cobertura e comprimento dos *reads* nem sempre estão disponíveis, como nos trabalhos de LUI et al. (2013) e LECHNER et al. (2013). A ausência destas informações levaa questionamentos sobre a confiabilidade dos resultados publicados, uma vez que as corridas de sequenciamento geram *reads* do genoma de forma randômica e independente. Em nosso estudo, a cobertura de mapeamento alcançou uma média de 48,7 vezes por amostra. Este dado é um importante controle de qualidade, uma vez que a alta cobertura permite avaliar as alterações genômicas com acurácia, diminuindo o número de falso positivos. Todas as amostras foram preparadas e sequenciadas nas mesmas condições. Para o sequenciamento, foram utilizadas quatro amostras por *lane*, o que de acordo com o kit utilizado para o preparo da biblioteca, gera uma cobertura em torno de 75x, como descrito previamente (Seção Material e Métodos). Regiões ricas em GC, tais como ilhas CpG, são particularmente propensas a baixa cobertura, em parte porque essas regiões permanecem aneladas durante a PCR realizada para o preparo das bibliotecas (SIMS et al. 2014). AJAY et al. (2011) relataram que uma cobertura média de 35x é o suficiente para a identificação de SNVs. NICHOLS et al. (2012a) reportaram uma cobertura de 28,1 vezes onde foram comparadas amostras HPV positivas e negativas de carcinomas de tonsila e língua

pareados com amostras de sangue periférico. Dessa forma, a média de cobertura alcançada nesse estudo foi superior a de outros, reforçando a qualidade e a credibilidade dos resultados obtidos.

O kit comercial para a captura do exoma utilizado neste estudo, é capaz de capturar regiões estendidas aos exons, como regiões intergênicas intrônicas e UTRs. Embora as mutações sinônimas ou silenciosas não modifiquem a sequência proteica e, assim, sendo consideradas funcionalmente irrelevantes, têm-se demonstrado seu envolvimento na tumorigênese. Acredita-se que estas mutações silenciosas em éxons podem estar envolvidas em diversos mecanismos moleculares e, geralmente, estão associados com alterações em sítios de *splicing* (SAPEKet al. 2014). No presente estudo, o enfoque foi dado às alterações não sinônimas e exclusivas dos tumores. Vale ressaltar que estudos posteriores serão desenvolvidos entre eles, a avaliação de sítios de *splicing* e a análise minuciosa das alterações germinativas.

Considerando o conjunto de resultados dos 34 casos pareados (caso CPJ-33 excluído, devido a baixa qualidade e cobertura, o que comprometeria os resultados), o perfil mutacional foi heterogêneo. Foi identificada uma média de 37.454 alterações, com uma variação de 158 a 20.525 SNVs em regiões de exons. Nossos resultados estão em conformidade com os descritos em literatura (STADLER et al. 2008; AGRAWAL et al. 2011; LEEMANS et al. 2011; STRANSKY et al. 2011). PICKERING et al. (2014a) sequenciaram o exoma total de amostras normais e tumorais pareadas de 39 pacientes de CCECP, encontrando uma variação de 86 a 12.112 alterações. Já em uma análise de sequenciamento do exoma de 12 diferentes tipos de câncer, o número de alterações somáticas encontradas variou de um a 14.204 para todos os 3.281

casos. Especificamente em tumores de cabeça e pescoço essa variação foi de seis a 2.146 mutações em 308 casos (KANDOTH et al. 2013).

Dentre as alterações mais comuns em CCECP, estão as descritas no gene *TP53*. O gene supressor tumoral *TP53* apresenta-se mutado em uma variedade de tumores humanos, incluindo uma alta frequência de alterações em carcinomas de cabeça e pescoço (AGRAWAL et al. 2011; STRANSKY et al. 2011; PICKERING et al. 2013; PICKERING et al. 2014a). Em nosso estudo, foi observada uma baixa frequência de alterações neste gene sendo identificadas apenas cinco variantes (Quadro 4). Nenhuma das variantes foi identificada em mais de um indivíduo e todas foram preditas como patogênicas. Dados em literatura evidenciam que, de um modo geral, o gene *TP53* (o qual é inativado em nível proteico pela E6) não se apresenta mutado nos tumores HPV positivos, enquanto nos tumores negativos o gene apresenta-se frequentemente mutado (SHI et al. 2009; VAN MONSJOU et al. 2013). Não foi possível estabelecer um perfil para os pacientes com as alterações em *TP53*, pois esses pacientes não compartilhavam características do estadiamento ao diagnóstico. Os pacientes com a mutação no gene *TP53* apresentaram idade variando de 35 a 49 anos, três casos positivos para infecção pelo HPV e dois casos negativos sendo quatro carcinomas orais e um de orofaringe. Estes dados não concordam, em sua totalidade, com os dados da literatura.

Considerando as variantes identificadas no gene *TP53*, duas (rs28934578 e rs121912656) foram previamente relacionadas à síndrome de Li-Fraumeni e uma ao carcinoma hepatocelular (rs28934571) no banco de dados COSMIC. A variante rs28934578 é caracterizada pela substituição de arginina por histidina no códon 175 de ilha CpG, induzindo a formação de

ligação de hidrogênio entre a histidina em 175 e histidina em 193. Essa mutação, G13203A (rs28934578), foi descrita por DAMINENI et al. (2014), como forte fator de risco de início precoce de câncer de mama relacionado à síndrome de Li-Fraumeni. Segundo os prontuários e os dados obtidos junto ao Departamento de Oncogenética, os pacientes incluídos neste estudo não possuem histórico familiar significativo e não se enquadram nos critérios da Síndrome de Li-Fraumeni [pacientes com sarcoma na infância ou idade jovem (< 45 anos) e parente de primeiro grau com qualquer câncer em idade jovem (< 45 anos) e parente de primeiro ou segundo grau com diagnóstico de câncer em idade jovem (< 45 anos) ou sarcoma em qualquer idade]. Ressaltando que, neste momento, as variantes discutidas em nosso estudo, foram identificadas na amostra tumoral, não sendo relacionadas aqui, à predisposição a síndromes hereditárias. As três variantes foram encontradas apenas em CEC de língua e ausentes nos demais subsítios analisados. As variantes rs28934578 e rs121912656 foram exclusivas dos pacientes HPV positivos enquanto rs28934571 foi identificada em um paciente HPV negativo, e ausente no grupo HPV positivo. A variante rs28934578 foi identificada apenas entre os não tabagistas enquanto que pacientes tabagistas apresentaram as variantes rs121912656 e rs28934571. A variante identificada no paciente CPJ-07 (c.527G>A/p.C176Y) e a deleção (c.902del1) identificada no paciente CPJ-27 ainda não foram descritas nos bancos de dados utilizados para este estudo.

AGRAWAL et al. (2011) e STRANSKY et al. (2011) publicaram simultaneamente, os primeiros dados de NGS em CCECP. Os dois estudos utilizaram amostras pareadas de tumor primário e sangue periférico (ou tecido normal) de indivíduos com carcinomas de diferentes subsítios anatômicos da

cabeça e pescoço. Como resultado desses estudos, foram confirmadas mutações em vários genes previamente descritas em CCECP, incluindo *TP53*, *CDKN2A*, *FAT1*, *PTEN*, *HRAS*, *PIK3CA* e *EGFR*. O principal achado dos dois grupos foram mutações nos genes *NOTCH1* e *FBXW7*, até aquele momento ainda não descritas em CCECP. No presente estudo não foi encontrada nenhuma alteração somática não-sinônima no gene *FBXW7*. Além disso, foi encontrado somente um caso com alteração em *NOTCH1*. Em uma avaliação dos genes mais comumente descritos em literatura como relacionados ao desenvolvimento de CCECP (Tabela 5), apenas o gene *CASP8* apresentou uma frequência de pelo menos 20% de alteração entre os casos aqui estudados. A análise de variação do número de cópias de DNA em CCECP do TCGA (2015) identificou um subgrupo de tumores com índice baixo de alterações no número de cópias genômicas, frequentemente com mutações em *CASP8* e *HRAS*, mas com *TP53* selvagem. Esses tumores, aparentemente apresentam uma via alternativa de tumorigênese, envolvendo RAS e/ou vias de sinalização NF-κB. Os autores não avaliaram se esses casos apresentavam correlação com idade ao diagnóstico (TCGA 2015).

Embora a frequência de casos com alterações nesses genes tenha sido menor do que a esperada, em nossa casuística, é importante ressaltar que diversas outras alterações em outros genes supressores e oncogenes também foram identificadas, e estas podem ter um papel relevante em CCECP de pacientes jovens. Das 28.206 SNVs não-sinônimas identificadas em 10.092 genes exclusivos das amostras tumorais, 705 SNVs em 506 genes que estavam presentes em pelo menos 20% dos casos e 47 foram descritos no banco de dados GSEA. Dentre esses, 6 genes foram classificados como oncogenes pelo

banco de dados GSEA (*BCL9*, *CASC5*, *EML4*, *THRAP3*, *TSHR* e *USP6*). Duas mutações presentes em 20,6% dos casos foram identificadas no gene *BCL9* (*B-cell lymphoma 9*). O seu envolvimento nos processos de crescimento tumoral, angiogênese, invasão e metástase se dá pela sua interação com a via de sinalização *Wnt/β-catenina*. A inibição de *BCL9* tem sido investigada com potencial terapêutico, o que indica que alterações nesse gene podem estar relacionadas a um maior risco de desenvolvimento tumoral (TAKADA et al. 2012). Outro oncogene alterado em 20,6% dos casos foi o *THRAP3* (*Thyroid hormone receptor associated protein 3*). Deleção nesse gene foi descrita em todos os carcinomas de células escamosas de cavidade oral em um estudo de sete amostras tumorais e adjacentes com displasia. O estudo mostrou que a alteração nesse gene já estava presente no tecido adjacente ao tumor com displasia (86% dos casos) (CHA et al. 2011). Em outro estudo de integração de dados de sequenciamento de genoma e transcriptoma total, o gene *THRAP3* foi indicado como um dos *drivers* em carcinoma hepatocelular (SHIRAISHI et al. 2014). A investigação do papel das alterações encontradas nesses genes para o desenvolvimento do CCECP em pacientes jovens deve ser validada em um grupo independente de casos. Além disso, o papel funcional variantes selecionadas será investigado em estudos posteriores. Essas análises não foram realizadas por inviabilidade de tempo e recursos.

Um item importante de se investigar é a correlação de desenvolvimento de CCECP em pacientes jovens com Anemia de Fanconi. A Anemia de Fanconi é uma síndrome hereditária que causa falência da medula óssea e está associada com o desenvolvimento precoce do CCECP na ausência dos fatores de risco comuns, como o consumo de tabaco e álcool. Pacientes com

AF têm um risco aumentado em 700 vezes de apresentar CCECP do que a população em geral e tendem a desenvolver a doença em idade jovem (ROSENBERG et al. 2003; ALTER et al. 2005). Nenhum paciente deste estudo foi diagnosticado com AF, entretanto, as alterações nos genes relacionados a esta síndrome foram investigadas. Entre os quatro pacientes que apresentaram a variante rs3864017 (presente na amostra pareada), apenas um foi HPV negativo e a idade com relação aos demais era aproximada, acima dos 40 anos. Porém é uma variante de significado clínico benigno (ClinVar), com última avaliação depositada no dbSNP em Junho de 2016. Entre os quatro amostras tumorais, nos quais foram identificadas a mesma variante, a infecção pelo HPV mostrou um comportamento diferente, apenas um caso era HPV positivo e apresentava idade inferior a 40 anos (39 anos). O paciente mais jovem do conjunto amostral (20 anos de idade), apresentava uma variante no gene *FANCA* (rs17233141), com alelo patogênico, de acordo com dbSNP, porém sem dados descritos em literatura. ZHANG et al. (2016), avaliaram seis casos pediátricos de AF por SET que revelaram alterações em *RPS19* (c.173-2A>G), *ELANE* (c.659G>A; (p.Arg220Gln), *ELANE* (c.182C>T; p.Ala61Val), (c.904C>T; p.Arg302Trp) e (c.2715+1G>A), *MYH9* (c.G283G>A; p.Ala95Thr) e *KRAS* (c.38G>A;p.Gly13Asp). Todas as alterações eram *missense* e estavam em heterozigose. A aplicação dessa técnica tem sido promissora, principalmente em casos nos quais o fenótipo não é bem esclarecido. Elucidar a relação entre AF e desenvolvimento de CCECP, teria grande impacto clínico.

Quando os tumores estão associados a infecção pelo HPV, estes constituem um perfil distinto com relação a epidemiologia, fatores de risco e comportamento clínico da doença. Dados em literatura apontam que pacientes

HPV- positivos tem prognóstico favorável, com melhor resposta ao tratamento e aumento da taxa de sobrevida, quando comparados aos casos HPV-negativos (LI et al. 2003; WEINBERGER et al. 2006). Não obstante, as mutações nesse grupo também apresentam um perfil distinto. O número de mutações identificadas em tumores HPV positivos foram relatadas com pelo menos metade das alterações identificadas em pacientes HPV negativos (LOYO et al. 2013). Essa informação alinha-se as variantes identificadas entre os pacientes HPV positivos e negativos do nosso estudo, onde a média de SNVs por paciente no grupo HPV positivo foi 336 SNVs em contraste com uma média de 2.252 SNVs em HPV negativos. Os casos HPV negativos apresentaram uma maior frequência e média de alterações.

Nas análises de vias e redes gênicas foi observado que vias relacionadas a componentes imunológicos estavam alteradas independentemente do status do HPV. Porém nos casos HPV positivos foi observado um maior número de vias relacionadas a imunidade adaptativa, incluindo a via sinalização do *PD1* e *TCR* e que são potenciais alvos imunoterapêuticos.

Recentes estudos evidenciam que a elevação de *PD-1* pode mediar a exaustão de células T, marcadas por diminuição da proliferação, produção de citocinas tipo I e baixa atividade citolítica (ZHANG et al. 2015). Quando *PD1* interage com seu ligante *PD-L1*, observa-se uma expressiva supressão de genes *TCR* e citocinas (LI R et al. 2015). Estudos em literatura reportam uma expressiva correlação entre a expressão de *PD-L1* em células tumorais e pior prognóstico entre vários tipos de tumores, incluindo carcinomas de esôfago, renal, adenocarcinoma pulmonar e CCECP (TAUBE et al. 2012; LIN et al.

2015). LIN et al. (2015) observaram que em pacientes do sexo masculino, fumantes e com carcinoma da cavidade oral, a alta expressão de *PD-L1* é um marcador de pior prognóstico, independente das características clínico-patológicas. Apenas um dos pacientes com a variante patogênica do *PD-L1* do nosso estudo apresentava carcinoma oral, era fumante e foi a óbito pela doença em um ano após o diagnóstico.

Avia de sinalização GPCR (*G-protein coupled receptors*) foi identificada entre as principais vias canônicas nos casos HPV negativos (Tabela 8). Essa família de receptores tem papel regulatório de vias de transdução de sinais, envolvendo quimiotaxia, processo de extrema importância no tráfego *in vivo* de células do sistema imune, incluindo linfócitos e neutrófilos, assim como células cancerosas relacionadas a metástase (JIN et al. 2008). MICHNA et al. (2016) utilizaram a linhagem celular CAL-33 de CCECP submetida a esquemas de radiação fracionada, similar ao tratamento por radioterapia, e analisaram o seu perfil transcriptômico. Os autores relacionaram os genes desta via GPCR como potenciais candidatos relacionados à resistência adquirida ao tratamento por radioterapia.

Dentre as comparações clínico-patológicas e as mutações encontradas, foi identificado um padrão heterogêneo de mutações em relação a localização do tumor. Os pacientes com CCECO (N=24) apresentaram 11.459 SNV em 6.087 genes, enquanto 7.082 SNVs em 4.469 genes foram identificados nos pacientes com diagnóstico de CCEOF. Os genes mais frequentemente mutados em CCECO (mais de 20% dos casos, N=23) foram *PDLIM5*, *NCOR1*, *INSL5* e *FOXD4*, enquanto os CCECOF tinham mutações mais frequentemente em *SIRPG*, *SEMA5B*, *POU5F1B*, *MICALL1*, *KIRDL2*,

HELQ e *GYPA*. Os dados avaliados pelo TCGA (2015) mostraram que os tumores de orofaringe possuem um menor número de mutações que os tumores da cavidade oral. Os genes mutados entre os diferentes sítios anatômicos foram *TP53*, *CASP8*, *NSD1* e *CDKN2A*. Os genes *TP53*, *PIK3CA* e *HLA-A* estavam significativamente mutados tanto nos carcinomas da cavidade oral quanto de orofaringe. No entanto, outros genes como *FAT1*, *CASP8*, *NOTCH1*, *HRAS* e *KMT2D* foram identificados como significativamente mutados somente nos tumores da cavidade oral, apresentando poucos ou nenhum caso mutado no grupo de tumores de orofaringe. Em contrapartida, os genes *B2M*, *CDH3* e *NPY5R* estavam frequentemente mutados em CCEOF, o que não foi observado nos tumores da cavidade oral. Os casos descritos no TCGA tinham idade ao diagnóstico variando de 19 a 90 anos, ainda que os autores tenham identificado essa variação no perfil mutacional de acordo com a localização tumoral, não foi avaliado como seria esse perfil somente nos casos mais jovens (TCGA 2015). Posteriormente, outros trabalhos usaram os dados gerados pelo TCGA para avaliar pacientes jovens, mas não dividiram os mesmos por sítio anatômico (MEUCCI et al. 2016).

FERLITO et al. (2001) relataram que o pulmão é o sítio de acometimento específico de 66% das metástases a distância dos CCECP. Porém, segundo WOO et al. (2009), metástases são raras em pacientes jovens. No nosso estudo, sete dos 34 pacientes desenvolveram metástase, seis delas em pulmão e uma metástase cervical. O perfil de mutações identificados por SET nos pacientes com metástase revelou 10 SNVs frequentemente mutados e patogênicos. Porém, não há dados sobre essas SNV sem literatura. Em três pacientes (CPJ05, CPJ27 e CPJ34) foram

identificadas praticamente as mesmas variantes exclusivas dos pacientes com metástase, sendo o estadiamento N2 a única característica em comum; entre todos os casos apenas um foi a óbito (o único com carcinoma de orofaringe e positivo para os três principais cofatores para o desenvolvimento de CCECP – tabagista, etilista e HPV-16). O gene *MGA* (*MAX dimerization protein*), mapeado em 15q15.1, é dos genes com alterações entre esses três pacientes.

ROMERO et al. (2014) observaram que a inativação da proteína de dimerização MAX/MGA, foi observada no câncer de pulmão de células pequenas e não pequenas. A inativação de MAX é mutuamente exclusiva a alterações de *MYC* e *BRG1*, onde a última codifica uma ATPase do complexo switch / sacarose não fermentável (SWI / SNF). Os autores relataram que *BRG1* regula a expressão de MAX por meio do recrutamento direto para o promotor MAX, e que a depleção de *BRG1* impede fortemente o crescimento celular, especificamente em células com deficiência de MAX, anunciando uma interação sintética letal. O que aponta um papel importante desse gene no desenvolvimento tumoral e possível relação com o desenvolvimento de metástases (ROMERO et al. 2014).

Neste estudo foi possível identificar relevantes alterações genéticas relacionadas ao desenvolvimento dos carcinomas de células escamosas de cabeça e pescoço de pacientes jovens. Foi realizado o sequenciamento do exoma total de amostras pareadas de sangue periférico e tumor desses pacientes sendo possível identificar alterações características de tumores pacientes jovens e sua correlação com dados clínicos e patológicos. As alterações aqui identificadas devem ser confirmadas em outros estudos e tem potencial para contribuir para a melhor entendimento da biologia de tumores

dos pacientes jovens. Foram identificadas alterações relevantes no gene da Anemia de Fanconi assim como alterações específicas encontradas em um grupo de pacientes com metástases. Entre as principais vias canônicas alteradas, foi observada que as relacionadas a componentes imunológicos tem relevância, independentemente do status do HPV. Estes achados podem contribuir a identificar pacientes que poderiam se beneficiar de um tratamento imunológico antes da quimio/ radioterapia. Os dados deste estudo tem potencial para apontar para marcadores diagnósticos, prognósticos e de tratamento para este grupo de pacientes.

6 CONCLUSÃO

- O perfil de alteração genética de tumores provenientes de pacientes com CCECP em idade jovens é heterogêneo. A frequência de alterações genéticas foi semelhante àquelas as descritas em literatura, mas o perfil mutacional foi distinto.
- O gene *TP53* apresentou alteração em cinco pacientes(14%), diferente ao descrito para pacientes mais velhos (30 dos 35 pacientes apresentaram uma cobertura considerada muito boa para este gene, acima de 60x).
- Os casos HPV-positivos apresentaram um menor número de mutações quando comparados com os casos HPV-negativos. Os casos HPV positivos apresentaram as via sinalização do PD1 e TCR alteradas. Essas vias são potenciais alvos imuno-terapêuticos e tem potencial para serem utilizadas na prática clínica.
- Foram detectados seis pacientes com variantes nos genes *FANCA* ou *FANCD2*, os quais estão relacionados a Anemia de Fanconi. Entretanto, nenhum paciente preenchia os critérios clínicos para esta síndrome.
- Foi detectada uma assinatura genética mutacional exclusiva de pacientes que apresentaram metástases. Dez SNVs preditos como patogênicos estavam presentes em pelo menos três dos sete pacientes que apresentaram metástase pulmonar. Estas variantes tem potencial para contribuir para o risco de desenvolvimento de metástases em pacientes jovens com tumores de cabeça e pescoço.

7 REFERENCIAS BIBLIOGRAFICAS

Adeel M, Suhail A, Zaman SU. Squamous cell carcinoma of oral tongue in young patients - A 10 years tertiary care experience. **J Pak Med Assoc** 2016; 66:155-8.

Agrawal N, Frederick MJ, Pickering CR, et al. Exome sequencing of head and neck squamous cell carcinoma reveals inactivating mutations in NOTCH1. **Science** 2011; 333:1154-7.

Ajay SS, Parker SC, Abaan HO, Fajardo KV, Margulies EH. Accurate and comprehensive sequencing of personal genomes. **Genome Res** 2011; 21:1498-505.

Alter BP, Joenje H, Oostra AB, Pals G. Fanconi anemia: adult head and neck cancer and hematopoietic mosaicism. **Arch Otolaryngol Head Neck Surg** 2005; 131:635-9.

Ambulos NP, Schumaker LM, Mathias TJ, et al. Next-generation sequencing-based hpv genotyping assay validated in formalin-fixed, paraffin-embedded oropharyngeal and cervical cancer specimens. **J Biomol Tech** 2016; 27:46-52.

Amit M, Yen TC, Liao CT, et al. Prognostic performance of current stage iii oral cancer patients after curative intent resection: evidence to support a revision of the American Joint Committee on Cancer Staging System. **Ann Surg Oncol** 2015; 22 Suppl 3:S985-91.

Ang KK, Harris J, Wheeler R, Weber R, et al. Human papillomavirus and survival of patients with oropharyngeal cancer. **N Engl J Med** 2010; 363:24-35.

Asangani IA, Dommeti VL, Wang X, et al. Therapeutic targeting of BET bromodomain proteins in castration-resistant prostate cancer. **Nature** 2014; 510:278-82.

Bamshad MJ, Ng SB, Bigham AW, et al. Exome sequencing as a tool for Mendelian disease gene discovery. **Nat Rev Genet** 2011; 12:745-55.

Baykul T, Yilmaz HH, Aydin U, Aydin MA, Aksoy M, Yildirim D. Early diagnosis of oral cancer. **J Int Med Res** 2010; 38:737-49.

Belcher R, Hayes K, Fedewa S, Chen AY. Current treatment of head and neck squamous cell cancer. **J Surg Oncol** 2014; 110:551-74.

Birkeland AC, Brenner JC. Personalizing medicine in head and neck squamous cell carcinoma: the rationale for combination therapies. **Med Res Arch** 2015; 3.

Biswas NK, Das S, Maitra A, Sarin R, Majumder PP. Somatic mutations in arachidonic acid metabolism pathway genes enhance oral cancer post-treatment disease-free survival. **Nat Commun** 2014; 5:5835.

Bodner L, Manor E, Friger MD, van der Waal I. Oral squamous cell carcinoma in patients twenty years of age or younger--review and analysis of 186 reported cases. **Oral Oncol** 2014; 50:84-9.

Brenner JC, Ateeq B, Li Y, et al. Mechanistic rationale for inhibition of poly(ADP-ribose) polymerase in ETS gene fusion-positive prostate cancer. **Cancer Cell** 2011; 19:664-78.

Cassidy M, Brett FM, Cryan J, Catháin NO, Curtis M. The contribution of alcohol to fatal traumatic head injuries in the forensic setting. **Ir Med J** 2010; 103:303-5.

Cha J-D, Kim HJ, Cha I-H. Genetic alterations in oral squamous cell carcinoma progression detected by combining array-based comparative genomic hybridization and multiplex ligation-dependent probe amplification. **Oral Surg**

Oral Med Oral Pathol Oral Radiol Endod 2011; 111:594-607.

Chaturvedi AK, Engels EA, Pfeiffer RM, et al. Human papillomavirus and rising oropharyngeal cancer incidence in the United States. **J Clin Oncol** 2011; 29:4294-301.

Chung CH, Guthrie VB, Masica DL, et al. Genomic alterations in head and neck squamous cell carcinoma determined by cancer gene-targeted sequencing. **Ann Oncol** 2015; 26:1216-23.

Curado MP, Hashibe M. Recent changes in the epidemiology of head and neck cancer. **Curr Opin Oncol** 2009; 21:194-200.

Dahlstrand H, Näsman A, Romanitan M, Lindquist D, Ramqvist T, Dalianis T. Human papillomavirus accounts both for increased incidence and better prognosis in tonsillar cancer. **Anticancer Res** 2008; 28:1133-8.

Damineni S, Rao VR, Kumar S, et al. Germline mutations of TP53 gene in breast cancer. **Tumour Biol** 2014; 35:9219-27.

Danecek P, Auton A, Abecasis G, et al. The variant call format and VCFtools. **Bioinformatics** 2011; 27:2156-8.

Deschler DG, Richmon JD, Khariwala SS, Ferris RL, Wang MB. The "new" head and neck cancer patient-young, nonsmoker, nondrinker, and HPV positive: evaluation. **Otolaryngol Head Neck Surg** 2014; 151:375-80.

D'Souza G, Cullen K, Bowie J, Thorpe R, Fakhry C. Differences in oral sexual behaviors by gender, age, and race explain observed differences in prevalence of oral human papillomavirus infection. **PLoS One** 2014; 9:e86023.

D'Souza G, Dempsey A. The role of HPV in head and neck cancer and review of the HPV vaccine. **Prev Med** 2011; 53:5-11.

D'Souza G, Kreimer AR, Viscidi R, et al. Case-control study of human papillomavirus and oropharyngeal cancer. **N Engl J Med** 2007; 356:1944-56.

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

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

Ferlito A, Shaha AR, Silver CE, Rinaldo A, Mondin V. Incidence and sites of distant metastases from head and neck cancer. **ORL J Otorhinolaryngol Relat Spec** 2001; 63:202-7.

Giefing M, Wierzbicka M, Szyfter K, et al. Moving towards personalised therapy in head and neck squamous cell carcinoma through analysis of next generation sequencing data. **Eur J Cancer** 2016; 55:147-57.

Gillison ML, Koch WM, Capone RB, et al. Evidence for a causal association between human papillomavirus and a subset of head and neck cancers. **J Natl Cancer Inst** 2000; 92:709-20.

Gourin CG, Kaboli KC, Blume EJ, Nance MA, Koch WM. Characteristics of participants in a free oral, head and neck cancer screening program. **Laryngoscope** 2009; 119:679-82.

Hashibe M, Brennan P, Benhamou S, 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** 2007; 99:777-89.

Hilly O, Shkedy Y, Hod R, et al. Carcinoma of the oral tongue in patients younger than 30 years: comparison with patients older than 60 years. **Oral Oncol** 2013; 49:987-90.

Hyam DM, Conway RC, Sathiyaseelan Y, et al. Tongue cancer: do patients younger than 40 do worse? **Aust Dent J** 2003; 48:50-4.

India Project Team of the International Cancer Genome Consortium. Mutational landscape of gingivo-buccal oral squamous cell carcinoma reveals new recurrently-mutated genes and molecular subgroups. **Nat Commun** 2013; 4:2873.

Jaber MA. Oral epithelial dysplasia in non-users of tobacco and alcohol: an analysis of clinicopathologic characteristics and treatment outcome. **J Oral Sci** 2010; 52:13-21.

Jemal A, Siegel R, Ward E, et al. Cancer Statistics 2006. **CA Cancer J Clin** 2006; 56:106-30.

Jin T, Xu X, Hereld D. Chemotaxis, chemokine receptors and human disease. **Cytokine** 2008; 44:1-8.

Jones S, Anagnostou V, Lytle K, et al. Personalized genomic analyses for cancer mutation discovery and interpretation. **Sci Transl Med** 2015; 7:283ra53.

Kaminagakura E, Caris A, Coutinho-Camillo C, Soares FA, Takahama-Júnior A, Kowalski LP. Protein expression of CYP1A1, CYP1B1, ALDH1A1, and ALDH2 in young patients with oral squamous cell carcinoma. **Int J Oral Maxillofac Surg** 2016; 45:706-12.

Kaminagakura E, Villa LL, Andreoli MA, et al. High-risk human papillomavirus in oral squamous cell carcinoma of young patients. **Int J Cancer** 2012; 130:1726–32.

Kandoth C, McLellan MD, Vandin F, et al. Mutational landscape and significance across 12 major cancer types. **Nature** 2013; 502:333-9.

Kang H, Kiess A, Chung CH. Emerging biomarkers in head and neck cancer in the era of genomics. **Nat Rev Clin Oncol** 2015; 12:11-26.

Ku CS, Cooper DN, Polychronakos C, Naidoo N, Wu M, Soong R. Exome sequencing: dual role as a discovery and diagnostic tool. **Ann Neurol** 2012; 71:5-14.

Kuriakose M, Sankaranarayanan M, Nair MK, et al. Comparison of oral squamous cell carcinoma in younger and older patients in India. **Eur J Cancer Part B Oral Oncol** 1992; 28:113-20.

Lechner M, Frampton GM, Fenton T, et al. Targeted next-generation sequencing of head and neck squamous cell carcinoma identifies novel genetic alterations in HPV+ and HPV- tumors. **Genome Med** 2013; 5:49.

Leemans CR, Braakhuis BJ, Brakenhoff RH. The molecular biology of head and neck cancer. **Nat Rev Cancer** 2011; 11:9-22.

Li H, Handsaker B, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools. **Bioinformatics** 2009; 25:2078-9.

Li H, Wawrose JS, Gooding WE, et al. Genomic analysis of head and neck squamous cell carcinoma cell lines and human tumors: a rational approach to preclinical model selection. **Mol Cancer Res** 2014;12:571-82.

Li R, Faden DL, Fakhry C, et al. Clinical, genomic, and metagenomic characterization of oral tongue squamous cell carcinoma in patients who do not smoke. **Head Neck** 2015; 37:1642-9.

Li Q, Prince ME, Moyer JS. Immunotherapy for head and neck squamous cell carcinoma. **Oral Oncol** 2015; 51:299-304.

Li W, Thompson CH, O'Brien CJ, et al. Human papillomavirus positivity predicts favourable outcome for squamous carcinoma of the tonsil. **Int J Cancer** 2003; 106:553-8.

Lin YM, Sung WW, Hsieh MJ, et al. High PD-L1 Expression Correlates with Metastasis and Poor Prognosis in Oral Squamous Cell Carcinoma. **PLoS One** 2015; 10:e0142656.

Liu YF, Chiang SL, Lin CY, et al. Somatic mutations and genetic variants of NOTCH1 in head and neck squamous cell carcinoma occurrence and development. **Sci Rep** 2016; 6:28409.

Llewellyn CD, Linklater K, Bell J, Johnson NW, Warnakulasuriya KA. An analysis of risk factors for oral cancer in young people: a case-control study. **Oral Oncol** 2004; 40:304-13.

Lowy DR, Munger K. Prognostic implications of HPV in oropharyngeal cancer. **N Engl J Med** 2010; 363:82-4.

Loyo M, Li RJ, Bettgowda C, et al. Lessons learned from next-generation sequencing in head and neck cancer. **Head Neck** 2013; 35:454-63.

Lui VW, Hedberg ML, Li H, et al. Frequent mutation of the PI3K pathway in head and neck cancer defines predictive biomarkers. **Cancer Discov** 2013; 3:761-9.

Majchrzak E, Szybiak B, Wegner A, et al. Oral cavity and oropharyngeal squamous cell carcinoma in young adults: a review of the literature. **Radiol Oncol** 2014; 48:1-10.

Majewski J, Schwartzentruber J, Lalonde E, Montpetit A, Jabado N. What can exome sequencing do for you? **J Med Genet** 2011; 48:580-9.

Manuel S, Raghavan SK, Pandey M, Sebastian P. Survival in patients under 45

years with squamous cell carcinoma of the oral tongue. **Int J Oral Maxillofac Surg** 2003;32:167-73.

Marta GN, Riera R, Bossi P, et al. Induction chemotherapy prior to surgery with or without postoperative radiotherapy for oral cavity cancer patients: systematic review and meta-analysis. **Eur J Cancer** 2015; 51:2596-603.

Marta GN, William WN, Feher O, Carvalho AL, Kowalski LP. Induction chemotherapy for oral cavity cancer patients: current status and future perspectives. **Oral Oncol** 2015; 51:1069-75.

Marur S, Forastiere AA. Head and neck squamous cell carcinoma: update on epidemiology, diagnosis, and treatment. **Mayo Clin Proc** 2016; 91:386-96.

McKenna A, Hanna M, Banks E, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. **Genome Res** 2010; 20:1297-303.

McLaughlin-Drubin ME, Münger K. Oncogenic activities of human papillomaviruses. **Virus Res** 2009; 143:195-208.

Meucci S, Keilholz U, Tinhofer I, Ebner O. Mutational load and mutational patterns in relation to age in head and neck cancer. **Oncotarget** 2016; 7:69188-99.

Michna A, Schötz U, Selmsberger M, et al. J. Transcriptomic analyses of the radiation response in head and neck squamous cell carcinoma subclones with different radiation sensitivity: time-course gene expression profiles and gene association networks. **Radiat Oncol** 2016; 26;11:94.

Ministério da Saúde. Instituto Nacional do Câncer José Alencar Gomes da Silva. **Estimativa/2016 Incidência de câncer no Brasil**. Rio de Janeiro: INCA; 2015.

Moody CA, Laimins LA. Human papillomavirus oncoproteins: pathways to transformation. **Nat Rev Cancer** 2010; 10:550-60.

Müller S, Pan Y, Li R, Chi AC. Changing trends in oral squamous cell carcinoma with particular reference to young patients: 1971-2006. The Emory University experience. **Head Neck Pathol** 2008; 2:60-6.

Myers JN, Elkins T, Roberts D, Byers RM. Squamous cell carcinoma of the tongue in young adults: increasing incidence and factors that predict treatment outcomes. **Otolaryngol Head Neck Surg** 2000; 122:44-51.

[NIH] National Cancer Institute. Surveillance, Epidemiology, and Results Program-SEER. **Cancer stat fact: oral cavity and pharynx cancer**. Available from: <URL:<https://seer.cancer.gov/statfacts/html/oralcav.html>> [2016 set 26].

Nguyen BC, Lefort K, Mandinova A, et al. Cross-regulation between Notch and p63 in keratinocyte commitment to differentiation. **Genes Dev** 2006; 20:1028-42.

Nichols AC, Chan-Seng-Yue M, Yoo J, et al. A pilot study comparing HPV-Positive and HPV-negative head and neck squamous cell carcinomas by whole exome sequencing. **ISRN Oncol** 2012a; 2012:809370.

Nichols AC, Yoo J, Palma DA, et al. Frequent mutations in TP53 and CDKN2A found by next-generation sequencing of head and neck cancer cell lines. **Arch Otolaryngol Head Neck Surg** 2012b; 138:732-9.

Okami K. Clinical features and treatment strategy for HPV-related oropharyngeal cancer. **Int J Clin Oncol** 2016;21:827-35.

Olsen MH, Bøje CR, Kjær TK, et al. Socioeconomic position and stage at diagnosis of head and neck cancer - a nationwide study from DAHANCA. **Acta Oncol** 2015; 54:759-66.

Pai SI, Westra WH. Molecular pathology of head and neck cancer: implications for diagnosis, prognosis, and treatment. **Annu Rev Pathol** 2009; 4:49-70.

Patel SC, Carpenter WR, Tyree S, et al. Increasing incidence of oral tongue squamous cell carcinoma in young white women, age 18 to 44 years. **J Clin Oncol** 2011; 29:1488-94.

Pickering CR, Zhang J, Yoo SY, et al. Integrative genomic characterization of oral squamous cell carcinoma identifies frequent somatic drivers. **Cancer Discov** 2013; 3:770-81.

Pickering CR, Zhang J, Neskey DM, et al. Squamous cell carcinoma of the oral tongue in young non-smokers is genomically similar to tumors in older smokers. **Clin Cancer Res** 2014a; 20:3842-8.

Pickering CR, Zhou JH, Lee JJ, et al. Mutational landscape of aggressive cutaneous squamous cell carcinoma. **Clin Cancer Res** 2014b; 20:6582-92.

Pitman KT, Johnson JT, Wagner RL, Myers EN. Cancer of the tongue in patients less than forty. **Head Neck** 2000; 22:297-302.

Popovtzer A, Shpitzer T, Bahar G, Marshak G, Ulanovski D, Feinmesser R. Squamous cell carcinoma of the oral tongue in young patients. **Laryngoscope** 2004; 114:915-7.

Robinson DR, Wu YM, Vats P, et al. Activating ESR1 mutations in hormone-resistant metastatic breast cancer. **Nat Genet** 2013; 45:1446-51.

Rodrigo JP, Kowalski LP, Silver CE, et al. When is reoperative surgery not indicated for recurrent head and neck squamous cell carcinoma? **Eur Arch Oto-Rhino-Laryngol** 2015; 272:259-62.

Rodriguez T, Altieri A, Chatenoud L, et al. Risk factors for oral and pharyngeal cancer in young adults. **Oral Oncol** 2004; 40:207-13.

Romero OA, Torres-Diz M, Pros E et al. MAX inactivation in small cell lung cancer disrupts MYC-SWI/SNF programs and is synthetic lethal with BRG1. **Cancer Discov** 2014; 4:292-303.

Rosenberg PS, Greene MH, Alter BP. Cancer incidence in persons with Fanconi anemia. **Blood** 2003; 101:822-6.

Ryan CJ, Smith MR, de Bono JS, et al. Abiraterone in metastatic prostate cancer without previous chemotherapy. **N Engl J Med** 2013; 368:138-48.

Samman M, Wood HM, Conway C, et al. A novel genomic signature reclassifies an oral cancer subtype. **Int J Cancer** 2015; 137:2364-73.

Supek F, Miñana B, Valcárcel J, Gabaldón T, Lehner B. Synonymous mutations frequently act as driver mutations in human cancers. **Cell** 2014; 156:1324-35.

Sepiashvili L, Bruce JP, Huang SH, O'Sullivan B, Liu FF, Kislinger T. Novel insights into head and neck cancer using next-generation "omic" technologies. **Cancer Res** 2015; 75:480-6.

Shalem O, Sanjana NE, Hartenian E, et al. Genome-scale CRISPR-Cas9 knockout screening in human cells. **Science** 2014; 343:84-7.

Shi W, Kato H, Perez-Ordóñez B, et al. Comparative prognostic value of HPV16 E6 mRNA compared with in situ hybridization for human oropharyngeal squamous carcinoma. **J Clin Oncol** 2009; 27:6213-21.

Shiboski CH, Schmidt BL, Jordan RC. Tongue and tonsil carcinoma: increasing trends in the U.S. population ages 20-44 years. **Cancer** 2005; 103:1843-9.

Shiraishi Y, Fujimoto A, Furuta M, et al. Integrated analysis of whole genome and transcriptome sequencing reveals diverse transcriptomic aberrations driven by somatic genomic changes in liver cancers. **PLoS One** 2014; 9:1-22.

Sims D, Sudbery I, Iltott NE, Heger A, Ponting CP. Sequencing depth and coverage: key considerations in genomic analyses. **Nat Rev Genet** 2014; 15:121-32.

Soudry E, Preis M, Hod R, et al. Squamous cell carcinoma of the oral tongue in patients younger than 30 years: clinicopathologic features and outcome. **Clin Otolaryngol** 2010; 35:307-12.

Stadler ME, Patel MR, Couch ME, Hayes DN. Molecular biology of head and neck cancer: risks and pathways. **Hematol Oncol Clin North Am** 2008; 22:1099-124, vii.

Stranneheim H, Wedell A. Exome and genome sequencing: a revolution for the discovery and diagnosis of monogenic disorders. **J Intern Med** 2016; 279:3-15.

Stransky N, Egloff AM, Tward AD, et al. The mutational landscape of head and neck squamous cell carcinoma. **Science** 2011; 333:1157-60.

Sun W, Califano JA. Sequencing the head and neck cancer genome: implications for therapy. **Ann N Y Acad Sci** 2014; 1333:33-42.

Tabatabaeifar S, Kruse TA, Thomassen M, Larsen MJ, Sørensen JA. Use of next generation sequencing in head and neck squamous cell carcinomas: a review. **Oral Oncol** 2014; 50:1035-40.

Takada K, Zhu D, Bird GH, et al. Targeted Disruption of the BCL9/ -Catenin Complex Inhibits Oncogenic Wnt Signaling. **Sci Transl Med** 2012; 4:148ra117-148ra117.

TCGA The Cancer Genome Atlas Network. Comprehensive genomic characterization of head and neck squamous cell carcinomas. **Nature** 2015; 517:576-82.

Tinhofer I, Budach V, Saki M, et al. Targeted next-generation sequencing of locally advanced squamous cell carcinomas of the head and neck reveals druggable targets for improving adjuvant chemoradiation. **Eur J Cancer** 2016; 57:78-86.

Toner M, O'Regan EM. Head and neck squamous cell carcinoma in the young: A spectrum or a distinct group? Part 1. **Head Neck Pathol** 2009; 3:246-8.

Toporcov TN, Znaor A, Zhang ZF, et al. Risk factors for head and neck cancer in young adults: a pooled analysis in the INHANCE consortium. **Int J Epidemiol** 2015; 44:169-85.

Torossian JM, Beziat JL, Philip T, Bejui FT. Squamous cell carcinoma of the tongue in a 13-year-old boy. **J Oral Maxillofac Surg** 2000; 58:1407-10.

Tsantoulis PK, Kastrinakis NG, Tourvas AD, Laskaris G, Gorgoulis VG. Advances in the biology of oral cancer. **Oral Oncol** 2007; 43:523-34.

Udeabor SE, Rana M, Wegener G, Gellrich NC, Eckardt AM. Squamous cell carcinoma of the oral cavity and the oropharynx in patients less than 40 years of age: a 20-year analysis. **Head Neck Oncol** 2012;4:28.

Upadhyay P, Nair S, Kaur E, et al. Notch pathway activation is essential for maintenance of stem-like cells in early tongue cancer. **Oncotarget** 2016; 7:50437-49.

Van Monsjou HS, Wreesmann VB, van den Brekel MW, Balm AJ. Head and neck squamous cell carcinoma in young patients. **Oral Oncol** 2013; 49:1097-102.

Vettore AL, Ramnarayanan K, Poore G, et al. Mutational landscapes of tongue carcinoma reveal recurrent mutations in genes of therapeutic and prognostic relevance. **Genome Med** 2015; 7:98.

Vidal L, Gillison ML. Human papillomavirus in HNSCC: recognition of a distinct disease type. **Hematol Oncol Clin North Am** 2008; 22:1125-42, vii.

Vogelstein B, Papadopoulos N, Velculescu VE, et al. Cancer genome landscapes. **Science** 2013; 339:1546-58.

Wang C, Liu XQ, Hou JS, Wang JN, Huang HZ. Molecular mechanisms of chemoresistance in oral cancer. **Chin J Dent Res** 2016; 19:25-33.

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

Wang T, Wei JJ, Sabatini DM, Lander ES. Genetic screens in human cells using the CRISPR-Cas9 system. **Science** 2014; 343:80-4.

Weinberger PM, Yu Z, Haffty BG, et al. Molecular classification identifies a subset of human papillomavirus–associated oropharyngeal cancers with favorable prognosis. **J Clin Oncol** 2006; 24:736-47.

Woo VL, Kelsch RD, Su L, Kim T, Zegarelli DJ. Gingival squamous cellcarcinoma in adolescence.**Oral Surg Oral Med Oral Pathol Oral Radiol Endod** 2009; 107:92-9.

Yap LF, Lee D, Khairuddin A, et al. The opposing roles of NOTCH signalling in head and neck cancer: a mini review. **Oral Dis** 2015; 21:850-7.

Zhang J, Barbaro P, Guo Y, et al. Utility of next-generation sequencing technologies for the efficient genetic resolution of haematological disorders. **Clin Genet** 2016; 89:163-72.

Zhang Y, Gallastegui N, Rosenblatt JD. Regulatory B cells in anti-tumor immunity. **Int Immunol** 2015; 27:521-30.

Zur Hausen H. Papillomavirus infections--a major cause of human cancers. **Biochim Biophys Acta** 1996; 1288:F55-78.

Zur Hausen H. Papillomaviruses and cancer: from basic studies to clinical application. **Nat Rev Cancer** 2002; 2:342-50.

Anexo 1 - Carta de aprovação do Comitê de Ética em Pesquisa-CEP.

 **Hospital
A.C. Camargo**
Centro de Tratamento, Ensino e Pesquisa em Câncer

**Comitê de Ética em
Pesquisa - CEP**

São Paulo, 07 de março de 2013.

À
Dra. Silvia Regina Rogatto.

Aluno: Priscila Mayrink de Miranda (Doutorado).

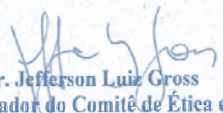
Ref.: Projeto de Pesquisa nº. 1727/12
“Rastreamento global de alterações genômicas em carcinomas de cabeça e pescoço de pacientes jovens por sequenciamento total do exoma”.

Os membros do Comitê de Ética em Pesquisa em Seres Humanos da Fundação Antonio Prudente – Hospital do Câncer - A.C. Camargo/SP, em sua última reunião de 05/02/2013, após analisarem as respostas aos questionamentos realizados em reunião de 11/12/2012, aprovaram a realização do projeto do estudo em referência (datado de janeiro de 2013) e tomaram conhecimento dos seguintes documentos:

- Folha de Rosto para Pesquisa Envolvendo Seres Humanos;
- Termo de Compromisso do Pesquisador com as Resoluções do Conselho Nacional de Saúde;
- Termo de Dispensa do Consentimento Livre e Esclarecido;
- Declaração sobre os Dados Coletados, Publicação dos Dados e Propriedade das Informações Geradas;
- Declaração de Ciência e Comprometimento do Biobanco do Hospital AC Camargo;
- Declaração sobre a publicação dos dados de pesquisa utilizando amostras fornecidas e processadas pelo Biobanco do Hospital A.C. Camargo;
- Declaração de Ciência e Comprometimento do Departamento de Zootecnia – Laboratório de Biotecnologia Animal (USP – Piracicaba);
- Declaração de Ciência e Comprometimento do Departamento de Cirurgia de Cabeça e Pescoço e Otorrinolaringologia do Hospital AC Camargo;
- Declaração de Ciência e Comprometimento do Laboratório de NeoGene do CIPE do Hospital AC Camargo;
- Declaração de Infraestrutura e Instalações do CIPE (Centro Internacional Pesquisa e Ensino);
- Orçamento Financeiro Detalhado.

Informações a respeito do andamento do referido projeto deverão ser encaminhadas ao CEP dentro de 12 meses.

Atenciosamente,


Dr. Jefferson Luiz Gross
1º Vice-Coordenador do Comitê de Ética em Pesquisa

1/1

Fundação Antonio Prudente – CNPJ/MFEN. 60.961.958/0001-06
Rua Prof. Antônio Prudente, 211 – Liberdade – São Paulo, SP – 01509-900
Telefone: (11) 2189-5000
www.accamargo.org.br