Universidade de São Paulo
Instituto de Medicina Tropical de São Paulo

LUCIANA REIS ROSA SACOMAN

AVALIAÇÃO DO PERFIL GENÉTICO DO HPV16 E SEU SÍTIO DE INTEGRAÇÃO EM CÉLULAS EPITELIAIS NORMAIS E NEOPLÁSICAS DA CÉRVIX E TONSILAS

Tese apresentada ao Instituto de Medicina Tropical de São Paulo da Universidade de São Paulo para obtenção do título de Doutora em Ciências.

Área de concentração: Doenças Tropicais e Saúde Internacional

Orientador: Dr. José Eduardo Levi

São Paulo

2018

LUCIANA REIS ROSA SACOMAN

AVALIAÇÃO DO PERFIL GENÉTICO DO HPV16 E SEU SÍTIO DE INTEGRAÇÃO EM CÉLULAS EPITELIAIS NORMAIS E NEOPLÁSICAS DA CÉRVIX E TONSILAS

Tese apresentada ao Instituto de Medicina Tropical de São Paulo da Universidade de São Paulo para obtenção do título de Doutora em Ciências.

Área de concentração: Doenças Tropicais e Saúde Internacional

Orientador: Dr. José Eduardo Levi

São Paulo

2018

Ficha catalográfica elaborada pela Biblioteca do Instituto de Medicina Tropical de São Paulo da Universidade de São Paulo – Bibliotecário Carlos José Quinteiro, CRB-8 5538

Rosa-Sacoman, Luciana Reis
Avaliação do perfil genético do HPV16 e seu sítio de integração em células epiteliais normais e neoplásicas da cérvix e tonsilas / Luciana Reis Rosa-Sacoman. – São Paulo, 2018.
Tese (Doutorado) – Instituto de Medicina Tropical de São Paulo da Universidade de São Paulo, para obtenção do título de Doutor em Ciências.
Área de concentração: Doenças Tropicais e Saúde Internacional
Orientador: José Eduardo Levi

Descritores: 1. HPV. 2. PREVALÊNCIA. 3. NEOPLASIAS DE CABEÇA E PESCOÇO. 4. NEOPLASIAS DO COLO UTERINO. 5. BIOLOGIA MOLECULAR. 6. SEQUENCIAMENTO GENÉTICO.

USP/IMTSP/BIB-13/2018.

SUA MELHOR CONTRIBUIÇÃO PARA A HUMANIDADE TALVEZ NÃO SEJA ALGO QUE VOCÊ FAÇA, MAS ALGUÉM QUE VOCÊ CRIE.

- AUTOR DESCONHECIDO -

Dedicatória

Dedico esse trabalho ao meu filho, **Joaquim Reis Rosa Sacoman,** o maior e melhor presente que eu poderia ter recebido!

Obrigada pelo amor gratuito e pelos ensinamentos diários! O que eu aprendi com você, filho, ao longo desses quase dois anos de vida, não está escrito em nenhum artigo científico, nem nas mais renomadas revistas científicas! Você me transformou na melhor versão de mim, para que eu pudesse ser a melhor versão para você!

Te amo infinitamente, gratuitamente, eternamente...

Agradecimentos

AGRADECIMENTOS

À minha família,

pela confiança, reconhecimento e apoio integral! Em principal ao meu marido, *Thiago Marchi Sacoman,* pela paciência de sempre;

aos meus pais e maiores facilitadores desse feito, *Luiz Cláudio e Tereza,* que nunca mediram esforços para que os nossos sonhos e objetivos fossem concretizados;

e às minhas irmãs, *Liliana e Luana*, pela presença, pura e integral!

Eu amo todos vocês, com o mais puro e sincero sentimento!

É *por* vocês e *pra* vocês!

Aos meus familiares e amigos,

sem a leveza da presença de vocês, tudo teria sido mais difícil! Meus avós, em especial minha avó Antônia (*in memorian*) que nos enche de saudades, tios, primos, afilhado, sogros e cunhados. Minha gratidão!

Aos pesquisadores que tornaram possível a execução desse trabalho,

Dr. José Eduardo Levi, meu orientador, por ter acreditado nas minhas propostas, apoiado, estimulado e entendido todos os momentos vividos ao longo desses anos de pós-graduação! Obrigada Dudi!

Profa. Dra. Eloiza Helena Tajara da Silva, pesquisadora responsável pelo GENCAPO (Genoma do Câncer de Cabeça e Pescoço), que nunca mediu esforços para contribuir e ceder as amostras pelo bem dos pacientes portadores desse tumor! Um verdadeiro exemplo de humildade e competência!

Profa. Dra. Luisa Villa, coordenadora do INCT-HPV (Instituto Nacional de Ciência e Tecnologia das doenças do Papilomavírus Humano), e todos os membros

do grupo, que participaram intensamente desse projeto fornecendo instrumentos didáticos, materiais (os reagentes de sequenciamento) e científicos. Dra. Luisa, a senhora é inspiração pura! Obrigada por nunca desistir da ciência!

Dra. Raquel Ajub Moyses e Dr. Ronaldo Frizzarini, Disciplina de Cabeça e Pescoço da Faculdade de Medicina da Universidade de São Paulo e Disciplina de Otorrinolaringologia da Faculdade de Medicina da Universidade de São Paulo, pela colaboração com a estrutura e coleta das amostras de tonsilas de pacientes sem câncer. Obrigada pela contribuição! Assim como a aluna de Mestrado, *Jéssica Boscariol da Silva,* que executou a análise dessa série de pacientes! Boa sorte nas próximas etapas!

Dra. Cristina Mendes de Oliveira, durante seu período de pós-doutoramento, com a coleta e análise das amostras de carcinoma cervical e suporte ao grupo de HPV do laboratório de Virologia do IMT (Toni, Célia, Cris C.), e a todos os colegas pós-graduandos e funcionários do IMT, em especial ao *Rodrigo Pêssoa* e à *Ana Carolina Mamana*! Obrigada pelas risadas em qualquer situação!

Às instituições de fomento FAPESP, CNPq e CAPES, por darem suporte financeiro ao projeto, congressos e treinamentos!

A todos, meu muito obrigada!

Resumo

RESUMO

Rosa-Sacoman LR. Avaliação do perfil genético do HPV16 e seu sítio de integração em células epiteliais normais e neoplásicas da cérvix e tonsilas (tese). São Paulo: Instituto de Medicina Tropical de São Paulo da Universidade de São Paulo; 2018.

O papilomavírus humano (HPV) atua como importante agente etiológico em um subgrupo de tumores humanos, sendo responsável pelo desenvolvimento de quase 100% dos tumores cervicais e uma porcentagem variável dos carcinomas de cabeça e pescoço, principalmente da orofaringe. Dentre os HPVs de alto risco, o HPV16 é o tipo mais prevalente nos tumores de orofaringe, encontrado em mais de 80% dos tumores HPV positivos, e 50% dos casos de carcinoma cervical. Em lesões de baixo grau, o genoma do HPV geralmente permanece na forma epissomal, em contraste, na maioria das lesões cancerígenas, o DNA dos HPVs de alto risco é frequentemente quebrado e integrado ao genoma da célula hospedeira. Há um enorme esforço da comunidade científica na identificação de regiões susceptíveis à integração viral, permitindo uma maior compreensão do processo de transformação causado por HPV. Este estudo teve como objetivo identificar, a partir do sequenciamento do genoma completo do HPV16, padrões moleculares capazes de responder pelo tropismo diferencial apresentado pelo HPV16 e do curso assintomático ou transformante da infecção pelo HPV16 na região da orofaringe. Amostras FFPE de CECP estavam disponíveis em 510 pacientes que foram recrutados em três hospitais brasileiros e 113 pacientes sem neoplasia tiveram amostras a fresco de tonsila não neoplásica analisadas. A detecção e genotipagem do HPV foi realizada pelo método *Inno-Lipa*, enquanto a reação de PCR com os iniciadores *PGMY09/11* foi empregada para tecidos não neoplásicos. Todas as amostras positivas para HPV tiveram a presença de HPV16 investigada por PCR em tempo real. As amostras com presença

exclusiva de HPV16 foram submetidas ao sequenciamento pela metodologia de NGS – na plataforma *Ion Torrent PGM* utilizando a tecnologia de *Target Seq* e a análise dos produtos sequenciados foram realizadas em colaboração com o laboratório da Dr Lisa Mirabello, no NIH. A frequência de DNA de HPV de alto risco em CEC de cabeça e pescoço foi de 10% (49/491), sendo 78% deles HPV16. Houve grande variabilidade na prevalência do HPV nos tumores segundo o sítio anatômico, variando de 3,4% (base da língua e hipofaringe) a 25% (orofaringe). Em contraste, não houve presença de hrHPV-DNA nas amígdalas não neoplásicas analisadas. A análise do sequenciamento do HPV16 foi viável em quatro amostras, que apresentaram 99-100% de identidade com o HPV16. Uma delas apresentou uma deleção de 300nt na região L2 possivelmente atribuída ao processo de integração. No presente estudo, a crescente frequência de DNA de HPV em CECP (principalmente na orofaringe) corrobora a hipótese de que o HPV está iniciando sua projeção nesse continente, apesar de ainda não estar circulante entre o tecido normal de tonsila da população brasileira. Estudos envolvendo a avaliação da expressão proteica associada à infecção por HPV devem ser conduzidos para reforçar se esse perfil crescente de DNA de HPV também reflete uma crescente participação etiológica do HPV 16 nos novos casos de tumores epidermoides de orofaringe no Brasil.

Descritores: HPV. Prevalência. Neoplasias de cabeça e pescoço. Neoplasia do colo uterino. Biologia molecular. Sequenciamento genético.

Abstract

ABSTRACT

Rosa-Sacoman, LR. Evaluation of the genetic profile of HPV16 and its integration site in normal and neoplastic epithelial cells from cervix and tonsils (thesis). São Paulo: Instituto de Medicina Tropical de São Paulo da Universidade de São Paulo; 2018.

Human papillomavirus (HPV) acts as an important etiologic agent in a subset of human tumors, responsible for the development of almost 100% of cervical tumors and a variable percentage of head and neck carcinomas, mainly of the oropharynx. Among high-risk HPV, HPV16 is the most prevalent type in oropharyngeal tumors, found in more than 80% of HPV positive tumors, and 50% of cases of cervical carcinoma. In low grade lesions, the HPV genome generally remains epissomal; in contrast, in most cancerous lesions, the DNA of high-risk HPVs is often disrupted and integrated into the host cell genome. There is a huge effort by the scientific community to identify regions susceptible to viral integration, allowing a greater understanding of the transformation process caused by HPV. The aim of this study was to identify, from the complete genome sequencing of HPV16, molecular patterns capable of responding to the differential tropism presented by HPV16 and to the asymptomatic or transforming course of HPV16 infection in the oropharynx region. FFPE samples of HNSCC were available from 510 patients who were recruited at three Brazilian hospitals and 113 patients with no neoplasia had fresh samples of non-neoplastic tonsil analyzed. HPV detection and genotyping was performed using the *Inno-Lipa* method, while the PCR reaction with *PGMY09/11* primers was used for non-neoplastic tissues. All HPV positive samples had the presence of HPV16 investigated by real-time PCR. Samples with exclusive HPV16 presence were submitted to sequencing by the NGS methodology - on the *Ion Torrent PGM* platform using *Target Seq* technology and the analysis of the sequenced products were performed in

collaboration with Dr. Lisa Mirabello's laboratory at NIH. The frequency of high-risk HPV DNA in HNCSS was 10% (49/491), 78% of which were HPV16. There was a huge variability in the prevalence of HPV in the tumors according to the anatomical site, ranging from 3.4% (base of the tongue and hypopharynx) to 25% (oropharynx). In contrast, no hrHPV-DNA was present in the non-neoplastic tonsils analyzed. HPV16 sequencing analysis was feasible in four samples, which showed 99-100% identity with HPV16. One of them presented a 300nt deletion in the L2 region possibly attributed to the integration process. In the present study, the increasing frequency of HPV DNA in HNSCC (mainly in the oropharynx) corroborates the hypothesis that HPV is initiating its projection in this continent, although it is not yet circulating among the normal tonsil tissue of the Brazilian population. Studies involving the evaluation of protein expression associated with HPV infection should be conducted to reinforce whether this increasing profile of HPV DNA also reflects an increasing etiological involvement of HPV 16 in the new cases of oropharyngeal squamous cell carcinomas in Brazil.

Descriptors: HPV. Prevalence. Head and neck neoplasia. Cervical neoplasia. Molecular biology. Genetic sequencing.

Lista de abreviaturas e siglas

LISTA DE ABREVIATURAS E SIGLAS

CCP	Carcinoma de Cabeça e Pescoço
CECP	Carcinomas Epidermoides de Cabeça e Pescoço
CEP	Comitê de Ética em Pesquisa
CONEP	Comissão Nacional de Ética em Pesquisa
DNA	do inglês, Ácido Desoxirribonucléico
DP	Desvio Padrão
GENCAPO	Grupo de Estudo de Câncer de Cabeça e Pescoço
HC	Hospital das Clínicas de São Paulo
HH	Hospital Heliópolis, São Paulo
HPV	do inglês, PapilomaVírus Humano
IACR	do inglês, Agência Internacional de Pesquisa em Câncer, Lyon, França
ICAVC	Instituto do Câncer Arnaldo Vieira de Carvalho, São Paulo
NCI	do inglês, Instituto Nacional do Câncer, Rockville, MD, EUA
ng	nanogramas
NGS	do inglês, Sequenciamento de Nova Geração
NIC	Neoplasia Intraepitelial Cervical
NIH	do inglês, Institutos Nacionais de Saúde, Rockville, MD, EUA
PCR	do inglês, Reação em Cadeia da Polimerase
PB	pares de base
QC	do inglês, Controle de qualidade
qPCR	do inglês, PCR quantitativa (PCR em Tempo Real)
RNA	do inglês, Ácido Ribonucléico
µL	microlitros

Sumário

SUMÁRIO

Introdução

1 INTRODUÇÃO

1.1 Papilomavírus Humano

O Papilomavírus Humano (HPV) foi inicialmente identificado em 1949 em células epiteliais[1] e, desde então, já foram caracterizados mais de 200 tipos desse vírus[2,3]. Esses vírus são partículas pequenas de DNA circular dupla fita com genoma composto por aproximadamente 8.000 pares de base (pb), sendo sua infecção restrita às células epiteliais[4].

Os tipos de HPV são divididos em dois grupos: os de alto e baixo risco, de acordo com seu potencial maligno[5]. Os HPV de alto risco compreendem os tipos HPV16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 68, 73 e 82, presentes em lesões escamosas de alto grau ou carcinomas invasivos, e os tipos HPV6, 11, 40, 42, 44, 54, 61, 70, 72, 81 e 89 são os considerados de baixo risco, podendo ser isolados de lesões epiteliais de baixo grau [6,7].

A maioria das infecções pelo HPV são transitórias, assintomáticas ou subclínicas, não resultando em doença e, geralmente, são suprimidas pelo sistema imunológico. Estudos epidemiológicos indicam que a prevalência do vírus na população masculina, independente do sítio, pode variar de 0 a 73%[8]. Além disso, estudos clínicos conduzidos com mulheres associaram o tabaco com uma maior persistência das infecções por tipos oncogênicos do HPV e com o risco aumentado de desenvolver carcinomas cervicais invasivos[9].

Estudos epidemiológicos recentes indicam que os HPV de alto risco, semelhantes aos que estão envolvidos no desenvolvimento e progressão dos carcinomas cervicais, estão também associados aos carcinomas epidermoides de cabeça e pescoço (CECP) de localizações anatômicas específicas. O sítio anatômico mais associado à infecção pelo HPV no trato aero-digestivo é a orofaringe, especialmente as tonsilas e base da língua[10,11,12,13,14].

Nos tumores cervicais, o HPV16, junto com o 18, são os tipos predominantes, sendo responsáveis por mais de 70% dos casos de câncer cervical e são os dois tipos de HPV mais frequentes em todas as regiões geográficas[15]. Dentre os HPV de alto risco, somente os tipos 16, 18, 31, 33 e 35 foram identificados como importantes na carcinogênese dos tumores de orofaringe, sendo o HPV16 o mais frequentemente detectado nesses tumores, responsável por aproximadamente 90% dos tumores HPV positivos[4,16,17].

A incidência crescente de carcinomas de orofaringe nos países desenvolvidos se deve principalmente a ocorrência da doença em pacientes jovens, não apresentando os fatores de risco clássicos, como tabaco e álcool. Esse fato leva a uma associação cada vez maior de um subgrupo desses tumores à infecção pelo papilomavírus[18,19,20].

A proporção de tumores HPV positivos varia substancialmente de acordo com o sítio anatômico tumoral. Este achado sugere que a proporção dos tumores HPV positivos pode ser dependente de fatores geográficos e genéticos, além de fatores comportamentais (como hábitos sexuais) e a ausência de padronização da metodologia para a identificação destes vírus em tumores humanos. Já nos tumores cervicais, propõe-se amplamente que todos os casos tem como etiologia uma infecção genital pelo HPV, independente de fatores paralelos[20].

A estratificação de risco para CECP é realizada pela localização anatômica, estágio e características histológicas do tumor. Com exceção do status de HPV, vários fatores de risco moleculares e clínicos que foram investigados têm utilidade clínica limitada[21]. As frações atributáveis ao HPV (HPV-AFs) para CECP, fora da orofaringe, são substancialmente menores.

O tipo HPV16 é particularmente ativo nos queratinócitos da mucosa oral sendo considerado o mais importante para o desenvolvimento dos carcinomas de orofaringe[22]. Tumores associados ao HPV16 apresentam alterações genéticas

características, entre elas o aumento da expressão da proteína supressora tumoral p16 e a ausência de mutações no gene supressor tumoral *TP53*[12,23,24,25,26].

Os HPV de alto risco produzem duas oncoproteínas virais, E6 e E7, necessárias para a replicação viral, estimulando a proliferação e atuando na transformação maligna e manutenção tumoral. Quando incorporados ao genoma humano, os genes *E6 e E7* encontram-se geralmente com expressão aumentada, levando a alterações nas vias $p14^{ARF}$/MDM2/p53 e $p16^{INK4a}$/ciclina D/pRb, respectivamente. Esta integração pode levar tanto a ativação como a inativação de genes celulares envolvidos no processo tumoral[27].

A inativação da proteína supressora tumoral p53 pela oncoproteína viral E6 altera sua via e leva ao descontrole do ciclo celular, enquanto a inativação de pRb (pela oncoproteína E7) leva à ativação do fator de transcrição E2F, resultando na perda do controle celular e no aumento da expressão da proteína p16[4,28,29]. Porém, nem sempre há uma correlação entre a presença do vírus e a ausência de mutação no gene *TP53*[27].

A proteína p16 é codificada pelo gene supressor tumoral *CDKN2A* e regula a atividade do complexo de ciclinas D-CDK4/6, os quais fosforilam pRb levando à liberação do fator de transcrição E2F. Essa liberação ativa a progressão do ciclo celular. O complexo protéico pRb-E2F atua como regulador negativo inibindo a expressão de *CDKN2A*. Portanto, a inativação funcional de *Rb* pela oncoproteína viral E7 leva a perda da inativação do transcrito p16[4].

Os tumores associados ao HPV apresentam um padrão molecular distinto, com um mecanismo de tumorigênese independente dos efeitos mutagênicos do tabaco e álcool. Esses tumores são caracterizados pela ausência ou baixo número de alterações genéticas e epigenéticas nas vias clássicas de tumorigênese, com expressão diminuída de *TP53* selvagem, devido a sua inativação e degradação pela oncoproteína viral E6[4,14].

A integração do DNA do vírus no genoma hospedeiro é observada na maioria dos tumores cervicais invasivos, mas raramente nas neoplasias intraepiteliais cervicais (NIC), sendo considerada como um importante divisor de águas na oncogênese[30]. O HPV está presente em formas integradas e não integradas nesses tumores, assim como em uma parte dos CCP[31]. Parfenov e colaboradores (2014) demonstraram que, além da expressão viral de E6 e E7, o HPV provavelmente dirige a tumorigênese pela alteração do genoma do hospedeiro em locais de integração. Mecanismos pelos quais a integração afeta os genes do hospedeiro incluem geração de transcritos alterados, ruptura de genes supressores de tumor, amplificações e rearranjos intercromossômicos[31]. E, além disso, eles sugerem que as ações do HPV no processo de tumorigênese provavelmente vão além da expressão de oncoproteínas virais (pelo menos em um subconjunto de CECP) e sugerem um papel mais sutil para o vírus do que a visão convencional de transformação mediada por E6/E7 e alteração dos eixos *TP53* e *CDKN2A/RB1*. Também pode ser o caso de que o aumento mundial de infecção pelo HPV, em comparação com os cânceres causados pelo HPV, seja explicada pela necessidade de eventos genômicos adicionais além da infecção pelo HPV para estimular a formação de tumores[31].

A análise de amostras do estudo *The Cancer Genome Atlas* mostrou que a integração do HPV ocorre em mais de 80% dos carcinomas cervicais [31,32]. Destas, 76% das amostras positivas para HPV16 têm HPV integrado, enquanto a integração é evidente em todas as amostras positivas para HPV18. Nos CE da orofaringe positivos para HPV, a incidência de integração viral é menor, e muitos tumores têm DNA extracromossômico ou misto, extracromossômico e integrado[31,32,33,34]. Segundo Gao e colaborados (2017), que analisaram sítios frágeis em CE de orofaringe pelo sequenciamento Sanger, as sequências nucleotídicas próximas aos locais de junção identificados continham sequências repetitivas e ricas em AT que demonstraram ter o potencial de formar estruturas secundárias de DNA de "stem-loop" que poderiam

protelar a progressão da forquilha de replicação. Isso poderia causar um aumento da instabilidade nessas regiões, o que poderia levar ao desenvolvimento de câncer nas células humanas[34]. Essas descobertas sugerem que os sítios frágeis e alguns genes específicos parecem desempenhar papéis importantes no CE de orofaringe.

Não está claro se a integração das sequências de DNA viral é uma causa ou uma consequência das alterações genômicas estruturais[30]. Estudos *in vitro* sugerem que a integração precede as alterações genômicas[32], entretanto, há relatos que demonstram que a integração é dependente do tipo de HPV e independente da idade dos pacientes, sugerindo que as quebras genômicas seriam uma consequência, e não a causa da integração viral[30].

As alterações genômicas mais associadas com a integração viral, principalmente nos carcinomas cervicais, são as co-amplificações das sequências virais e regiões flanqueadoras e/ou perda das sequências no sítio de integração. Porém, há poucos relatos de análises sistemáticas da estrutura genômica em sítios de integração, até o presente momento.

Novos estudos tem utilizado sequenciamento paralelo massivo (*next generation sequence* - NGS), como uma ferramenta mais robusta para as análises de genomas completos. Conway e colaboradores (2012)[35] descreveram a utilização do NGS para a identificação de HPV em tecido parafinado de CCP, demonstrando sua eficácia para a determinação da carga viral de HPV, dos genótipos de HPV, em conjunto às alterações cromossômicas nos tumores analisados. Além disso, os autores afirmaram o potencial da técnica para gerar dados e contribuir amplamente como ferramenta de diagnóstico, prognóstico ou estratificação clínica das amostras de CCP HPV positivas ou negativas.

Recentemente, a revisão de Tuna e Amos, 2016[36], levantando dados de NGS, determinou os locais de integração do HPV, seus genes e vias concomitantemente interrompidos e suas consequências funcionais nos cânceres cervicais e de cabeça

e pescoço. A integração dos dados do NGS com outros dados “ômicos” e clínicos é crucial para entender melhor a fisiopatologia de cada malignidade individual e, com base nisso, selecionar alvos e projetar opções eficazes de tratamento personalizado.

1.2 Carcinomas Cervicais

Os carcinomas cervicais apresentam a terceira maior incidência global de câncer em mulheres, com uma estimativa de 569.847 novos casos em 2018 e 311.365 mortes[37]. É o único câncer ginecológico que pode ser prevenido por meio de testes de rastreamento regulares e possui fator etiológico bem estabelecido.

Em 1941, Papanicolaou e Traut[38] relataram que os carcinomas cervicais poderiam ser detectados pela esfoliação das células epiteliais do colo do útero, incluindo a zona de transformação. Essa metodologia pode identificar a presença de células tumorais anteriormente aos achados clínicos da doença, garantindo uma melhor sobrevida das pacientes por meio da intervenção após diagnóstico precoce dos carcinomas cervicais.

Na década de 50, foi descoberto que as lesões cervicais não invasivas precediam os carcinomas cervicais. Essas lesões não invasivas, denominadas neoplasia intraepitelial cervical (NIC), contém células com alterações morfológicas características que podem ser detectadas pela citologia oncótica. Apesar dos benefícios que essa técnica proporcionou, deve-se ressaltar que esse exame funciona como um rastreamento para identificar pacientes em risco para o desenvolvimento de doenças cervicais, e não como um exame diagnóstico, o qual necessita de confirmação por meio de biópsia. Além disso, esses exames requerem a disponibilidade de citopatologistas devidamente treinados e o acesso dos pacientes ao sistema de saúde com profissionais qualificados para realizar o exame[39].

A infecção genital pelo HPV é extremamente comum e a maioria delas, assintomática, sendo o principal fator de risco para o desenvolvimento dos

carcinomas cervicais. Aproximadamente 90% das infecções pelos HPV de alto risco são eliminadas pelo sistema imunológico feminino dentro de 16 meses. Algumas das mulheres que não eliminam os HPV de alto risco desenvolverão NIC grau II, III e câncer[40], associado a fatores de risco adicionais como uso de anticoncepcionais e as condições imunológicas do hospedeiro[41]. Além disso, estudos clínicos conduzidos com mulheres associaram o uso de tabaco a uma maior persistência das infecções por tipos oncogênicos do HPV, e o risco aumentado de desenvolver carcinomas cervicais invasivos[42].

O aprimoramento dos testes de HPV abriram um novo paradigma para o rastreamento do câncer cervical, sendo associados com uma redução significativa no número de carcinomas cervicais avançados e mortes pela doença[39,43].

Atualmente, o teste de DNA para HPV foi incorporado na prática clínica não só para a triagem de casos com resultados citológicos indeterminados, mas também para o acompanhamento clínico após o tratamento, acompanhamento de pacientes sem lesões citológicas ou apenas lesões menores que foram identificadas por colposcopia; resolução de resultados discordantes de citologia, colposcopia ou histologia. Mais recentemente tem sido empregado como uma ferramenta de rastreio primário,[45,46] em substituição a citologia.

Os testes para a detecção de HPV de alto risco, principalmente métodos baseados na reação em cadeia da polimerase (PCR), tem de 90 a 99% de sensibilidade para lesões precursoras de alto grau e são altamente reprodutíveis[39,45]. Essas características tornam os testes moleculares para identificação de HPV uma metodologia de rastreamento primário mais efetivo, principalmente para mulheres com idade superior a 30 anos, quando comparados à citologia cervical, como é praticada hoje em dia[46].

1.3 Carcinomas de Cabeça e Pescoço

Os carcinomas de cabeça e pescoço (CCP) representam um grupo heterogêneo de tumores agrupados devido a sua localização anatômica. Acometem mais frequentemente homens com idade superior a 50 anos e histórico de tabagismo e etilismo crônicos.

O CCP é um termo amplo que engloba alterações epiteliais malignas entre elas as provenientes dos seios paranasais, orofaringe, cavidade oral, laringe e hipofaringe. O principal subtipo é o carcinoma epidermoide (CE), respondendo por mais de 90% dos casos. A localização mais frequente é a cavidade oral (40% dos casos), seguida pela laringe (25%) e faringe (15%)[47]. Dados do Instituto Nacional do Câncer (INCA), 2018, mostram que, na população brasileira, há uma proporção de mais de três casos no gênero masculino (11.200 novos casos) para cada caso do gênero feminino (3.500 novos casos), localizados com maior frequência na cavidade oral[47].

O desenvolvimento desta doença está relacionado com o acúmulo de danos envolvendo genes que controlam a proliferação, senescência, invasão, motilidade e sobrevivência celular[48]. Estas alterações genéticas quase sempre ocorrem como resultado à exposição excessiva ao tabaco, álcool e outros elementos carcinogênicos que desencadeiam danos irreversíveis ao DNA, ocasionando alterações nas regiões codificadoras e regulatórias de genes supressores tumorais e oncogenes. O acúmulo destas mutações possibilita que as células adquiram características tumorais, incluindo resistência à morte e proliferação descontrolada, conferindo vantagem seletiva ao crescimento[49]. As células tumorais podem formar clones com capacidade de invasão e estímulo aumentado da angiogênese, levando ao estabelecimento de tumores mais invasivos e com maior potencial para desenvolver metástases[21].

As taxas mais elevadas de carcinomas orais são encontradas no Brasil, Melanésia, Centro-Sul da Ásia e Europa Central (Espanha e França) e Leste

Europeu. Em contrapartida, as menores taxas são encontradas na África, América Central e Leste da Ásia (Japão e China) para ambos os sexos[37].

O consumo de tabaco e álcool são os fatores de risco clássicos e bem estabelecidos para CECP, incluindo os da orofaringe. As tendências na prevalência do tabagismo diferem significativamente entre as regiões geográficas, e uma diminuição notável de 1980 a 2012 foi observada em homens e mulheres na América do Norte e Norte da Europa[50]. Além disso, o comportamento sexual é agora estabelecido como um fator de risco para o CE de orofaringe relacionado ao HPV, com o número de parceiros de sexo oral ao longo da vida como o fator mais fortemente associado ao desenvolvimento do câncer. Além disso, assim como para o tabagismo, a prevalência do sexo oral difere dramaticamente entre as populações[51]. Outros fatores de risco incluem beijo de boca aberta, parceiro sexual vaginal (ou outros), uso de maconha e história de infecção por HPV cervical[21]. Assim, as diferenças na exposição dos fatores de risco e no comportamento sexual poderiam explicar, em parte, as diferenças nas frações atribuídas ao HPV observadas entre as regiões e ao longo das décadas.

O CCP tem elevada chance de cura quando diagnosticado precocemente, mas a maioria dos pacientes se apresenta com doença inicialmente avançada (aproximadamente 2/3 dos casos). Tem sido assumido que, nas últimas duas décadas, a qualidade de vida dos pacientes com CCP aumentou como resultado do uso de técnicas cirúrgicas e radioterápicas mais elaboradas[52,53,54]. Em contrapartida, a sobrevida desses pacientes não tem melhorado significativamente, uma vez que os mesmos desenvolvem frequentemente recorrência locorregional, metástases à distância e segundos tumores primários[54,55].

No Brasil, estima-se uma incidência (taxa ajustada pela idade - ASR) de 7,35 novos casos de carcinoma de orofaringe por 100.000 habitantes por ano, atingindo

pouco mais de 1.000 novos casos por ano[37]. Para os carcinomas de cavidade oral, essa taxa chega a mais de 14.000 novos casos por ano[47].

O carcinoma epidermoide de cabeça e pescoço (CECP) tem uma incidência global estimada em mais de 500.000 casos e 300.000 mortes ao ano. A incidência de CECP causada pelo papilomavírus humano (HPV) está aumentando em muitos países desenvolvidos, enquanto em outras áreas, como o Brasil, permanece indefinida a associação entre o HPV e o desenvolvimento desses tumores. Questões sobre o papel das diferenças geográficas no CECP associado ao HPV ainda não foram elucidadas [55,56,57,58].

Além das evidências do envolvimento do HPV na etiologia dos CE de orofaringe[13], parece haver uma correlação significativa entre a presença de infecção e a resposta mais efetiva ao tratamento dos pacientes com esse tumor[21,59,60]. O DNA do HPV está presente em aproximadamente 25% dos CECP e, nos EUA, cerca de 50% dos CE de orofaringe contém DNA de HPV de alto risco[16].

Vários relatos na literatura indicam que o prognóstico dos pacientes com tumores de orofaringe HPV positivos é melhor do que dos casos HPV negativos, independentemente do status linfonodal, idade, estádio clínico, diferenciação tumoral ou gênero[4,61,62,63,64,65]. Além disso, esses tumores respondem melhor a diversos tratamentos, entre eles a cirurgia, radioterapia convencional e fracionada, quimioterapia de indução, quimioterapia adjuvante e quimioterapia de indução associada à quimioterapia adjuvante[59,60,65,66].

Os mecanismos exatos que levam a um melhor prognóstico nos casos associados ao HPV ainda não foram plenamente elucidados. Estudos em literatura apontam a correlação inversa entre a infecção pelo HPV e o status de TP53 como um dos fatores de bom prognóstico, sendo uma possível explicação à resposta apoptótica mediada pelo TP53 intacta nos tumores HPV positivos. Outra possibilidade seria a presença de fatores imunológicos relacionados aos tumores

cujo fator carcinogênico principal é a infecção pelo HPV, devido a linfócitos infiltrantes de tumor (TILs) conferindo um efeito protetor através de uma resposta imune adaptativa do hospedeiro dirigida contra antígenos virais [4,21,67,68,69].

Justificativa

2 JUSTIFICATIVA

Sendo o HPV16 sabidamente relacionado a uma fração dos tumores de orofaringe, principalmente das tonsilas palatinas, a causa do tropismo desse tipo viral por esses epitélios ainda não foi elucidada. Os estudos de tropismo e de outras características biológicas dentre os HPV 16, como sua maior capacidade oncogênica, tem focado em regiões específicas do genoma viral, como a região regulatória ou os oncogenes E6 e E7. O sequenciamento completo, aqui proposto, poderá dar uma visão completa das semelhanças e diferenças entre cepas distintas, eventualmente identificando padrões correlacionáveis ao tropismo viral, entre outras particularidades.

Objetivos

3 OBJETIVO GERAL

Determinar a prevalência de HPV em carcinomas de cabeça e pescoço e nas tonsilas não neoplásicas e realizar a genotipagem dos tipos de HPV em amostras de cabeça e pescoço.

3.1 Objetivos específicos

Sequenciar o genoma completo do HPV, a fim de identificar alterações específicas características dos vírus presentes nos tumores de orofaringe;

Sequenciar as regiões do genoma humano que flanqueiam os sítios de integração do HPV16, em ambos os grupos, para verificar se há regiões preferenciais de inserção viral.

Metodologia

4 METODOLOGIA

4.1 Casuística

As amostras de carcinoma de cabeça e pescoço avaliadas no presente estudo fazem parte dos projetos colaborativos do Grupo de Estudos de Câncer de Cabeça e Pescoço - GENCAPO 1 e 2 (Registro CONEP 16491). Foram incluídas 510 amostras do projeto, todos os tecidos fixados em formalina e emblocados em parafina.

As amostras não neoplásicas de tonsila foram coletadas em cirurgias de uvulopalatofaringoplastias, conhecida como "cirurgia do ronco" em adultos apneicos, sob responsabilidade do Dr. Ronaldo Frizzarini, da disciplina de otorrinolaringologia da Faculdade de Medicina da Universidade de São Paulo, além de amostras de pacientes submetidos à tonsilectomia sem alterações inflamatórias na tonsila. Essa coleta faz parte do projeto de mestrado da aluna Jéssica Boscariol da Silva (IMT-USP, do mesmo grupo de pesquisa/orientador). A prevalência de HPV neste tipo de amostra é pouco conhecida, mas estudos na literatura apontam variações de 0 a 15% para amostras de biópsias normais, e de 2,8 a 25% para amostras de enxague bucal, dependendo da metodologia avaliada[71]. Desta forma, foram incluidos 113 pacientes provenientes dos dois hospitais participantes, que geraram 452 amostras (2 alíquotas de cada amostra, lado esquerdo e direito), no período de 30 de janeiro de 2013 a 14 de setembro de 2016.

Esse estudo tem a aprovação da Comissão de Ética para Análise de Projetos de Pesquisa do HC (CAPPesq) sob o parecer 173.181.

4.2 Extração de DNA

O isolamento do DNA das amostras normais de tonsila, obtidas a fresco e mantidas congeladas a -70ºC, foi realizado utilizando-se o kit *Qiagen DNeasy blood*

& tissue kit (Qiagen, São Paulo, Brasil), seguindo o protocolo recomendado pelo fabricante. Às amostras, foi adicionado 20µL de Proteinase K (10mg/mL) para promover a digestão enzimática e lise do tecido, seguido da adição de 180µL de Tampão ATL. As amostras foram incubadas à temperatura de 56°C, overnight. Em seguida, foi adicionado 200µL de Tampão AL para completar a lise dos tecidos e etanol na concentração entre 96%-100%, para agrupar as moléculas de DNA. As amostras lisadas foram transferidas para a coluna QIAamp MinElute em um tubo de coleta de 2mL, em seguida foi adicionado na coluna 500µL de Tampão AW1 e 500µL de Tampão AW2 para lavagem. Após esse processo, os tubos seguiram para centrifugação a 20.000xg e a coluna transferida para um tubo final de 1,5mL onde aplicou-se 20-100µL de Tampão AE de eluição no centro da membrana e uma nova centrifugação a 20.000xg por 1 min, separando apenas o DNA.

Para verificação da concentração e pureza, todas as amostras extraídas foram quantificadas no comprimento de onda de 260nm e sua pureza atestada pela razão 260/280nm e 260/230nm. As medidas foram realizadas no espectrofotômetro NanoDrop 2000/2000C (*Spectrophotometer, Thermo Scientific, Wilmington, Delawere, EUA*). Após esse processo, todas as amostras de DNA foram diluídas a 50ng/µL e armazenadas em freezer a temperatura de -20°C.

Os tumores fixados em formalina e incluídos em blocos de parafina de amostras de carcinoma de cabeça e pescoço foram tratados com solução de octano e etanol 100% para desparafinização e digeridos com solução de proteinase K. Para posterior extração do DNA genômico, foi utilizado o *QIAamp DNA FFPE Tissue Kit* (Qiagen, São Paulo, Brasil) seguindo especificações do fabricante acima e a quantificação das amostras foi realizada em espectrofotômetro seguindo os parâmetros de pureza e qualidade sugeridos pelo fabricante e especificadas acima.

Além disso, para verificar a possibilidade de se amplificar as amostras de DNA extraídas, foi realizada a reação em cadeia da polimerase (PCR) para o gene da

beta-globina, a partir de 100ng de DNA com os iniciadores PCO3 e PCO4[72] (PCO3+ CTT CTG ACA CAA CTG TGT TCA CTA GC e PCO4+ TCA CCA CAA CTT CAT CCA CGT TCA CC), amplificando um fragmento de 107pb. A reação de PCR foi realizada sob as seguintes condições: 0,20mM de dNTP, 0,1µM de cada iniciador, 1.25U/µL de *Taq DNA polimerase* (Invitrogem, USA), tampão de PCR 1X (Invitrogem, USA), 2µg de DNA genômico e água destilada deionizada para um volume final de 25µL. A mistura de reação foi amplificada pelo programa: 1 ciclo de 94°C por 7 minutos, 35 ciclos (94°C por 1minuto, 55°C por 1minuto e 72°C por 1 minuto) e 1 ciclo de 72°C por 10 minutos.

Após a realização da PCR, os produtos amplificados foram visualizados em gel de agarose a 2% (*Agarose Ultrapure TM – Invitrogen Life Technologies*) em tampão TAE 1X [*DNA Typing Grade TAE Buffer* – GIBCO – Invitrogen Corporation (2mM Tris-Acetato e 50mM de ácido etilenodiaminotetracético - EDTA)], corado com SYBR Safe™ 1X (*Invitrogen*, USA). Os produtos da amplificação e o padrão de tamanho molecular de 100pb (*Invitrogen*, USA) foram aplicados no gel de agarose e submetidos a uma corrente constante de 90 amperes por 50 minutos, e visualizados sob luz ultravioleta (UV).

4.3 Genotipagem de HPV

Os casos positivos para beta-globina foram submetidos a genotipagem, para determinar aqueles que continham a presença do HPV. Para essa genotipagem foi utilizado o kit *Inno-Lipa HPV* (Innogenetics, Gent, Bélgica), capaz de identificar 28 tipos de HPV, sendo 18 de alto risco - HR-HPV e 10 de baixo risco - LR-HPV. Esse teste se baseia na amplificação do DNA alvo por PCR utilizando iniciadores para regiões consenso para os HPV avaliados e hibridação reversa dos produtos amplificados em tiras que contem sondas específicas para os mesmos.

Para as amostras não tumorais de tonsila, foi realizada a PCR com os iniciadores PGMY09/11 que amplificam fragmentos de 450pb da região L1 de vários tipos de HPV, sendo considerado padrão ouro na detecção do vírus e descritos em estudos prévios[26]. A PCR foi realizada no termociclador com volume final de 25µL, dos quais foram utilizados 60nM de iniciadores PCO4 e GH20 (PCO4+ TCA CCA CAA CTT CAT CCA CGT TCA CC e GH20+ GAA GAG CCA AGG ACA GGT AC), 200nM de iniciadores PGMY09/11, 250ng de DNA extraído, 2U de *Platinum Taq DNA Polymerase* (Invitrogem, USA), 0,2mM de cada dNTP e 2mM de $MgCl^2$. Os ciclos das reações foram realizados seguindo o protocolo para amplificação: 1 ciclo de 94°C por 7 minutos, 40 ciclos (95°C por 5 minutos, 95°C por 1 minuto, 55°C por 1 minuto, 72°C por 1 minuto) e 1 ciclo de 72°C por 7 minutos).

Após a PCR, os produtos da reação foram analisados por meio de eletroforese submetido a uma corrente constante de 100 amperes por 30 minutos. O gel de agarose foi realizado a 1,5% (Agarose Ultrapure TM – Invitrogen Life Technologies) com tampão TAE 1X, e visualizados sob luz ultravioleta (UV).

4.4 PCR em tempo real para HPV16

A genotipagem das amostras não tumorais de tonsila, assim como a confirmação dos casos positivos para HPV16 no *Inno-Lipa*, foi realizada por PCR quantitativo (em tempo real), específico para HPV16. A PCR em tempo real foi realizada com o sistema *TaqMan Universal PCR Master Mix* 2x (*Applied Biosystems*, Inc., EUA) com par de iniciadores específicos da região de E7 do HPV16 descrito por Walboomers e colaboradores[73] além de uma sonda também específica para essa região por nós desenhada[74] (E7F+ GAT GAA ATA GAT GGT CCA GC; E7R+ GCT TTG TAC GCA CAA CCG AAG e sonda FAM /CAA GCA GAA CCG GCA AG/ MGB). Este protocolo já foi validado em nosso laboratório apresentando sensibilidade e especificidade adequadas. As reações de PCR em tempo real foram realizadas no

termociclador *ABI Prism*® 7300 HT SDS (*Applied Biosystems*, Inc., EUA) para volume final de 25µL, dos quais foram utilizados 5µL de DNA na concentração final de 100ng/µL, 400nM de iniciadores HPV16-E7 senso e anti-senso, 200nM da sonda específica para HPV16, completando o volume final com água livre de DNAse RNAse (Gibco/BRL, *Life Technologies*, USA). A amplificação foi realizada segundo o protocolo pré-estabelecido: 50°C por 2 minutos, 95°C por 10 minutos, seguido de 40 ciclos de (95°C por 15 segundos, 55°C por 1 minuto e 60°C por 1 minuto).

As análises das curvas de amplificação foram realizadas no software 7300 System Software (*Applied Biosystems,* Inc., EUA).

4.5 Sequenciamento de nova geração (NGS) no IARC - Lyon

Como parte da colaboração entre o GENCAPO e a Agência Internacional de Pesquisa sobre Câncer - IARC, Lyon, França, o treinamento no sequenciamento de amostras de cabeça e pescoço para identificar mutações em genes específicos associados ao desenvolvimento de CCP foi realizado com o grupo de Epidemiologia, chefiado pelo Dr. Paul Brennan e o grupo de Susceptibilidade, chefiado pelo Dr. James McKay.

A metodologia de sequenciamento envolve a amplificação de sequências específicas do genoma humano, conhecido como *Target Seq*. Essa técnica provem do desenvolvimento de iniciadores customizados para o estudo dos genes específicos (Figura 1). Para essa etapa do estudo foram amplificadas as sequências dos genes *TP53, CDKN2A, PTEN, NOTCH1, CASP8*, conhecidos por terem um papel importante no desenvolvimento dos CCP, frequentemente relatado com alterações moleculares[21].

As etapas seguintes da amplificação envolveram a purificação do material por meio de esferas magnéticas (*beads*), construção da biblioteca dos fragmentos alvo, PCR de emulsão baseada nas partículas esféricas exclusivas da tecnologia do NGS

do *Ion Torrent PGM*, enriquecimento e sequenciamento do material, seguindo o protocolo do fabricante e, por fim, a análise dos dados gerados.

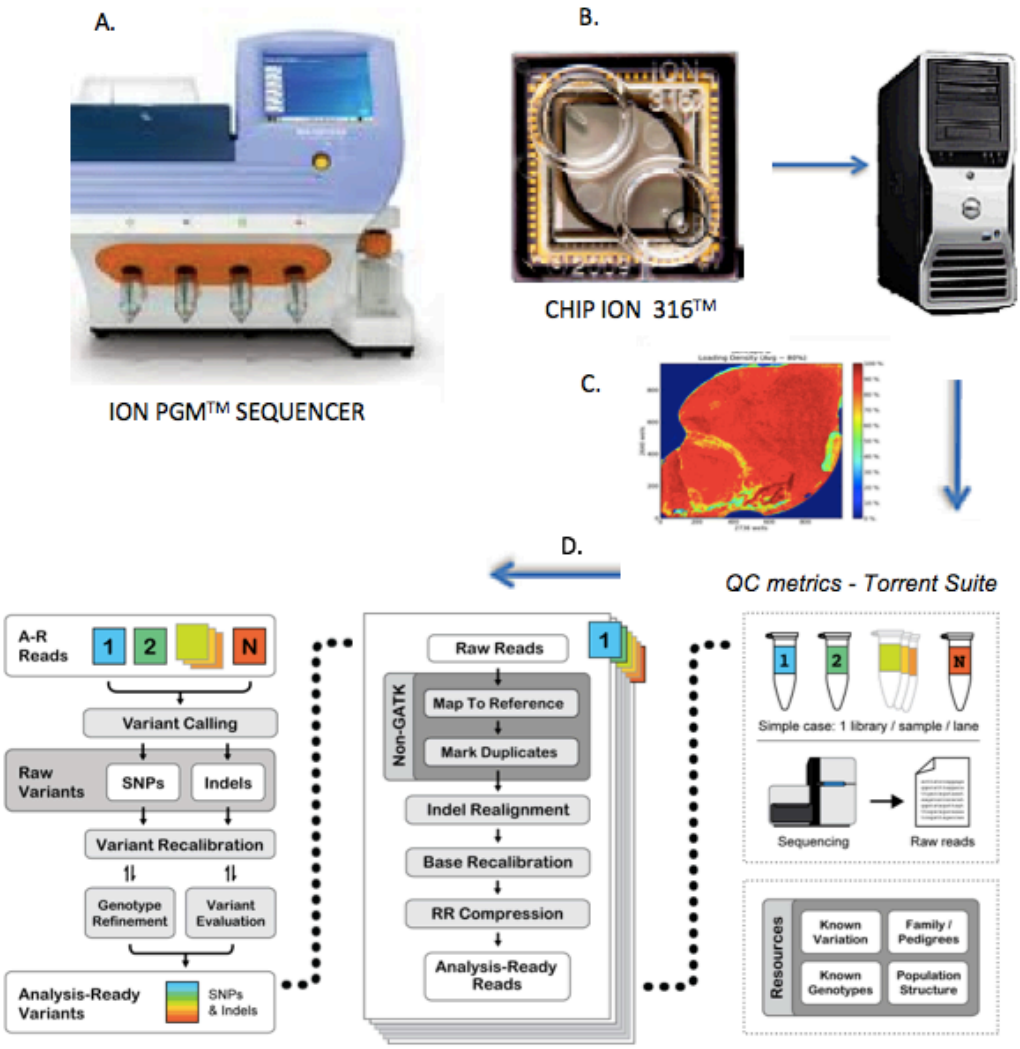

Figura 1 - Fluxograma do sequenciamento realizado. A. Equipamento *Ion Torrent PGM*, B. Chip utilizado (Ion 316™). C. Validação do Chip com o material a ser sequenciado (um dos critérios avaliados no *QC metrics - Torrent Suite™ Software 3.6.2*). D. Fluxograma da análise das sequencias geradas.

4.6 Sequenciamento de nova geração (NGS) HPV16

Como o objetivo desse projeto foi identificar os pontos de inserção do DNA viral no genoma do hospedeiro, foi utilizada uma metodologia que se baseia na captura do material genômico viral a partir de amostras de DNA do hospedeiro. As amostras

que tinham infecção por HPV16 exclusivo foram selecionadas para o sequenciamento. Para isso, a etapa inicial foi a construção de uma biblioteca de DNA genômico utilizando o Kit de preparo de biblioteca *Ion Xpress* e também o Kit de adaptadores Ion com "*barcode*" (código de barras), permitindo a análise em paralelo de até 16 amostras em uma mesma corrida.

Com as bibliotecas prontas, foi realizada a captura da porção do DNA que contém sequências do DNA viral. Para isso, foi sintetizado um kit de captura de DNA customizado a partir de sondas de oligonucleotídeos marcados com biotina, com a tecnologia *SureSelect* (*Life Technologies*). A captura contém entre 20 e 500 Kb, que é suficiente para cobrir o genoma de subtipos distintos do vírus HPV. Nesse caso, foram customizadas sequências para os HPV de maior importância clínica, além do HPV16, que são o HPV 18, 33, 35, 31, 45, 52, 56, 68 e o HPV 11 de baixo risco. A sequência dos oligonucleotídeos biotinilados são complementares ao DNA viral e, após a incubação com o DNA extraído da amostra biológica, o material hibridizado é capturado com esferas magnéticas cobertas com estreptavidina. Esse material enriquecido com sequências do DNA viral deveria conter as bordas das regiões de inserção no genoma humano, uma vez que o desenho das "iscas" foi realizado de modo que permitisse a captura, além das sequências alvo do HPV, sequencias ajusantes de DNA humano (em torno de 100pb – "*near target*") (Figura 2).

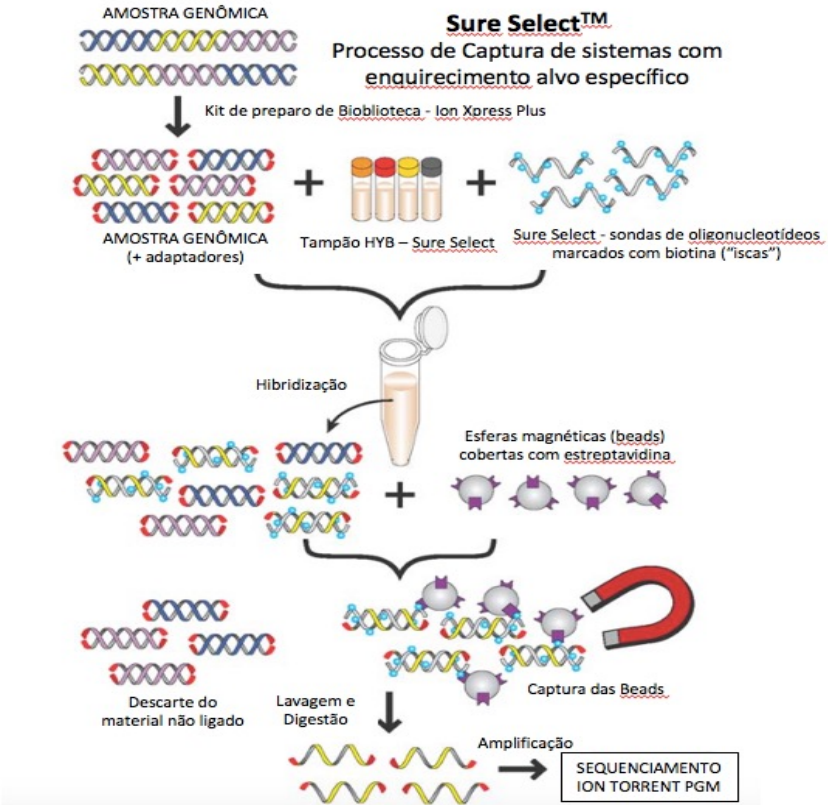

Figura 2 - Processo de preparo da biblioteca (adaptado do Manual da *Agilent Technologies, Inc. 2014, 2016*), desde o preparo da amostra genômica (HPV16) até o produto amplificado que será, posteriormente, sequenciado.

O material enriquecido foi sequenciado no Chip 316, que permite a geração de dados entre 300Mb a 1Gb, com a química Hi-Q no *Ion PGM* (*Life Technologies*). Essa química de sequenciamento permite leituras de 400pb, que devem ser longas o suficiente para que algumas das leituras abranjam regiões tanto do genoma viral quanto do genoma humano, o que permitiria a definição dos pontos de inserção.

O sequenciador Ion PGM™ da *Life Technologies* utilizado está no sistema multiusuário no Instituto de Medicina Tropical – IMT/USP. É um sequenciador de nova geração com alta capacidade de geração de dados. Todo o sequenciamento é feito através da medição da liberação de íons H^+ durante a ação da polimerase, os quais ativam uma camada sensora nos chips semicondutores, produzindo um sinal elétrico.

4.7 Análise das sequências

As sequências obtidas foram analisadas pelo grupo da Dr. Lisa Mirabello, da Divisão de Epidemiologia do Câncer e Genética, NCI, NIH, Rockville, MD, USA, alinhadas utilizando o programa *ClustalX*[75], juntamente com sequências referência dos diferentes tipos de HPV, obtidas no GenBank. O alinhamento foi visualizado e inspecionado no programa *BioEdit* (www.mbio.ncsu.edu/BioEdit/BioEdit.html). A confirmação do tipo de HPV por análises filogenéticas, realizadas por meio dos critérios de distância (*Neighbor Joining*) e máxima verossimilhança implementados no programa *MAFFT*[76], utilizando o modelo de substituição nucleotídica definido pelo *ModelTest3.7*[77] e com verificação de suporte de ramos por *bootstrap* com 1000 réplicas.

Resultados e Discussão

5 RESULTADOS E DISCUSSÃO

5.1 Casuística de Cabeça e Pescoço

As amostras de carcinoma de cabeça e pescoço foram provenientes dos projetos Gencapo 1 e 2 (Registro CONEP 16491). Como parte do estudo colaborativo entre o Gencapo e o IARC (Agência Internacional de Pesquisa sobre Câncer, Lyon, França), 510 amostras de material parafinado provenientes de três diferentes instituições da cidade de São Paulo (369 amostras do Hospital das Clínicas de São Paulo - HC; 84 amostras do Hospital Heliópolis - HH e 57 amostras do Instituto do Câncer Arnaldo Vieira de Carvalho – ICAVC). O fluxograma abaixo (Figura 3) apresenta, de uma maneira geral, as amostras e as metodologias aplicadas nesse estudo.

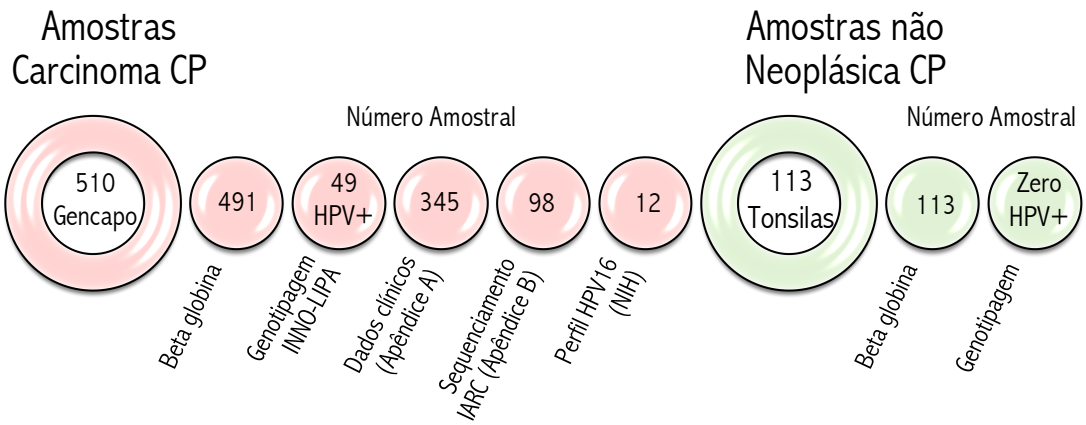

Figura 3 – Fluxograma das amostras avaliadas nesse estudo. Em vermelho estão representadas o coorte de amostras de carcinoma de cabeça e pescoço e em verde, amostras não neoplásica. Dentro de cada círculo está apresentado o número de amostras que resultaram das análises realizadas na legenda.

Das amostras não neoplásicas de orofaringe analisadas, como parte do projeto de mestrado da aluna Jéssica Boscariol da Silva, nenhum dos 113 pacientes apresentou o HPV no tecido analisado. Foram incluídas 47 mulheres de 18 a 58 anos, com idade média de 28,5 anos (DP 9,5 e mediana de 27 anos) e 66 homens de 18 a 68 anos, com idade média de 34 anos (DP 10,5 e mediana 33 anos).

As amostras de CCP que amplificaram a beta-globina (491 amostras no total: 350 HC, 84 HH e 57 ICAVC) foram genotipadas pelo kit *Inno-Lipa HPV*, das quais 49 amostras foram positivas para HPV (10% do total), sendo 38 amostras exclusivamente HPV 16 (78% das amostras positivas), quatro amostras positivas para HPV11 (8% das amostras positivas), duas amostras HPV 66 (4% das amostras positivas) e cinco amostras apresentaram infecção múltipla (10% das amostras positivas).

Apesar de se demonstrar em crescimento, comparada a estudos prévios, a baixa prevalência do HPV em populações brasileiras de CCP tem sido um achado recorrente na literatura[37,47,55,56,57,58], discordando dos dados de países da Europa e América do Norte (Tabela 1). Ao que tudo indica, países em desenvolvimento ainda resguardam as características convencionais da carcinogênese dos tumores de cabeça e pescoço, associados majoritariamente com o consumo abusivo de tabaco e álcool.

Tabela 1 – Prevalência de HPV em amostras brasileiras de CCP de orofaringe.

Autor e ano	n. de casos	% HPV	Método	Período da coleta
Rosa-Sacoman, 2018*	24	25	Inno Lipa DNA e HPV16 RT-PCR	2002-2010
Barros-Filho et al., 2018	40	58	Linear array	< 2011
Anantharaman et al., 2017	171	4,1	HPV16 DNA e expressão de p16INK4a	2002-2015
Betiol et al., 2016	28	14,3	Inno Lipa DNA	1991-2012
López et al., 2014	91	6,6	Inno Lipa DNA e PGMY09/11	1998-2008
Ribeiro et al., 2011	68	4,4	PGMY09/11 seguido de HPV16 PCR	1998-2003

*artigo em preparação; Referências extraídas do artigo em preparação - Apêndice A.

Para os 369 tumores de cabeça e pescoço do Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo incluídos neste estudo, foram também analisados alguns dados clínicos e epidemiológicos dessa população.

Esta casuística consiste basicamente de homens (quase 90% da população) com idade entre 27 e 90 anos (média: 61; DP: 10), e 92% dos tumores eram carcinomas de células epidermoides. A taxa de positividade para o DNA de HPV foi de mais de 7% dos tumores de cabeça e pescoço (25/350 amostras genotipadas), dentre estes 80% foram HPV16, seguido pelo HPV11, HPV18/39 e HPV31/33/44/52. Quando estratificada por sítio anatômico, a prevalência de HPV varia de 3,4% (língua e hipofaringe) para 25% (câncer de orofaringe - todos HPV16).

Esses dados resultaram no artigo apresentado no **APÊNDICE A** (em preparação) - *"Prevalence of HPV in head and neck carcinomas and normal tissue from São Paulo, Brazil"*. ***Luciana Reis Rosa Sacoman****, Jéssica Boscariol da Silva, Luisa Lina Villa, Eloiza Helena Tajara da Silva, GENCAPO Group, José Eduardo Levi.*

5.2 Sequenciamento de nova geração (NGS) no IARC - Lyon

As amostras que foram utilizadas no sequenciamento eram provenientes de tecidos fixados e emblocados em parafina do GENCAPO 1, e o material foi extraído no laboratório de Virologia do Instituto de Medicina Tropical da Universidade de São Paulo e enviado ao IARC.

No total, 98 amostras que possuíam dados clínicos completos e amostras de sangue periférico pareadas (para posterior análise de mutações germinativas) foram quantificadas por fluorometria no equipamento Qubit (*Thermo Fisher Scientific*) e apresentaram padrão ótimo de concentração inicial para a metodologia (pelo menos 4ng/µL).

A investigação de como alterações somáticas no CECP interagem com fatores ambientais e fatores de risco do hospedeiro para o desenvolvimento do tumor é de extrema importância para elucidar o processo de carcinogênese. Nesse quesito, a avaliação de alterações no número de cópias se mostra uma ferramenta usual para

a identificação de eventuais alterações nos genes responsáveis pela transformação maligna.

Os dados parciais exclusivos da análise do grupo brasileiro estão apresentados abaixo (Figura 4). O perfil mutacional mostrou uma associação tanto com exposições ambientais quanto com características clínicas, sendo essas alterações importantes como preditores da sobrevida global do CECP. Os principais fatores avaliados foram: presença de HPV, consumo de tabaco e álcool, localização anatômica e as mutações identificadas (SNV – variação de nucleotídeo único, DEL – deleção, INS – inserção).

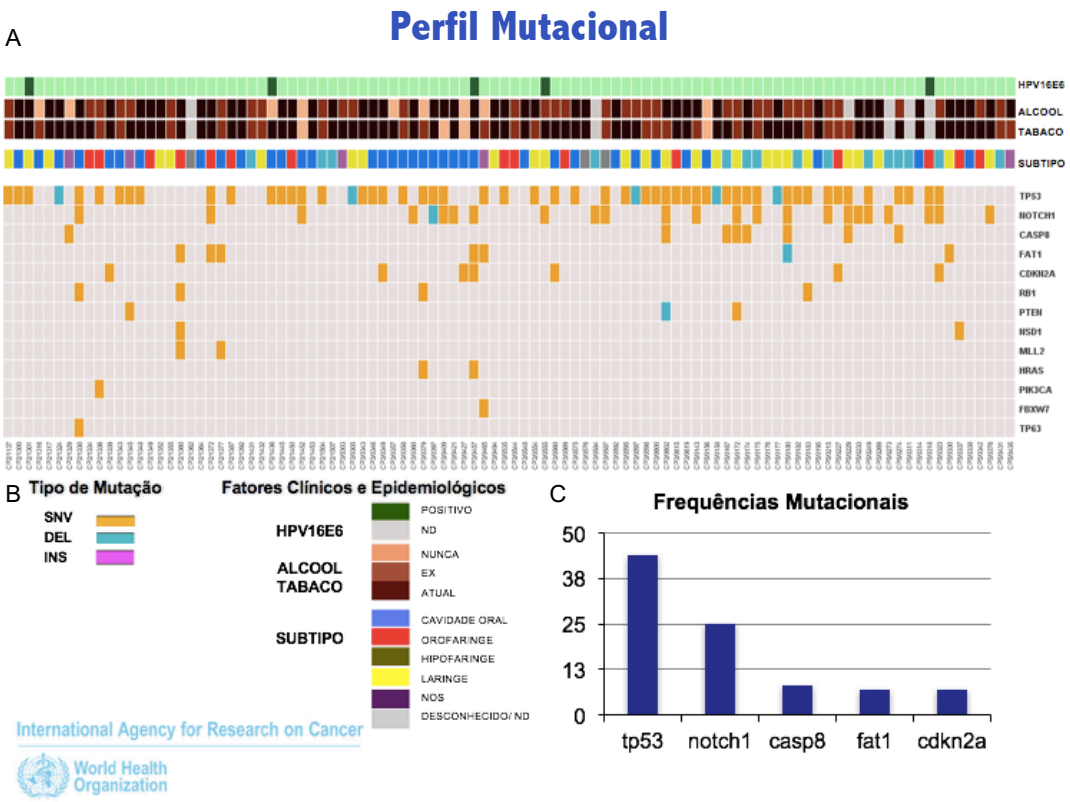

Figura 4 – Perfil mutacional das amostras do GENCAPO avaliadas no IARC. A. Clusterização dos critérios clínicos das amostras, como presença de HPV (verde), uso de álcool e tabaco (marrom), e o tipo do CCP (colorido) e abaixo, os genes (linhas) e o tipo de mutação apresentada em cada paciente (colunas). B. Legenda das cores e associações com o perfil clínico e epidemiológico. C. Gráfico com a frequência mutacional pra cada um dos genes analisados na população. ND – não determinado, SNV – variação de nucleotídeo único (amarelo), DEL – deleção (azul), INS – inserção (rosa).

Os genes mais frequentemente mutados nessa população foram *TP53* e *NOTCH1*, sendo o status do HPV um dado sem expressão clínica (poucos casos sem correlação epidemiológica). Esses dados, assim como a análise dos resultados completos que envolve a comparação de dois grupos amostrais de CCP de regiões geográficas diferentes, está no artigo apresentado no **APÊNDICE B** - *"Genomic analysis of head and neck cancer cases from two high incidence regions" - Perdomo S, Anantharaman D, Foll M, Abedi-Ardekani B, Durand G,* ***Reis Rosa LA****, et al. (2018). PLoS ONE 13(1): e0191701. https://doi.org/10.1371/journal. pone.0191701*[81].

A análise de dados integrados, genômicos e transcriptômicos, é uma ferramenta bastante utilizada para determinar potenciais genes "drivers" que direcionam a transformação maligna e estão diretamente associados ao prognóstico e resposta ao tratamento, tendo sua importância na identificação de novos alvos terapêuticos nos CECP, principalmente aqueles localizados na orofaringe, que sabidamente tem um comportamento celular singular.

Casos positivos de HPV de alto risco, detectados frequentemente nos tumores epidermoides da orofaringe, foram associados a melhores prognósticos. Diferentes perfis de expressão são observados quando comparados tumores HPV positivos e negativos. A identificação de potenciais fatores moleculares associados ao prognóstico e resposta ao tratamento em carcinomas epidermoides de orofaringe avançados está no artigo apresentado no **APÊNDICE C** - *"Oncogenic drivers in 11q13 associated with prognosis and response to therapy in advanced oropharyngeal carcinomas". M.C. Barros-Filho*,* ***L.A. Reis-Rosa****, M. Hatakeyama et al. Oral Oncology 83 (2018) 81–90 (both authors contributed equaly)*[82].

5.3 Sequenciamento de nova geração (NGS) HPV16

Como o objetivo desse projeto era identificar os pontos de inserção do DNA viral no genoma do hospedeiro, a metodologia utilizada foi inteiramente customizada para essa finalidade, que se baseia na captura do material genômico viral e das sequencias de DNA humano que flanqueiam o vírus. Das amostras de orofaringe positivas para HPV16 identificadas e sequenciadas, apenas quatro delas demonstraram qualidade nas sequências após a definição dos parâmetros iniciais, avaliados pelo laboratório da Dra. Lisa Mirabello[83], no NIH/EUA – (Tabela 2), enquanto nenhumas das 113 amostras de orofaringe normal revelaram a presença de HPV16 para que pudessem ser submetidas ao sequenciamento.

As quatro amostras que tiveram cobertura suficiente no sequenciamento foram analisadas de maneira descritiva, uma vez que não foi possível realizar uma análise mais robusta com esse número amostral. A baixa eficiência apresentada na reação de sequenciamento pode ser devido ao tipo de material usado, apesar de materiais emblocados em parafina já terem demonstrado reprodutibilidade com essa mesma técnica, ou a baixa carga viral de HPV nas amostras. Entre as amostras analisadas, ouve uma correlação direta entre o Ct (ciclo threshold) de amplificação por PCR em tempo real do HPV16 e a quantidade de “reads” sequenciadas. Amostras com Ct acima de 30 não obtiveram dados sequenciáveis (Tabela 3).

Tabela 2 - Cobertura do sequenciamento, por amostra (1 a 12 – em ordem decrescente na quantidade de “reads”) e por amplicon (fragmentos do genoma do HPV). Em verde estão representadas os amplicons com quantidades substanciais de “reads” e, em vermelho, os amplicons com pouca ou nenhuma sequencia identificada.

				Amostras - Reads TOTAL - Reads por amplicom											
	HPV16_Ref		Reads por amplicon	012.hpv.bam	009.hpv.bam	005.hpv.bam	008.hpv.bam	011.hpv.bam	006.hpv.bam	003.hpv.bam	001.hpv.bam	010.hpv.bam	007.hpv.bam	002.hpv.bam	004.hpv.bam
Amplicon (reads) TOTAL	Inicio	Final	1458345	994599	310813	115555	31705	4403	765	212	189	47	42	10	5
HPV16_48_B	0	238	9844	5304	2722	1142	676	0	0	0	0	0	0	0	0
HPV16_1	227	459	24873	12022	8456	2700	1416	196	81	0	0	0	2	0	0
HPV16_2	440	652	40076	28934	5089	4450	1555	25	23	0	0	0	0	0	0
HPV16_3	623	830	47653	34361	8013	4322	909	38	4	4	0	0	2	0	0
HPV16_4	783	968	26698	15779	7375	2629	908	0	2	0	0	3	1	1	0
HPV16_5	920	1080	26192	16273	6560	2660	696	0	1	0	0	0	1	1	0
HPV16_6_v2	1061	1292	33168	20289	7994	3151	1704	0	0	0	0	30	0	0	0
HPV16_7	1246	1464	38182	25729	8553	3234	653	0	1	0	0	5	7	0	0
HPV16_8_v3	1439	1669	29317	19834	6273	2612	597	0	0	1	0	0	0	0	0
HPV16_9	1590	1784	30820	24179	4535	1901	193	7	0	5	0	0	0	0	0
HPV16_10_v3	1782	1929	30460	21457	6300	2414	40	244	2	0	0	0	3	0	0
HPV16_11	1887	2069	29331	22070	4768	1794	699	0	0	0	0	0	0	0	0
HPV16_12	2020	2290	48549	38464	7171	2285	613	0	1	14	0	0	1	0	0
HPV16_13_v3	2173	2443	37267	28749	5375	2480	356	299	3	4	0	1	0	0	0
HPV16_14	2424	2646	33901	24603	6617	2017	615	0	47	2	0	0	0	0	0
HPV16_15	2630	2854	49922	37698	8908	3015	260	0	3	0	38	0	0	0	0
HPV16_16	2770	2964	42476	34357	5290	2226	519	50	26	7	0	1	0	0	0
HPV16_17	2949	3112	31106	23352	4908	1753	1091	0	0	0	0	0	2	0	0
HPV16_18_v3	3113	3313	33000	25475	5596	1721	206	0	0	1	0	0	1	0	0
HPV16_19	3282	3454	31860	17358	8423	6066	10	2	0	0	0	0	0	1	0
HPV16_20	3366	3543	55884	29203	20155	5556	911	57	0	0	0	1	1	0	0
HPV16_21_v4	3520	3707	32885	19917	9733	2759	328	147	0	1	0	0	0	0	0
HPV16_22	3639	3868	24065	17843	4318	1299	199	406	0	0	0	0	0	0	0
HPV16_23_v4	3783	3942	23957	17096	4993	1430	368	67	0	0	0	2	1	0	0
HPV16_24_v3	3928	4222	14568	12034	2046	454	0	34	0	0	0	0	0	0	0
HPV16_25_v3	4016	4272	6475	4613	1357	499	0	2	4	0	0	0	0	0	0
HPV16_26	4268	4490	37945	29325	6353	1925	4	268	67	2	0	0	1	0	0
HPV16_27_v4	4447	4687	33779	21988	8282	2663	213	454	13	165	0	0	0	1	0
HPV16_28	4668	4874	41838	28509	9232	3313	346	400	37	0	0	0	1	0	0
HPV16_29_v3	4871	5007	16813	9251	5303	1750	374	133	0	0	0	0	1	1	0
HPV16_30_v4	4972	5109	9173	6666	1632	399	235	238	3	0	0	0	0	0	0
HPV16_31	5088	5293	21225	12514	6730	1835	16	129	0	0	0	1	0	0	0
HPV16_32_v4	5174	5497	27509	17527	7117	2426	205	167	66	0	0	0	1	0	0
HPV16_33	5438	5668	35503	23406	7794	3213	978	105	3	2	0	1	1	0	0
HPV16_34	5616	5869	33361	20952	9108	2033	1124	142	0	1	0	0	0	1	0
HPV16_35	5784	5963	24144	16123	5838	1776	366	15	22	1	0	0	1	1	1
HPV16_36	5951	6119	37553	27174	6225	2008	2138	0	2	0	0	1	1	0	4
HPV16_37	6065	6285	35881	26468	5051	2849	1010	267	236	0	0	0	0	0	0
HPV16_38_v3	6265	6465	33558	26208	4927	1797	623	0	0	0	0	0	3	0	0
HPV16_39	6459	6662	37255	28016	5205	2628	1330	72	0	0	0	0	4	0	0
HPV16_40	6619	6833	25878	18441	4246	2574	571	43	0	2	0	0	1	0	0
HPV16_41	6792	7017	19306	9507	6889	1907	766	237	0	0	0	0	0	0	0
HPV16_42_v3	6895	7091	22010	11942	7141	1967	883	0	73	0	0	1	0	3	0
HPV16_43	7081	7290	38604	23349	10326	2715	2021	0	42	0	151	0	0	0	0
HPV16_44_v4	7223	7435	23172	16518	3693	1972	988	0	1	0	0	0	0	0	0
HPV16_46_v3	7456	7697	39307	22808	11337	4131	991	35	2	0	0	0	3	0	0
HPV16_47	7683	7906	16570	11510	3361	1407	283	9	0	0	0	0	0	0	0
HPV47	7876	8144	2407	1493	337	351	226	0	0	0	0	0	0	0	0
HPV16_48_A	7876	7906	13025	7911	3158	1347	492	115	0	0	0	0	2	0	0

Tabela 3 – Correlação entre o Ct da amplicação (por qPCR), a quantidade de "reads" (fragmentos do genoma do HPV) e dados clínicos dos pacientes.

AMOSTRA	Ct qPCR HPV16	READS TOTAL	SEXO	IDADE AO DIAGNOSTICO	LOCALIZAÇAO ESPECIFICA
012_HPV16	ND	994599	ND	ND	ND
009_HPV16	21,63	310813	Masculino	54	trigono retromolar
005_HPV16	21,41	115555	Masculino	66	ND
008_HPV16	24,32	31705	Feminino	47	amigdala direita

ND: não determinado; Ct: ciclo threshold; Idade em anos; qPCR: PCR em tempo real.

Dentre as características clínicas das amostras, a proporção de 2:1 entre homens e mulheres apresentada para a prevalência de CCP descritos na literatura[47] se manteve, com idade média de 57,7 anos e desvio padrão 9,6.

Novos experimentos para determinar o limite de detecção real do método de NGS devem ser realizados a partir de amostras com diferentes concentrações de HPV previamente conhecidas, o que, apesar de ser o ideal, precisa ser planejado adequadamente devido ao alto custo dos reagentes para o sequenciamento.

As amostras analisadas pertencem ao gênero *Alphapapillomavirus*, com 100% de identidade, com exceção da amostra (008 – 99,14%) como mostra a análise filogenética, um padrão da análise exploratória de sequências de DNA, capaz de mostrar o caminho evolutivo e epidemiológico de agentes etiológicos de doenças e/ou o caminho evolutivo de seus hospedeiros. As amostras analisadas são pertencentes à linhagens específicas (amostras 012, 009 e 005 – A2 e amostra 008 – linhagem C). Essas análises foram realizadas no programa *online* MAFFT (*Multiple alignment program for amino acid or nucleotide sequences*)[76] e no laboratório da Dra. Lisa Mirabello, NIH (Figura 5).

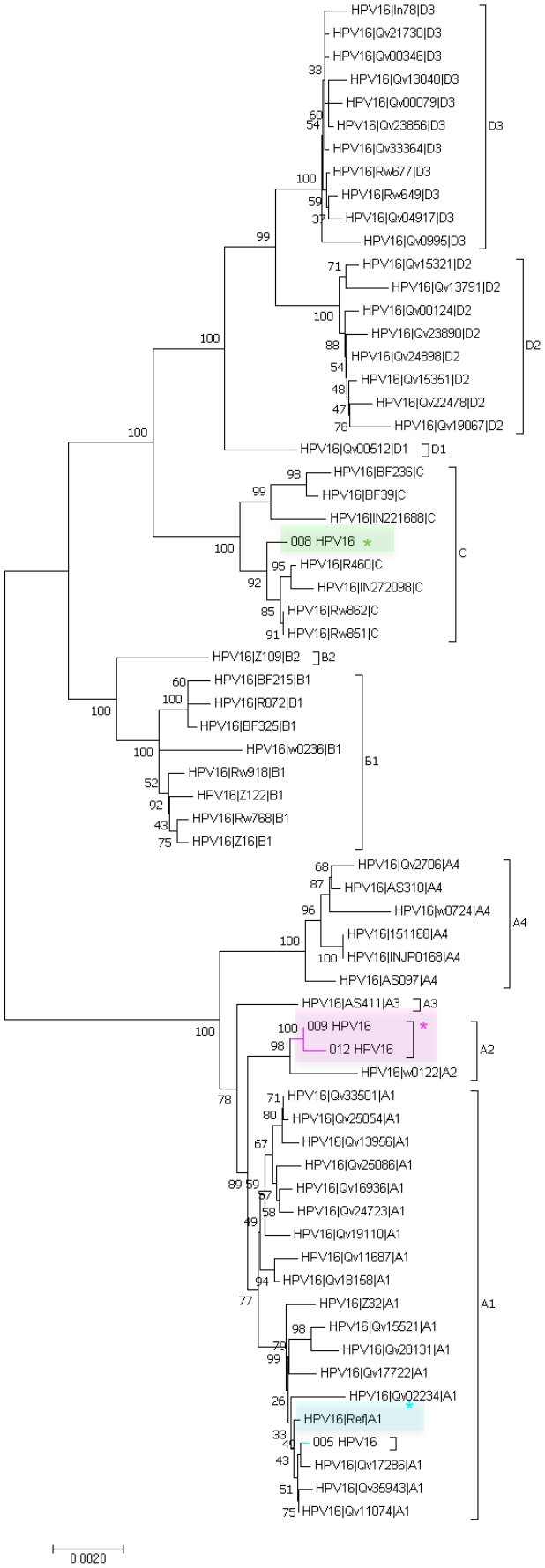

Figura 5 - Árvore filogenética produzida pelo laboratório da Dra. Lisa Mirabello a partir das sequencias geradas e analisadas. No asterisco destacado na caixa (*) rosa temos representada as amostras de CCP 009 e 012, no asterisco azul, a amostra 005, e no asterisco verde a amostra 008, de acordo com a proximidade do genoma referência.

Após a aplicação dos “pipelines” desenvolvidos pelo laboratório da Dra. Lisa Mirabello (NIH), as sequências obtidas de cada amostra foram alinhadas com o genoma de referência para HPV16 (NC_001526.4) no programa *CLC Genomics Workbench 6* (Qiagen Bioinformatics) ferramenta desenvolvida para analisar e visualizar os dados de NGS (Figura 6).

Dentre as amostras analisadas, a amostra 008 possui uma grande região sem cobertura comparada as demais amostras sequenciadas (aproximadamente 300pb) na região L2 do genoma viral. Essa abordagem pode ser vista como uma deleção de parte do gene codificando a proteína L2 constituinte minoritário do capsídeo viral e de expressão tardia, possivelmente devido ao processo de integração do DNA viral no genoma humano do hospedeiro, apesar de não ser uma região de integração descrita como oncogênica[32] (Figura 6 – C).

Outra região a se destacar foi um “contig” de sequências não detectadas em nenhuma das quatros amostras analisadas, localizada na região E6 do vírus HPV, que codifica uma das oncoproteínas virais essenciais para a transformação maligna (Figura 6 – D).

Maiores esforços serão necessários para analisar os produtos sequenciados, principalmente nas regiões de deleção. Vale a pena ressaltar que o sequenciamento das amostras extraídas de blocos de parafina no Instituto de Medicina Tropical – USP, não interferiu na qualidade das análises, que se mostrou tão eficiente quanto o sequenciamento realizado previamente em amostras frescas de tecido congelado, de outras casuísticas do grupo.

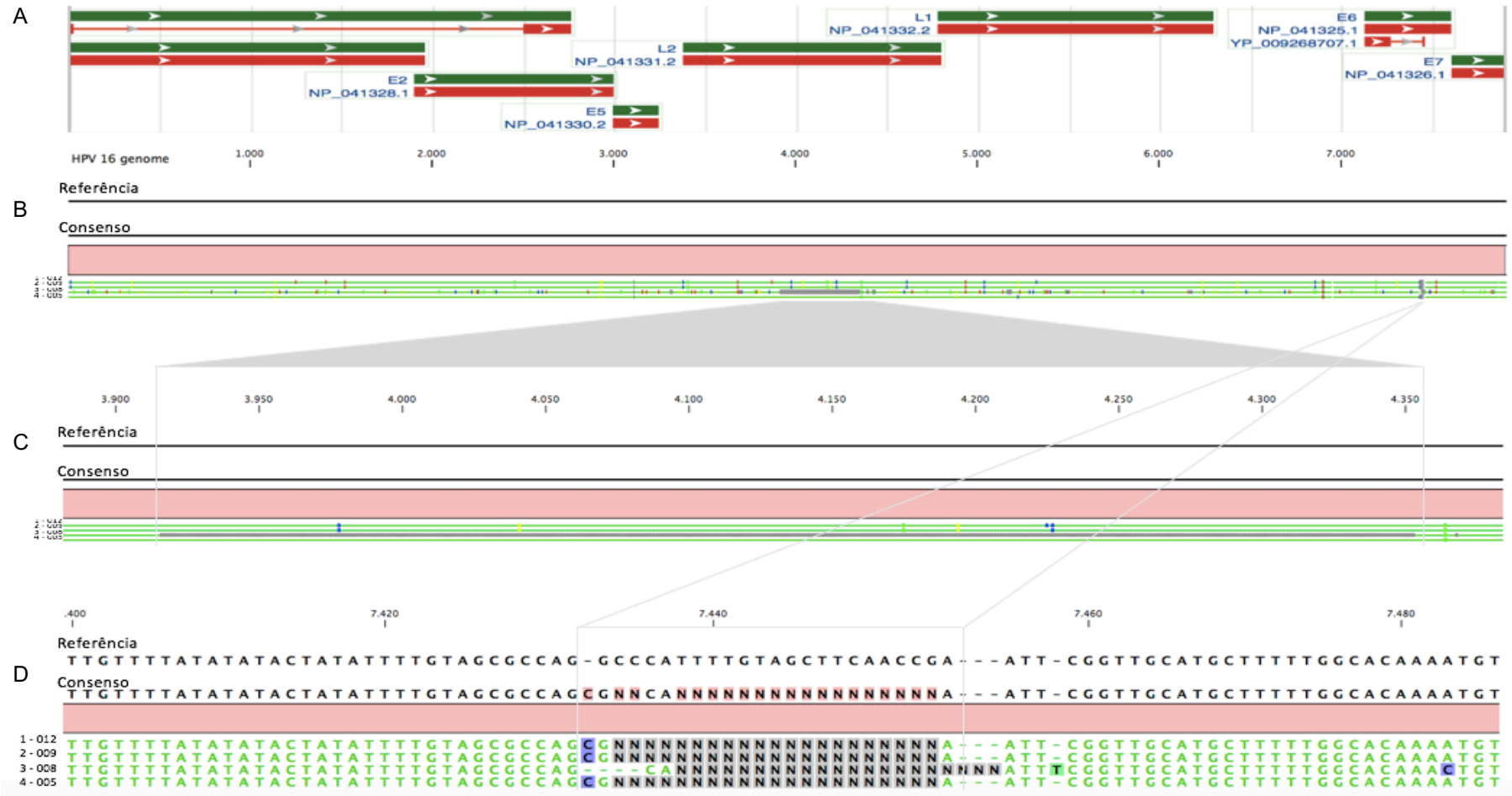

Figura 6 – Comparação dos produtos amplificados das quatro amostras com reads detectados, assim como o genoma referencia do HPV16 (NC_001526.4). A. Genoma completo do HPV16 referência com os genes destacados e B. as amostras de CECP analisadas. C. Destaque da imagem B na região de deleção da amostra 008, entre 3900 e 4200 nucleotídeos e D. Destaque da imagem B em um segmento com ausência de nucleotídeo nas quatro amostras analisadas.

Conclusões

6 CONCLUSÕES

6.1 Prevalência de HPV em tumores de Cabeça e Pescoço

A prevalência de HPV em pacientes com CECP de tecido emblocado em parafina correspondeu a 10% (49 HPV+ / 491 amostras). Apesar de ser uma prevalência baixa quando comparada a outras localizações, como EUA e Europa, que chegam a ultrapassar 60%, para o Brasil é uma das maiores taxas já reportadas[25], apesar de apresentar dados concordantes na prevalência de HPV16, que corresponde a 78% das amostras positivas.

Devido à diversidade de sítios anatômicos de carcinomas de cabeça e pescoço incluídos nessa casuística, essa prevalência pode ser ainda maior se isolados por sítio anatômico específico, como é o caso da orofaringe, na qual a prevalência atinge 25% dos casos desse sítio anatômico nas amostras do Hospital das Clínicas de SP, sendo todos os casos positivos para HPV 16, confirmando sua alta prevalência dentre os tumores dessa região associados ao papilomavírus humano.

Observou-se um aumento significativo da prevalência do DNA-HPV em tumores de orofaringe quando comparado a série anterior, com HPV16 desempenhando o principal papel da carcinogênese desses tumores (Tabela 3).

A divergência nas taxas de DNA do HPV observada pode ser atribuída não só as diferenças nas metodologias utilizadas para a detecção do HPV nos diferentes estudos, mas também às características dos indivíduos em estudo, incluindo práticas sexuais, status econômico e social. Infelizmente, a ausência de informações sócio demográficas e de comportamento sexual dos pacientes avaliados impede uma análise adicional.

Em conclusão, a crescente frequência de DNA de HPV em CECP (principalmente na orofaringe) corrobora a hipótese de que o HPV está iniciando sua projeção nesse continente, apesar de não estar circulante entre o tecido normal de

tonsila da população brasileira. Estudos envolvendo a avaliação da expressão proteica associada à infecção por HPV devem ser conduzidos para reforçar se esse perfil crescente de DNA de HPV também reflete uma crescente participação etiológica do HPV 16 nos novos casos de tumores epidermoides de orofaringe no Brasil.

6.2 Análise Mutacional em tumores de Cabeça e Pescoço

A análise piloto de 98 casos de tumores de cabeça e pescoço do Gencapo em colaboração com o IARC levou ao sequenciamento de genes específicos associados ao desenvolvimento de tumores de cabeça e pescoço. O status HPV desta amostra foi baseado em sorologia anti-HPV16-E6. Essa casuística apresentou 8% de positividade para o HPV, no entanto, 75% dos casos apresentaram mutação em pelo menos um gene dos cinco genes analisados. Esta evidência aumenta a discussão do papel do HPV como um indutor de carcinogênese ou simplesmente um agente passivo no processo natural de carcinogênese dos tumores de cabeça e pescoço.

Foram identificados casos de CEC com poucas alterações que se diferenciam como um subconjunto de tumores de cabeça e pescoço dirigidos predominantemente por mutações gênicas e alterações focais em vez de eventos de instabilidade cromossômica, sendo caracterizados por uma maior sobrevida geral. Avaliações mais específicas são necessárias para estabelecer esta associação entre as alterações “drivers” e “passengers” nas amostras HPV positivas, assim como as informações relatadas no artigo apresentado no Apêndice B, que mostrou uma clara associação com ambas as exposições ambientais (consumo de álcool e fumo e infecção pelo HPV) e às características clínicas. Mais estudos integrando dados genômicos, clínicos e epidemiológicos, especialmente em populações de alto risco, são necessários para melhor identificar a estratificação de alto risco e caracterizar o prognóstico de casos de câncer de cabeça e pescoço.

6.3 Sequenciamento de nova geração (NGS) HPV16

O sequenciamento do HPV16 em amostras de CCP, principalmente aqueles da orofaringe, apresentou um desafio importante no desenvolvimento dessa tese. Esse foi o objetivo inicial e principal apresentado, com o intuito de correlacionar o sequenciamento dessas amostras com o sítio de integração ao DNA humano, em regiões conhecidas de instabilidade cromossômica, e comparar com o HPV16 que igualmente infecta a região da cérvix uterina.

O protocolo desenvolvido foi inteiramente customizado, elaborado em conjunto com o suporte técnico da empresa, porém, devido a questões técnicas, muitos foram os contratempos encontrados nessa elaboração. Infelizmente, por esse motivo, não tivemos sucesso na detecção das regiões de DNA humano flanqueadoras, além do baixo rendimento obtido com as amostras sequenciadas, assim como a ausência de amostras representativas da cérvix uterina para realizar a comparação entre o vírus nos diferentes sítios anatômicos, como proposto inicialmente.

Apesar disso, a análise em colaboração com o NIH apresentou dados interessantes que precisam ser confirmados posteriormente, principalmente nas regiões de deleção.

Referências

REFERÊNCIAS

1. Strauss Mj, Shaw Ew. Crystalline virus-like particles from skin papillomas characterized by intranuclear inclusion bodies. Proc Soc Exp Biology and Medicine. 1949:72;46-50.

2. Pannone G; Santoro A; Papagerakis S; et al. The role of Human Papillomavirus in the pathogenesis of Head & Neck Squamous Cell Carcinoma: an overview; Infectious Agents and Cancer 2011, 6:4 doi:10.1186/1750-9378-6-4

3. de Sanjose S, Brotons M, Pavon MA. The natural history of human papillomavirus infection. Best Pract Res Clin Obstet Gynaecol. 2018;47:2-13.

4. Marklund L, Hammarstedt L. Impact of HPV in Oropharyngeal Cancer. J Oncol. 2011;2011:509036.

5. de Villiers, E.M., 2013. Cross-roads in the classification of papillomaviruses. Virology 445, 2–10.

6. Muñoz N, Bosch Fx, De Sanjosé S, et al. Epidemiologic classification of human papillomavirus types associated with cervical cancer. International Agency for Research on Cancer Multicenter Cervical Cancer Study Group. N Engl J Med 2003;348:518-27.

7. Shukla S, Bharti Ac, Mahata S, et al. Infection of human papillomaviruses in cancers of different human organ sites. Indian J Med Res. 2009;130(3):222-33.

8. Dunne Ef, Nielson Cm, Stone KM. Prevalence of HPV infection among men: A systematic review of the literature. J Infect Dis. 2006;194:1044-57.

9. Schabath Mb, Villa Ll, Lazcano-Ponce E. Smoking and Human Papillomavirus (HPV) Infection in the HPV in Men (HIM) Study. Cancer Epidemiol Biomarkers Prev. 2012;21:102-110.

10. Andl T, Kahn T, Pfuhl A, et al. Etiological involvement of oncogenic human papillomavirus in tonsillar squamous cell carcinomas lacking retinoblastoma cell cycle control. Cancer Research. 1998;58:5-13.

11. Frisch M, Biggar Rj. A etiological parallel between tonsillar and anogenital squamous-cell carcinomas. Lancet. 1999;354:1442-3.

12. Weinberg 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.

13. 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.

14. Jung Ac, Briolat J, Millon R, et al. Biological and clinical relevance of transcriptionally active human papillomavirus (HPV) infection in oropharynx squamous cell carcinoma. Int J Cancer. 2010;126:1882-94.

15. Li Lw, Li Yy, Li Xy, et al. A novel tumor suppressor gene ECRG4 interacts directly with TMPRSS11A (ECRG1) to inhibit cancer cell growth in esophageal carcinoma. BMC Cancer. 2011;11:52-9.

16. Kreimer Ar, Clifford Gm, Boyle P, et al. Human papillomavirus types in head and neck squamous cell carcinomas worldwide: a systematic review. Cancer Epidemiol Biomarkers Prev. 2005;14:467-75.

17. Dayyani F, Etzel CJ, Liu M, et al. Meta-analysis of the impact of human papillomavirus (HPV) on cancer risk and overall survival in head and neck squamous cell carcinomas (HNSCC). Head Neck Oncol. 2010;29:2-15.

18. D'Souza G, Agrawal Y, Halpern J. Oral sexual behaviors associated with prevalent oral human papillomavirus infection. J Infect Dis. 2009;199:1263-1269.

19. Jemal A, Bray F, Center MM, et al. Global Cancer Statistics. Ca Cancer J Clin 2011;61:69-90.

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

21. Taberna M, Mena M, Pavón MA, et al. Human papillomavirus-related oropharyngeal cancer. Annals of Oncology. 2017;28:2386–2398.

22. Duvvuri U, Myers JN. Contemporary management of oropharyngeal cancer: anatomy and physiology of the oropharynx. Curr Probl Surg. 2009;46:119-84.

23. Westra WH, Taube JM, Poeta ML et al. Inverse Relationship between Human Papillomavirus-16 Infection and Disruptive p53 GeneMutations in Squamous Cell Carcinoma of the Head and Neck. Clin Cancer Res. 2008;14:366-9.

24. Kumar B, Cordell KG, Lee JS, et al. EGFR, p16, HPV Titer, Bcl-xL and p53, sex, and smoking as indicators of response to therapy and survival in oropharyngeal cancer. J Clin Oncol 2008;26:3128-37.

25. O'Regan EM, Toner ME, Finn SP, et al. p16(INK4A) genetic and epigenetic profiles differ in relation to age and site in head and neck squamous cell carcinomas. Hum Pathol. 2008;39:452-8.

26. López RVM, Levi JE, Eluf-Neto J, et al. Human papillomavirus (HPV) 16 and the prognosis of head and neck cancer in a geographical region with a low prevalence of HPV infection. Cancer Causes Control. 2014; 25:461-71.

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

28. Havre PA, Yuan J, Hedrick L, et al. p53 inactivation by HPV16 E6 results in increased mutagenesis in human cells. Cancer Research. 1995:55;4420-4.

29. Chung CH, Gillison ML. Human papillomavirus in head and neck cancer: its role in pathogenesis and clinical implications. Clinical Cancer Research. 2009;15:6758-62.

30. Peter M, Stransky N, Couturier J, et al. Frequent genomic structural alterations at HPV insertion sites in cervical carcinoma. J Pathol. 2010;221:320-30.

31. Parfenov M, Pedamallu CS, Gehlenborg N, Freeman SS, Danilova L, Bristow CA, et al. Characterization of HPV and host genome interactions in primary head and neck cancers. Proc Natl Acad Sci U S A. 2014; 111(43):15544–9.

32. McBride AA, Warburton A. The role of integration in oncogenic progression of HPV-associated cancers. PLoS Pathog. 2017;13(4):e1006211.

33. Vojtechova Z, Sabol I, Salakova M, Turek L, Grega M, Smahelova J, et al. Analysis of the integration of human papillomaviruses in head and neck tumours in relation to patients' prognosis. Int J Cancer. 2016; 138(2):386–95.

34. Gao G, Johnson SH, Vasmatzis G, Pauley CE, Tombers NM, Kasperbauer JL, et al. Common fragile sites (CFS) and extremely large CFS genes are targets for

human papillomavirus integrations and chromosome rearrangements in oropharyngeal squamous cell carcinoma. Genes Chromosomes Cancer. 2017;56(1):59-74.

35. Conway C, Chalkley R, High A, et al. Next-generation sequencing for simultaneous determination of human papillomavirus load, subtype, and associated genomic copy number changes in tumors. J Mol Diagn. 2012;14:104-11.

36. Tuna M and Amos CI. Next generation sequencing and its applications in HPV associated cancers. 2016

37. Bray F, Ferlay J, Soerjomataram I, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: A Cancer Journal for Clinicisnd. 2018;0:1-31.

38. Papanicolaou GN, Traut HF. The diagnostic value of vaginal smears in carcinoma of the uterus. Am J Obstet Gynecol. 1941;42:193-206.

39. Belinson SE, Belinson JL. Human papillomavirus DNA testing for cervical cancer screening: practical aspects in developing countries. Mol Diagn Ther. 2010;14:215-22.

40. Parkin DM, Bray F, Ferlay J, Pisani P. Global cancer statistics, 2002. CA Cancer J Clin. 2005;55(2):74-108.

41. Castellsagué X, Muñoz N. Cofactors in human papillomavirus carcinogenesis - role of parity, oral contraceptives, and tobacco smoking. J Natl Cancer Inst Monogr. 2003;31:20-8.

42. Schabath MB, Villa LL, Lazcano-Ponce E. Smoking and Human Papillomavirus (HPV) Infection in the HPV in Men (HIM) Study. Cancer Epidemiol Biomarkers Prev. 2012;21:102-110.

43. Sankaranarayanan R, Nene BM, Shastri SS, et al. HPV screening for cervical cancer in rural India. N Eng J Med. 2009;360:1385-94.

44. Martins TR, Oliveira, CM, Rosa LR, et al. HPV genotype distribution in Brazilian women with and without cervical lesions: correlation to cytological data. Virology Journal. 2016 Aug 12;13(1):138.

45. Cuzick J, Clavel C, Petry KU, et al. Overview of the European and North American studies on HPV testing in primary cervical cancer screening. Int J Cancer. 2006;119:1095-101.

46. Martins TR, Longatto-Filho A, Cohen D, et al. Influence of prior knowledge of human papillomavirus status on the performance of cytology screening. Am J Clin Pathol. 2018;7:149(4):316-323.

47. INCA, Estimativa 2018: incidência de câncer no Brasil / Instituto Nacional de Câncer José Alencar Gomes da Silva. Coordenação de Prevenção e Vigilância. – Rio de Janeiro: INCA,2017.[http://www.inca.gov.br/estimativa/2018/estimativa-2018.pdf]

48. Ha PK, Califano JA: The role of human papillomavirus in oral carcinogenesis. Crit Rev Oral Biol Med 2004;15:188-96.

49. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell. 2011;144(5):646-74.

50. Ng M, Freeman MK, Fleming TD et al. Smoking prevalence and cigarette consumption in 187 countries, 1980–2012. JAMA 2014; 311(2):183–192.

51. Gillison ML, Chaturvedi AK, Anderson WF, Fakhry C. Epidemiology of human papillomavirus-positive head and neck squamous cell carcinoma. J Clin Oncol 2015; 33(29): 3235–3242.

52. Smith, RB, Sniezek JC, Weed DT et al. Utilization of free tissue transfer in head and neck surgery. Otolaryngol. Head Neck Surg. 2007:137;182–191.

53. Vergeer MR, Doornaert PA, Rietveld DH, et al. Intensity-modulated radiotherapy reduces radiation-induced morbidity and improves health-related quality of life: results of a nonrandomized prospective study using a standardized follow-up program. Int. J. Radiat. Oncol. Biol. Phys. 2009;74:1–8.

54. Leemans CR, Braakhuis BJ, Brakenhoff RH. The molecular biology of head and neck cancer. Nat Rev Cancer. 2011;11(1):9-22.

55. D'Souza G, Anantharaman D, Gheit T, et al. Effect of HPV on head and neck cancer patient survival, by region and tumor site: A comparison of 1362 cases across three continentes. Oral Oncol. 2016; 62: 20–27.

56. Torre LA, Bray F, Siegel RL. Global cancer statistics, 2012. CA Cancer J Clin. 2015; 65:87–108.

57. Chaturvedi AK, Anderson WF, Lortet-Tieulent J, et al. Worldwide trends in incidence rates for oral cavity and oropharyngeal cancers. J Clin Oncol. 2013; 31:4550–4559.

58. Ribeiro KB, Levi JE, Pawlita M, et al. Low human papillomavirus prevalence in head and neck cancer: results from two large case-control studies in high-incidence regions. Int J Epidemiol. 2011; 40:489–502.

59. Lassen P, Eriksen JG, Hamilton-Dutoit S, et al. Effect of HPV-associated p16INK4A expression on response to radiotherapy and survival in squamous cell carcinoma of the head and neck. J Clin Oncol. 2009;27:1992–8.

60. Ang KK, Harris J, Wheeler R et al. “Human papillomavirus and survival of patients with oropharyngeal cancer,” New England Journal of Medicine, vol. 363, no. 1, pp. 24–35, 2010

61. Mellin H, Friesland S, Lewensohn R, Dalianis T, Munck-Wikland E. Human papillomavirus (HPV) DNA in tonsillar cancer: clinical correlates, risk of relapse, and survival. Int J Cancer. 2000;89(3):300-4.

62. Gillison ML, Lowy DR. A causal role for human papillomavirus in head and neck cancer. Lancet. 2004;363(9420):1488-9.

63. Weinberg 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.

64. Lindquist D, Romanitan M, Hammarstedt L, et al. Human papillomavirus is a favourable prognostic factor in tonsillar cancer and its oncogenic role is supported by the expression of E6 and E7. Mol Oncol. 2007;1(3):350-5.

65. Fakhry C, Westra WH, LI S, et al. Improved survival of patients with human papillomavirus-positive head and neck squamous cell carcinoma in a prospective clinical trial. J Natl Cancer Inst. 2008;100:261-9.

66. Licitra L, Perrone F, Bossi P, et al. High-risk human papillomavirus affects prognosis in patients with surgically treated oropharyngeal squamous cell carcinoma. J Clin Oncol. 2006;24:5630–6.

67. 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.

68. Vu HL, Sikora AG, Fu S, Kao J. HPV-induced oropharyngeal cancer, immune response and response to therapy. Cancer Lett 2010; 288(2):149–155.

69. Spanos WC, Nowicki P, Lee DW et al. “Immune response during therapy with cisplatin or radiation for human papillomavirus-related head and neck cancer,” Archives of Otolaryngology—Head and Neck Surgery, vol. 135, no. 11, pp. 1137–1146, 2009.

70. Oliveira CMd, Bravo IG., Souza NCSe, et al. High-level of viral genomic diversity in cervical cancers: A Brazilian study on human papillomavirus type 16. Infection, Genetics and Evolution. 34:2015;44–51.

71. Turner Do, Williams-Cocks Sj, Bullen R, et al. High-risk human papillomavirus (HPV) screening and detection in healthy patient saliva samples: a pilot study. BMC Oral Health. 2011;11:28.

72. Saiki RK, Scharf S, Faloona F, Mullis KB, Horn GT, Erlich HA e Arnheim N: Enzymatic amplification of beta‐globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. Science. 1985;230:1350-1354.

73. Walboomers JM, Jacobs MV, Manos MM et al. Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. J Pathol. 1999;189:12-9.

74. Scapulatempo-Neto C, Veo C, Fregnani JHTG, Lorenzi A, Mafra A, Melani AGF, et al. Characterization of topoisomerase II alpha and minichromosome maintenance protein 2 expression in anal carcinoma. Oncol Lett. 2017;13(3):1891-1898.

75. Higgins DG, Thompson JD, Gibson TJ. Using CLUSTAL for multiple sequence alignments. Methods Enzymol. 1996;266:383-402.

76. Katoh, Rozewicki, Yamada. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. 2017. Briefings in Bioinformatics, in press.

77. Posada D, Crandall KA. MODELTEST: testing the model of DNA substitution. Bioinformatics. 1998;14:817-8.

78. Pond SL, Frost SD, Muse SV. HyPhy: hypothesis testing using phylogenies. Bioinformatics. 2005;21:676-9.

79. Kumar S, Tamura K, Jakobsen IB, Nei M. MEGA4: Molecular Evolutionary Genetics Analysis Software. Bioinformatics. 2004;17:1244-45.

80. Librado P, Rozas J. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. Bioinformatics. 2009; 25: 1451-2.

81. Perdomo S, Anantharaman D, Foll M, Abedi-Ardekani B, Durand G, Reis Rosa LA, et al. Genomic analysis of head and neck cancer cases from two high incidence regions. PLoS ONE. 2018;13(1):e0191701.

82. Barros-Filho MC, Reis-Rosa LA, Hatakeyama M, et al. Oncogenic drivers in 11q13 associated with prognosis and response to therapy in advanced oropharyngeal carcinomas. Oral Oncology. 2018:83;81–90

83. Mirabello L, Yeager M, Yu K, et al. HPV16 E7 Genetic Conservation Is Critical to Carcinogenesis. Cell. 2017; 170(6): 1164–1174.

Apêndice

APÊNDICE A - *"Prevalence of HPV in head and neck normal tissue and carcinomas from São Paulo, Brazil**". (em preparação)***

***Luciana Reis Rosa Sacoman**, Jéssica Boscariol da Silva, Luisa Lina Villa, Eloiza Helena Tajara da Silva, GENCAPO Group, José Eduardo Levi.*

Prevalence of HPV in head and neck carcinomas and normal tissue from Sao Paulo, Brazil

Luciana Reis Rosa Sacoman, Jessica Boscariol da Silva, Luisa Lina Villa, Eloiza Helena Tajara da Silva, GENCAPO Group, Jose Eduardo Levi.

Background: Human papillomavirus (HPV) is an important etiologic agent in a subset of human tumors and HPV16 is the most prevalent type in tumors of the oropharynx, but varies widely according to tumor site and geographical region. The aim of this study is to investigate the frequency of high-risk HPV (hrHPV) DNA in a large Brazilian cohort of patients with HNSCC and a comparative group of non-neoplasic tonsils. **Methods**: Formalin-fixed paraffin-embedded specimens of HNSCC were available from 510 patients who were recruited at three Brazilian hospitals and fresh tonsillar tissue was obtained from 113 patients without neoplasia that underwent surgery for non-neoplasic causes. HPV detection and genotyping from FFPE HNSCCs was performed by the Inno-Lipa method while PGMY09/11 PCR was employed for non-neoplastic frozen tonsillar tissues. All HPV positive samples had the HPV16 presence investigated by type-specific real-time PCR. **Results**: The frequency of high-risk HPV DNA in HNCSS was 10% (49/491), 78% of which were HPV16. There was a huge variability in the prevalence of HPV in the tumors according to the anatomical site, ranging from 3.4% (base of the tongue and hypopharynx) to 25% (oropharynx). In contrast, no hrHPV-DNA was present in the non-neoplastic tonsils analyzed. **Conclusion**: In the current study, the increasing frequency of HPV DNA in HNSCC (mainly in the oropharynx) corroborates the hypothesis that HPV is initiating its projection in this continent, although it is not yet circulating among the normal tonsil tissue of the Brazilian population. Studies involving the evaluation of protein expression associated with HPV infection should be conducted to reinforce whether this increasing profile of HPV DNA also reflects an increasing etiological involvement of HPV16 in the new cases of oropharyngeal squamous cell carcinomas in Brazil.

Keywords: Head and Neck cancer, oropharyngeal cancer, normal tonsils, HPV prevalence, high-risk HPV

Background

Head and neck cancer (HNC) is a broad term that encompasses malignant epithelial changes including those from the paranasal sinuses, oropharynx, oral cavity, larynx and hypopharynx, comprising a heterogeneous group of tumors associated with these anatomical locations. They often affect men over the age of 50 years and a history of smoking and drinking habits. The main histological subtype is squamous cell carcinoma (SCC), accounting for over 95% of the cases. The most frequent location is the oral cavity (40% of cases), followed by larynx (25%) and pharynx (15%). Data from the Brazilian National Cancer Institute (INCA, 2018) shows that, in the Brazilian population, there is a ratio of 3.3 cases in males to each case in females.

Worldwide, lip and oral cavity cancers accounted 300,000 cases in 2012, with 145,000 deaths recorded, 77% of them in developing countries. The highest rates of these cancers are found in Brazil, Melanesia, south-central Asia and Central (Spain and France) and Eastern Europe (Ferlay et al, 2015 - GLOBOCAN 2012).

Development of HNC is associated with accumulation of damage in genes that control cell proliferation, senescence, invasion, motility and cell survival (Califano et al., 1996; Forastiere et al., 2001; Ha and Califano, 2006; Perdomo et al., 2018). These genetic alterations often occur as a result of excessive exposure to tobacco, alcohol and other carcinogenic elements that trigger irreversible damage to DNA, causing changes in coding and regulatory regions of tumor suppressor genes and oncogenes. The accumulation of these mutations enable the tumor cells to acquire characteristics, including resistance to death and uncontrolled proliferation, conferring selective growth advantage (Hanahan and Weinberg, 2010). Tumor cells can form clones with increased invasiveness and stimulation of angiogenesis, leading to the establishment of more invasive tumors with greater potential to develop metastasis (Duvvuri and Myers, 2009).

In Brazil, it is estimated an incidence (age standardized rate - ASR) of 7.35 new cases of oropharyngeal carcinoma per 100,000 inhabitants per year, reaching just over 1,000 new cases per year (Ferlay et al, 2015 - GLOBOCAN 2012).

The survival rate at five years for the oral and oropharyngeal carcinoma in Brazil is still below 50%, with a trend for better survival observed among patients with HPV positive tumors independent of age (Carvalho et al., 2004; Kaminagakura et al., 2012).

Prognostic factors that were examined in relation to the survival of tumors of the head and neck include clinical tumor stage, education and patient hygiene habits, smoking and alcohol consumption (of Graeff et al., 2001; Dikshit et al., 2005). More recently, some genetic markers have been studied in relation to the survival of tumors of the head and neck, including gene polymorphisms related to invasion and metastasis, inflammation and DNA repair (Chiang et al., 2008; Shimizu et al., 2008; Hopkins et al., 2008; López et al., 2011; Barros-Filho et al., 2018).

In addition to the evidence of the involvement of HPV in the etiology of oropharyngeal SCC (D'Souza et al., 2007), there seems to be a significant correlation between the presence of infection and the most effective response to treatment of patients with this tumor (Duvvuri and Myers, 2009; Lassen et al., 2009; Ang et al., 2010). Recently, D'Souza et al. (2016) analyzed 1,362 HNSCC diagnosed between 2002-2011, including Brazilian samples (388) and, after testing for p16INK4a and HPV16 DNA (by PCR) the risk of mortality was significantly lower among HPV-related (i.e., p16+/HPV16+) compared to HPV-unrelated (p16– and/or HPV16–) OPSCC cases (HR=0.25, 95%CI=0.18-0.34). The mortality was reduced among HPV-related OPSCC cases from the U.S., Europe, and Brazil (each $p \leq 0.01$) and after adjustment, remained significantly lower (hare=0.34, 95%CI=0.24-0.49). Among non-OP HNSCC, neither p16 (aHR=0.83, 95%CI=0.60-1.14), HPV16 DNA (aHR=1.20, 95%CI=0.89-1.63), or p16+/HPV16+ (aHR=0.59, 95%CI=0.32-1.08) was a significantly predictor of mortality.

HPV types are divided into two groups: high and low risk, according to their malignant potential. High-risk HPVs comprise HPV16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 68, 73 and 82 types present in high-grade squamous lesions or invasive carcinomas, and the types HPV6, 11, 40, 42, 44, 54, 61, 70, 72, 81 and 89 are considered low risk and are associated with low-grade epithelial lesions (Muñoz et al., 2003; Shukla et al., 2009). More recently, types with oncogenic potential not well established, known as "potentially high risk" have been identified, among them HPV26, 53 and 66 (Muñoz et al., 2003; Pannone et al., 2011).

Most HPV infections are transient, asymptomatic or subclinical, not resulting in disease and are usually suppressed by the immune system. Epidemiological studies indicate that the prevalence of the virus in the male population, regardless of site, can vary from 0 to 73% (Dunne et al., 2006; Schabath et al., 2012).

HPV infections are the cause of almost all cases of cervical cancer (~99%) and anal cancer (~88%), and a subgroup of tumors of the vagina (~70%), penis (~50%), and vulva (~43%), in addition to being increasingly associated with the development of squamous cell carcinomas of the head and neck (Marur et al., 2010; de Martel et al., 2012).

HPV DNA is present in approximately 25% of HNSCC, and in the US, about 50% of the oropharynx SCC contain high risk HPV DNA (Kreimer et al., 2005). The increased incidence of oropharyngeal carcinomas in developed countries is mainly due to the occurrence of the disease in young patients, not presenting classic risk factors, such as prolonged use of tobacco and alcohol. This fact leads to an increasing association of a subgroup of these tumors with papillomavirus infection (D'Souza et al., 2009; Jemal et al., 2011).

Ndiaye et al. (2014), in a systematic review and meta-analysis, provided an update on HPV-associated oropharynx, oral cavity and laryngeal carcinomas, questioning the variability of methodologies for viral DNA detection. The authors reported that 45.8% of oropharyngeal carcinomas, 24.2% of oral cavity carcinomas, and 22.1% of laryngeal carcinomas were HPV DNA-positive. When compared to the E6 / E7 mRNA expression data, the degree of overestimation of HPV-associated tumors by PCR varied according to the anatomical site, with the lowest rates of overestimation for oropharynx (~13%) carcinomas and higher for carcinomas of the oral cavity (~33%) and larynx (~61%).

The proportion of HPV positive tumors varies substantially according to the tumor anatomical site. This finding suggests that the proportion of HPV-positive tumors may be dependent on geographic and genetic factors, behavioral factors (such as sexual habits) and the lack of standardization of the methodology for the identification of these viruses in human tumors. In cervical tumors, it is widely proposed that all cases have as their etiology a genital HPV infection, despite of the importance of parallel factors in the carcinogenic process.

The presence of HPV in oropharyngeal precancerous lesions is much greater than in normal tissue samples. This fact suggests that the virus infection is a factor for the development of the disease. Studies have highlighted the need to understand not only the prevalence of oral HPV, but also the duration and persistence of the infection,

due to its potential for malignant progression (Turner et al., 2011). It is estimated that HPV types, considered to be at high risk in infected and asymptomatic individuals, are around 3.5% (Kreimer et al., 2010).

To date, few studies have reported data on the detection of HPV DNA in normal tissue in the head and neck region. For these studies, non-invasive methods of cell sampling, such as scraping and cytological brushing, are usually used to test the presence of HPV DNA by Polymerase Chain Reaction (PCR). A disadvantage of these methods is that they mainly collect superficial epithelial cells. For better detection of the presence of HPV in tonsillar tissue, biopsy or tonsillectomy samples are required in combination with PCR, or immunoblotting for $p16^{INK4a}$ (Klingenberg et al., 2010).

A study conducted in the United States with tonsils from tonsillectomy, filed at a university hospital from 2012 to 2015, found in 102 healthy adults an overall prevalence of 4.9% (n = 5) and type 16 or 18, 3.9% (n = 4). The study concludes that HPV exists in the biofilm and the extracellular space of the tonsils, configuring an oncogenic virus reservoir (Rieth et al., 2018).

In Brazil, a study was conducted to determine the presence of HPV DNA in the oropharynx region of individuals without cancer. The anatomical sites of the study were the palatine tonsils, soft palate, base of the tongue and posterior wall of the pharynx. After the surgical procedures pertinent to the treatment of the patients, scraping was performed under general anesthesia, to collect the samples. The methodologies used in the study were the conventional PCR and dot-blot hybridization techniques. This study identified HPV DNA in 14.0% of the individuals and the most prevalent sites found were the tonsils and the soft palate. The data from this study were consistent with those from other countries that reported a prevalence of HPV DNA in normal tissue with a range of 0% to 20% (do Sacramento et al., 2006).

Although it is not fully elucidated whether HPV infection in healthy individuals will or not develop HNSCC, the tonsils appear to act as reservoir for the virus (do Sacramento et al. 2006).

Material and methods

Casuistic HNSCC

Formalin-fixed paraffin embedded (FFPE) tissues from 510 head and neck carcinoma samples were obtained from the biobank and archives of the collaborative projects named GENCAPO 1 and 2 (CONEP 16491 - Brazilian Head and Neck Genome Project - complete list of members and affiliations presented at http://www.gencapo.famerp.br). The patients were attended at three different institutions (369 patients from Hospital das Clínicas de São Paulo - HC, 84 patients from Heliópolis Hospital - HH and 57 patients from Instituto do Câncer Arnaldo Vieira de Carvalho - ICAVC). Clinical data was available exclusively from the HC patients that occurred during 2002 to 2010.

Casuistic Non Neoplasic Oropharynx

Patients with age above 18 years old, from both genders, without suspicion of neoplasia that were submitted to surgical procedures of uvulopalatopharyngoplasty or tonsillectomy at the Otorhinolaryngology Service of the Hospital das Clinicas of the Faculty of Medicine of the University of São Paulo and Hospital Paulista of São Paulo were invited to this study (CAPPesq - Ethics Committee for Analysis of Research Projects under opinion 173.181). This arm of the study included 113 patients that underwent surgery from 2013 to 2016.

After signing the consent form, a life-style interview was conducted, whose questionnaire included the use of tobacco, alcohol, and sexual habits. The questionnaire used is part of the thematic project: "Environmental, clinical, histopathological and molecular factors associated with the development and prognosis of squamous cell carcinomas of the head and neck", carried out by the Gencapo (Head and Neck Cancer Genome) funded by FAPESP (Foundation for Research Support of the State of São Paulo), under number 2010/51168-0, with the approval of CONEP (National Commission for Research Ethics) registration number 16491.

DNA Extraction

FFPE-DNA isolation from the FFPE head and neck carcinoma samples was

performed by 100% octane and subsequently ethanol washes for deparaffinization followed by digestion with proteinase K solution, according to López et al. (2014). Subsequent extraction of the genomic DNA was done employing QIAamp DNA FFPE Tissue Kit (Qiagen). Spectrophotometry (NanoDrop ND-1000 Spectrophotometer v.3.0.1, Labtrade) was used to quantify extracted DNA and verify the purity and quality parameters.

Fresh tissues from non neoplasic tonsils removed during the surgeries went through a cut, to obtain a duplicate sample of each tonsil, right and left, generating four samples from each patient. For extraction, the Qiagen DNeasy Blood & Tissue Kit (Qiagen, São Paulo, Brazil) was used according to the manufacturer's instructions. For concentration and purity checking, all extracted samples were quantified on the NanoDrop 2000/2000C spectrophotometer (Spectrophotometer, Thermo Scientific, Wilmington, Delawere, USA). After this process, DNA solutions were diluted to 50ng/µL, and stored at -20 °C.

PCR for constitutive Beta-globin gene

To determine the integrity of the extracted DNA from normal tissue, samples (250ng of DNA) were submitted to amplification of the cellular beta-globin gene, using the primers PCO4 and GH20 (Gravitt et al., 2000) spanning a 268bp fragment. For the HNSCC samples, the set of beta-globin primers were PCO3 and PCO4 (Saiki et al. 1985), which amplifies 107bp fragments, starting from 100ng of DNA. The difference on the size of the beta-globin amplicons is due to the well-known reduced integrity of DNA on FFPE samples, less likely to amplify larger DNA fragments.

Detection of HPV by conventional PCR

PCR was performed in normal HN samples with primers PGMY09/11 that amplify 450bp fragment of HPV L1 region from different viral types, and is considered gold standard in mucosal HPVs detection (Gravitt et al., 2000). PCR was performed according to López et al. (2014).

Amplification products and the 100bp molecular size standard (Invitrogen, USA) were applied to the agarose gel 2% and visualized under ultraviolet light (UV).

INNO-LIPA HPV Genotyping

Beta-globin positive FFPE samples were genotyped by the Inno-Lipa HPV kit (Innogenetics, Gent, Belgium), capable of identifying 28 types of HPV (18 high-risk HPVs - HR-HPV and 10 low-risk - LR-HPV). This test is based on the amplification of the target DNA by PCR using primers from a consensus region spanning a short amplicon from the L1 gene for the evaluated HPVs and reverse hybridization of the amplified products into strips containing specific probes therefore (Kleter et al., 1999).

HPV16 Real Time PCR

Genotyping of all HPV positive samples (by PGMY PCR or INNO-LIPA) was performed by qPCR-specific HPV16 as a confirmative method, especially for those positive in the generic PGMY PCR and those from INNO-LIPA test with multiple HPV infections. Real time PCR was performed with the TaqMan Universal PCR Master Mix 2x (Applied Biosystems, Inc., USA) with a primer pair specific to the HPV16 E7 region, as described by Walboomers et al. (1999) in addition to a specific probe for this region as described by Scapulatempo-Neto et al. (2017). This protocol has already been validated in our laboratory presenting high sensitivity and specificity.

Positive controls consisted of DNA extracted from SiHa (1 copie of HPV16 per cell) and CaSki cell lines (600 copies of HPV 16 per cell) and water as negative control. Real-time PCR reactions were performed on the ABI Prism® 7300 HT SDS (Applied Biosystems, Inc., USA) and analysis of the amplification curves was performed using its own software.

Results

HNSCC Casuistic

From the 510 FFPE samples, more than 95% (491) were considered adequate due to proper beta-globin PCR amplification, and were, then, tested by the Inno-Lipa HPV kit. From the 491 samples genotyped, 49 were positive for HPV (10%), being 38 samples exclusively HPV16 positive (78% of the HPV positive samples), four samples were positive for HPV11 (8% of the positive samples), two HPV66 samples (4% of the positive samples) and five samples harbored multiple infections (10% of the positive samples) (Figure 1).

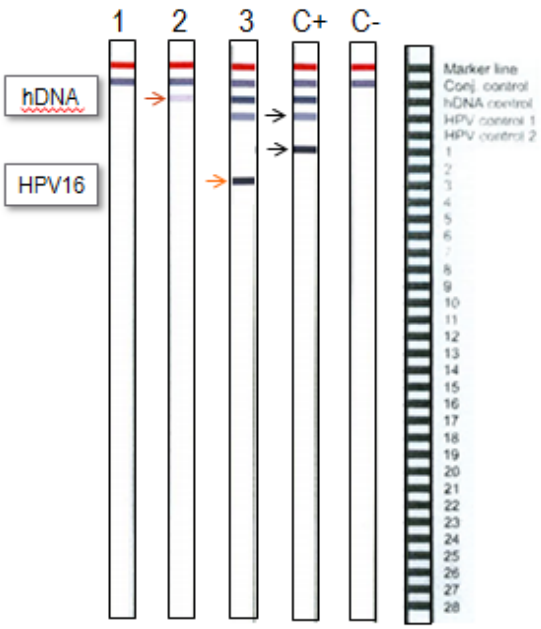

Figure 1. HPV genotyping by INNO-LiPA Genotyping Extra (Innogenetics), based on reverse hybridization with specific probes. In this image are the results obtained with the FFPE samples: Strips 1 and 2 represents, respectively, negative samples without and with the control DNA presence (red arrow). Strip 3, followed by the positive and negative controls (C + and C-). The orange arrow shows the probe result of a positive sample for HPV 16 (strip 3), and finally, the strip used as a guide for annotation.

For the 369 FFPE head and neck tumors of the Hospital das Clínicas of the University of São Paulo School of Medicine included in this study, almost all of them (345) had clinical and epidemiological data, that were analyzed for this population (Table 1).

This series consists mainly of men (almost 90%), aging between 27 to 90 years (average: 59; SD: 11), and 92% of the tumors were squamous cell carcinomas. The

HPV positivity rate was 7.1% of head and neck tumors (25/350 samples), of which 80% were HPV16, followed by HPV11, HPV18/39 and HPV31/33/44/52. When stratified by anatomical site (Table 1), the prevalence of HPV ranges from 3.4% (tongue and hypopharynx) to 25% (oropharyngeal cancer - all HPV16).

Table 1. HPV-DNA frequency and genotype by anatomical site (n = 369 samples, Hospital das Clínicas de São Paulo)

Anatomical site	n. of Patients	Genotyped samples	HPV+	HPV+ %	HPV Type
Base of tongue	29	29	1	3.4	18/39 (1)
Oropharynx	24	24	6	25.0	16 (6)
Oral cavity	115	108	7	6.5	16 (5), 11 (2)
Hypopharynx	30	29	1	3.4	11 (1)
Larynx	143	134	9	6.7	16 (7), 16/56, 31/33/44/52
ND	28	26	1	3.8	16 (1)
TOTAL	369	350	25	**7.1%**	

ND: not determined; n.: number; HPV+: positive samples

Non Neoplasic Oropharynx Casuistic

The study included 113 patients from the two participating hospitals. For the analysis, one aliquot of right and left tonsils of each patient, totaling 226 samples, were evaluated. This casuistic consists of 47 women aged from 18 to 58 years (mean age 28.5 years, SD 9.5, median age 27 years) and 66 men between 18 and 68 years (mean age 34 years, SD 10.5, median 33 years).

All 226 samples were positive for the beta-globin amplicon, however, no samples showed the presence of HPV-DNA either by the generic PCR with PGMY09/11 primers or the real-time HPV16 specific assay. All precautions and tests to ensure the reliability of the results were performed, such as the evaluation of the eventual presence of inhibitors in the samples and determination of the detection limit of the method from a known standard. The analytical sensitivity of the reaction was determined by Probit analysis, establishing the limit of detection (95%), which was

0.0002 copies per cell (using 4.77µL DNA extracted from CaSki cells - 55.3ng/µL - to 95.3µL of the negative tonsil DNA - 50ng/µL – in a total volume of 100µL) by subsequent serial dilutions until $3x10^{-7}$ copies of HPV16 per cell.

Discussion

This study presented a cohort with a significant number of patients with HNSCC. Our results demonstrate that the frequency of high-risk HPV types in HNSCC was 7.1%, that is a low when compared to worldwide data, despites the divergent results reported (Kreimer et al., 2005, Marur et al., 2010, Ndiaye et al., 2014, Anantharaman et al., 2017, de Abreu et al., 2018).

According to the anatomical site, in this study, the HPV prevalence ranges from 3.4% (base of tongue and hypopharynx) to 25% (oropharyngeal cancer). This variability occurs due to the tumors location, being the oropharyngeal known to be more associated with HPV than the others sites.

Studies in the US suggest that the majority of oropharyngeal squamous cell carcinoma (OPSCC) are nowadays caused by HPV16, as reported by Chaturvedi et al. (2008, 2013) and Nasman et al. (2009), although frequencies of <10% have been reported in the few studies conducetd in South America, by Herrero et al. (2003), Ribeiro et al. (2011) and López et al. (2014) with European estimates being in between, as reported by Reimers et al. (2007), Smeets et al. (2007) and Hoffmann et al. (2010). (Table 2).

In contrast, Betiol et al. (2016) evaluated HNSCC samples displaying an overall HPV prevalence of 19.4%, being highest in the oral cavity as compared to the other anatomical sites, including the oropharyngeal. The authors concluded that these results are consistent with other studies conducted in São Paulo and Rio de Janeiro in which HPV prevalence observed was 19.2% (22/114) and 15.5% (11/71) in oral SCC, respectively (Hauck et al., 2015; Kaminagakura et al., 2012), but is otherwise much higher than described in a large study involving 132 oral tumor samples from four countries in Latin America including Brazil (HPV prevalence was 0.0%) (Ribeiro et al., 2011). HPV-16 was the most commonly detected viral type, identified in 69.5% (25/36) of the HPV positive specimens. Nevertheless, for all anatomical sites, Betiol et al. has detected lower HPV DNA prevalence when compared to studies conducted in the USA and Europe (over 30%).

Table 2. Prevalence of HPV in non-oropharyngeal and oropharyngeal SCC studies

	Author and year	Country or Continent	n. of cases	HPV %	Methods	Date	Sites
Non-OPSCC	Present study	Brazil (SP)	510	10	Inno Lipa DNA	2002-2010	HNSCC FFPE
	de Abreu et al., 2018	Brazil (ES)	101 OC	3.3	MY09/MY11 and GP5+/GP6+ as well as PGMY09/11 L1 + Sequencing	2012-2015	OC
	Anantharaman et al., 2017	Brazil (SP) US Europe	233 234 420	0 6.8 5.5	HPV16 DNA and p16INK4a overexpression	2002-2015 2002-2006 ND	OC, L, H
	Castellsagué et al., 2016	29 countries	1374 P 1264 OC 1042 L	22.4 º(18,5) 4.4 º(3.0) 3.5 º(1.5)	HPV-DNA, HPVE6*mRNA or p16(INK4a) º(HPV-DNA and HPVE6*mRNA and p16(INK4a))	>1990	P, OC, L
	Betiol et al., 2016	Brazil (SP)	92 OC 66 L	25.0 13.7	Inno Lipa DNA	1991-2010 2009-2012	HNSCC FFPE
	López et al., 2014	Brazil (SP)	398	8,8	Inno Lipa DNA and PGMY09/11	1998-2008	HNSCC FFPE
	Ndiaye et al., 2014	44 countries	12163	29.5	Systematic review and meta-analysis	2004-2012*	OC, L, OPSCC
	Ribeiro et al., 2011	CE and LA	196	3.1	PGMY09/11 followed by HPV16 PCR	1998-2003	HNSCC Frozen tumor
	Kaminagakura et al., 2012	Brazil (SP)	114	19.2	GP5+/GP6+ PCR	1970-2006	OC SCC FFPE
OPSCC	Present study	Brazil (SP)	24	25	Inno Lipa DNA and HPV16 RT-PCR	2002-2010	
	Barros-Filho et al., 2018	Brazil (SP)	40	58	Linear array	< 2011	
	Anantharaman et al., 2017	Brazil (SP) US Europe	171 243 119	4.1 59,3 31.1	HPV16 DNA and p16INK4a overexpression	2002-2015 2002-2006 ND	
	Castellsagué et al., 2016	29 countries	1090	22.4 18.5	HPV-DNA, HPVE6*mRNA or p16(INK4a) HPV-DNA and HPVE6*mRNA and p16(INK4a)	>1990	
	Betiol et al., 2016	Brazil (SP)	28	14.3	Inno Lipa DNA	1991-2012	
	López et al., 2014	Brazil (SP)	91	6.6	Inno Lipa DNA and PGMY09/11	1998-2008	
	Ndiaye et al., 2014	44 countries	3946	45.8	Systematic review and meta-analysis	2004-2012*	
	Ribeiro et al., 2011	Brazil (SP)	68	4.4	PGMY09/11 followed by HPV16 PCR	1998-2003	

CE - Central Europe; LA - Latin America; SP - São Paulo; ES - Espirito Santo; OC - Oral Cavity; P- Pharynx; L- Larynx; H – Hypopharynx; º- results from "end" group; * publish data

Further, the role of HPV16 in HNC outside the oropharynx remains unclear (Ndiaye et al., 2014). A recent review has estimated that the probability of an HPV-attributable cancer of the oral cavity, larynx and hypopharynx could be up to five times lower than that of oropharyngeal cancer (Combes and Franceschi, 2014). Whether these divergent geographic results represent important differences in the etiology of HNC or whether they are explained by differences in laboratory practices is unclear. The recent publication on global HPV prevalence in HNC is of note where, based on a comparison of over 3,000 tumors tested for HPV DNA and subsequently, following triage for HPV16 E6*mRNA and $p16^{INK4a}$, the authors report that nearly 22% of OPSCC could be attributed to HPV infection, while fewer than 5% of oral cavity and laryngeal cancers were HPV-positive (Castellsagué et al., 2016). This study also reported strongly divergent results from South and Central America (OPSCC HPV prevalence of 37%), and did not provide estimates for North America. Although the largest study to date, this report was unable to explore lifestyle factors associated with HPV-positivity.

Anantharaman et al. (2017), in a study of 1,420 HNCs from three world regions demonstrated dramatic differences in the prevalence of HPV16-positive HNCs. While nearly 60% of OPSCCs in the US are HPV16-positive, this proportion is only about 4% in Brazil, and OPSCCs in Europe have intermediate HPV16 prevalence (31%). Less than 4% of laryngeal and oral cavity cancers were HPV16-positive. That similar results were noted for the marginal incident fractions of HPV16-positive OPSCC and non-OPSCC lends support to these conclusions.

In a recent study from Brazilian oral cancer, de Abreu et al. (2018) demonstrate that the frequency of high-risk HPV types in oral cavity SCC is very low and is less than 4%. Furthermore, the frequency of HPV in their study is close to the HPV DNA rates found in healthy individuals.

The natural history of HPV infection has been extensively investigated in epidemiologic studies by PCR-based methods, HPV serology and DNA/RNA in situ hybridization (Lopes et al., 2011; Lee et al., 2012). A review about epidemiological investigation on oral HPV prevalence in healthy individuals, published by Shigeishi & Sugiyama (2016), reported that HPV frequency in saliva of healthy individuals have shown low and variable rates in a period of time, which is related to each patient's

immune response and can, therefore, be inconstant. In addition, rates of oncogenic HPV infection in the oral cavity of healthy people are also known to be low (around 2%) and the natural history of HPV in this anatomical site suggests HPV acquisition and persistence to be rare events compared to genital or anal infections (Kreimer et al., 2013; Shigeishi & Sugiyama, 2016).

Regardless to the absence of HPV infection in normal tissue from the tonsils reported on this study, Combes et al. (2017) used extensive ex vivo brushing and gargling from patients undergoing tonsillectomy for benign indications to evaluate human papillomavirus (HPV) prevalence in the tonsil in a large age-stratified sample of cancer-free patients (from 2012 to 2016) in France – a high prevalence OPSCC HPV-related region. The HPV-DNA detection used Luminex to evaluate 21 HPV types and were applied to 692 patients aged 1–70 years and gargles from 268 adults. Overall HPV prevalence in adults was 3.6% in tonsil brushings and 13.1% in gargles and HPV16 prevalence was 2.2% and 4.1%, respectively. Tonsil brushings were also positive in two girls from 139 children (1.4%). HPV prevalence in gargles significantly varied by sex (prevalence ratio in men vs women = 2.1). They related a low agreement (9.5%) in paired tonsil brushings and gargles in adults, suggesting that gargle is not representative of HPV prevalence in the tonsil.

In this study, as we are in a low prevalence HPV-associated HNSCC, we expected an even lower HPV rate on normal tissue, besides the use of surgical samples (uvulopalatopharyngoplasty and tonsillectomy).

Conclusion

Regarding paraffin embedded tissue samples, this prevalence corresponds to 10% of HPV positive head and neck samples (49 HPV + / 491 samples). Despite being a low prevalence when compared to countries such as USA and Europe, which exceed 60%, for Brazilian studies, it is one of the highest rates ever reported (Table 2), being the HPV16 the most representative HPV type (78% of the positive samples), as widely verified.

The overall prevalence of HPV-DNA in HNSCC (non-oropharyngeal tumors) ranges from 3.4% to 6.7% due to the diversity of anatomical sites of head and neck carcinomas. In the oropharyngeal SCC, the HPV prevalence reaches 25% of this anatomical site samples from Hospital das Clínicas in São Paulo, all of them HPV16, confirming its high prevalence among the tumors of this site associated with human papillomavirus. Besides our OPSCC HPV-related rate remains lower when compared to high incidence countries, like the US (Chaturvedi et al. 2011), in the past years, we saw a progressive trend in HPV positive tumors, going from 4.4% (Ribeiro et al., 2011), 6.6% (López et al., 2014), 14.3% (Betiol et al., 2016) to 25% in the present study, including a higher incidence report of 58% (Barros-Filho et al., 2018) in a young population, with HPV16 playing the major role in the carcinogenesis of these tumors.

Divergence in HPV DNA rates observed could be attributed not only to differences in the methodologies used for HPV detection in the different studies, but also to the characteristics of the individuals under study, including sexual practices, economic and social status. Unfortunately, the absence of sociodemographic and sexual behavior information of the patients evaluated here prevents an additional analysis, that could also elucidate the absence of HR-HPV in our population.

In conclusion, in the current study we observed an increasing frequency of HPV DNA in HNSCC (especially the oropharyngeal site), supporting the evidence that HPV is involved in the development of a growing subset of HNSCC in the Brazilian population. Studies involving the evaluation of protein expression associated with HPV infection should be conducted to reinforce whether this increasing profile of HPV DNA also reflects an increasing etiological involvement of HPV 16 in the new cases of oropharyngeal squamous cell carcinomas in Brazil.

References

ANANTHARAMAN D, ABEDI-ARDEKANI B, BEACHLER DC. Geographic heterogeneity in the prevalence of human papillomavirus in head and neck cancer. Int. J. Cancer: 2017;140, 1968–1975.

ANG KK, HARRIS J, WHEELER R, et al. Human papillomavirus and survival of patients with oropharyngeal cancer. N Engl J Med. 2010;363(1):24-35.

BARROS-FILHO MC, REIS-ROSA LA, HATAKEYAMA M, et al. Oncogenic drivers in 11q13 associated with prognosis and response to therapy in advanced oropharyngeal carcinomas". Oral Oncology. 2018:83;81–90

BETIOL JC, SICHERO L, COSTA HOO, et al. Prevalence of human papillomavirus types and variants and p16(INK4a) expression in head and neck squamous cells carcinomas in São Paulo, Brazil. Infect Agent Cancer. 2016;11:20.

CALIFANO J, VAN DER RIET P, WESTRA W, et al. Genetic progression model for head and neck cancer: implications for field cancerization. Cancer Res. 1996;56:2488-92.

CARVALHO AL, IKEDA MK, MAGRIN J, KOWALSKI LP. Trends of oral and oropharyngeal cancer survival over five decades in 3267 patients treated in a single institution. Oral Oncol 2004;40:71-6.

CASTELLSAGUÉ X, ALEMANY L, QUER M et al. HPV involvement in head and neck cancers: comprehensive assessment of biomarkers in 3680 patients. J Natl Cancer Inst. 2016;108(6): djv403.

CHATURVEDI AK, ANDERSON WF, LORTET-TIEULENT J, et al. Worldwide trends in incidence rates for oral cavity and oropharyngeal cancers. J Clin Oncol 2013;31:4550–9.

CHATURVEDI AK, ENGELS EA, ANDERSON WF, et al. Incidence trends for human papillomavirus-related and -unrelated oral squamous cell carcinomas in the United States. J Clin Oncol 2008;26:612–9.

CHATURVEDI AK, ENGELS EA, PFEIFFER RM, et al. Human papillomavirus and rising oropharyngeal cancer incidence in the United States. J Clin Oncol. 2011;29:4294-4301.

CHIANG YY, TSAI MH, LIN TY, CHIANG IP. Expression profile of metastasis-related genes in invasive oral cancers. Histol Histopathol 2008;23:1213-22.

COMBES JD, DALSTEIN V, GHEIT T, et al. Prevalence of human papillomavirus in tonsil brushings and gargles in cancer-free patients: The SPLIT study. Oral Oncol. 2017;66:52–57.

COMBES JD, FRANCESCHI S. Role of human papillomavirus in non-oropharyngeal head and neck cancers. Oral Oncol 2014;50:370–9.

D'SOUZA G, AGRAWAL Y, HALPERN J. Oral sexual behaviors associated with prevalent oral human papillomavirus infection. J Infect Dis. 2009;199:1263-1269.

D'SOUZA G, ANANTHARAMAN D, GHEIT T, et al. Effect of HPV on head and neck cancer patient survival, by region and tumor site: A comparison of 1362 cases across three continents. Oral Oncol. 2016;62:20-27.

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.

de ABREU PM, CÓ ACG, AZEVEDO PL, et al. Frequency of HPV in oral cavity squamous cell carcinoma. BMC Cancer.2018: 18:324.

de GRAEFF A, DE LEEUW JR, ROS WJ, *et al*. Sociodemographic factors and quality of life as prognostic indicators in head and neck cancer. Eur J Cancer 2001;37:332-9.

de MARTEL C, FERLAY J, FRANCESCHI S, et al. Global burden of cancers attributable to infections in 2008: a review and synthetic analysis. Lancet Oncol. 2012 Jun;13:607-15.

DIKSHIT RP, BOFFETTA P, BOUCHARDY C*, et al.* Lifestyle habits as prognostic factors in survival of laryngeal and hypopharyngeal cancer: a multicentric European study. Int J Cancer 2005;117:992-5.

do SACRAMENTO PR, BABETO E, COLOMBO J, Ruback MJC, Bonilha JL, Fernandes AM, et al. The prevalence of human papillomavirus in the oropharynx in healthy individuals in a Brazilian population. Journal of Medical Virology. 2006;78(5):614-8.

DUNNE EF, NIELSON CM, STONE KM. Prevalence of HPV infection among men: A systematic review of the literature. J Infect Dis. 2006;194:1044-57.

DUVVURI U, MYERS JN. Contemporary management of oropharyngeal cancer: anatomy and physiology of the oropharynx. Curr Probl Surg. 2009;46:119-84.

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.

FORASTIERE A, KOCH W, TROTTI A, et al. Head and neck cancer. N Engl J Med. 2001;345:1890-900.

GRAVITT PE, PEYTON CL, ALESSI TQ, Wheeler CM, Coutlee F, Hildesheim A, et al. Improved amplification of genital human papillomaviruses. J Clin Microbiol. 2000;38(1):357-61.

HA PK, CALIFANO JA. Promoter methylation and inactivation of tumour-suppressor genes in oral squamous-cell carcinoma. Lancet Oncol. 2006;7:77-82.

HANAHAN D, WEINBERG RA. Hallmarks of cancer: the next generation. Cell. 2011;144(5):646-74.

HAUCK F, OLIVEIRA-SILVA M, DREYER JH, et al. Prevalence of HPV infection in head and neck carcinomas shows geographical variability: a comparative study from Brazil and Germany. Virchows Arch. 2015;466:685–93.

HERRERO R, CASTELLSAGUE X, PAWLITA M, et al. Human papillomavirus and oral cancer: thevInternational Agency for Research on Cancer multicenter study. J Natl Cancer Inst 2003;95: 1772–83.

HOFFMANN M, IHLOFF AS, GOROGH T, et al. p16(INK4a) overexpression predicts translational active human papillomavirus infection in tonsillar cancer. Int J Cancer 2010;127:1595–602.

HOPKINS J, CESCON DW, TSE D, *et al*. Genetic polymorphisms and head and neck cancer outcomes: a review. Cancer Epidemiol Biomarkers Prev. 2008;17:490-9.

INCA, Estimativa 2018: incidência de câncer no Brasil / Instituto Nacional de Câncer José Alencar Gomes da Silva. Coordenação de Prevenção e Vigilância. – Rio de Janeiro: INCA, 2017. [Available from: http://www.inca.gov.br/estimativa/2018/estimativa-2018.pdf]

JEMAL A, BRAY F, CENTER MM, *et al*. Global Cancer Statistics. Ca Cancer J Clin 2011;61:69-90.

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.

KLETER B, VAN DOORN LJ, SCHRAUWEN L, et al. Development and clinical evaluation of a highly sensitive PCR-reverse hybridization line probe assay for detection and identification of anogenital human papillomavirus. 1999;J Clin Microbiol 37:2508–2517

KLINGENBERG B, HAFKAMP HC, HAESEVOETS A, et al. p16 INK4A overexpression is frequently detected in tumour-free tonsil tissue without association with HPV. Histopathology. 2010;56(7):957-67.

KREIMER AR, BHATIA RK, MESSEGUER AL, et al. Oral human papillomavirus in healthy individuals: a systematic review of the literature. Sex Transm Dis. 2010;37(6):386-91.

KREIMER AR, CLIFFORD GM, BOYLE P, et al. Human papillomavirus types in head and neck squamous cell carcinomas worldwide: a systematic review. Cancer Epidemiol Biomarkers Prev. 2005;14:467-75.

KREIMER AR, PIERCE CAMPBELL CM, LIN H-Y, et al. Incidence and clearance of oral human papillomavirus infection in men: the HIM cohort study. Lancet. 2013;382:877–87.

LASSEN P, ERIKSEN JG, HAMILTON-DUTOIT S, et al. Effect of HPV-associated p16INK4A expression on response to radiotherapy and survival in squamous cell carcinoma of the head and neck. J Clin Oncol. 2009;27:1992–8.

LEE LA, HUANG CG, LIAO CT, et al. Human papillomavirus-16 infection in advanced oral cavity cancer patients is related to an increased risk of distant metastases and poor survival. PLoS One. 2012;7:7–9.

LOPES V, MURRAY P, WILLIAMS H, et al. Squamous cell carcinoma of the oral cavity rarely harbours oncogenic human papillomavirus. Oral Oncol. 2011;47:698–701.

LÓPEZ RVM, LEVI JE, ELUF-NETO J, *et al.* Human papillomavirus (HPV) 16 and the prognosis of head and neck cancer in a geographical region with a low prevalence of HPV infection. Cancer Causes Control. 2014; 25:461-71.

LÓPEZ RVM, ZAGO MA, ELUF-NETO J, *et al.* Education, tobacco smoking, alcohol consumption, and IL-2 and IL-6 gene polymorphisms in the survival of head and neck cancer. Braz J Med Biol Res. 2011;44:1006-12.

MARUR S, D'SOUZA G, WESTRA WH, et al. HPV- associated head and neck cancer: a virus-related cancer epidemic. Lancet Oncol 2010;11:781–9.

MUÑOZ N, BOSCH FX, DE SANJOSÉ S, *et al*. Epidemiologic classification of human papillomavirus types associated with cervical cancer. International Agency for Research on Cancer Multicenter Cervical Cancer Study Group. N Engl J Med 2003;348:518-27.

NASMAN A, ATTNER P, HAMMARSTEDT L, et al. Incidence of human papillomavirus (HPV) positive tonsillar carcinoma in Stockholm, Sweden: an epidemic of viral-induced carcinoma? Int J Cancer 2009;125:362–6.

NDIAYE C, MENA M, ALEMANY L, *et al*. HPV DNA, E6/E7 mRNA, and p16INK4a detection in head and neck cancers: a systematic review and meta-analysis. Lancet Oncol. 2014;15:1319-31.

PANNONE G; SANTORO A; PAPAGERAKIS S; *et al.* The role of Human Papillomavirus in the pathogenesis of Head & Neck Squamous Cell Carcinoma: an overview; Infectious Agents and Cancer 2011, 6:4 doi:10.1186/1750-9378-6-4

PERDOMO S, ANANTHARAMAN D, FOLL M, et al. Genomic analysis of head and neck cancer cases from two high incidence regions. PLoS ONE. 2018;13(1): e0191701.

REIMERS N, KASPER HU, WEISSENBORN SJ, et al. Combined analysis of HPV-DNA, p16 and EGFR expression to predict prognosis in oropharyngeal cancer. Int J Cancer 2007;120:1731–8.

RIBEIRO KB, LEVI JE, PAWLITA M, et al. Low human papillomavirus prevalence in head and neck can- cer: results from two large case-control studies in high-incidence regions. Int J Epidemiol 2011;40: 489–502.

RIETH KKS, GILL SR, LOTT-LIMBACH AA, et al. Prevalence of High-Risk Human Papillomavirus in Tonsil Tissue in Healthy Adults and Colocalization in Biofilm of Tonsillar Crypts. JAMA Otolaryngol Head Neck Surg. 2018;144(3):231-7.

SAIKI RK, SCHARF S, FALOONA F, et al. Enzymatic amplification of beta‐globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. Science. 1985;230:1350-1354.

SCAPULATEMPO-NETO C, VEO C, FREGNANI JHTG, et al. Characterization of topoisomerase II alpha and minichromosome maintenance protein 2 expression in anal carcinoma. Oncol Lett. 2017;13(3):1891-1898.

SCHABATH MB, VILLA LL, LAZCANO-PONCE E. Smoking and Human Papillomavirus (HPV) Infection in the HPV in Men (HIM) Study. Cancer Epidemiol Biomarkers Prev. 2012;21:102-110.

SHIGEISHI H, SUGIYAMA M. Risk factors for oral human papillomavirus infection in healthy individuals: a systematic review and meta-analysis. J Clin Med Res. 2016;8:721–9.

SHIMIZU Y, KONDO S, SHIRAI A, *et al*. A single nucleotide polymorphism in the matrix metalloproteinase-1 and interleukin-8 gene promoter predicts poor prognosis in tongue cancer. Auris Nasus Larynx. 2008;35:381-9.

SHUKLA S, BHARTI AC, MAHATA S, *et al*. Infection of human papillomaviruses in cancers of different human organ sites. Indian J Med Res. 2009;130(3):222-33.

SMEETS SJ, HESSELINK AT, SPEEL EJ, et al. A novel algorithm for reliable detection of human papillomavirus in paraffin embedded head and neck cancer specimen. Int J Cancer 2007;121:2465–72.

TURNER DO, WILLIAMS-COCKS SJ, BULLEN R, *et al.* High-risk human papillomavirus (HPV) screening and detection in healthy patient saliva samples: a pilot study. BMC Oral Health. 2011;11:28.

WALBOOMERS JM, JACOBS MV, MANOS MM *et al*. Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. J Pathol. 1999;189:12-9.

APÊNDICE B - *"Genomic analysis of head and neck cancer cases from two high incidence regions"*

Perdomo S, Anantharaman D, Foll M, Abedi-Ardekani B, Durand G, ***Reis Rosa LA****, et al. (2018). PLoS ONE 13(1): e0191701.*

PLOS | ONE

RESEARCH ARTICLE

Genomic analysis of head and neck cancer cases from two high incidence regions

Sandra Perdomo[1,2], Devasena Anantharaman[1¤a], Matthieu Foll[1], Behnoush Abedi-Ardekani[1], Geoffroy Durand[1], Luciana Albina Reis Rosa[3], Reetta Holmila[1¤b], Florence Le Calvez-Kelm[1], Eloiza H. Tajara[4], Victor Wünsch-Filho[5], José Eduardo Levi[3], Marta Vilensky[6], Jerry Polesel[7], Ivana Holcatova[8], Lorenzo Simonato[9], Cristina Canova[9], Pagona Lagiou[10], James D. McKay[1], Paul Brennan[1]*

1 International Agency for Research on Cancer (IARC), Lyon, France, **2** Institute of Nutrition, Genetics and Metabolism Research, Faculty of Medicine, Universidad El Bosque, Bogotá, Colombia, **3** Instituto de Medicina Tropical de SP Universidade de São Paulo- USP, São Paulo, Brazil, **4** School of Medicine of São José do Rio Preto, São José do Rio Preto, Brazil, **5** Faculdade de Saúde Pública, Universidade de São Paulo, São Paulo, Brazil, **6** Instituto Angel Roffo, Buenos Aires, Argentina, **7** Centro di Riferimento Oncologico (CRO), Aviano National Cancer Institute, Aviano, Italy, **8** Charles University of Prague, Prague, Czech Republic, **9** Laboratory of Public Health and Population Studies, Padova, Italy, **10** University of Athens Medical School, Athens, Greece

¤a Current address: Rajiv Gandhi Centre for Biotechnology, Thycaud PO, Trivandrum, India
¤b Current address: Molecular Medicine section, Wake Forest School of Medicine, Winston Salem, North Carolina, United States of America
* gep@iarc.fr

OPEN ACCESS

Citation: Perdomo S, Anantharaman D, Foll M, Abedi-Ardekani B, Durand G, Reis Rosa LA, et al. (2018) Genomic analysis of head and neck cancer cases from two high incidence regions. PLoS ONE 13(1): e0191701. https://doi.org/10.1371/journal.pone.0191701

Editor: Scott M. Langevin, University of Cincinnati College of Medicine, UNITED STATES

Received: September 14, 2017

Accepted: December 4, 2017

Published: January 29, 2018

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This work has been funded by the International Agency for Research on Cancer (IARC). The GENCAPO study was supported by grants from FAPESP (São Paulo Research Foundation) (04/12054-9 and 10/51168-0), which had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Abstract

We investigated how somatic changes in HNSCC interact with environmental and host risk factors and whether they influence the risk of HNSCC occurrence and outcome. 180-paired samples diagnosed as HNSCC in two high incidence regions of Europe and South America underwent targeted sequencing (14 genes) and evaluation of copy number alterations (SCNAs). *TP53*, *PIK3CA*, *NOTCH1*, *TP63* and *CDKN2A* were the most frequently mutated genes. Cases were characterized by a low copy number burden with recurrent focal amplification in 11q13.3 and deletion in 15q22. Cases with low SCNAs showed an improved overall survival. We found significant correlations with decreased overall survival between focal amplified regions 4p16, 10q22 and 22q11, and losses in 12p12, 15q14 and 15q22. The mutational landscape in our cases showed an association to both environmental exposures and clinical characteristics. We confirmed that somatic copy number alterations are an important predictor of HNSCC overall survival.

Introduction

Head and neck squamous cell carcinomas (HNSCC) constitute a heterogeneous group of cancers, which include cancers arising at the oral cavity, nasopharynx, oropharynx, hypopharynx, and larynx. Collectively, these cancers are the seventh most common malignancy diagnosed worldwide [1], with areas of high incidence including Mediterranean Europe and South

Competing interests: The authors declare that no competing interests exist.

America [2]. Despite current therapeutic approaches, the prognosis is quite poor, with a 5-year survival ranging from approximately 25% to 60%, according to cancer subsite [3].

Cigarette smoking and alcohol abuse are the major risk factors, consistently associated with the incidence of head and neck cancers [4]. Additionally, human papillomavirus (HPV) infection is strongly associated with oropharyngeal cancer risk and prognosis, alongside a small number of other HNSCC [5]. Recent studies have highlighted the association between numerous differential genomic features and these exposures as well as clinical factors, providing insights for potentially improving prognostic risk stratification for HNSCC[6, 7]. The Cancer Genome Atlas TCGA has conducted the largest comprehensive genomic study of 528 HNSCC cases, consisting of an integrative analysis of multi-genomic data including somatic mutations, gene expression, methylation and miRNAs expression in a clinically and pathologically characterized dataset. The complete data analysis of a subset of 279 patients has allowed the description of the landscape of somatic genomic alterations and the identification of the principal molecular pathways involved in HNSCC development. Particularly, HNSCC are characterized by mutation of *TP53*, whole genome duplications and multiple recurrent chromosomal gains and losses associated to increased genomic disruption affecting cell cycle checkpoints and PI3K-AKT signaling[8, 9]. Increased rates of somatic copy number alterations (SCNAs) across the tumour genome are associated with poor prognosis and therefore it becomes important to identify SCNAs that might be functionally driving progression and outcome. In addition, genomic studies have revealed how differential genomic patterns among cases could identify various subgroups of tumours showing specific associations with histological subtypes, smoking, HPV status and overall survival[6, 10]

The principal objective of this study was to investigate whether somatic genetic changes identified in two large comprehensive case series in Europe and South America could influence the risk of HNSCC occurrence and outcome from those areas. A second objective was to investigate how somatic changes interact with environmental and host risk factors including HPV infection, alcohol and smoking. We selected 180 paired samples diagnosed as HNSCC from three multicentre studies representative of high incidence regions in Europe (ARCAGE study), Brazil (GENCAPO study) and Argentina (LA study); from which both tumour and blood samples were available in the IARC biorepository, along with complete epidemiological data.

Materials and methods

Study population and risk factor data collection

A total of 240 HNSCC cases were selected from three multicentre studies: two conducted in South America (LA study) between 1998 and 2002, and (GENCAPO study) between 1998 and 2008; and one completed in Europe (ARCAGE study) between 2002 and 2005. Selection of cases was based on availability for biological samples along with complete epidemiological and clinical data. However, no treatment information was obtained from most of these cases as this variable was not included in the original protocols. Extensive details of the three-large multicentre case-control studies are included elsewhere [11–13]. Briefly, all subjects underwent personal interviews to collect information on lifestyle exposures and hospital records were reviewed to obtain clinical and pathological information. All cases had biological samples collected at diagnosis and before any treatment [11–13]. Centralized HPV testing was completed for the three participating studies determined on serology testing as described before [14]. HPV positivity was defined based on HPV16 E6 status, which has been shown to be a highly sensitive and specific marker of HPV16-related oropharyngeal tumours [15–17]. Immunohistochemical evaluation of $P16^{INK4a}$ expression and HPV DNA genotyping were also completed

for a subset of samples using protocols previously described [12, 14], and these data were also used to confirm HPV status.

Informed consent was obtained from all participants in the three studies, and the analysis was approved by the Ethical Review Committee of the International Agency for Research on Cancer. All experiments were performed in accordance with relevant guidelines and regulations.

Targeted sequencing

A customized gene panel of 14 genes (GeneRead DNAseq Custom Panels, Qiagen®) was used for targeted sequencing of tumour-blood pair cases. Gene selection was based on an independent analysis of TCGA data on HNSCC using MutsigCV algorithm complemented with the list of the most frequently mutated genes reported in the literature. Briefly, 20ng of DNA were used in multiplex PCR reactions using Qiagen® recommended protocol. For library preparation, 100 ng of multiplex pools and the NEBNext End Repair Module (New England Biolabs, Ipswich, MA, USA) following manufacturer's instructions. Individual barcodes (designed in-house and produced by Eurofins MWG Operon, Ebersberg, Germany) were ligated to each multiplex pool for sequencing. Both tumour and blood samples were sequenced at an average depth of 250X and 50X respectively using the PGM/PROTON™ Systems (Life Technologies, Carlsbad, CA, USA); sequences used for mutational calling had on target sequencing of 85%, and uniformity of 80–85%

Mutational calling

Identification of somatic variants was performed using a recently developed statistical model called Needlestack[18] based on the idea that analysing several samples together can help estimate the distribution of sequencing errors to accurately identify variants. At each position and for each candidate variant, we model sequencing errors using a robust Negative-Binomial regression with a linear link and a zero intercept [19]. We calculate for each sample a p-value for being a variant (outlier from the regression) that we further transform into q-values to account for multiple testing. Needlestack has a detection limit of variant allelic fractions between 0.05% and 0.5% depending on the error rate at the base change considered (ranging from 0.001% to >10% at homopolymers) and the sequencing depth. Needlestack is free and open-source and is available publicly as a beta version under https://github.com/IARCbioinfo/needlestack. A detailed description of the *Needlestack* variant caller has been previously published [20, 21]. Variant calls were annotated using ANNOVAR [22] and indels, nonsense, splicing, or missense variants were only kept for subsequent analyses if reported in COSMIC-76 and/or classified as deleterious, disease causing or damaging in at least one of the five variant classification databases (SIFT, Polyphen, MutationTaster, MutationAssessor, FATHMM, LR) (S3 Table).

Filtering of VCF calls was done using a threshold of 0.5% allelic fraction, minimal read depth of 100X and minimal phred-scaled q-value of 30. Removal of germline variants was additionally confirmed by comparison of corresponding paired blood sequences; all filtered variants were manually curated by inspection of BAM files using the Integrative Genomics Viewer (IGV) 2.3 (Broad Institute, Cambridge, MA, USA).

Internal technical validation of both the sequencing procedure and the mutational calling was done by including 10% of samples as technical replicates in each library preparation. Additionally, an independent library preparation including a random selection of 20 tumour samples was sequenced and analysed independently and results were 100% concordant. All cases from the GENCAPO study had been previously sequenced for *TP53* mutations using Sanger

sequencing, which we used to further validate our mutational calls and compare them with a different calling method (GeneRead Panel Variant Calling analysis tool from Qiagen®) (S1 Fig).

Somatic copy number alterations (SCNAs)

DNA from each tumour was hybridized to Illumina HumanCytoSNP-12v2.1 arrays using standard manufacturer's protocol. Formalin-fixed paraffin-embedded (FFPE) samples underwent a quality control assay using the Illumina FFPE QC Kit, samples were selected based on a ΔCq below or equal to 2 and then restored using the Infinium HD FFPE Restore Protocol. We included 10% of technical and biological replicates for quality control and validation. Microarray data are available in the ArrayExpress database (www.ebi.ac.uk/arrayexpress) under accession number E-MTAB-4863. The R package crlmm [23] was used for pre-processing, genotyping and calculation of circular binary segmentation to estimate the normalized copy number. Germline copy number alterations were removed using the Database of Genomic Variants [24]. Identification of significant amplified or deleted regions was performed by using GISTIC 2.0 [25] using 99% confidence level and q-value threshold 0.25. Focal amplification or deletion for all the 14 genes sequenced was determined only using the GISTIC copy number value 2 or -2 respectively as the true value. OncoPrinter and MutationMapper tools were used for visualization of mutational data [26, 27]. Integrative cluster analysis of mutation and copy number data was performed using the R package iClusterPlus [28].

Statistical analysis

Mutual exclusion and co-occurrence test for mutations (including both single nucleotide variants and copy number alterations) found in the 14 genes evaluated, were based on weighted permutations assessing the deviation of the observed coverage compared to expected obtained by permuting events [29]. Fisher exact test was used to determine the relationship of clinical characteristics in the 3 studies. For each patient, time at risk was calculated from cancer diagnosis to death or end of follow up (Last Follow up date: 30/01/2013 for the ARCAGE study, 30/06/2009 for GENCAPO and 30/06/2006 for the LA study). Follow-up was censored at 5 years, given that most cancer related events occur before that time. The Kaplan-Meier estimator was used to estimate the distribution of the 5-year survival. Multivariate Cox proportional hazard models were used to estimate HRs and their corresponding p values for all candidate risk factors and genomic biomarkers. Age, subsite, stage, nodal status (defined by pathological nodal stage), smoking and alcohol status were used as covariates. A correction for multiple-hypothesis testing was employed using the method of Benjamini and Hochberg [30] Log-rank test was used to compare the different survival distributions.

Results

Epidemiological description of the three studies

A total of 180 cases had complete sequencing and copy number information (Fig 1). Clinical and pathologic characteristics of cases in the three studies are described in Table 1. Consistent with previous reports the majority of the cases were males (82%), current smokers (67%) and current drinkers (70%). Mean age at diagnosis was 59 years (range 18–88 years). Thirty-three-percent of all cases were diagnosed with oral cavity cancer, 25% with oropharyngeal cancer, 18% with laryngeal cancer, 7% with hypopharyngeal cancer and 16% with overlapping topographies. Seventy-percent of all cases presented advanced disease (stages III-IV). The majority of non-smokers (80%) and oropharyngeal cases (67%) were part of the European study

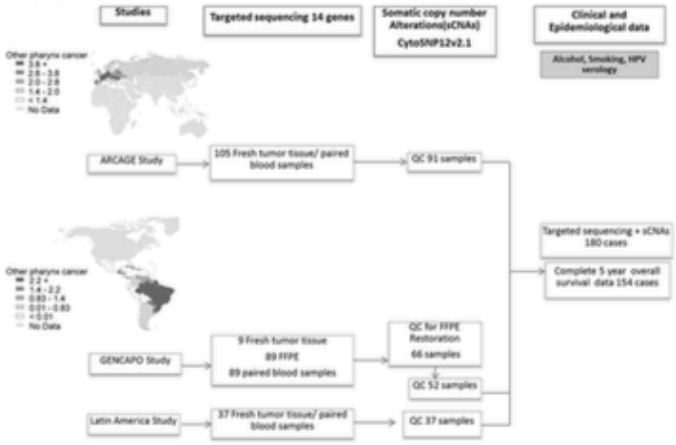

Fig 1. Workflow of processing and analysis of HNSCC samples from the three different studies. QC for copy number evaluation: Quality control of samples based on signal to noise ratio>5.0. Maps show estimated age-standardized incidence rates for HNSCC (other pharynx sites) in Europe and South America. [31].

https://doi.org/10.1371/journal.pone.0191701.g001

(ARCAGE). Fifteen cases out of 180 (8%) were classified as HPV16 positive, 73% of which were oropharyngeal cases.

Mutational profile of the 14 gene panel in cases

Ninety four-percent of all sequenced cases had at least one alteration (single nucleotide variants (SNVs) or amplification/deletion) in any of the 14 genes selected (Fig 2) (S3 Table). The overall frequency of alterations for the 14 genes was similar to previous publications with a higher enrichment of alterations in the *TP53*, *NOTCH1* and *CDKN2A* genes [10, 32–34]. Among the 10 cases without alteration in the 14 genes, 4 corresponded to HPV positive cases (S4 Fig).

TP53, *FAT1*, *MLL2* and *NOTCH1* were the genes more frequently altered by single nucleotide variants (SNVs) (Fig 2). As previously described [35, 36], *TP53* mutation was mostly prevalent in HPV negative tumours (only three out of 15 HPV16 positive tumours harboured a *TP53* mutation, and all three cases were current smokers) (S4 Fig). *TP53* mutations clustered predominantly in DNA binding domains, particularly in hotspot codons 175, 248, 249, 273 and 282 (S3 Fig). Forty-four-percent of all mutations were classified as disruptive mutations according to the definition by Poeta and colleagues[37]. Fifty-five-percent of all *TP53* SNVs were missense mutations and from those 64% were classified as high-risk mutations based on the evolutionary action score EAP53[38]. *FAT1*, *MLL2* and *NOTCH1* mutations (missense and truncating mutations) were distributed along the gene coding region and did not show mutational enrichment of specific protein domains (S3 Fig).

Mutual exclusive alterations were identified between genes with recognized activity in the same signalling pathway, suggesting overlapping functional consequences of those mutations. This included *TP53* and *PIK3CA* ($p<0.001$), both involved in cell cycle control and survival, and *NOTCH1* and *TP63* ($p = 0.003$) genes, which play important functions of squamous cell differentiation (S2 Fig).

Table 1. Clinical and epidemiological description of 180 HNSCC cases from the three studies.

	STUDY							
	ARCAGE (Czech Republic, Italy, Greece)		GENCAPO (Brazil)		LA (Argentina)		Total	
Sex*	n	%	n	%	n	%	n	%
Female	26	28.57	2	3.85	5	13.51	33	18.33
Male	65	71.43	50	96.15	32	86.49	147	81.67
Age group								
18 to 50	18	19.78	9	17.31	7	18.92	34	18.89
51 to 60	25	27.47	28	53.85	14	37.84	67	37.22
61 to 70	28	30.77	10	19.23	9	24.32	47	26.11
>70	20	21.98	5	9.62	7	18.92	32	17.78
Subsite*								
Oral cavity	32	35.16	14	26.92	13	35.14	59	32.78
Oropharynx	30	32.97	10	19.23	5	13.51	45	25
Hypopharynx	2	2.2	11	21.15	0	0	13	7.22
Larynx	17	18.68	3	5.77	12	32.43	32	17.78
Overlapping	10	10.99	12	23.08	7	18.92	29	16.11
No information	0	0	2	3.85	0	0	2	1.11
Stage*								
T1	8	8.79	1	1.92	0	0	9	5
T2	25	27.47	9	17.31	0	0	34	18.89
T3	16	17.58	17	32.69	3	8.11	36	20
T4	40	43.96	19	36.54	34	91.89	93	51.67
No information	2	2.2	6	11.54	0	0	8	4.44
Nodal Status*								
N0	50	54.95	14	26.92	3	8.11	67	37.22
N+	38	41.76	27	51.92	8	21.62	73	40.56
No information	3	3.30	11	21.11	26	70.27	40	22.22
Smoking*								
Non-smoker	18	19.78	1	1.92	3	8.11	22	12.22
Former smoker	13	14.29	14	26.92	6	16.22	33	18.33
Current smoker	60	65.93	32	61.54	28	75.68	120	66.67
No information	0	0	5	9.62	0	0	5	2.78
Alcohol*								
Non-drinker	7	7.69	2	3.85	6	16.22	15	8.33
Former drinker	11	12.09	18	34.62	4	10.81	33	18.33
Current drinker	73	80.22	26	50	27	72.97	126	70
No information	0	0	6	11.54	0	0	6	3.33
HPV Status (HPV16E6 serology)								
Negative	82	90.11	47	90.38	34	91.89	163	91.57
Positive	9	9.89	3	5.77	3	8.11	15	8.43
No information	0	0	2	3.85	0	0	2	1.11

*p value<0.05

https://doi.org/10.1371/journal.pone.0191701.t001

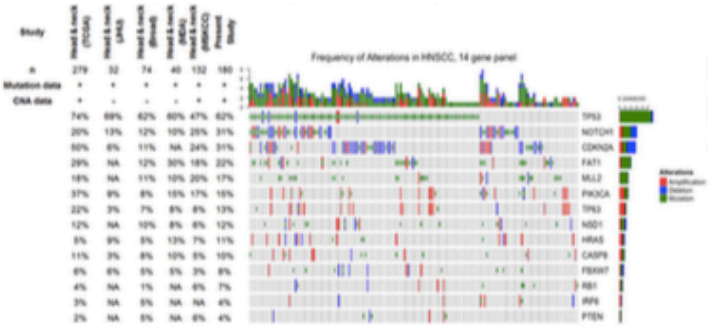

Fig 2. OncoPrint diagram of mutational frequencies and types of alterations of the 14 genes sequenced. Only altered samples are shown. Rows are sorted based on the frequency of the alterations in all samples and columns are sorted to visualize the mutual exclusivity across genes. Frequency of mutations for the following Head and Neck cancer publications are shown: Head & Neck (TCGA)[10], Head & Neck (JHU)[39], Head & Neck (Broad)[32], Head & Neck (MDA)[40], Head & Neck (MSKCC)[41]. NA: Not available.

https://doi.org/10.1371/journal.pone.0191701.g002

Significant co-occurring alterations were found principally in the *TP63* and *PIK3CA* genes ($p<0.001$), both genes located on a frequently amplified region (3q) along with concomitant alterations in *HRAS* and *NOTCH1* genes ($p<0.001$).

Somatic copy number alterations (SCNAs)

Overall, cases were characterized by low chromosomal instability represented by a low copy number burden (mean 23 alterations included amplifications and deletions) compared to the TCGA dataset [10]. We found a total of 47 significantly recurrent amplified regions and 69 deleted regions (q-value<0.1) (Fig 3 and S1 and S3 Tables). The most recurrent focal amplified region was 11q13.3 including the *CCND1* and *FGF3* genes amplified in 40% of samples (60/66 with smoking history), consistent with a region preferentially amplified on smoking related tumours [10, 42]. In addition, we identified regions harbouring oncogenes frequently activated in HNSCC as previously described [10, 32, 33, 39, 40]: 11q22 (*BIRC2*), 3q26 (*SOX2*, *PIK3CA*), 3q28 (*TP63*), 7p11 (*EGFR*), 17q12 (*ERBB2*), along with amplification of regions 8p11, 13q22 and 7q22.

The most frequently deleted region was 15q22, including the locus of the *ANXA2* gene that has been previously found to be downregulated in both head and neck dysplasia and HNSCC [43, 44]. Additionally, recurrent focal deletions were present in cases, particularly at three regions on chromosome 11 (11p15-p15.5, 11q13-q13.3 and 11q23-q24) previously identified as being of frequent microsatellite instability and/or loss of heterozygosity (MSI/LOH) in HNSCC. We also identified deletions in regions of commonly described transcription regulators and tumour suppressor genes in HNSCCs [10, 45]: 5q35.2 (*NSD1*), 20p11 (*NKX2-2*), 8p22.2 (*CSMD1*), 9q34.3 (*NOTCH1*); together with loss of 9p21.3 containing the *CDKN2A* gene which was found almost exclusively in HPV negative tumours (deletion in 1 out of 15 HPV positive cases) (S4 Fig).

Comparison of copy number alterations based on HPV16 status showed a lower proportion of significantly altered regions in HPV positive cases. In particular, the 11q24.3 region (containing the *ATM* and *APLP2* genes) was differentially lost in HPV positive cases (S4 Fig). Additional losses in the 6p region, close to the HLA class I genes loci, were also identified in HPV positive cases.

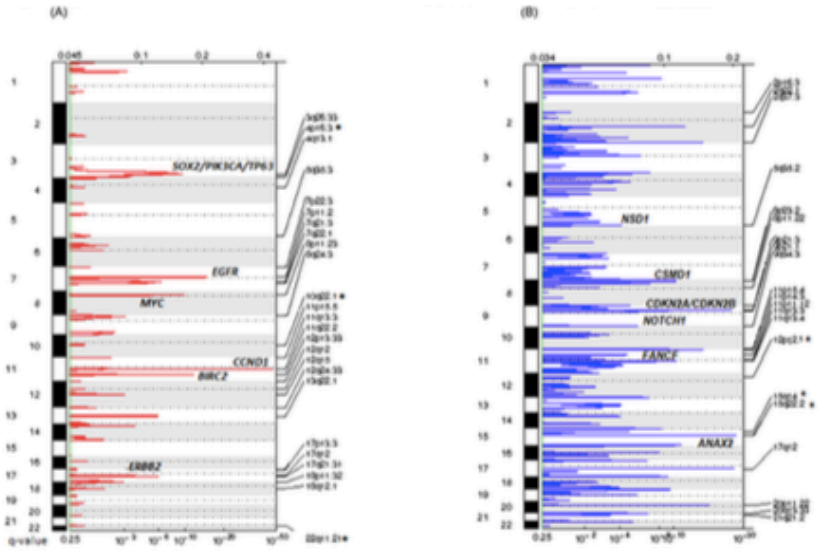

Fig 3. Diagram of significant focal copy number alterations. FDR (Top) and q-values of the alterations are shown in each panel. **(A)** Copy number gains **(B)** Copy number losses. Selected associated genes in some regions are shown. (*) Regions significantly associated with overall survival.

https://doi.org/10.1371/journal.pone.0191701.g003

Integrated analysis

Integrative cluster analysis of both mutational and copy number data identified three distinct clusters with major genomic features including *TP53*, *FAT1* and *FBXW7* SNVs and low, intermediate and high genomic instability. The *FBXW7* gene was significantly mutated in both groups with high and intermediate SCNAs (Fig 4). Eighty-percent of total cases were clustered in the low SCNAs group (mean copy number events = 19). The intermediate SCNAs group (mean copy number events = 39) had only advanced cases (11) and the high SCNAs group (mean copy number events = 43) clustered only cases from Brazil with history of alcohol and smoking exposure (23 cases).

Survival analysis

Survival data was available for 154 cases (Fig 1). Age and nodal status were the only clinical or demographic variables significantly associated to overall survival ($p = 0.01$) (Fig 5 and S2 Table). Multivariate analysis including each of the 14 genes sequenced showed no association with overall survival. Further analysis of *TP53* mutational status showed no association between mutation type (either disruptive/non-disruptive or EAP53 score of missense mutations) and overall survival (S2 Table).

Analysis of the most frequently focal SCNAs showed significant associations between the amplified regions 4p16, 10q22 and 22q11 and a reduction in overall survival. We found additional associations between losses in regions 12p12, 15q14 and15q22 and decreased overall survival (Fig 5 and S2 Table). Although individual candidate genes in these regions were difficult to identify due to the large number of enclosed genes (>20), we identify some genes that

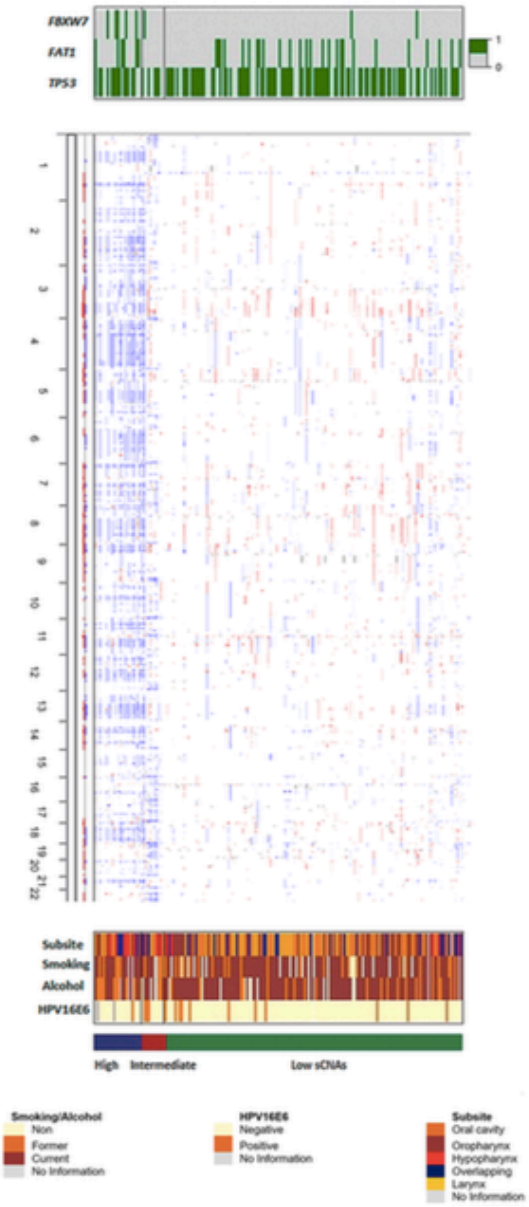

Fig 4. Integrative cluster analysis plot. Cases are grouped by mutation and SCNA status. Top panel: only significant clustering genes are shown (0 = non-mutated, 1 = mutated), middle panel: SCNAs. Amplified (red) and deleted (blue) chromosomal regions. Altered regions are arranged vertically and sorted by genomic locus, with chromosome 1 at the top of the panel and chromosome 22 at the bottom, lower panel: colour coded clinical and epidemiological characteristics.

https://doi.org/10.1371/journal.pone.0191701.g004

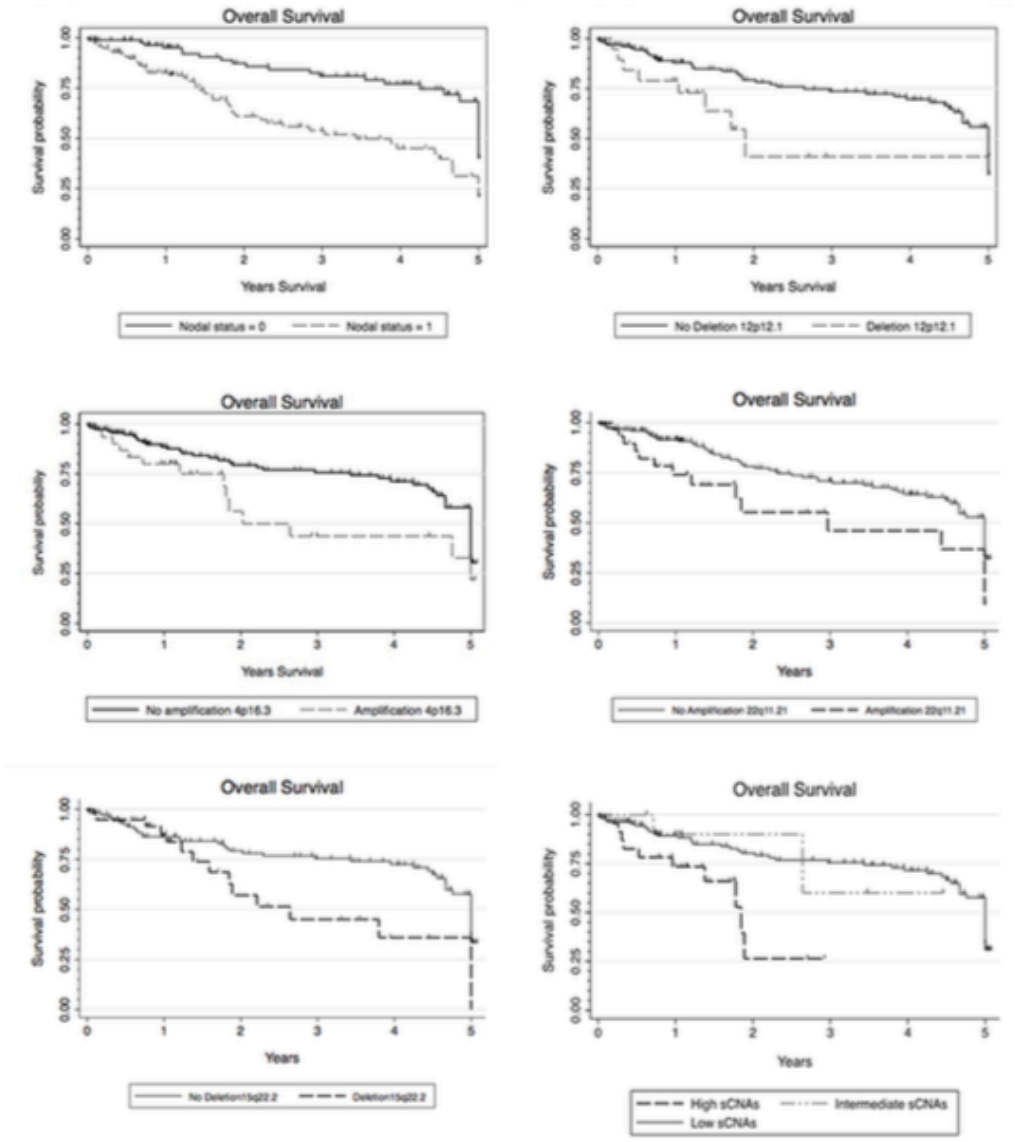

Fig 5. Kaplan-Meier curves showing overall survival outcome for nodal status, significant focal copy number alterations in 22q11.2,15q22 and 12p12 regions associated to smoking and advanced stage, amplification in 4p16.3 and for the three SCNAs clusters.

https://doi.org/10.1371/journal.pone.0191701.g005

have been previously altered in HNSCC (S2 Table) and have been included in our discussion below.

Our integrative clustering approach based on copy number events was also associated with improved overall survival for cases clustered in the low copy number group ($p = 0.01$) (Fig 5 and S2 Table).

Discussion

Head and neck carcinomas show common genomic features determined by SNVs and copy number events in driver genes and cellular pathways associated to the common histology of squamous cell types. However, there is broader genomic heterogeneity due to the variability in anatomic subsite location and the interaction of multiple risk factors such as alcohol and tobacco exposure as well as HPV infection.

Even though we limited our sequencing study to 14 genes, our results showed that most of the mutations described in these genes are representative of the mutational profile of head and neck cancer cases (mutations in 94% of cases). Additionally, the mutational frequency in all 14 genes was comparable to the frequencies observed in previous publications from the largest sequencing projects of Head and Neck cancer cases. In future studies, inclusion of some additional genes such as *AJUBA*, *HLA-A/B*, *NFE2L2*, *KRAS*, *FGFR2/3* and *TRAF3* could improve mutation detection and better capture the mutational landscape of HPV positive tumours, as well as favour the understanding of additional cellular and molecular mechanisms involved in tumour development such as the oxidative stress pathway.

The predominance of low SCNAs in our cases confirms previous studies that differentiate subsets of head and neck tumours (described as M-mutational class tumours) characterized predominantly by mutations rather than chromosomal instability events [8]. A subclass of these low SCNAs group is enriched with alterations in the *PIK3CA-AKT* and p53-mediated apoptosis pathways, in agreement with the number of alterations in *TP53*, *CDKN2A* and *PIK3CA* we observed in our cases.

Eight percent of all cases were HPV 16 positive and 73% corresponded to oropharyngeal tumours. The reduced number of oropharyngeal tumours in the study (25%) and the predominance of older cases, current smokers and drinkers, characteristics preferentially associated to non-related HPV HNSCC[46], might account for the low number of HPV positive cases. In addition, half of our study cases were from Brazil and Argentina which could contribute to the low percentage of HPV positive HNSCC, as it has been previously described in South America [12, 47]. Despite of the limited number of HPV positive cases in our series, we established that HPV positive tumours remain a distinct subset characterized by lower somatic copy number events and differential mutation patterns [36, 48, 49]. Loss of the 11q24.3 region which contains both *ATM* and *APLP2* genes, is a frequent alteration in HPV positive cases[48]. Moreover, the *APLP2* gene is related to tumour immunology as it regulates surface expression of the MHC class I molecules[50, 51]. These results suggest that alterations related to immunological responses might differentiate infection related HNSCC tumours. Further characterization should however, be performed for this group particularly to address the associations between genomic alterations and smoking and alcohol exposure and a differentiated analysis by histological subsite.

Our results confirmed that somatic copy number alterations are an important predictor of overall survival. We have described an improved overall survival for those cases with low SCNAs. These results are in agreement with recent observations showing the direct association between low copy number events, intratumour heterogeneity and clonality with genomic instability and how the joint effect of these factors might influence survival [10, 52–54].

Recently, Andor and colleagues analysed clonality across 12 cancer types from the TCGA dataset, including head and neck cancer cases, and showed that intratumour heterogeneity levels above or below an intermediate measure of clonality were associated with significantly reduced risk of mortality. Moreover, they used copy number alteration abundance as a surrogate measure of genomic instability and found that when SCNAs were present either in a low or a high fraction of the tumours, cases had an improved survival [53]. The high SCNAs group in our study showed the lowest overall survival and clustered only samples from Brazil, all characterized by higher stage and history of both smoking and alcohol exposure. These results give additional evidence to support the rise in mortality due to this malignancy in this country [55].

The mutational profile described in our series of cases showed a clear association to both environmental exposures and clinical characteristics including associations with overall survival. We found that both mutational and focal copy number alterations were correlated with genetic alterations previously described for smoking related head and neck cancers as well as for biomarkers of late stage tumours [10, 32]. Alterations exclusively found in cases with history of both smoking and alcohol consumption included 5q35.3 amplification and 11p14.3 deletion. This last region is of interest as it encloses the *FANCF* gene, involved in the Fanconi anemia pathway and commonly associated to squamous cell carcinoma susceptibility. In addition, *FANCF* inactivation has been previously related to chromosomal instability on sporadic HNSCC [56].

Additionally, focal copy number alterations were found to be significant prognostic markers: 22q11.2 amplification and deletions in 15q22 and 12p12 regions have been associated to smoking related tumours and advanced stage. The 22q11 region contains the *CRKL* gene, which has been characterized as an oncogene in lung SCC [57] and as a promoter of cell growth, motility and adhesion during HNSCC tumorigenesis [58]. Decreased survival in cases with loss of 12p12.1 region, locus of the *PIK3C2G* gene, showed a HR of 3.0 95% CI [1.2; 7.77]. Advanced stage HNSCC tumours have shown mutations in more than one PI3K pathway molecule: *PIK3CA*, *PTEN* and described alterations in *PI3C2G* [59, 60]. Moreover, the 15q22 region, locus of the *ANXA2* gene, has been previously shown to be associated with poorly differentiated tumours in advanced cases. Decreased *ANXA* expression has not however, been formerly shown to be an independent prognostic factor for disease-specific survival in HNSCC [43, 44]. We report for the first time an association between decreased overall survival and amplification of the region 4p16.3, locus of the *FGFR3* gene. High expression levels of FGFR3 contribute of tumour initiation and early-stage progression in HNSCC[61]. More importantly, preclinical studies have demonstrated that FGFR inhibition reduced cell proliferation and increased cell apoptosis in head and neck cancer *in vitro* and *in vivo*[62], highlighting the potential prognostic and therapeutic role of *FGFR3* in HNSCC.

Most studies on HNSCC have documented a decreased overall survival associated to *TP53* mutations[6, 37, 63]. Our study, however, did not find any association between the mutational status of the 14 genes sequenced and overall survival. A specific analysis based on *TP53* mutation type (disruptive vrs nondisruptive or EAP53 score of missense mutations) showed no association to overall survival, either. In agreement to our results, Kim and colleagues, found that patients diagnosed with oral squamous cell carcinoma of the gingivo-buccal region (GBSCC) from the Indian Team project of the International Cancer Genome Consortium (ICGC), did not showed an association between *TP53* mutation status and overall survival [64]. Similar to the epidemiological and clinical characteristics of our study cases, GBSCC patients from the ICGC study were most exposed to tobacco and/or alcohol, presented advanced stage (III/IV) and half of the cases had confirmed nodal metastasis [33].

One of the main limitations of our study is the reduced number of HNSCC cases with early stage tumours. It would be important to further characterize the genomic alterations in early

stages of head and neck cancer cases in order to identify biomarkers for early detection and prognostic stratification especially for the high-risk groups in regions of increase incidence. In addition, our survival analysis was limited due to the lack of complete treatment information for most cases. Treatment regimens have an important association with Head and Neck cancer overall survival and should be included in future analysis specially those involving multicentre studies[65].

In summary, we have identified HNSCC cases with low SCNAs that differentiate as a subset of head and neck cancers driven predominantly by gene mutations and focal alterations rather than chromosomal instability events and are characterized by an improved overall survival. The mutational landscape described in our series of cases showed a clear association to both environmental exposures (alcohol and smoking consumption and HPV infection) and clinical characteristics. Further studies integrating genomic, clinical and epidemiological data, especially in high-risk populations, are necessary to better identify high-risk stratification and characterize prognosis of head and neck cancer cases.

Supporting information

S1 Fig. Mutation calling validation. (A) Venn diagram of number of *TP53* mutations detected in the Gencapo Series. Example: *TP53*Asn239Asp mutation previously detected by Sanger sequencing (B) Plots of mutational calling showing an example of independent libraries sequenced from the same case.
(PDF)

S2 Fig. (A) Mutually exclusive alterations between the 14 genes sequenced (Significance p-value of mutual exclusivity derived from the Z-score) (B) Co-occurrence of alterations (Significance p-value of co-occurrence derived from the Z-Score). Z score based on deviation of the observed mutations compared to expected, obtained by permuting events.
(PDF)

S3 Fig. Diagrams of mutation distribution in genes with frequent SNVs. Mutation colours represent: Green: Missense Mutations; red: Truncating Mutations (Nonsense, Nonstop, Frameshift deletion, Frameshift insertion, Splice site), black: Inframe Mutations (Inframe deletion, Inframe insertion). Circles colored with purple indicate residues that are affected by different mutation types at the same proportion.
(PDF)

S4 Fig. Mutational Profile and copy number losses in HPV positive cases. (A) Mutational frequencies of the 14 genes sequenced in 15 HPV16E6 positive cases. (B) Comparison of Significant Focal copy number losses between HPV positive and HPV negative cases. (*) Regions significantly associated with overall survival.
(PDF)

S1 Table. GISTIC list of focal copy number amplifications and deletions. In red regions significantly associated with overall survival and head and neck cancer related genes.
(XLSX)

S2 Table. Survival analysis of main demographic, clinical and genomic variables.
(XLSX)

S3 Table. List of filtered and annotated somatic mutations (SNVS).
(XLSX)

Acknowledgments

The authors wish to acknowledge Dr Javier Oliver for his valuable collaboration in the standardization of the targeted sequencing assays and Valérie Gaborieau for her contribution in database curation and data homogenization from the three studies. The authors acknowledge all patients who donated their biological specimens and the contribution of GENCAPO (Brazilian Head and Neck Genome Project) for clinical samples and for clinical and pathological data collection (complete list of members and affiliations presented at http://www.gencapo.famerp.br).

This work was undertaken during the tenure of a Postdoctoral Fellowship to Dr Sandra Perdomo from The International Agency for Research on Cancer, partially supported by the European Commission FP7 Marie Curie Actions—People—Co-funding of regional, national and international programmes (COFUND).

Author Contributions

Conceptualization: Paul Brennan.

Formal analysis: Sandra Perdomo, Matthieu Foll.

Investigation: Sandra Perdomo, Devasena Anantharaman, Matthieu Foll, Behnoush Abedi-Ardekani, Geoffroy Durand, Luciana Albina Reis Rosa, Reetta Holmila, Florence Le Calvez-Kelm, Eloiza H. Tajara, Victor Wünsch-Filho, José Eduardo Levi, Marta Vilensky, Jerry Polesel, Ivana Holcatova, Lorenzo Simonato, Cristina Canova, Pagona Lagiou, James D. McKay.

Methodology: Sandra Perdomo, Devasena Anantharaman, Matthieu Foll, Geoffroy Durand, Luciana Albina Reis Rosa, Reetta Holmila, Florence Le Calvez-Kelm, Eloiza H. Tajara, Victor Wünsch-Filho, José Eduardo Levi, Marta Vilensky, Jerry Polesel, Ivana Holcatova, Lorenzo Simonato, Cristina Canova, Pagona Lagiou, James D. McKay.

Resources: Devasena Anantharaman, Matthieu Foll, Eloiza H. Tajara, Victor Wünsch-Filho, José Eduardo Levi, Marta Vilensky, Jerry Polesel, Ivana Holcatova, Lorenzo Simonato, Cristina Canova, Pagona Lagiou, James D. McKay.

Software: Matthieu Foll.

Supervision: Paul Brennan.

Writing – original draft: Sandra Perdomo, Paul Brennan.

Writing – review & editing: Sandra Perdomo, Devasena Anantharaman, Matthieu Foll, Behnoush Abedi-Ardekani, Geoffroy Durand, Luciana Albina Reis Rosa, Reetta Holmila, Florence Le Calvez-Kelm, Eloiza H. Tajara, Victor Wünsch-Filho, José Eduardo Levi, Marta Vilensky, Jerry Polesel, Ivana Holcatova, Lorenzo Simonato, Cristina Canova, Pagona Lagiou, James D. McKay.

References

1. Ferlay J SI, Ervik M, Dikshit R, Eser S, Mathers C, Rebelo M, Parkin DM, Forman D, Bray F. GLOBOCAN 2012 v1.0. Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 11 2014. http://globocan.iarc.fr/.
2. Forman D BF, Brewster DH, Gombe Mbalawa C, Kohler B, Piñeros M, Steliarova-Foucher E, Swaminathan R and Ferlay J. Cancer Incidence in Five Continents, Vol. X (electronic version). Lyon: IARC; 2013.

3. Gatta G, Botta L, Sanchez MJ, Anderson LA, Pierannunzio D, Licitra L, et al. Prognoses and improvement for head and neck cancers diagnosed in Europe in early 2000s: The EUROCARE-5 population-based study. Eur J Cancer. 2015. https://doi.org/10.1016/j.ejca.2015.07.043 PMID: 26421817.

4. Humans IWGotEoCRt. Personal habits and indoor combustions. Volume 100 E. A review of human carcinogens. IARC Monogr Eval Carcinog Risks Hum. 2012; 100(Pt E):1–538. PMID: 23193840.

5. Castellsagué X, Alemany L, Quer M, Halec G, Quirós B, Tous S, et al. HPV Involvement in Head and Neck Cancers: Comprehensive Assessment of Biomarkers in 3680 Patients. J Natl Cancer Inst. 2016; 108(6). https://doi.org/10.1093/jnci/djv403 PMID: 26823521.

6. Gross AM, Orosco RK, Shen JP, Egloff AM, Carter H, Hofree M, et al. Multi-tiered genomic analysis of head and neck cancer ties TP53 mutation to 3p loss. Nat Genet. 2014; 46(9):939–43. Epub 2014/08/05. https://doi.org/10.1038/ng.3051 PMID: 25086664.

7. Reddy RB, Bhat AR, James BL, Govindan SV, Mathew R, Dr R, et al. Meta-Analyses of Microarray Datasets Identifies ANO1 and FADD as Prognostic Markers of Head and Neck Cancer. PLoS One. 2016; 11(1):e0147409. https://doi.org/10.1371/journal.pone.0147409 PMID: 26808319.

8. Ciriello G, Miller ML, Aksoy BA, Senbabaoglu Y, Schultz N, Sander C. Emerging landscape of oncogenic signatures across human cancers. Nat Genet. 2013; 45(10):1127–33. https://doi.org/10.1038/ng.2762 PMID: 24071851.

9. Zack TI, Schumacher SE, Carter SL, Cherniack AD, Saksena G, Tabak B, et al. Pan-cancer patterns of somatic copy number alteration. Nat Genet. 2013; 45(10):1134–40. https://doi.org/10.1038/ng.2760 PMID: 24071852.

10. The Cancer Genome Atlas N. Comprehensive genomic characterization of head and neck squamous cell carcinomas. Nature. 2015; 517(7536):576–82. https://doi.org/10.1038/nature14129 PMID: 25631445

11. Lagiou P, Georgila C, Minaki P, Ahrens W, Pohlabeln H, Benhamou S, et al. Alcohol-related cancers and genetic susceptibility in Europe: the ARCAGE project: study samples and data collection. Eur J Cancer Prev. 2009; 18(1):76–84. Epub 2008/10/03. https://doi.org/10.1097/CEJ.0b013e32830c8dca PMID: 18830131.

12. López RVM, Levi JE, Eluf-Neto J, Koifman RJ, Koifman S, Curado MP, et al. Human papillomavirus (HPV) 16 and the prognosis of head and neck cancer in a geographical region with a low prevalence of HPV infection. Cancer Causes & Control. 2014; 25(4):461–71. https://doi.org/10.1007/s10552-014-0348-8 PMID: 24474236

13. Szymańska K, Hung RJ, Wünsch-Filho V, Eluf-Neto J, Curado MP, Koifman S, et al. Alcohol and tobacco, and the risk of cancers of the upper aerodigestive tract in Latin America: a case–control study. Cancer Causes & Control. 2011; 22(7):1037–46. https://doi.org/10.1007/s10552-011-9779-7 PMID: 21607590

14. Anantharaman D, Gheit T, Waterboer T, Abedi-Ardekani B, Carreira C, McKay-Chopin S, et al. Human Papillomavirus Infections and Upper Aero-Digestive Tract Cancers: The ARCAGE Study. JNCI Journal of the National Cancer Institute. 2013; 105(8):536–45. https://doi.org/10.1093/jnci/djt053 PMID: 23503618

15. Reuschenbach M, Waterboer T, Wallin KL, Einenkel J, Dillner J, Hamsikova E, et al. Characterization of humoral immune responses against p16, p53, HPV16 E6 and HPV16 E7 in patients with HPV-associated cancers. Int J Cancer. 2008; 123(11):2626–31. https://doi.org/10.1002/ijc.23837 PMID: 18785210.

16. Herrero R. Human Papillomavirus and Oral Cancer: The International Agency for Research on Cancer Multicenter Study. CancerSpectrum Knowledge Environment. 2003; 95(23):1772–83. https://doi.org/10.1093/jnci/djg107

17. Kreimer AR, Johansson M, Waterboer T, Kaaks R, Chang-Claude J, Drogen D, et al. Evaluation of human papillomavirus antibodies and risk of subsequent head and neck cancer. J Clin Oncol. 2013; 31(21):2708–15. https://doi.org/10.1200/JCO.2012.47.2738 PMID: 23775966.

18. Oh JE, Ohta T, Satomi K, Foll M, Durand G, McKay J, et al. Alterations in the NF2/LATS1/LATS2/YAP Pathway in Schwannomas. J Neuropathol Exp Neurol. 2015; 74(10):952–9. https://doi.org/10.1097/NEN.0000000000000238 PMID: 26360373.

19. Aeberhard WH, Cantoni E, Heritier S. Robust inference in the negative binomial regression model with an application to falls data. Biometrics. 2014; 70(4):920–31. https://doi.org/10.1111/biom.12212 PMID: 25156188.

20. Fernandez-Cuesta L, Perdomo S, Avogbe PH, Leblay N, Delhomme TM, Gaborieau V, et al. Identification of Circulating Tumor DNA for the Early Detection of Small-cell Lung Cancer. EBioMedicine. 2016. https://doi.org/10.1016/j.ebiom.2016.06.032 PMID: 27377626.

21. Le Calvez-Kelm F, Foll M, Wozniak MB, Delhomme TM, Durand G, Chopard P, et al. KRAS mutations in blood circulating cell-free DNA: a pancreatic cancer case-control. Oncotarget. 2016; 7(48):78827–40. https://doi.org/10.18632/oncotarget.12386 PMID: 27705932.

22. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res 2010. p. e164. https://doi.org/10.1093/nar/gkq603 PMID: 20601685

23. Scharpf RB, Irizarry RA, Ritchie ME, Carvalho B, Ruczinski I. Using the R Package crlmm for Genotyping and Copy Number Estimation. J Stat Softw. 2011; 40(12):1–32. PMID: 22523482.

24. MacDonald JR, Ziman R, Yuen RK, Feuk L, Scherer SW. The Database of Genomic Variants: a curated collection of structural variation in the human genome. Nucleic Acids Res. 2014; 42(Database issue): D986–92. https://doi.org/10.1093/nar/gkt958 PMID: 24174537.

25. Mermel CH, Schumacher SE, Hill B, Meyerson ML, Beroukhim R, Getz G. GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. Genome Biol. 2011; 12(4):R41. https://doi.org/10.1186/gb-2011-12-4-r41 PMID: 21527027.

26. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Sci Signal. 2013; 6(269):pl1. https://doi.org/10.1126/scisignal.2004088 PMID: 23550210.

27. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov. 2012; 2(5):401–4. https://doi.org/10.1158/2159-8290.CD-12-0095 PMID: 22588877.

28. Mo Q, Wang S, Seshan VE, Olshen AB, Schultz N, Sander C, et al. Pattern discovery and cancer gene identification in integrated cancer genomic data. Proc Natl Acad Sci U S A. 2013; 110(11):4245–50. https://doi.org/10.1073/pnas.1208949110 PMID: 23431203.

29. Perez-Llamas C, Lopez-Bigas N. Gitools: analysis and visualisation of genomic data using interactive heat-maps. PLoS One. 2011; 6(5):e19541. https://doi.org/10.1371/journal.pone.0019541 PMID: 21602921.

30. Benjamini Y, and Hochberg Yosef. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society Series B (Methodological). 1995; 57(1). Epub 300.

31. Ervik M.FL, Ferlay J., Mery L., Soerjomataram I., Bray F.. Cancer Today Lyon, France: International Agency for Research on Cancer; 2016 [cited 2016 10 May]. http://goo.iarc.fr/today.

32. Stransky N, Egloff AM, Tward AD, Kostic AD, Cibulskis K, Sivachenko A, et al. The Mutational Landscape of Head and Neck Squamous Cell Carcinoma. Science. 2011; 333(6046):1157–60. https://doi.org/10.1126/science.1208130 PMID: 21798893

33. India Project Team of the International Cancer Genome C. Mutational landscape of gingivo-buccal oral squamous cell carcinoma reveals new recurrently-mutated genes and molecular subgroups. Nat Commun. 2013; 4:2873. https://doi.org/10.1038/ncomms3873 PMID: 24292195.

34. Morris LG, Chandramohan R, West L, Zehir A, Chakravarty D, Pfister DG, et al. The Molecular Landscape of Recurrent and Metastatic Head and Neck Cancers: Insights From a Precision Oncology Sequencing Platform. JAMA Oncol. 2016. https://doi.org/10.1001/jamaoncol.2016.1790 PMID: 27442865.

35. Dai M, Clifford GM, le Calvez F, Castellsague X, Snijders PJ, Pawlita M, et al. Human papillomavirus type 16 and TP53 mutation in oral cancer: matched analysis of the IARC multicenter study. Cancer Res. 2004; 64(2):468–71. Epub 2004/01/28. PMID: 14744758.

36. Seiwert TY, Zuo Z, Keck MK, Khattri A, Pedamallu CS, Stricker T, et al. Integrative and comparative genomic analysis of HPV-positive and HPV-negative head and neck squamous cell carcinomas. Clin Cancer Res. 2015; 21(3):632–41. https://doi.org/10.1158/1078-0432.CCR-13-3310 PMID: 25056374.

37. Poeta ML, Manola J, Goldwasser MA, Forastiere A, Benoit N, Califano JA, et al. TP53 mutations and survival in squamous-cell carcinoma of the head and neck. N Engl J Med. 2007; 357(25):2552–61. https://doi.org/10.1056/NEJMoa073770 PMID: 18094376.

38. Neskey DM, Osman AA, Ow TJ, Katsonis P, McDonald T, Hicks SC, et al. Evolutionary Action Score of TP53 Identifies High-Risk Mutations Associated with Decreased Survival and Increased Distant Metastases in Head and Neck Cancer. Cancer Res. 2015; 75(7):1527–36. https://doi.org/10.1158/0008-5472.CAN-14-2735 PMID: 25634208.

39. Agrawal N, Frederick MJ, Pickering CR, Bettegowda C, Chang K, Li RJ, et al. Exome Sequencing of Head and Neck Squamous Cell Carcinoma Reveals Inactivating Mutations in NOTCH1. Science. 2011; 333(6046):1154–7. https://doi.org/10.1126/science.1206923 PMID: 21798897

40. Pickering CR, Zhang J, Yoo SY, Bengtsson L, Moorthy S, Neskey DM, et al. Integrative genomic characterization of oral squamous cell carcinoma identifies frequent somatic drivers. Cancer Discov. 2013; 3(7):770–81. https://doi.org/10.1158/2159-8290.CD-12-0537 PMID: 23619168.

41. Hedberg ML, Goh G, Chiosea SI, Bauman JE, Freilino ML, Zeng Y, et al. Genetic landscape of metastatic and recurrent head and neck squamous cell carcinoma. J Clin Invest. 2016; 126(1):169–80. https://doi.org/10.1172/JCI82066 PMID: 26619122.

42. !! INVALID CITATION!!! {Pattle, 2017 #4150;Pattle, 2017 #4150}.

43. Rodrigo JP, Lequerica-Fernández P, Rosado P, Allonca E, García-Pedrero JM, de Vicente JC. Clinical significance of annexin A2 downregulation in oral squamous cell carcinoma. Head & Neck. 2011; 33 (12):1708–14. https://doi.org/10.1002/hed.21661 PMID: 21500302

44. Pena-Alonso E, Rodrigo JP, Parra IC, Pedrero JM, Meana MV, Nieto CS, et al. Annexin A2 localizes to the basal epithelial layer and is down-regulated in dysplasia and head and neck squamous cell carcinoma. Cancer Lett. 2008; 263(1):89–98. https://doi.org/10.1016/j.canlet.2007.12.029 PMID: 18262347.

45. Imai FL, Uzawa K, Miyakawa A, Shiiba M, Tanzawa H. A detailed deletion map of chromosome 20 in human oral squamous cell carcinoma. Int J Mol Med. 2001; 7(1):43–7. PMID: 11115607.

46. Taberna M, Mena M, Pavon MA, Alemany L, Gillison ML, Mesia R. Human papillomavirus related oropharyngeal cancer. Ann Oncol. 2017. https://doi.org/10.1093/annonc/mdx304 PMID: 28633362.

47. Anantharaman D, Abedi-Ardekani B, Beachler DC, Gheit T, Olshan AF, Wisniewski K, et al. Geographic heterogeneity in the prevalence of human papillomavirus in head and neck cancer. Int J Cancer. 2017; 140(9):1958–75. https://doi.org/10.1002/ijc.30608 PMID: 28108990.

48. Hayes DN, Van Waes C, Seiwert TY. Genetic Landscape of Human Papillomavirus-Associated Head and Neck Cancer and Comparison to Tobacco-Related Tumors. J Clin Oncol. 2015; 33(29):3227–34. https://doi.org/10.1200/JCO.2015.62.1086 PMID: 26351353.

49. Keck MK, Zuo Z, Khattri A, Stricker TP, Brown CD, Imanguli M, et al. Integrative analysis of head and neck cancer identifies two biologically distinct HPV and three non-HPV subtypes. Clin Cancer Res. 2015; 21(4):870–81. https://doi.org/10.1158/1078-0432.CCR-14-2481 PMID: 25492084.

50. Peters HL, Tuli A, Sharma M, Naslavsky N, Caplan S, MacDonald RG, et al. Regulation of major histocompatibility complex class I molecule expression on cancer cells by amyloid precursor-like protein 2. Immunol Res. 2011; 51(1):39–44. https://doi.org/10.1007/s12026-011-8238-6 PMID: 21826533.

51. Feenstra M, Veltkamp M, van Kuik J, Wiertsema S, Slootweg P, van den Tweel J, et al. HLA class I expression and chromosomal deletions at 6p and 15q in head and neck squamous cell carcinomas. Tissue Antigens. 1999; 54(3):235–45. https://doi.org/10.1034/j.1399-0039.1999.540304.x PMID: 10519360

52. Mroz EA, Tward AD, Hammon RJ, Ren Y, Rocco JW. Intra-tumor genetic heterogeneity and mortality in head and neck cancer: analysis of data from the Cancer Genome Atlas. PLoS Med. 2015; 12(2): e1001786. https://doi.org/10.1371/journal.pmed.1001786 PMID: 25668320.

53. Andor N, Graham TA, Jansen M, Xia LC, Aktipis CA, Petritsch C, et al. Pan-cancer analysis of the extent and consequences of intratumor heterogeneity. Nat Med. 2015. https://doi.org/10.1038/nm.3984 PMID: 26618723.

54. Smeets SJ, Brakenhoff RH, Ylstra B, van Wieringen WN, van de Wiel MA, Leemans CR, et al. Genetic classification of oral and oropharyngeal carcinomas identifies subgroups with a different prognosis. Cell Oncol. 2009; 31(4):291–300. https://doi.org/10.3233/CLO-2009-0471 PMID: 19633365.

55. Chatenoud L, Bertuccio P, Bosetti C, Malvezzi M, Levi F, Negri E, et al. Trends in mortality from major cancers in the Americas: 1980–2010. Annals of Oncology. 2014; 25(9):1843–53. https://doi.org/10.1093/annonc/mdu206 PMID: 24907637

56. Stoepker C, Ameziane N, van der Lelij P, Kooi IE, Oostra AB, Rooimans MA, et al. Defects in the Fanconi Anemia Pathway and Chromatid Cohesion in Head and Neck Cancer. Cancer Res. 2015; 75 (17):3543–53. https://doi.org/10.1158/0008-5472.CAN-15-0528 PMID: 26122845.

57. Kim YH, Kwei KA, Girard L, Salari K, Kao J, Pacyna-Gengelbach M, et al. Genomic and functional analysis identifies CRKL as an oncogene amplified in lung cancer. Oncogene. 2010; 29(10):1421–30. https://doi.org/10.1038/onc.2009.437 PMID: 19966867.

58. Yanagi H, Wang L, Nishihara H, Kimura T, Tanino M, Yanagi T, et al. CRKL plays a pivotal role in tumorigenesis of head and neck squamous cell carcinoma through the regulation of cell adhesion. Biochem Biophys Res Commun. 2012; 418(1):104–9. https://doi.org/10.1016/j.bbrc.2011.12.142 PMID: 22244889.

59. Lui VWY, Hedberg ML, Li H, Vangara BS, Pendleton K, Zeng Y, et al. Frequent Mutation of the PI3K Pathway in Head and Neck Cancer Defines Predictive Biomarkers. Cancer Discovery. 2013; 3(7):761–9. https://doi.org/10.1158/2159-8290.CD-13-0103 PMID: 23619167

60. Giudice FS, Squarize CH. The determinants of head and neck cancer: Unmasking the PI3K pathway mutations. J Carcinog Mutagen. 2013; Suppl 5. https://doi.org/10.4172/2157-2518.S5-003 PMID: 25126449.

61. Vairaktaris E, Ragos V, Yapijakis C, Derka S, Vassiliou S, Nkenke E, et al. FGFR-2 and -3 play an important role in initial stages of oral oncogenesis. Anticancer Res. 2006; 26(6B):4217–21. PMID: 17201136.

62. Sweeny L, Liu Z, Lancaster W, Hart J, Hartman YE, Rosenthal EL. Inhibition of fibroblasts reduced head and neck cancer growth by targeting fibroblast growth factor receptor. Laryngoscope. 2012; 122 (7):1539–44. Epub 2012/03/27. https://doi.org/10.1002/lary.23266 PMID: 22460537.

63. Zhou G, Liu Z, Myers JN. TP53 Mutations in Head and Neck Squamous Cell Carcinoma and Their Impact on Disease Progression and Treatment Response. J Cell Biochem. 2016; 117(12):2682–92. https://doi.org/10.1002/jcb.25592 PMID: 27166782.

64. Kim KT, Kim BS, Kim JH. Association between FAT1 mutation and overall survival in patients with human papillomavirus-negative head and neck squamous cell carcinoma. Head Neck. 2016; 38 Suppl 1:E2021–9. https://doi.org/10.1002/hed.24372 PMID: 26876381.

65. Pulte D, Brenner H. Changes in survival in head and neck cancers in the late 20th and early 21st century: a period analysis. Oncologist. 2010; 15(9):994–1001. https://doi.org/10.1634/theoncologist.2009-0289 PMID: 20798198.

APÊNDICE C - *"Oncogenic drivers in 11q13 associated with prognosis and response to therapy in advanced oropharyngeal carcinomas".*

M.C. Barros-Filho, **L.A. Reis-Rosa***, M. Hatakeyama* et al. Oral Oncology 83 (2018) 81–90. (*these authors contributed equally)*

Oral Oncology 83 (2018) 81–90

Contents lists available at ScienceDirect

Oral Oncology

journal homepage: www.elsevier.com/locate/oraloncology

Oncogenic drivers in 11q13 associated with prognosis and response to therapy in advanced oropharyngeal carcinomas

M.C. Barros-Filho[a,1], L.A. Reis-Rosa[b,1], M. Hatakeyama[a,1], F.A. Marchi[a], T. Chulam[c], C. Scapulatempo-Neto[d], U.R. Nicolau[e], A.L. Carvalho[f], C.A.L. Pinto[g], S.A. Drigo[h], L.P. Kowalski[c,i], S.R. Rogatto[i,j,*]

[a] International Research Center (CIPE), A. C. Camargo Cancer Center, São Paulo, Brazil
[b] Instituto de Medicina Tropical de SP Universidade de São Paulo-USP, São Paulo, Brazil
[c] Department of Head and Neck Surgery and Otorhinolaryngology, A. C. Camargo Cancer Center, São Paulo, Brazil
[d] Molecular Oncology Research Center, Barretos, and Diagnósticos da América (DASA), São Paulo, Brazil
[e] Department of Oncology, A. C. Camargo Cancer Center, São Paulo, Brazil
[f] Department of Head and Neck Surgery, Barretos Cancer Hospital, Barretos, Brazil
[g] Department of Pathology, A. C. Camargo Cancer Center, São Paulo, Brazil
[h] Department of Surgery and Orthopedics, School of Medicine and Department of Veterinary Clinic, School of Veterinary Medicine and Animal Science, UNESP, Botucatu, Brazil
[i] National Institute of Science and Technology in Oncogenomics (INCiTO), São Paulo, Brazil
[j] Department of Clinical Genetics, Vejle Hospital, Institute of Regional Health Research, University of Southern Denmark, Denmark

ARTICLE INFO

Keywords:
Oropharyngeal cancer
Driver alterations
Human papilloma virus
Prognostic factors
Predictive factors
Transcriptome profiling
Array comparative genomic hybridization
Reverse transcriptase polymerase chain reaction

ABSTRACT

Objectives: To identify potential molecular drivers associated with prognosis and response to treatment in advanced oropharyngeal squamous cell carcinomas (OPSCC).
Materials and methods: Thirty-three OPSCC biopsies from untreated Brazilian patients were evaluated for human papilloma virus genotyping, genome wide copy number alterations and gene expression profiling. Data were integrated using CONEXIC algorithm. Validation with TCGA dataset and confirmation by RT-qPCR of candidate genes were performed.
Results: High-risk HPV positive cases, detected in 55% of advanced OPSCC, were associated with better outcome. Losses of 8p11.23-p11.22, 14q11.1-q11.2 and 15q11.2, and gains of 11q13.2 and 11q13.2-q13.3 were detected as recurrent alterations. Gains of 3q26.31 and 11q13.2 and losses of 9p21.3 were exclusively detected in HPV-negative tumors. Two clusters of expression profiles were observed, being one composed mostly by HPV positive cases (83%). HPV-positive enriched cluster showed predominantly immune response-related pathways. Integrative analysis identified 10 modulators mapped in 11q13, which were frequently cancer-related. These 10 genes showed copy number gains, overexpression and an association with worse survival, further validated by TCGA database analyses. Overexpression of four genes (*ORAOV1, CPT1A, SHANK2* and *PPFIA1)* evaluated by RT-qPCR confirmed their association with poor survival. Multivariate analysis showed that *PPFIA1* overexpression and HPV status are independent prognostic markers. Moreover, *SHANK2* overexpression was significantly associated with incomplete response to treatment.
Conclusion: The integrative genomic and transcriptomic data revealed potential driver genes mapped in 11q13 associated with worse prognosis and response to treatment, giving fundamentals for the identification of novel therapeutic targets in OPSCC.

Introduction

In the last decades, several epidemiological studies have revealed decreased incidence of head and neck squamous cell carcinomas (HNSCC) in oral cavity and larynx as a consequence of lower exposure to the tobacco products. Nevertheless, an increasing incidence of

* Corresponding author at: Department of Clinical Genetics, Vejle Hospital, Institute of Regional Health Research, University of Southern Denmark, Beriderbakken 4, 7100 Vejle, Denmark.
E-mail address: silvia.regina.rogatto@rsyd.dk (S.R. Rogatto).
[1] These authors contributed equally.

https://doi.org/10.1016/j.oraloncology.2018.06.010
Received 5 March 2018; Received in revised form 26 April 2018; Accepted 12 June 2018
1368-8375/

oropharyngeal squamous cell carcinomas (OPSCC) mainly associated with oncogenic human papillomavirus (HPV) has been reported [1–6].

In general, HPV-positive OPSCC is associated with good prognosis, presenting better survival in comparison with HPV-negative cases [7–11]. De-escalation of radiation and chemotherapy for HPV-positive cases has been proposed and tested in different clinical trials aiming to avoid overtreatment and long-term toxicities [8,12,13]. However, accurate identification of cases with good prognosis and treatment-responsive tumors are critical findings, since distant metastasis may occur in a set of HPV-positive OPSCC [8,14]. In contrast, few advances have been made for treatment of HPV-negative OPSCC patients and a large number of them will present loco-regional recurrence [9].

The molecular mechanisms underlying oropharyngeal carcinogenesis have been investigated and potential biomarkers were reported, however the data are still unclear and controversial [15–19]. The integration of genomic and transcriptomic analysis can be used to identify cancer-driver genes and disrupted pathways, which can be drug targetable [20]. This strategy has revealed functionally relevant drivers involved in the carcinogenic process in different tumor types, including oral carcinoma [21], ovarian cancer [22], penile carcinoma [23,24], uterine leiomyoma [25] and leiomyosarcoma [26].

Integration of genomic, transcriptomic and epigenomic data of 279 HNSCC, including oral (n = 172, 62%), oropharyngeal (n = 33, 12%) and laryngeal (n = 72, 26%) carcinomas was reported by The Cancer Genome Atlas (TCGA) [27]. Distinct genetic alterations were observed between HPV-positive HNSCCs (68% in oropharynx) and HPV-negative cases. Recurrent deletions and truncating mutations of *TRAF3* found in HPV-positive tumors were associated with anti-viral immune response. Conversely, HPV-negative HNSCCs presented loss of 9p21.3 (including *CDKN2A* gene) and co-amplifications of 11q13 and 11q22, which contain genes implicated in cell death/NF-kB and Hippo pathways [27]. A distinct genetic subgroup of HPV-negative tumors is also being reported, characterized by low frequency of copy number alterations (CNA), wild-type *TP53*, mutation in *HRAS* and *CASP8* and more favorable prognosis [13].

Nevertheless, prognostic and predictive biomarkers in advanced OPSCC are still limited and need to be further investigated. In this study, we integrated DNA CNA and gene expression analyses to identify drivers in advanced OPSCC according to HPV status. *In silico* functional analysis was performed to identify genes and pathways associated with oropharyngeal carcinogenesis, which can reveal potential drug targets.

Patients and methods

Patients and samples

Fresh-frozen tumor biopsy samples from 40 OPSCC patients naive of treatment were obtained from A.C. Camargo Cancer Center and Barretos Cancer Hospital, Brazil. Eligibility criteria included patients harboring locally advanced clinical stages III, IVA and IVB according to AJCC (7Th Edition). The follow-up time ranged from 0.5 to 190 months (mean of 53 months). The study was approved by the Human Research Ethics Committee from both Institutions (A.C. Camargo Cancer Center #1249/09 and Barretos Cancer Hospital #139/2008). All patients provided written informed consent. Patients underwent curative therapy according to standard clinical protocol taking into account the medical decisions for each patient, which included induction chemotherapy (IC) followed by radiotherapy with concurrent chemotherapy (CT-RT), and upfront surgical resection followed by radiotherapy with or without chemotherapy. Pretreatment risk stratification was defined as low, intermediate and high, based on HPV status, smoking history and tumor/node stage [7].

Treatment and response assessment

The majority of patients included in this study were treated by IC followed by CT-RT (n = 24). Patients were treated with a combination of docetaxel, cisplatin and 5-fluorouracil (TPF) as IC followed by radiotherapy and weekly carboplatin or cisplatin concurrent to radiotherapy. Cetuximab was employed in eight cases as concurrent therapy to radiation (one case received IC before). Six patients were treated by surgery followed by adjuvant therapy (RT or RT + CT); one patient deceased surgery and one patient lost the follow-up before completing the treatment (supplementary Table S1). Six patients were treated only by surgery followed by adjuvant therapy (RT and or CT), one patient deceased soon after surgery and one patient lost the follow-up before treatment begins.

Nucleic acids extraction and HPV genotyping

OPSCC samples (80% of tumor cells) and surrounding normal tissues were macrodissected for DNA (Qiagen DNeasy Blood & Tissue Kit; Qiagen, Valencia, CA) and total RNA extraction (RNeasy MiniKit; Qiagen, Valencia, CA). The Linear Array HPV Genotyping Test Kit (Roche Molecular Systems, Inc., Branchburg, NJ, USA) was used for HPV detection.

CNA analysis by array-based Comparative genomic Hybridization

OPSCC (n = 33) and normal commercial DNA (Promega) samples were differentially labeled (Genomic DNA Enzymatic Labeling Kit; Agilent Technologies) and the hybridized on Agilent Human CGH 180K Oligo Microarrays. Genomic data were extracted by Feature Extraction 10.1.1.1 software (Agilent Technologies) and analyzed using the Nexus Copy Number software (v.6.0, Biodiscovery, El Segundo, CA, USA). CNA was defined as exceeding the significance threshold of 1×10^{-6} and containing at least five consecutive altered probes per segment. The thresholds were defined as the average $\log_2$ CGH fluorescence ratio for copy gains ≥ 0.6, high copy number gains ≥ 1.4, losses ≤ −0.6 and homozygous losses ≤ −1.25. Genomic variants detected in control individuals from worldwide populations and classified as common (> 1%) according to DGV database (http://dgv.tcag.ca/dgv/app/home) were excluded. Alterations detected in at least 20% of the cases were selected for further analysis. The unsupervised hierarchical clustering analysis was performed using complete linkage and Euclidian distance.

Gene expression microarray

Total RNA from OPSCC (n = 33) and surrounding non-neoplastic oropharyngeal tissues (n = 3) were labeled and hybridized using the Two-Color Human GE 4x44K microarray platform (Agilent Technologies), following the manufacturer instructions. Data processing, quality control filtering and normalization (Lowess) were performed using the Feature Extraction v.10.1.1.1 software (Agilent Technologies) and an in-house pipeline. Gene expression analysis was performed using R *version* 2.15 (http://www.bioconductor.org/) and BRB ArrayTools software (v.4.4.0). An unsupervised hierarchical clustering analysis was employed with the most variable probes (interquartile range > 0.1) using complete linkage and Euclidian distance. Transcriptomic variations among clusters were identified by significance analysis of microarray (SAM) (false discovery ratio < 1%).

The CNA and expression microarray data are available at the Gene Expression Omnibus (GEO) (GSE111395).

Integrative analysis

Paired CNA and gene expression data of 33 OPSCC was integrated using COpy Number and EXpression In Cancer (CONEXIC) algorithm to identify drivers, which results in a ranked list with high scores modulators [28]. In this analysis, unbalanced expressed genes are correlated with the expression of group of genes (modules), and genomic regions

with significant alterations indicate a greater possibility of the gene has some adaptive advantage on the tumor phenotype. Parameters used in this analysis were described elsewhere [25]. Mann–Whitney test (Prismv.5.0, GraphPad Software, La Jolla, CA) was applied to analyze the relationship between gene expression and CNA.

External data validation and in silico functional analysis of driver candidates

Driver candidates were validated in 78 OPSCC available in the TCGA public database (SNP arrays and RNAseq), using cBio Cancer Genomics Portal (http://www.cbioportal.org/public-portal/index.do) assessed in February 2018). Ingenuity® Pathway Analysis (IPA v2.3; http://www.ingenuity.com) and KOBAS (v.3.0; http://kobas.cbi.pku.edu.cn/) software was additionally used to highlight enriched canonical pathways modulations.

Data validation by reverse transcription quantitative polymerase chain reaction (RT-qPCR)

Thirty OPSCC (23 array-dependent and 7 independent samples) and four non-neoplastic oropharyngeal tissues were assessed to evaluate the *ORAOV1*, *CPT1A*, *PPFIA1* and *SHANK2* transcript levels by RT-qPCR (primer sequences are showed in supplementary Table S2). *HMBS* and *HPRT1* were selected as reference genes using geNorm algorithm (tested in combination with *ACTB*, *GAPDH* and *GUSB*) [29]. Robotic pipetting was carried out using QIAgility (QIAGEN) in a total volume of 12.5 uL containing Power SYBR Green PCR Master Mix (Applied Biosystems), 20 ng of cDNA and 200 nM of each primer. All samples were analyzed in duplicate. Normalized relative gene expression was obtained according to Pffafl method [30].

Statistical analysis

Fisher exact test was used to associate the clustering analysis and copy number alterations with clinicopathological features. Transcript levels were evaluated according to CNA and response to therapy using Mann–Whitney test. High expression levels were defined as at least two-fold in tumor tissues compared to normal tissues. Overall and disease-free survival probabilities were calculated using Kaplan-Meyer method and statistically compared with log rank. Multivariate analysis was applied using Cox proportional hazards including all variables with p value < 0.1 in the univariate analysis. Receiver Operating Characteristic (ROC) curve was applied to verify the predictive value of the variables in relation to response to treatment, defined by the Area Under the ROC Curve (AUC). Statistical analysis was performed with GraphPad Prismv5.0 (GraphPad Software Inc., La Jolla, CA) and SPSS v.21.0 (SPSS; Chicago, IL).

Results

HPV infection status and prognosis

Clinical and pathological features of the Brazilian patients included in this study are presented on Table 1. Fifty-eight percent (23/40 cases) of the advanced OPSCC were HPV-positive, being HPV16 the most prevalent subtype (19 cases; 1 HPV16/18; 1 HPV16/33; 1 HPV18 and 1 HPV33) (supplementary Table S1). Tonsil was the most common affected site by HPV (57%) (Table 1). Longer overall survival was observed for HPV-positive compared to HPV-negative cases ($p = 0.008$, supplementary Fig. S1A). Risk stratification based on HPV status, smoking history and tumor/node stage [7] was performed. High-risk patients (n = 15) showed shorter survival compared with low (n = 15) and intermediate-risk patients (n = 10) ($p = 0.013$, supplementary Fig. S1B).

Table 1
Patient distribution according to the demographic and clinical variables and 11q13 amplification status.

Characteristics	Number of patients (%), $N = 40$	11q13 amplification* (%), $N = 33$	p value
Gender			
Female	2 (5)	0/2 (0)	0.542
Male	38 (95)	11/31 (35.5)	
Age			
Median (interquartile range)	55.5 (47.8–62.3)		
< 60 years	26 (65)	6/20 (30)	0.714
≥ 60 years	14 (35)	5/13 (38.5)	
Anatomic site			
Base of the tongue	16 (40)	4/14 (28.6)	1.000
Soft palate	1 (2.5)		
Tonsil	23 (57.5)	6/19 (31.8)	
HPV infection status			
Negative	17 (42.5)	7/15 (46.7)	0.163
HPV16	19 (47.5)	4/18 (22.2)	
Other HPV (18/33)	2 (5)		
Co-infection (HPV16 and HPV18/33)	2 (5)		
Tobacco consumption			
No	9 (22.5)	1/8 (12.5)	0.218
Yes	31 (77.5)	10/25 (40)	
Alcohol consumption			
No	14 (35)	1/10 (10)	0.109
Yes	26 (65)	10/23 (43.5)	
Clinical T stage			
T1	1 (2.5)	7/23 (30.4)	0.696
T2	7 (17.5)		
T3	21 (52.5)		
T4	11 (27.5)	4/10 (40)	
Clinical N stage			
N0	11 (27.5)	2/9 (22.2)	0.681
N1	5 (12.5)	9/24 (37.5)	
N2	18 (45)		
N3	6 (15)		
Stratification risk*			
Low	15 (37.5)	4/20 (20)	0.065
Intermediate	10 (25)		
High	15 (37.5)	7/13 (53.8)	
Treatment response (RXT + CT)			
Stable disease/ progression	3 (9.7)	5/9 (55.6)	0.097
Partial	7 (22.6)		
Complete	21 (67.7)	4/19 (21.1)	
Not available#	9	5	
Clinical outcome			
Free of disease	25 (62.5)	8/19 (42.1)	0.278
Local recurrence	11 (27.5)	3/14 (21.4)	
Distant Metastasis	4 (10)		
Death			
No	26 (65)	5/20 (25)	0.270
Yes	14 (35)	6/13 (46.2)	

* Risk stratification according to Ang et al. [7].
\# Nine patients not submitted to concurrent radiotherapy and chemotherapy (CT-RT).
* Amplification in at least one gene located in 11q13; p value: Fisher exact test.

Genomic and transcriptomic analyses in OPSCCs

Fig. 1 summarizes the genomic and transcriptomic analysis approaches and the main results. Recurrent CNAs included losses of 8p11.23-p11.22, 14q11.1-q11.2, 15q11.2 and gains of 11q13.2 and 11q13.2-q13.3. After the comparison of all significant CNA ($p < 0.05$) with DGV database, only gains of 11q13.2 and 11q13.2-q13.3 were

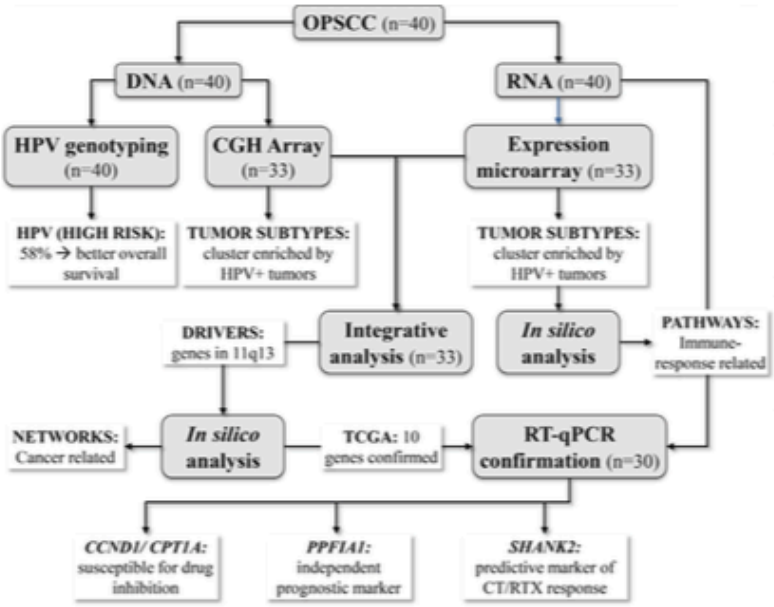

Fig. 1. Flowchart illustrating the methodologies and major findings of the study.

classified as new alterations. Genes and ORFs mapped in these regions are shown in supplementary Table S3.

CNA profiles revealed two clusters of samples, being the cluster 1 enriched by HPV-positive cases (cluster 1: 67% and cluster 2: 28%, $p = 0.038$), low/intermediate risk stratification (cluster 1: 77% and cluster 2: 25%, $p = 0.005$), complete chemoradiation response (cluster 1: 89% and cluster 2: 26%, $p = 0.004$) and lower mortality (cluster 1: 27% and cluster 2: 82%, $p = 0.008$) (Fig. 2A). CNA analysis according to HPV status revealed gains of 3q26.31 and 11q13.2 and losses of 9p21 exclusively detected in HPV-negative cases ($p < 0.05$, supplementary Table S4).

The gene expression analysis also demonstrated a cluster of samples enriched by HPV-positive cases (cluster 1: 87% and cluster 2: 44%, $p = 0.027$). A trend of low/intermediate risk enrichment (cluster 1: 85% and cluster 2: 50%, $p = 0.067$) was also observed in the cluster 1, which presented a partial overlapping with CNA cluster 1 (Fig. 2B). The comparison between these two gene expression clusters revealed 423 differentially expressed transcripts (supplementary Table S5). The pathways identified in HPV-enriched cluster were related to immune response, including interferon-gamma (IFNG) and programmed cell death-1 (PD-1) signaling (supplementary Table S6) (IPA and KOBAS v3.0). The main molecule predicted as activated by IPA upstream regulator analysis was *IFNG* (supplementary Table S7), which was detected as overexpressed in HPV-enriched cluster.

Integrated genomic and gene expression data

The integrative analysis using CONEXIC revealed 16 putative modulators mapped at 11q13 (*ANO1, CCND1, CPT1A, FADD, FGF3, FGF4, FGF19, IGHMBP2, MRGPRD, MRGPRF, MRPL21, ORAOV1, PPFIA1, PPP6R3, SHANK2* and *TPCN2*). Ten of 16 modulators showed statistically significant association between copy number gains and higher expression levels (Fig. 3A). Interestingly, copy number gains of these genes were associated with shorter overall survival (Fig. 3B). *In silico* analysis showed interactions between these 10 driver candidates and their association with cell cycle, cancer, organismal injury and abnormalities (supplementary Fig. S2). In addition, interactions with recognized cancer-related molecules (calmodulin and p21) were observed, being CCND1 and CPT1A susceptible for drug inhibition (arsenic trioxide and perhexiline, respectively) (drugbank: https://www.drugbank.ca/. Accession in February 2018).

Validation of candidate driver genes

The putative drivers were further tested using the TCGA database of OPSCC, confirming the increased gene expression levels and genomic copy number gains of 10 genes (supplementary Fig. S3A). Similarly to our data, these genomic alterations were also associated with shorter overall survival (supplementary Fig. S3B).

In 30 OPSCC samples with available tissues for RNA extraction RT-qPCR confirmed a significant association of genomic gains in 11q13 and overexpression of *ORAOV1*, *CPT1A*, *PPFIA1* and *SHANK2* genes (Fig. 4A). High expression levels of these genes were associated with shorter survival (Fig. 4B). The multivariate analysis revealed *PPFIA1* expression and HPV status as independent prognostic markers in advanced OPSCC (Table 2).

Among the four potential markers identified, significant decreased expression level of *SHANK2* was detected as predictive marker of complete response (Fig. 4C) and complete response without recurrence in at least five years (Fig. 4D) after concurrent radiotherapy and chemotherapy ($p = 0.010$ and $p = 0.012$, respectively). Compared to risk stratification (low, intermediate or high) or HPV status (negative or positive), *SHANK2* relative expression presented a higher performance in predicting concurrent CT-RT complete response (AUC = 0.733, AUC = 0.646, AUC = 0.833; respectively) (Fig. 4C) and complete response without recurrence (AUC = 0.659, AUC = 0.606, AUC = 0.811; respectively) (Fig. 4D).

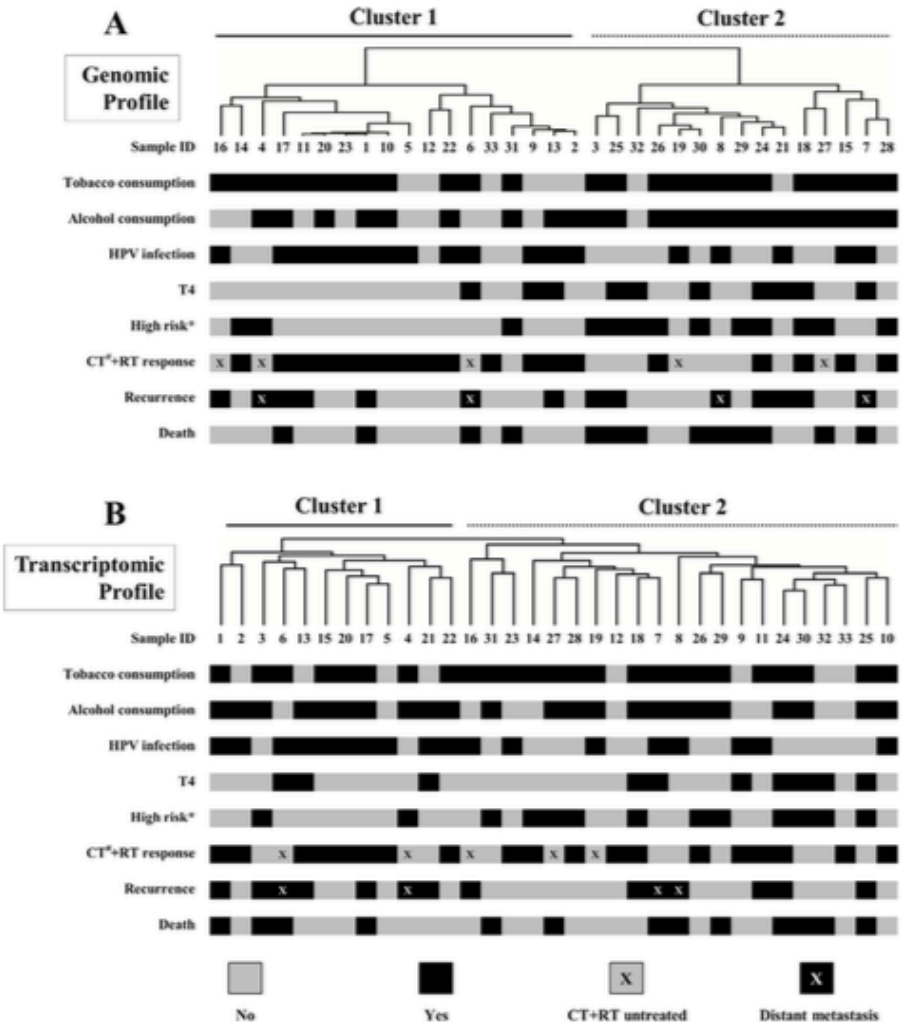

Fig. 2. Unsupervised hierarchical clustering analysis of (A) array-CGH and (B) global gene expression data according to the clinical features. CT: chemotherapy; RT: radiotherapy; *Risk stratification (Ang et al. [7]), #Patients submitted to induction chemotherapy followed by concurrent platinum/cetuximab and radiotherapy or only concurrent platinum/cetuximab and radiotherapy.

Discussion

This study was designed to reveal potential drivers in locally advanced OPSCC, aiming to identify new therapeutic targets and biomarkers. About half of OPSCC cases were positive for high-risk HPV, which was associated with improved survival, corroborating previous reports [7,8,10,11,31,32]. Among the HPV positive cases, 95% (22/23) were HPV16/18, similar to recent findings described in HNSCC [33]. Moreover, worse prognosis was observed for high-risk patients (according to HPV, TNM and smoking), supporting the hypothesis that the clinical course of HPV-positive tumors is modified by tobacco usage [7]. Curiously, the distribution of HPV infection was uneven between the patient recruitment centers. Tumors collected from Sao Paulo (the largest metropolis in Brazil) presented 77% (23/30) of positivity compared to 0% (0/10) in the Barretos cohort (an inland city in Brazil southeast). The Brazilian OPSCC patients were recently characterized as having 4% of HPV positivity compared to 59% and 31% in USA and Europe, respectively [34]. However, the predominance of HPV in Brazilian patients can drastically differ due to geographic heterogeneity of our population, which can vary socioeconomically, in alcohol-tobacco consumption and in sexual behavior.

Significant recurrent gains of 11q and two clusters based on CNA similarities were observed, being the cluster 1 enriched by HPV-positive tumors. Particularly, gains of 11q13.2 (*LRP5*, *PPP6R3* and *SAPS3* genes) were exclusively detected in HPV-negative samples. Amplification of 11q13 is commonly reported in different tumors, including HNSCC [20,27,35–38], particularly in HPV-negative tumors [39,40] and OPSCC [15]. Mutually exclusive amplifications of 11q13 (*CCND1* and *FADD*) and 11q22 (*BIRC2* and *YAP1*) were detected in HPV-negative HNSCC [27]. These data emphasizes that distinct molecular alterations

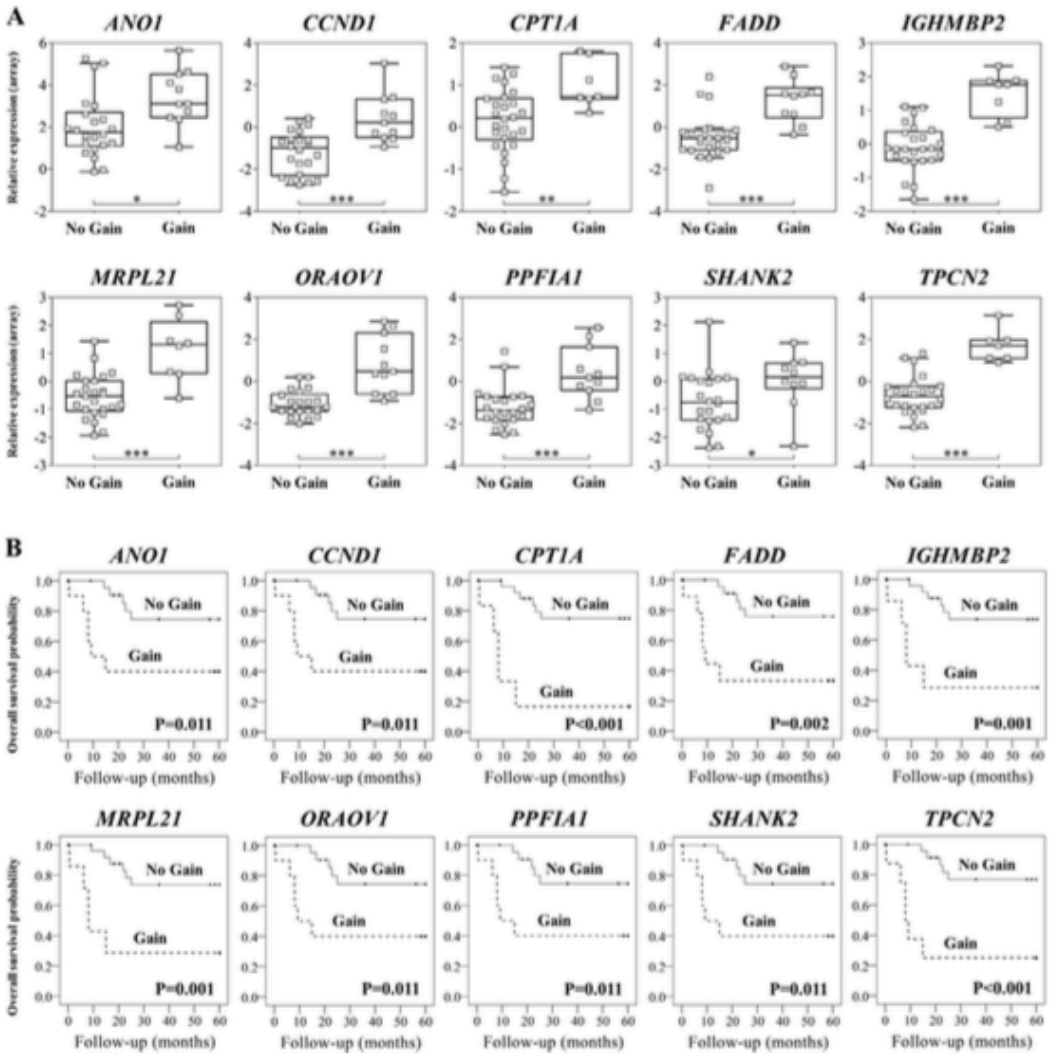

Fig. 3. A. Association of DNA copy number gains with expression levels of ten genes identified as driver candidates in OPSCC. B. Overall survival analysis according to DNA copy number gains in 10 candidate driver genes. $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$ (Mann-Whitney test); Overall survival analysis performed by Kaplan Meyer method with log rank test.

drive the oropharynx carcinogenesis according to the HPV status.

Large-scale gene expression analysis also showed distinct clusters according to HPV status. HPV-enriched cluster presented predominantly immune response-related genes, mainly involving negative regulators of T-cell immune function, as cytotoxic T-lymphocyte–associated antigen 4 (CTLA-4) and programmed death 1 (PD-1) and its ligand PD-L1 immune checkpoints pathway members. The PD-1/PD-L1 and CTLA-4, have shown promising clinical success as a cancer immunotherapy target [41]. Nonetheless, many factors influence in the response to immunotherapies, as the degree of tumor lymphocyte infiltration and the checkpoint proteins levels [13]. A previous study showed that 70% of HPV-positive HNSCCs presents PD-L1 positive immunostaining [42]. The authors reported increased transcript levels of *IFNG* in oropharyngeal tumors HPV + /PD-L1 + , suggesting that IFN-γ might induce PD-L1 in tumor cells. In our study, IFN-γ was highly overexpressed in the HPV-positive enriched cluster and was predicted by IPA as the one of the main upstream regulator. Tumor infiltrating lymphocytes overexpressing PD-1 presented better clinical outcome, mostly in head and neck tumors HPV-positives [43]. Furthermore, CD8 + tumor infiltrating lymphocytes showing high expression levels of PD-1 were found more frequently in HPV-negative tumors with compromised IFNG secretion. These patients presented worse prognosis. Conversely, HPV-positive tumors presenting low PD-1 expression in T cells were associated with better response to immunotherapy (anti-PD-1) and outcome [43]. In melanoma, the immunostaining positivity for MHC (Major Histocompatibility Complex) class II molecules was

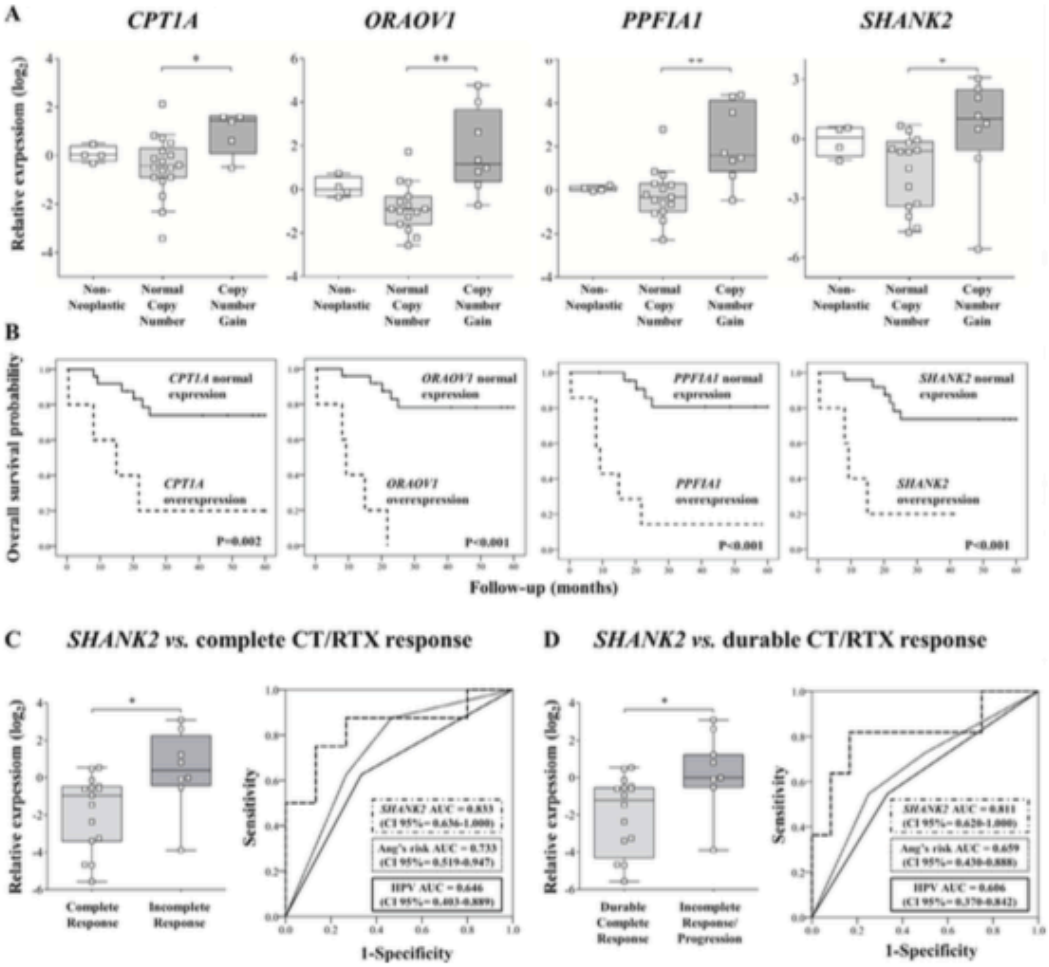

Fig. 4. A. Relative gene expression analysis using RT-qPCR in non-neoplastic samples and OPSCC with two copies (normal) and 11q13 copy number gains. B. Disease-free survival curves of tumors in normal range of expression and overexpression of the genes assessed by RT-qPCR (p value obtained from log rank test). C and D. Expression level of *SHANK2* (RT-qPCR) as a predictive marker of complete response (C) and durable complete response (D) to concurrent radiotherapy and chemotherapy, comparing HPV status and risk stratification (Ang et al. [7] $^{*}p < 0.05$; $^{**}p < 0.01$ (Mann Whitney non-parametric test); AUC: area under the ROC (receiver operating characteristic) curve.

associated with therapeutic response and better clinical outcome in anti-PD-1-treated patients [44]. In our study, overexpression of MHC class II genes was detected in HPV-related cluster (supplementary Table S4). Overall, our findings give additional evidence that PD-1/PD-L1 is a target for new therapeutic strategies in OPSCC, particularly in HPV-positive cases.

Correlation between CNAs mapped in 11q13.2-11q13.3 and corresponding genes expression levels was previously described in HNSCC [21]. In our study, 10 cancer-related genes (*ANO1, CCND1, CPT1A, FADD, IGHMBP2, MRPL21, ORAOV1, PPFIA1, SHANK2* and *TPCN2*) mapped in 11q13 showed copy number gains and overexpression. These findings were further validated using the TCGA, suggesting that other genes besides *CCND1* may confer an advantage to OPSCC cells [45]. Moreover, gains involving these genes were associated with worse survival in our study and in TCGA, thus supporting their relevance in OPSCC outcomes.

Further analysis using RT-qPCR confirmed an association among *ORAOV1, CPT1A, PPFIA1* and *SHANK2* overexpression and shorter survival. A meta-analysis comprising 15 publically available microarray gene expression data assessing 140 normal and 277 HNSCC samples revealed a list of 181 differentially expressed genes. *ANO1* and *FADD* (both mapped in 11q13) were overexpressed in at least 30% of the TCGA samples (528 HNSCC), and associated with higher risk of recurrence and death [46]. Gains of 11q13.2, including *CPT1A*, were also associated with worse survival in patients with esophageal carcinoma [47]. *CPT1A* (*Carnitine palmitoyl transferase 1A*) gene encodes a rate-limiting enzyme for fatty acid transport inside the mitochondria for β-oxidation. Fatty acid oxidation is essential for ATP production in metabolic stress conditions and particularly in cancer [48]. Enzymes involved in fatty acid oxidation, such as CPT1, are promising targets for

Table 2
Univariate and multivariate analysis of OPSCC overall survival.

Variables	Univariate analysis		Multivariate analysis	
	HR ($CI_{95\%}$)	*p* value	HR ($CI_{95\%}$)	*p* value
HPV infection				
No	1		1	
Yes	0.3 (0.1–1)	0.056	0.1 (0.02–0.56)	0.001
Ang's risk stratification				
Low/Intermediate	1			
High	3.7 (1–14.5)	0.056	–	–
CPT1A expression				
Normal range	1			
Overexpression	6 (1.7–21.9)	0.006	–	–
ORAOV1 expression				
Normal range	1			
Overexpression	18.3 (4.1–80.9)	< 0.001	–	–
PPFIA1 expression				
Normal range	1		1	
Overexpression	11.9 (3.2–44.1)	< 0.001	30.9 (5.3–179.6)	< 0.001
SHANK2 expression				
Normal range	1			
Overexpression	7.6 (2–28.1)	0.003	–	–

HR: hazard ratio; $CI_{95\%}$: confidence interval of 95%. *p* value: Cox proportional-hazards regression.

cancer therapy [48] and might be considered as a potential target in OPSCC. Interestingly, CPT1 enzyme can be pharmacologically inhibited by perhexiline [49].

Although increased expression levels of *ORAOV1* (*oral cancer overexpressed* 1) have been reported in several tumors [50–52], its biological function is not completely characterized. In oral carcinomas, reduced *ORAOV1* expression was associated with delayed cell growth *in vitro* and inhibition of tumor growth and angiogenesis *in vivo* [51]. A protective response to reactive oxygen species was related to *ORAOV1* overexpression, a common feature of many cancers [53], including HNSCC [54]. Despite the small sample size and number of events (10 patients died during a 5-year follow-up), *PPFIA1* overexpression was identified as an independent prognostic marker (independent of the HPV status and/or tobacco consumption). *PPFIA1* was reported as the highest up-regulated gene associated with copy number gains of 11q13 in HNSCC cell lines [55]. In breast cancer cells, the silencing of *PPFIA1* promotes inhibition of migration and invasion [56]. Overall, our data suggest that *PPFIA1* acts as an oncogene in OPSCC, although its function in these tumors requires further studies.

SHANK2 overexpression was associated with incomplete response to therapy, being able to distinguish treatment-unresponsive from responsive tumors more efficiently than HPV status. SHANK2 is a member of scaffold protein family involved in excitatory synapses in the central nervous system. In oral carcinomas, co-amplification and overexpression of *SHANK2* and *CTTN* was reported, suggesting a cooperative role of these proteins on tumor cell motility and invasiveness [57]. Amplification and overexpression of *SHANK2* was also associated with poor prognosis in esophageal squamous cell carcinoma [58]. In oral carcinomas, *SHANK2* was described as one of seven prognostic markers evaluated by DNA methylation [59]. These finding suggest that CNA and/or epigenetic mechanisms may result in *SHANK2* increased expression and consequent HNSCC aggressiveness. To our knowledge, the association between genes mapped in 11q13 and chemoradiation response was not previously reported. The role of these genes in oropharyngeal carcinogenesis, particularly in tumor aggressiveness and their ability to predict therapy response, should be further investigated.

Considering that only locally advanced OPSCC biopsies obtained from patients subsequently treated with specific therapeutic strategies were included in our study, the major limitation was the small sample size (n = 40). Nonetheless, with the addition of an independent sample set (n = 78) from TCGA (not treatment-standardized), we were able to confirm the potential clinical relevance of genes mapped in 11q13.

In conclusion, we described a higher frequency of HPV infection than previous Brazilian OPSCC reports, which varied greatly according to the geographic region and tobacco-alcohol consumption. HPV-positive tumors showed an immune response-related gene signature, which can have an impact on immunotherapy response prediction leading to a more precise treatment indication. In addition to *CCND1* previously described, novel driver genes including *ORAOV1*, *CPT1A*, *PPFIA1* and *SHANK2* genes mapped in 11q13 presented a role in the prognosis and chemoradiation response prediction, regardless of the HPV status.

Conflict of interest statement

None declared.

Acknowledgments

The authors would like to thank the Macromolecules Laboratory at the A.C. Camargo Cancer Center and Barretos Cancer Hospital. This study was supported by grants from the National Institute of Science and Technology in Oncogenomics (FAPESP #2008/57887–9 and CNPq #573589/08–9).

Authors' contributions

SRR and LPK conceived and designed the experiments; MCBF, MH and LARR conducted the experiments; MCBF, FAM, LARR, SAD analyzed the data; SRR and LPK contributed with reagents/materials; SRR, LPK and SAD supervised the study; CALP performed the histopathological evaluation; URN, ALC and LPK selected the cases and obtained the clinical data; MCBF, SAD and SRR wrote and edited the manuscript. All authors read and approved the final version of the manuscript.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.oraloncology.2018.06.010.

References

[1] Chaturvedi AK, Anderson WF, Lortet-Tieulent J, Curado MP, Ferlay J, Franceschi S, et al. Worldwide trends in incidence rates for oral cavity and oropharyngeal cancers. J Clin Oncol 2013;31:4550–9. http://dx.doi.org/10.1200/JCO.2013.50.3870.
[2] Panwar A, Batra R, Lydiatt WM, Ganti AK. Human papilloma virus positive oropharyngeal squamous cell carcinoma: a growing epidemic. Cancer Treat Rev 2014;40:215–9. http://dx.doi.org/10.1016/j.ctrv.2013.09.006.
[3] Simard EP, Torre LA, Jemal A. International trends in head and neck cancer incidence rates: differences by country, sex and anatomic site. Oral Oncol 2014;50:387–403. http://dx.doi.org/10.1016/j.oraloncology.2014.01.016.
[4] D'Souza G, Westra WH, Wang SJ, van Zante A, Wentz A, Kluz N, et al. Differences in the prevalence of Human Papillomavirus (HPV) in head and neck squamous cell cancers by sex, race, anatomic tumor site, and HPV detection method. JAMA Oncol 2016:169–77. http://dx.doi.org/10.1001/jamaoncol.2016.3067.
[5] Hussein AA, Helder MN, de Visscher JG, Leemans CR, Braakhuis BJ, de Vet HCW, et al. Global incidence of oral and oropharynx cancer in patients younger than 45 years versus older patients: a systematic review. Eur J Cancer 2017;82:115–27. http://dx.doi.org/10.1016/j.ejca.2017.05.026.
[6] Shield KD, Ferlay J, Jemal A, Sankaranarayanan R, Chaturvedi AK, Bray F, et al. The global incidence of lip, oral cavity, and pharyngeal cancers by subsite in 2012. CA Cancer J Clin 2017;67:51–64. http://dx.doi.org/10.3322/caac.21384.
[7] Ang KK, Harris J, Wheeler R, Weber R, Rosenthal DI, Nguyen-Tân PF, et al. Human papillomavirus and survival of patients with oropharyngeal cancer. N Engl J Med 2010;363:24–35. http://dx.doi.org/10.1056/NEJMoa0912217.
[8] Psyrri A, Rampias T, Vermorken JB. The current and future impact of human papillomavirus on treatment of squamous cell carcinoma of the head and neck. Ann Oncol 2014;25:2101–15. http://dx.doi.org/10.1093/annonc/mdu265.
[9] Urban D, Corry J, Rischin D. What is the best treatment for patients with human papillomavirus-positive and -negative oropharyngeal cancer? Cancer 2014;120:1462–70. http://dx.doi.org/10.1002/cncr.28595.

[10] Fakhry C, Westra WH, Wang SJ, van Zante A, Zhang Y, Rettig E, et al. The prognostic role of sex, race, and human papillomavirus in oropharyngeal and non-oropharyngeal head and neck squamous cell cancer. Cancer 2017;123:1566–75. http://dx.doi.org/10.1002/cncr.30353.

[11] Lacau St Guily J, Rousseau A, Baujat B, Périé S, Schultz P, Barry B, et al. Papillophar group. Oropharyngeal cancer prognosis by tumor HPV status in France: the multicentric Papillophar study. Oral Oncol 2017;67:29–36. http://dx.doi.org/10.1016/j.oraloncology.2017.01.012.

[12] Chera BS, Amdur RJ. Current status and future directions of treatment deintensification in human papilloma virus-associated oropharyngeal squamous cell carcinoma. Semin Radiat Oncol 2018;28:27–34. http://dx.doi.org/10.1016/j.semradonc.2017.08.001.

[13] Leemans CR, Snijders PJF, Brakenhoff RH. The molecular landscape of head and neck cancer. Nat Rev Cancer 2018. http://dx.doi.org/10.1038/nrc.2018.11.

[14] O'Sullivan B, Huang SH, Siu LL, Waldron J, Zhao H, Perez-Ordonez B, et al. Deintensification candidate subgroups in human papillomavirus-related oropharyngeal cancer according to minimal risk of distant metastasis. J Clin Oncol 2013;31:543–50. http://dx.doi.org/10.1200/JCO.2012.44.0164.

[15] Klussmann JP, Mooren JJ, Lehnen M, Claessen SM, Stenner M, Huebbers CU, et al. Genetic signatures of HPV-related and unrelated oropharyngeal carcinoma and their prognostic implications. Clin Cancer Res 2009;15:1779–86. http://dx.doi.org/10.1158/1078-0432.CCR-08-1463.

[16] Jung AC, Briolat J, Millon R, de Reyniès A, Rickman D, Thomas E, et al. Biological and clinical relevance of transcriptionally active human papillomavirus (HPV) infection in oropharynx squamous cell carcinoma. Int J Cancer 2010;126:1882–94. http://dx.doi.org/10.1002/ijc.24911.

[17] 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. http://dx.doi.org/10.1038/nrclinonc.2014.192.

[18] Papillon-Cavanagh S, Lu C, Gayden T, Mikael LG, Bechet D, Karamboulas C, et al. Impaired H3K36 methylation defines a subset of head and neck squamous cell carcinomas. Nat Genet 2017;49:180–5. http://dx.doi.org/10.1038/ng.3757.

[19] Boscolo-Rizzo P, Furlan C, Lupato V, Polesel J, Fratta E. Novel insights into epigenetic drivers of oropharyngeal squamous cell carcinoma: role of HPV and lifestyle factors. Clin Epigenetics 2017;109:24. http://dx.doi.org/10.1186/s13148-017-0424-5.

[20] Pickering CR, Zhang J, Yoo SY, Bengtsson L, Moorthy S, Neskey DM, et al. Integrative genomic characterization of oral squamous cell carcinoma identifies frequent somatic drivers. Cancer Discov 2013;3:770–81. http://dx.doi.org/10.1158/2159-8290.CD-12-0537.

[21] Xu C, Liu Y, Wang P, Fan W, Rue TC, Upton MP, et al. Integrative analysis of DNA copy number and gene expression in metastatic oral squamous cell carcinoma identifies genes associated with poor survival. Mol Cancer 2010;9:143–54. http://dx.doi.org/10.1186/1476-4598-9-143.

[22] Etemadmoghadam D, deFazio A, Beroukhim R, Mermel C, George J, Getz G, et al. AOCS Study Group. Integrated genome-wide DNA copy number and expression analysis identifies distinct mechanisms of primary chemoresistance in ovarian carcinomas. Clin Cancer Res 2009;15:1417–27. http://dx.doi.org/10.1158/1078-0432.CCR-08-1564.

[23] Busso-Lopez AF, Marchi FA, Kuasne H, Scapulatempo-Neto C, Trindade-Filho JC, de Jesus CM, et al. Genomic profiling of human penile carcinoma predicts worse prognosis and survival. Cancer Prev Res (Phila) 2015;8:149–56. http://dx.doi.org/10.1158/1940-6207.CAPR-14-0284.

[24] Marchi FA, Martins DC, Barros-Filho MC, Kuasne H, Busso Lopes AF, Brentani H, et al. Multidimensional integrative analysis uncovers driver candidates and biomarkers in penile carcinoma. Sci Rep 2017;7:6707–17. http://dx.doi.org/10.1038/s41598-017-06659-1.

[25] Cirilo PD, Marchi FA, Barros Filho Mde C, Rocha RM, Domingues MA, Jurisica I, et al. An integrative genomic and transcriptomic analysis reveals potential targets associated with cell proliferation in uterine leiomyomas. PLoS One 2013;8:e57901. http://dx.doi.org/10.1371/journal.pone.0057901.

[26] Chudasama P, Mughal SS, Sanders MA, Hübschmann D, Chung I, Deeg KI, et al. Integrative genomic and transcriptomic analysis of leiomyosarcoma. Nat Commun. 2018;9:144. http://dx.doi.org/10.1038/s41467-017-02602-0.

[27] Cancer Genome Atlas Network. Comprehensive genomic characterization of head and neck squamous cell carcinomas. Nature 2015;517:576–82. http://dx.doi.org/10.1038/nature14129.

[28] Akavia UD, Litvin O, Kim J, Sanchez-Garcia F, Kotliar D, Causton HC, et al. An integrated approach to uncover drivers of cancer. Cell 2010;143:1005–17. http://dx.doi.org/10.1016/j.cell.2010.11.013.

[29] Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, et al. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol 2002;3. RESEARCH0034.

[30] Pfaffl MW. A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res 2001;29:e45.

[31] Lindel K, Beer KT, Laissue J, Greiner RH, Aebersold DM. Human Papillomavirus positive squamous cell carcinoma of the oropharynx. A radiosensitive subgroup of head and neck carcinoma. Cancer 2001;92:805–13.

[32] Fischer CA, Zlobec I, Green E, Probst S, Storck C, Lugli A, et al. Is the improved prognosis of p16 positive oropharyngeal squamous cell carcinoma dependent of the treatment modality? Int J Cancer 2010;126:1256–62. http://dx.doi.org/10.1002/ijc.24842.

[33] de Martel C, Plummer M, Vignat J, Franceschi S. Worldwide burden of cancer attributable to HPV by site, country and HPV type. Int J Cancer 2017;141:664–70. http://dx.doi.org/10.1002/ijc.30716.

[34] Anantharaman D, Abedi-Ardekani B, Beachler DC, Gheit T, Olshan AF, Wisniewski K, et al. Geographic heterogeneity in the prevalence of human papillomavirus in head and neck cancer. Int J Cancer 2017;140:1968–75. http://dx.doi.org/10.1002/ijc.30608.

[35] Meredith SD, Levine PA, Burns JA, Gaffey MJ, Boyd JC, Weiss LM, et al. Chromosome 11q13 amplification in head and neck squamous cell carcinoma. Association with poor prognosis. Arch Otolaryngol Head Neck Surg 1995;121:790–4.

[36] Ambatipudi S, Gerstung M, Gowda R, Pal P, Borges AM, Schäffer AA, et al. Genomic profiling of advanced-stage oral cancers reveals chromosome 11q alterations as markers of poor clinical outcome. PLoS One 2011;6:e17250. http://dx.doi.org/10.1371/journal.pone.0017250.

[37] Ambrosio EP, Silveira CG, Drigo SA, Sacomano Vde S, Molck MC, Rocha RM, et al. Chromosomal imbalances exclusively detected in invasive front area are associated with poor outcome in laryngeal carcinomas from different anatomical sites. Tumor Biol 2013;34:3015–26. http://dx.doi.org/10.1007/s13277-013-0866-0.

[38] Morris LG, Chandramohan R, West L, Zehir A, Chakravarty D, Pfister DG, et al. The molecular landscape of recurrent and metastatic head and neck cancers: insights from a precision oncology sequencing platform. JAMA Oncol 2017;3:244–55. http://dx.doi.org/10.1001/jamaoncol.2016.1790.

[39] Ragin CC, Taioli E, Weissfeld JL, White JS, Rossie KM, Modugno F, et al. 11q13 amplification status and human papillomavirus in relation to p16 expression defines two distinct etiologies of head and neck tumors. Br J Cancer 2006;95:1432–8.

[40] Seiwert TY, Zuo Z, Keck MK, Khattri A, Pedamallu CS, Stricker T, et al. Integrative and comparative genomic analysis of HPV-positive and HPV-negative head and neck squamous cell carcinomas. Clin Cancer Res 2015;21:632–41. http://dx.doi.org/10.1158/1078-0432.CCR-13-3310.

[41] Gotwals P, Cameron S, Cipolletta D, Cremasco V, Crystal A, Hewes B, et al. Prospects for combining targeted and conventional cancer therapy with immunotherapy. Nat Rev Cancer 2017;17:286–301. http://dx.doi.org/10.1038/nrc.2017.17.

[42] Lyford-Pike S, Peng S, Young GD, Taube JM, Westra WH, Akpeng B, et al. Evidence for a role of the PD-1:PD-L1 pathway in immune resistance of HPV-associated head and neck squamous cell carcinoma. Cancer Res 2013;73:1733–41. http://dx.doi.org/10.1158/0008-5472.CAN-12-2384.

[43] Kansy BA, Concha-Benavente F, Srivastava RM, Jie HB, Shayan G, Lei Y, et al. PD-1 Status in CD8(+) T Cells associates with survival and anti-PD-1 therapeutic outcomes in head and neck cancer. Cancer Res 2017;77:6353–64. http://dx.doi.org/10.1158/0008-5472.CAN-16-3167.

[44] Johnson DB, Estrada MV, Salgado R, Sanchez V, Doxie DB, Opalenik SR, et al. Melanoma-specific MHC-II expression represents a tumor-autonomous phenotype and predicts response to anti-PD-1/PD-L1 therapy. Nat Commun 2016;7:10582. http://dx.doi.org/10.1038/ncomms10582.

[45] Wilkerson PM, Reis-Filho JS. The 11q13-q14 amplicon: clinicopathological correlations and potential drivers. Genes Chromosomes Cancer 2013;52:333–55. http://dx.doi.org/10.1002/gcc.22037.

[46] Reddy RB, Bhat AR, James BL, Govindan SV, Mathew R, Ravindra DR, et al. Meta-analyses of microarray datasets identifies ANO1 and FADD as prognostic markers of head and neck cancer. PLoS One 2016;11:e0147409. http://dx.doi.org/10.1371/journal.pone.0147409.

[47] Shi ZZ, Liang JW, Zhan T, Wang BS, Lin DC, Liu SG, et al. Genomic alterations with impact on survival in esophageal squamous cell carcinoma identified by array comparative genomic hybridization. Genes Chromosomes Cancer 2011;50:518–26. http://dx.doi.org/10.1002/gcc.20875.

[48] Carracedo A, Cantley LC, Pandolfi PP. Cancer metabolism: fatty acid oxidation in the limelight. Nat Rev Cancer 2013;13:227–32. http://dx.doi.org/10.1038/nrc3483.

[49] Kennedy JA, Unger SA, Horowitz JD. Inhibition of carnitine palmitoyltransferase-1 in rat heart and liver by perhexiline and amiodarone. Biochem Pharmacol 1996;52:273–80.

[50] Jiang L, Zeng X, Wang Z, Ji N, Zhou Y, Liu X, et al. Oral cancer overexpressed 1 (ORAOV1) regulates cell cycle and apoptosis in cervical cancer HeLa cells. Mol Cancer 2010;9:20. http://dx.doi.org/10.1186/1476-4598-9-20.

[51] Jiang L, Zeng X, Yang H, Wang Z, Shen J, Bai J, et al. Oral cancer overexpressed 1 (ORAOV1): a regulator for the cell growth and tumor angiogenesis in oral squamous cell carcinoma. Int J Cancer 2008;123:1779–86. http://dx.doi.org/10.1002/ijc.23734.

[52] Li M, Cui X, Shen Y, Dong H, Liang W, Chen Y, et al. ORAOV1 overexpression in esophageal squamous cell carcinoma and esophageal dysplasia: a possible biomarker of progression and poor prognosis in esophageal carcinoma. Hum Pathol 2015;46:707–15. http://dx.doi.org/10.1016/j.humpath.2015.01.009.

[53] Zhai C, Li Y, Mascarenhas C, Lin Q, Li K, Vyrides I, et al. The function of ORAOV1/LTO1, a gene that is overexpressed frequently in cancer: essential roles in the function and biogenesis of the ribosome. Oncogene 2014;33:484–94. http://dx.doi.org/10.1038/onc.2012.604.

[54] Choudhari SK, Chaudhary M, Gadbail AR, Sharma A, Tekade S. Oxidative and antioxidative mechanisms in oral cancer and precancer: a review. Oral Oncol 2014;50:10–8. http://dx.doi.org/10.1016/j.oraloncology.2013.09.011.

[55] Tan KD, Zhu Y, Tan HK, Rajasegaran V, Aggarwal A, Wu J, et al. Amplification and overexpression of PPFIA1, a putative 11q13 invasion and suppressor gene, in head and neck squamous cell carcinoma. Genes Chromosomes Cancer 2008;47:353–62. http://dx.doi.org/10.1002/gcc.20539.

[56] Astro V, Asperti C, Cangi MG, Doglioni C, de Curtis I. Liprin-α1 regulates breast cancer cell invasion by affecting cell motility, invadopodia and extracellular matrix degradation. Oncogene 2011;30:1841–9. http://dx.doi.org/10.1038/onc.2010.562.

[57] Freier K, Sticht C, Hofele C, Flechtenmacher C, Stange D, Puccio L, et al. Recurrent coamplification of cytoskeleton-associated genes EMS1 and SHANK2 with CCND1

in oral squamous cell carcinoma. Genes Chromosomes Cancer 2006;45:118–25.
[58] Qin HD, Liao XY, Chen YB, Huang SY, Xue WQ, Li FF, et al. Genomic characterization of esophageal squamous cell carcinoma reveals critical genes underlying tumorigenesis and poor prognosis. Am J Hum Genet 2016;98:709–27. http://dx.doi.org/10.1016/j.ajhg.2016.02.021.
[59] Shen S, Wang G, Shi Q, Zhang R, Zhao Y, Wei Y, et al. Seven-CpG-based prognostic signature coupled with gene expression predicts survival of oral squamous cell carcinoma. Clin Epigenetics 2017;9:88–98. http://dx.doi.org/10.1186/s13148-017-0392-9.

Anexo

ANEXO A – Artigo 1

Martins *et al. Virology Journal* (2016) 13:138
DOI 10.1186/s12985-016-0594-3

Virology Journal

RESEARCH Open Access

HPV genotype distribution in Brazilian women with and without cervical lesions: correlation to cytological data

Toni Ricardo Martins[1,2*], Cristina Mendes de Oliveira[1], Luciana Reis Rosa[1], Cristiane de Campos Centrone[1], Célia Luiza Regina Rodrigues[1], Luisa Lina Villa[3,4] and José Eduardo Levi[1,4]

Abstract

Background: Human Papillomavirus (HPV) genotype distribution varies according to the method of assessment and population groups. This study analyzed type-specific HPV infections among women ranging from 14–95 years old, displaying normal and abnormal cytology, from São Paulo and Barretos cities, Brazil.

Methods: Women found positive for High Risk-HPVs DNA by either the Hybrid Capture 2 (HC2) or Cobas HPV Test ($n = 431$) plus a random sample of 223 negative by both assays and 11 samples with indeterminate results, totalizing 665 samples, were submitted to HPV detection by the PapilloCheck test. Cytological distribution included 499 women with a cytological result of Negative for Intraepithelial Lesion or Malignancy and 166 with some abnormality as follows: 54 Atypical Squamous Cells of Undetermined Significance; 66 Low-Grade Squamous Intraepithelial Lesion; 43 High-Grade Squamous Intraepithelial Lesion and 3 (0.5 %) Invasive Cervical Cancer.

Results: From the 323 samples (48.6 %) that had detectable HPV-DNA by the PapilloCheck assay, 31 were HPV negative by the cobas HPV and HC2 assays. Out of these 31 samples, 14 were associated with HR-HPVs types while the remaining 17 harbored exclusively low-risk HPVs. In contrast, 49 samples positive by cobas HPV and HC 2 methods tested negative by the PapilloCheck assay (19.8 %). Overall, the most frequent HR-HPV type was HPV 16 (23.2 %), followed by 56 (21.0 %), 52 (8.7 %) and 31 (7.7 %) and the most frequent LR-HPV type was HPV 42 (12.1 %) followed by 6 (6.2 %). Among the HR-HPV types, HPV 56 and 16 were the most frequent types in NILM, found in 19.1 and 17.7 % of the patients respectively while in HSIL and ICC cases, HPV 16 was the predominant type, detected in 37.2 and 66.7 % of these samples.

Conclusions: In the population studied, HPV 16 and 56 were the most frequently detected HR-HPV types. HPV 56 was found mainly in LSIL and NILM suggesting a low oncogenic potential. HPV 16 continues to be the most prevalent type in high-grade lesions whereas HPV 18 was found in a low frequency both in NILM and abnormal smears. Surveillance of HPV infections by molecular methods is an important tool for the development and improvement of prevention strategies.

Keywords: Cervical Cancer, Cytology, HPV, HPV types, PapilloCheck

* Correspondence: trmartins@usp.br
[1]Institute of Tropical Medicine, Universidade de São Paulo, Virology Laboratory / LIM 52, Dr. Enéas Carvalho de Aguiar, 470, CEP: 05403-000 São Paulo, SP, Brazil
[2]Department of Infectious Diseases, Universidade de São Paulo School of Medicine, São Paulo, Brazil
Full list of author information is available at the end of the article

ANEXO B – Artigo 2

ONCOLOGY LETTERS

Characterization of topoisomerase II α and minichromosome maintenance protein 2 expression in anal carcinoma

CRISTOVAM SCAPULATEMPO-NETO[1], CARLOS VEO[2], JOSÉ HUMBERTO T.G. FREGNANI[1,3], ADRIANA LORENZI[3], ALLINI MAFRA[3], ARMANDO G.F. MELANI[2], EDGAR ANTONIO ALEMÁN LOAIZA[2], LUCIANA ALBINA REIS ROSA[4], CRISTINA MENDES DE OLIVEIRA[4], JOSÉ EDUARDO LEVI[4] and ADHEMAR LONGATTO-FILHO[3,5,6,7]

Departments of [1]Pathology (CSN) and [2]Surgery; [3]Molecular Oncology Research Center, Barretos Cancer Hospital, Pio XII Foundation, Barretos, SP 14784-400; [4]Laboratory of Virology, Institute of Tropical Medicine São Paulo, University of São Paulo, São Paulo, SP 05403-000; [5]Laboratory of Medical Investigation 14, Faculty of Medicine, University of São Paulo, São Paulo, SP 1246-903, Brazil; [6]Life and Health Sciences Research Institute (ICVS), School of Health Sciences, University of Minho, 4710-057 Braga; [7]ICVS/3B's - PT Government Associate Laboratory, 4806-909 Caldas das Taipas/Guimarães, Portugal

Received November 30, 2015; Accepted August 25, 2016

DOI: 10.3892/ol.2017.5650

Abstract. The present study aimed to ascertain the significance of topoisomerase II α (TOP2A) and minichromosome maintenance protein (MCM) 2 expression in anal carcinoma. A total of 75 anal lesions were retrieved from the files of the Department of Pathology of Barretos Cancer Hospital (Barretos, Brazil) in order to verify the human papillomavirus (HPV) statuses of these lesions and characterize the immunohistochemical expression levels of TOP2A and MCM2 in anal carcinoma, as these are important markers for cervical HPV-induced lesions; their expression was also compared with respect to p16 and Ki-67. The vast majority of the cases tested positive for HPV16 (84%); 1 case tested positive for both HPV16 and HPV18. Positive HPV16 status was more frequent in early stages than in advanced stages (P=0.008). Positive immunohistochemical reactivity for MCM2 and TOP2A protein was observed in 71.6 and 100% of cases, respectively. Positive reactivity for p16 was significantly associated (P=0.001) with histological grade, and was more commonly expressed in squamous cell carcinoma than adenocarcinomas. HPV16 was strongly associated with positive p16 protein expression (76.6%). However, the high expression of Ki-67 combined with the high expression of p16 was predominantly observed in Stage III-IV cases. MCM2, TOP2A, p16 and Ki-67 exhibited intense positive staining in the anal lesions, indicating that these markers were significantly and constantly expressed in anal carcinoma.

Correspondence to: Dr Adhemar Longatto-Filho, Laboratory of Medical Investigation 14, Faculty of Medicine, University of São Paulo, 455 Dr. Arnaldo Avenue, São Paulo, SP 01246-903, Brazil
E-mail: longatto16@hotmail.com

Key words: anal carcinoma, p16, Ki-67, minichromosome maintenance protein 2, topoisomerase II α, human papillomavirus

Introduction

Human papillomavirus (HPV) is thought to be the carcinogenic agent responsible for all cases of cervical cancer, and for carcinomas of other anatomical sites, including anal carcinomas. Currently, >85% of anal carcinomas are thought to be associated with oncogenic HPV, and, among all high-risk HPV types, type 16 is recognized as the most common, with prevalence rates estimated at ~70% of all cases (1). In contrast to cervical carcinoma, the incidence of anal carcinoma is gradually increasing, accounting for ~2.2% of all gastrointestinal tract malignancies in the United States, with 6,230 cases newly diagnosed each year (2). Similar data have been also documented in other countries, such as Denmark (3). Populations at increased risk include women with cervical HPV-related neoplasia, immunosuppressed transplant patients and human immunodeficiency virus (HIV)-positive individuals. Risk factors for anal carcinoma acquisition also comprise history of smoking, history of condylomata (due to HPV exposition), and history of anal intercourse, indicating HPV infection in the anal canal (4). Notably, anal carcinoma is most frequent in men who have sex with men (MSM), who are ~20 times more likely than heterosexual men to develop the disease. Furthermore, MSM with HIV are at increased risk for anal cancer development (5). The estimated rate of anal carcinoma among HIV-positive persons is 174/100,000, as compared to an incidence of 2/100,000 among the HIV-negative population. MSM represent ~75% of the population at risk for the development of anal carcinoma (6).

ANEXO C – Artigo 3

MOLECULAR MEDICINE REPORTS 14: 3791-3797, 2016

Low mutation percentage of *KRAS* and *BRAF* genes in Brazilian anal tumors

LUCAS TADEU BIDINOTTO[1,2], CARLOS A R VÉO[1], EDGAR ALEMAN LOAIZA[1], ALESSANDRA PAULINO SANTOS DE FRANÇA[1], ADRIANA TARLA LORENZI[1], LUCIANA ALBINA REIS ROSA[3], CRISTINA MENDES DE OLIVEIRA[3], JOSÉ EDUARDO LEVI[3], CRISTOVAM SCAPULATEMPO-NETO[1,4], ADHEMAR LONGATTO-FILHO[1,5,6] and RUI MANUEL REIS[1,5,6]

[1]Molecular Oncology Research Center, Barretos Cancer Hospital, Barretos, São Paulo 14784 400; [2]Barretos School of Health Sciences, Dr. Paulo Prata-FACISB, Barretos, São Paulo 14785-002; [3]Laboratory of Virology, Institute of Tropical Medicine, University of São Paulo, São Paulo 05403 000; [4]Department of Pathology, Barretos Cancer Hospital, Barretos, São Paulo 14784 400, Brazil; [5]Life and Health Sciences Research Institute (ICVS), School of Health Sciences, University of Minho, Braga 4710-057; [6]ICVS/3B's-PT Government Associate Laboratory, Braga 4710-057, Portugal

Received January 26, 2016; Accepted June 1, 2016

DOI: 10.3892/mmr.2016.5684

Abstract. Anal cancer is a rare type of digestive tract disease, which has had a crescent incidence in a number of regions. Carcinomas are most frequently found, with squamous cell carcinoma (SCC) comprising ~95% of all anal tumors. The major risk factor for development of this type of tumor is human papillomavirus (HPV) infection. However, previous studies have identified patients with anal cancer that are HPV-/p16-and observed that they have a poorer outcome compared with HPV+/p16+ patients. This suggests that molecular profile may drive anal cancer progression. The aim of the present study was to evaluate the mutational status of two important oncogenes, *KRAS* and *BRAF*, in a series of anal cancer lesions. Resected tumors of the anal canal (n=43) were evaluated, nine of these were high-grade squamous intra-epithelial lesion cases (HSIL), 11 were adenocarcinomas, and 23 SCCs. Direct sequencing of KRAS proto-oncogene, GTPase (*KRAS*; codons 12 and 13) and B-Raf proto-oncogene, serine/threonine kinase (*BRAF*; codon 600) was performed and associated with patient clinicopathological and molecular features. There was a trend of poorer prognosis of adenocarcinoma compared with HSIL and SCC. Analysis indicated one SCC patient (2.3%) exhibited a *KRAS* p.G13D mutation, and one adenocarcinoma patient (2.3%) exhibited a *BRAF* p.V600E mutation. It was observed that, these mutations are rare in anal tumors, and certain patients may be at a disadvantage using targeted therapies based on *KRAS* and *BRAF* mutational status. As there is a low mutation percentage in SCCs, adenocarcinomas and HSIL, there may exist other underlying molecular alterations that result in anal cancer development, which require further elucidation.

Correspondence to: Dr Rui Manuel Reis, Molecular Oncology Research Center, Barretos Cancer Hospital, 1331 Rua Antenor Duarte Villela, Barretos, São Paulo 14784 400, Brazil
E-mail: ruireis.hcb@gmail.com

Key words: anal cancer, squamous cell carcinoma, adenocarcinoma, high-grade squamous intra-epithelial lesion, *KRAS, BRAF*

Introduction

Anal cancer is a rare type of digestive tract disease, which has had an increasing incidence in a number of regions (1-3). It is estimated a total of >7,200 new cases were diagnosed in the United States in 2015, with ~1,000 anal cancer-associated mortalities (4). Tumors in this site are classified, according to World Health Organization (WHO), as intraepithelial neoplasias, carcinomas and carcinoid tumors (5). Carcinomas are most frequently identified, with squamous cell carcinoma (SCC) comprising~95% of all the anal tumors (6), and ~5% of the lesions are adenocarcinomas (7). The age-standardized incidence is <1/100,000 people, and the mortality is 0.2/100,000 (1). In men who practice anal receptive intercourse, the incidence of anal cancer increases up to 35/100,000 (5,8). This is predominantly due to increased risk of human papillomavirus (HPV) infection (1), HPV16 is most frequently observed in anal SCC (9). HPV infection leads to intraepithelial neoplasia that progresses from low-grade to high-grade dysplasia and, finally, to invasive cancer. The regression of high-grade dysplastic lesions is rare (5). HPV infection results in high expression of cyclin-dependent kinase inhibitor 2A (p16), and disrupts the association between retinoblastoma protein and the E2F family of transcription factors, ultimately leading to cellular proliferation (10). Recently, it was demonstrated that a high frequency of women with cervical cancer also have infection of the anal canal by HPV16 (11). In addition to HPV infection, other known risk factors of anal cancer are immunodeficiency due to human immunodeficiency virus seropositivity, low cluster of differentiation 4 T cell count,

ANEXO D – Aprovação CEP

HOSPITAL DAS CLÍNICAS DA FACULDADE DE MEDICINA DA USP - HCFMUSP

Plataforma Brasil

PARECER CONSUBSTANCIADO DO CEP

DADOS DO PROJETO DE PESQUISA

Título da Pesquisa: AVALIAÇÃO DO PERFIL GENÉTICO DO HPV16 E SEU SÍTIO DE INTEGRAÇÃO EM CÉLULAS EPITELIAIS NORMAIS E NEOPLÁSICAS DA CÉRVIX E TONSILAS

Pesquisador: JOSÉ EDUARDO LEVI

Área Temática:

Versão: 2

CAAE: 01353812.0.0000.0068

Instituição Proponente: HOSPITAL DAS CLINICAS DA FACULDADE DE MEDICINA DA U S P

DADOS DO PARECER

Número do Parecer: 173.181

Data da Relatoria: 05/12/2012

Apresentação do Projeto:

Trata-se do segundo relato do projeto de pesquisa acima especificado.

O papilomavírus humano (HPV) constitui o principal agente etiológico de alguns tipos de tumores humanos, responsável pelo desenvolvimento de quase 100% dos tumores de colo uterino, e 25% dos carcinomas de cabeça e pescoço. Os carcinomas de colko uterino apresentam a terceira maior incidência global de câncer em mulheres. Os carcinomas de cabeça e pescoço (CCP) representam um grupo heterogêneo de tumores agrupados devido a sua localização anatômica. Acometem mais frequentemente homens com idade superior a 50 anos e histórico de tabagismo e etilismo crônicos. Dentre os HPV de alto risco, o HPV16 é o tipo mais detectado nos tumores de orofaringe, (90 a 95%), e em mais de 50% dos casos de carcinoma de colo uterino. Em infecções persistentes, o genoma do HPV permanece na forma epissomal, isto é, não integrado ao genoma humano. No entanto, na maioria das lesões cancerígenas, o DNA de HPV de

Endereço: Rua Ovídio Pires de Campos, 225 5º andar
Bairro: Cerqueira Cesar **CEP:** 05.403-010
UF: SP **Município:** SAO PAULO
Telefone: (11)2661-6442 **Fax:** (11)2661-6442 **E-mail:** marcia.carvalho@hc.fm.usp.br

alto risco
é frequentemente quebrado e integrado ao genoma da célula hospedeira, alterando a expressão dos genes virais e de genes do hospedeiro. Os principais genes expressos pelos vírus são as oncoproteínas E6 e E7, necessárias para a replicação viral, estimulando a proliferação e atuando na transformação maligna e na manutenção dos tumores. A integração pode levar tanto a ativação, quanto a inativação de genes celulares envolvidos no processo tumoral. As alterações genéticas quase sempre ocorrem como resultado à exposição excessiva ao tabaco, álcool e outros elementos carcinogênicos que desencadeiam danos irreversíveis ao DNA, ocasionando alterações nas regiões codificadoras e regulatórias de genes supressores tumorais e oncogenes. O acúmulo destas mutações possibilita que as células adquiram características tumorais, incluindo resistência à morte programada (apoptose), adquirindo características de proliferação descontrolada. Tendo em vista que o HPV16 é o
principal agente etiológico de carcinomas de colo uterino, e também de uma parte dos tumores de orofaringe, principalmente de tonsilas palatinas, os mecanismos que regulam o tropismo desses HPV por este tipo de epitélio ainda não foram elucidados. Sendo assim, tentativas de se identificar regiões susceptíveis à integração viral, permitindo uma maior compreensão do processo de transformação causado por HPV são benvindas. Este projeto tem como objetivo identificar, a partir do sequenciamento do genoma completo do vírus, padrões moleculares capazes de responder pelo tropismo diferencial apresentado por cepas de HPV16 e do curso assintomático ou transformante da infecção pelo HPV16 nos mesmos sítios. Além disso, avaliar a diversidade genética do HPV16 nos tumores de colo uterino e de orofaringe, e analisar os sítios de integração viral no genoma humano, para verificar se há uma região preferencial para a inserção do HPV em cada tumor. Vários relatos na literatura indicam que o prognóstico dos pacientes com tumores de orofaringe HPV positivos é melhor do que dos casos HPV negativos, independentemente do

4) Comparar a estrutura molecular do vírus HPV16 em 30 amostras de carcinomas epidermoides cervicais, relacionando com as sequências de 10 amostras normais de cérvix uterina com infecção do vírus HPV16;
5) Comparar a estrutura molecular do vírus HPV16 em 30 amostras de carcinomas epidermoide de orofaringe, relacionando com as sequências de 10 amostras normais de orofaringe com infecção do vírus HPV16;
6) Sequenciar o genoma completo do vírus nos dois grupos amostrais, a fim de identificar alterações específicas características dos vírus infectantes nos tumores de orofaringe e cérvix uterina;
7) Sequenciar as regiões do genoma humano que flanqueiam os sítios de integração do HPV16, em ambos os grupos, para verificar se há regiões preferenciais de inserção viral;
8) Realizar análise filogenética das sequências do vírus HPV16 e comparar com os dados clínicos dos pacientes envolvidos no estudo.

Avaliação dos Riscos e Benefícios:
Aparentemente, não existem riscos suplementares aos sujeitos da pesquisa pelo fato das amostras a serem analisadas fazerem parte de protocolo assistencial tanto das mulheres seguidas no Departamento de Ginecologia, quantos dos indivíduos portadores de neoplasias de cabeça e pescoço. O risco dos procedimentos foi considerado baixo, entretanto é preciso ter certeza que não existe aumento de risco para os grupos controle.

Comentários e Considerações sobre a Pesquisa:
Trata-se de pesquisa que lida com tema importante e que pretende usar metodologia laboratorial molecular de ponta.

Considerações sobre os Termos de apresentação obrigatória:
No primeiro relato foram sugeridas as seguintes alterações:

- adequação do TCLE (linguagem, destino das amostras e garantias dos sujeitos de pesquisa), e elaboração do TCLE do grupo controle: pendência atendida. A linguagem, apesar de ainda poder

ser simplificada, é agora compreensível aos sujeitos de pesquisa.

- solicitação de dispensa da aplicação de novo TCLE para uso das amostras da Ginecologia (tanto grupo de estudo quanto grupo controle): pendência atendida.

- adequação do orçamento da pesquisa (R$35.000,00 ou R$111.176,42 ?): na nova versão consta o valor de R$111.176,42. Pendência atendida.

- anexar a aprovação da pesquisa pelo Departamento de Otorrinolaringologia. Foram anexadas as aprovações de vários departamentos, todos parceiros na pesquisa (Ginecologia, Otorrinolaringologia). Pendência atendida.

Além disso, foram anexadas as aprovações do projeto mãe, ao qual a pesquisa atual encontra-se atrelada. Trata-se de pesquisa aprovada pela CONEP em 2 de março de 2012 e explicitada abaixo.

PARECER Nº. 128/2012
Registro CONEP 16491 (Este nº deve ser citado nas correspondências referentes a este projeto)
Folha de Rosto: 405501 Processo nº 25000.084206/2011-76
Projeto de Pesquisa: "Fatores ambientais, clínicos, histopatológicos e moleculares associados ao desenvolvimento e ao prognóstico de carcinomas epidermóides de cabeça e pescoço" (GENCAPO Fase II).
Pesquisador Responsável: Dra. Eloiza Helena Tajara da Silva
Instituição: Faculdade de Medicina de São Jose do Rio Preto/SP (1º CENTRO)
CEP de origem: Faculdade de Medicina de São José do Rio Preto (FAMERP/SP)
Área Temática Especial: Genética Humana, Biossegurança, Pesquisa com Cooperação Estrangeira
Patrocinador: Fundação de Amparo a Pesquisa do Estado de São Paulo.

Recomendações:
Não existem outras recomendações.

Conclusões ou Pendências e Lista de Inadequações:
Não existem outras pendências.

Situação do Parecer:
Aprovado

Necessita Apreciação da CONEP:
Não

HOSPITAL DAS CLÍNICAS DA
FACULDADE DE MEDICINA DA
USP - HCFMUSP

Considerações Finais a critério do CEP:

Sem pendências.

SAO PAULO, 14 de Dezembro de 2012

Assinador por:
Luiz Eugênio Garcez Leme
(Coordenador)

Endereço: Rua Ovídio Pires de Campos, 225 5º andar
Bairro: Cerqueira Cesar **CEP:** 05.403-010
UF: SP **Município:** SAO PAULO
Telefone: (11)2661-6442 **Fax:** (11)2661-6442 **E-mail:** marcia.carvalho@hc.fm.usp.br

ANEXO E – Termo de consentimento livre e esclarecido

1

HOSPITAL DAS CLÍNICAS DA FACULDADE DE MEDICINA DA UNIVERSIDADE DE SÃO PAULO-HCFMUSP

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

DADOS DE IDENTIFICAÇÃO DO SUJEITO DA PESQUISA OU RESPONSÁVEL LEGAL

1. NOME: .....

DOCUMENTO DE IDENTIDADE Nº: ..... SEXO: M ☐ F ☐

DATA NASCIMENTO: ...../...../.....

ENDEREÇO ..... Nº ..... APTO: .....

BAIRRO: ..... CIDADE .....

CEP..... TELEFONE: DDD (.....) .....

DADOS SOBRE A PESQUISA

2. TÍTULO DO PROTOCOLO DE PESQUISA: "AVALIAÇÃO DO PERFIL GENÉTICO DO HPV16 E SEU SÍTIO DE INTEGRAÇÃO EM CÉLULAS EPITELIAIS NORMAIS E NEOPLÁSICAS DA CÉRVIX E TONSILAS" .

PESQUISADOR: DR. JOSÉ EDUARDO LEVI .

CARGO/FUNÇÃO: PROFESSOR COLABORADOR, PESQUISA E DESENVOLVIMENTO, INSTITUTO DE MEDICINA TROPICAL DE SÃO PAULO, VIROLOGIA INSCRIÇÃO CONSELHO REGIONAL Nº 23.407/01-D (CRBio)

MÉDICO RESPONSÁVEL: DR. RONALDO FRIZZARINI .

UNIDADE DO HCFMUSP: DISCIPLINA DE OTORRINOLARINGOLOGIA .

3. AVALIAÇÃO DO RISCO DA PESQUISA:

RISCO MÍNIMO ☐ RISCO MÉDIO ☐

RISCO BAIXO ■ RISCO MAIOR ☐

4. DURAÇÃO DA PESQUISA: três anos .

Rubrica do sujeito de pesquisa ou responsável_____

Rubrica do pesquisador_____

HOSPITAL DAS CLÍNICAS DA FACULDADE DE MEDICINA DA UNIVERSIDADE DE SÃO PAULO-HCFMUSP

Convidamos o(a) Sr(a) para participar da Pesquisa intitulada *"Avaliação do perfil genético do HPV 16 e seu sítio de integração em **células epiteliais normais** e neoplásicas da cérvix e tonsilas"*, sob a responsabilidade do pesquisador *Dr. José Eduardo Levi* e do médico responsável *Dr. Ronaldo Frizzarini*, a qual pretende comparar características moleculares da infecção pelo vírus HPV em pacientes saudáveis (controles) e pacientes portadores de câncer de cabeça e pescoço e do colo do útero.

1 – Essas informações estão sendo fornecidas para sua participação voluntária neste estudo, que visa a utilização de **amostras normais de tonsila (amígdala)** provenientes de cirurgias de uvulopalatofaringoplastias, conhecida como "cirurgia do ronco" em adultos com transtornos do sono (apneicos) que serão identificadas para a presença do papilomavírus humano (HPV) por técnicas laboratoriais. As amostras que contenham o vírus HPV serão usadas como **controle normal** para a análise molecular comparadas com amostras de câncer de cabeça e pescoço (orofaringe);

2 – O procedimento para a obtenção das amostras não será diferente do rotineiro, sendo realizada a extração, total ou parcial, das tonsilas (amígdala), como o médico achar conveniente. As amostras serão armazenadas em freezers específicos para análises posteriores, que compreendem a extração do material genético e genotipagem de HPV 16 específico. Caso as amostras sejam positivas, será realizada a análise molecular do vírus;

3 – Os procedimentos rotineiros pré-operatórios incluem: Hemograma completo – permite avaliar se a paciente possui anemia, se a contagem de plaquetas, que auxiliam na contenção do sangramento está normal, e se o número e divisão percentual das células de defesa estão normais; Coagulograma, TAP e PTT – permitem avaliar a qualidade da coagulação, necessária para controlar os sangramentos; Bioquímica: glicose, sódio, potássio, creatinina, uréia – detecção de sinais de diabetes, alterações do metabolismo, assim como da função renal; Exames de imagem: Radiografia de Cavum,Videofaringolaringoscopia,Tomografia Computadorizada de Seios da Face e Resonância Magnética de seios da face são utilizados quando se faz necessário.

4 – Após a cirurgia, os procedimentos que deverão ser seguidos não diferem dos rotineiros, que são: alimentação líquido-pastosa fria ou gelada nos três primeiros dias, no quarto dia alimentos mais consistentes, ainda frios ou mornos, no sétimo dia em diante alimentação normal, evitando ingerir alimentos que possuírem pontas que possam machucar. Líquido a vontade. Repouso em casa por de 7 a 14 dias, de acordo com a evolução de cada paciente e orientação do médico. Nos três primeiros dias escovar apenas os dentes da frente. Não gargarejar ou bochechar com força. Os pontos da cirurgia, quando dados, cairão por conta própria. Febre de até 38 graus pode ocorrer nos dois primeiros dias, sem que seja sinal de infecção. Salivação e discreta eliminação de secreção nasal com raias de sangue podem ocorrer nas primeiras 24 horas. O médico deverá ser comunicado em caso de sangramento ativo (sangue vivo). Em alguns casos podem ocorrer vômitos com ou sem coágulos (sangue pisado,

Rubrica do sujeito de pesquisa ou responsável_____

Rubrica do pesquisador_____

preto) nas primeiras 24 horas, que são decorrentes da deglutição de sangue da cirurgia. Dores na garganta e de ouvido, semelhantes as que ocorrem na amigdalite, são comuns. Conseguimos reduzi-las com analgésicos em doses regulares;

5 – O benefício para o participante é o mesmo proveniente da cirurgia para a desobstrução dos seios paranasais e amígdalas, resultando na eliminação da apneia (ronco). Não há benefício adicional proporcionado pelo protocolo de pesquisa. Trata-se de estudo experimental testando a hipótese do porque pacientes infectados pelo HPV desenvolvem ou não carcinomas de orofaringe. O benefício poderá ser dado, posteriormente, para os pacientes que desenvolveram o tumor, ou para a identificação ou para a prevenção dos pacientes com risco aumentado para essa neoplasia;

6 – Não há procedimentos alternativos que possam ser vantajosos, pelos quais o paciente pode optar.

7 – Garantia de acesso: em qualquer etapa do estudo, você terá acesso aos profissionais responsáveis pela pesquisa para esclarecimento de eventuais dúvidas. O principal investigador é o Dr ***José Eduardo Levi,*** que pode ser encontrado no endereço Rua Dr. Enéas de Carvalho Aguiar 470, 2º andar, CEP 05403-000 - Laboratório de Virologia do Instituto de Medicina Tropical da Universidade de São Paulo, e-mail dudilevi@usp.br Telefone(s) 11 3061-8666. Se você tiver alguma consideração ou dúvida sobre a ética da pesquisa, entre em contato com o Comitê de Ética em Pesquisa (CEP) – Rua Ovídio Pires de Campos, 225 – 5º andar – tel: 3069-6442 ramais 16, 17, 18 ou 20 – e-mail: cappesq@hcnet.usp.br

8 – É garantida a liberdade da retirada de consentimento a qualquer momento e deixar de participar do estudo, sem qualquer prejuízo à continuidade de seu tratamento na Instituição.

9 – Direito de confidencialidade – As informações obtidas serão analisadas em conjunto com outros pacientes, não sendo divulgado a identificação de nenhum paciente.

10 – Direito de ser mantido atualizado sobre os resultados parciais das pesquisas, quando em estudos abertos, ou de resultados que sejam do conhecimento dos pesquisadores.

11 – Despesas e compensações: não há despesas pessoais para o participante em qualquer fase do estudo, incluindo exames e consultas. Também não há compensação financeira relacionada à sua participação.

12 - Compromisso do pesquisador de utilizar os dados e o material coletado somente para esta pesquisa. Esse material biológico coletado ficará disponível para a retirada do sujeito de pesquisa, caso seja de seu interesse, ou, alternativamente, as amostras poderão ser destruídas de acordo com as normas de biossegurança vigentes no departamento de virologia da instituição em questão, de acordo com a resolução do CNS 441, de 12 de maio de 2011.

Rubrica do sujeito de pesquisa ou responsável_____

Rubrica do pesquisador_____

HOSPITAL DAS CLÍNICAS DA FACULDADE DE MEDICINA DA UNIVERSIDADE DE SÃO PAULO-HCFMUSP

Acredito ter sido suficientemente informado a respeito das informações que li ou que foram lidas para mim, descrevendo o estudo "*AVALIAÇÃO DO PERFIL GENÉTICO DO HPV16 E SEU SÍTIO DE INTEGRAÇÃO EM CÉLULAS EPITELIAIS NORMAIS E NEOPLÁSICAS DA CÉRVIX E TONSILAS*".

Eu discuti com o ***Dr. Ronaldo Frizzarini*** sobre a minha decisão em participar nesse estudo. Ficaram claros para mim quais são os propósitos do estudo, os procedimentos a serem realizados, seus desconfortos e riscos, as garantias de confidencialidade e de esclarecimentos permanentes. Ficou claro também que minha participação é isenta de despesas e que tenho garantia do acesso a tratamento hospitalar quando necessário. Concordo voluntariamente em participar deste estudo e poderei retirar o meu consentimento a qualquer momento, antes ou durante o mesmo, sem penalidades ou prejuízo ou perda de qualquer benefício que eu possa ter adquirido, ou no meu atendimento neste Serviço.

_____

Assinatura do paciente/representante legal Data ___/___/___

_____

Assinatura da testemunha Data ___/___/___

para casos de pacientes menores de 18 anos, analfabetos, semi-analfabetos ou portadores de deficiência auditiva ou visual.

(Somente para o responsável do projeto)

Declaro que obtive de forma apropriada e voluntária o Consentimento Livre e Esclarecido deste paciente ou representante legal para a participação neste estudo.

_____

Assinatura do responsável pelo estudo Data ___/___/___

Rubrica do sujeito de pesquisa ou responsável_____

Rubrica do pesquisador_____