



**GUSTAVO JACOB LOURENÇO**

**IDENTIFICAÇÃO DE GENES DE SUSCEPTIBILIDADE  
HERDADA PARA O CARCINOMA DE CÉLULAS ESCAMOSAS  
DE BASE DE LÍNGUA POR GENOTIPAGEM EM  
LARGA ESCALA**

**CAMPINAS**

**UNICAMP**

**2011**



**GUSTAVO JACOB LOURENÇO**

**IDENTIFICAÇÃO DE GENES DE SUSCEPTIBILIDADE  
HERDADA PARA O CARCINOMA DE CÉLULAS ESCAMOSAS  
DE BASE DE LÍNGUA POR GENOTIPAGEM EM  
LARGA ESCALA**

Tese de Doutorado apresentada à Comissão de Pós-Graduação da Faculdade de Ciências Médicas da Universidade Estadual de Campinas para a obtenção do título de Doutor em Fisiopatologia Médica, área de concentração em Biologia Estrutural, Celular, Molecular e do Desenvolvimento

**Orientadora:** Profa. Dra. Carmen Silvia Passos Lima

**Co-orientador:** Prof. Dr. Fernando Ferreira Costa

**CAMPINAS**

**UNICAMP**

**2011**

## Ficha catalográfica

FICHA CATALOGRÁFICA ELABORADA POR  
ROSANA EVANGELISTA PODEROSO – CRB8/6652  
BIBLIOTECA DA FACULDADE DE CIÊNCIAS MÉDICAS  
UNICAMP

L934i Lourenço, Gustavo Jacob 1978 -  
Identificação de genes de susceptibilidade herdada  
para o carcinoma de células escamosas de base de  
língua por genotipagem em larga escala. / Gustavo  
Jacob Lourenço. – Campinas, SP : [s.n.], 2011.

Orientador : Carmen Silvia Passos Lima  
Coorientador : Fernando Ferreira Costa  
Tese (Doutorado) - Universidade Estadual de  
Campinas, Faculdade de Ciências Médicas.

1. Neoplasias de cabeça e pescoço. 2. Análise de  
sequência com séries de oligonucleotídeos. 3.  
Polimorfismo genético. 4. Polimorfismo de um único  
nucleotídeo. 5. Variações do número de cópias de  
DNA. I. Lima, Carmen Silvia Passos. II. Costa,  
Fernando Ferreira. III. Universidade Estadual de  
Campinas. Faculdade de Ciências Médicas. VI. Título.

### Informações para Biblioteca Digital

**Título em inglês:** Identification of inherited susceptibility genes for squamous cell carcinoma of base of tongue by large scale genotyping

**Palavras-chave em inglês:**

Head and neck neoplasms

Oligonucleotide Array Sequence Analysis

Genetic polymorphism

DNA copy number variation

**Área de concentração:** Biologia Estrutural, Celular, Molecular e do Desenvolvimento

**Titulação:** Doutor em Fisiopatologia Médica

**Banca examinadora:**

Carmen Silvia Passos Lima [Orientador]

Luiz Paulo Kowalski

Marilda de Souza Gonçalves

Carmen Silvia Bertuzzo

Monica Barbosa de Mello

**Data da defesa:** 20-12-2011

**Programa de Pós-Graduação:** Fisiopatologia Médica

## Folha de aprovação

---

### Banca examinadora de Tese de Doutorado

---

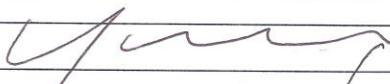
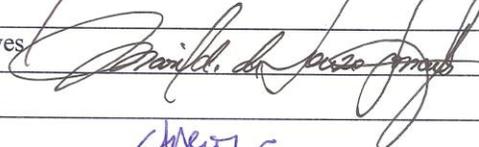
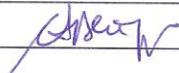
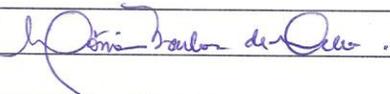
Gustavo Jacob Lourenço

---

**Orientador: Prof. Dr. Carmen Silvia Passos Lima**

---

---

<b>Membros:</b>	
Professor (a) Doutor (a) Luiz Paulo Kowalski	
Professor (a) Doutor (a) Marilda de Souza Gonçalves	
Professor (a) Doutor (a) Carmen Silvia Bertuzzo	
Professor (a) Doutor (a) Monica Barbosa de Mello	
Professor (a) Doutor (a) Carmen Silvia Passos Lima	

Curso de pós-graduação em Fisiopatologia Médica da Faculdade de Ciências Médicas da Universidade Estadual de Campinas.

---

**Data: 20/12/2011**

---



## **DEDICATÓRIA**

*Dedico este trabalho à minha família amada, pelo apoio incondicional, sem o qual eu não teria chegado até aqui*



## AGRADECIMENTOS

---

À Profa. Dra. Carmen Silvia Passos Lima, pelo profissionalismo, apoio e amizade, sem os quais eu não teria finalizado este trabalho.

Ao Prof. Dr. Fernando Ferreira Costa pelo incentivo e apoio durante o desenvolvimento deste projeto.

Ao Prof. Dr. Carlos Takahiro Chone do Departamento de Otorrinolaringologia da FCM da UNICAMP, pela ajuda na obtenção das amostras dos pacientes.

À Dra. Letícia Khater, Eugênia Camargo, Milena Simioni e Profa. Dra. Mônica Melo pela sincera ajuda no desenvolvimento da técnica de microarranjos de DNA.

Ao Prof. Dr. Benilton Carvalho do Grupo de Biologia Computacional do Departamento de Oncologia da Universidade de Cambridge, Inglaterra, e à Renata Pellegrino, do Departamento de Psicobiologia da Universidade Federal de São Paulo, pela dedicação nas análises dos dados referentes aos nossos resultados.

Ao Meredith e à Caroline Lloyd-Evans, de Cambridge na Inglaterra, pela atenção e por disponibilizarem a própria casa para a minha confortável estadia.

À Manoela, minha amiga e irmã, presente em todos os momentos.

Aos meus amigos Anderson Cunha, Luciana, Cristiane, Helvia, Iramaia, Cintia e Gabriela, pela amizade e ajuda inestimáveis, fundamentais no desenvolvimento do meu trabalho.

Ao pessoal do Laboratório de Genética do Câncer, Ana Paula, Érika, Anderson Zanetti, Stela, Aline, Camila e Leisa, pela pronta ajuda em todos os momentos.

À Dulcinéia e todo o pessoal do Laboratório de Hemoglobina e Genoma do HEMOCENTRO da UNICAMP, pela ajuda em diversos momentos.

Ao meu amigo Giovany, pela amizade especial.

À Claudete, ao Shizuo e todo pessoal do GAAC, pelo apoio e amizade.

À todos os pacientes que aceitaram participar deste estudo.



*Compreendi, então,  
que a vida não é uma sonata que,  
para realizar a sua beleza,  
tem de ser tocada até o fim.  
Dei-me conta, ao contrário,  
de que a vida é um álbum de mini-sonatas.  
Cada momento de beleza vivido e amado,  
por efêmero que seja,  
é uma experiência completa  
que está destinada à eternidade.  
Um único momento de beleza e amor  
justifica a vida inteira.*

**Rubem Alves**



A realização deste estudo contou com o apoio financeiro das agências de fomento:

**Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP)**

Bolsa de doutoramento: processo no. 2006/06231-0

Auxílio pesquisa: processo no. 2006/07054-5

**Financiadora de Estudos e Projetos (FINEP)**

Auxílio pesquisa: processo no. 01/07/0665/01

**Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq)**

Bolsa de apoio técnico: processo no. 502010/2008-8



## LISTA DE ABREVIATURAS

---

BL	Base de língua
brlmm	Algoritmo <i>bayesian robust linear model with mahalanobis</i>
BSA	Albumina bovina sérica
cbs	Algoritmo <i>circular binary segmentation</i>
CEC	Carcinoma espinocelular
CGH	Hibridização genômica comparativa
CNV	Variação no número de cópias
CQ	Controle de qualidade
cr1mm	Algoritmo <i>corrected robust linear model with maximum likelihood classification</i>
DAVID	<i>Database for annotation, visualization and integrated discovery</i>
DMSO	Dimetilsulfóxido
DNA	Ácido desoxirribonucleico
dNTP	Desoxirribonucleotídeo trifosfato
EDTA	Ácido etilenodiamino tetracético
EGFR	Receptor do fator de crescimento epidermal
GELE	Genotipagem em larga escala
HCl	Ácido clorídrico
hmm	Algoritmo <i>hidden Markov model</i>
HPV	Papiloma vírus humano
HSDNA	DNA de esperma de salmão
LOH	Perda de heterozigosidade
M	Molar
MES	Ácido <i>4-morpholineethanesulfonic</i>
mg	Miligrama
ml	Mililitro
mM	Milimolar
NCBI	<i>National Center for Biotechnology</i>
NfκB	Fator de transcrição nuclear kappa de células B
ng	Nanograma
OCR	<i>Oligo control reagent</i>
OR	Risco de ocorrência



<i>P53</i>	Gene supressor de tumor <i>P53</i>
pb	Pares de bases
PCR	Reação em cadeia da polimerase
<i>RAS</i>	Oncogene <i>RAS</i>
Rb	Proteína retinoblastoma
RFLP	Digestão enzimática
RNA <sub>m</sub>	Ácido ribonucléico mensageiro
rpm	Rotação por minuto
rs	Número de referência do SNP
SAPE	Estreptavidina e ficoeritrina
SNP	Polimorfismo gênico de base única
SSPE	Solução de hibridização
STAT	Sinal de transdução e transcrição
TdT	Enzima <i>terminal deoxynucleotidyl transferase</i>
TE	Tampão Tris EDTA
TGF-β	Fator de crescimento transformante beta
TMA <sub>CL</sub>	Cloreto de amônio tetrametil
<i>TP16</i>	Gene supressor de tumor <i>TP16</i>
U	Unidades
Wnt	Proteína transmembrana <i>wingless</i>
μl	Microlitro



## LISTA DE TABELAS

---

	Página
Tabela 1 Sequências de lavagens e marcações dos microarranjos contidos em lâminas por fluoróforos para a genotipagem em larga escala de polimorfismos gênicos de base única e identificação do número de cópias do DNA em pacientes com tumores de células escamosas de base de língua e controles .....	67
Tabela 2 Frequências das distribuições dos 55 pacientes com carcinoma de células escamosas de base de língua de acordo com a idade, o sexo, a raça e o hábito de fumar e de ingerir bebidas alcoólicas .....	76
Tabela 3 Frequências da distribuição dos 55 pacientes com carcinoma de células escamosas de base de língua de acordo com o padrão alimentar .....	77
Tabela 4 Frequências das distribuições dos 55 controles de acordo com a idade, o sexo, a raça e os hábitos de fumar e ingerir bebida alcoólica ..	78
Tabela 5 Frequências da distribuição dos 55 controles de acordo com o padrão alimentar .....	79
Tabela 6 Frequências das distribuições dos 55 pacientes com carcinoma de células escamosas de base de língua de acordo com a diferenciação e o estágio do tumor .....	80
Tabela 7 Concentrações de DNA em ng/ $\mu$ l de amostras de sangue periférico dos 55 pacientes com carcinoma de células escamosas de base de língua .....	83
Tabela 8 Concentrações de DNA em ng/ $\mu$ l em amostras de sangue periférico dos 55 controles .....	84
Tabela 9 Taxas dos controles de qualidade das amostras de DNA dos 49 pacientes com carcinoma de células escamosas de base de língua hibridizadas em lâminas com microarranjos de DNA .....	88



Tabela 10	Taxas dos controles de qualidade das amostras de DNA dos 49 controles hibridizadas em lâminas com microarranjos de DNA .....	89
Tabela 11	Regiões do DNA que possuem os polimorfismos gênicos de base única identificados em 49 pacientes com carcinoma de células escamosas de base de língua e 49 controles por meio dos algoritmos brlmm e crlmm .....	91
Tabela 12	Polimorfismos de base única localizados na região codificadora de aminoácidos do DNA identificados em 49 pacientes com carcinoma de células escamosas de base de língua e 49 controles por meio do algoritmo brlmm de acordo com o número do cromossomo, a região cromossômica, o símbolo do gene, o número de referência e os alelos distintos .....	92
Tabela 13	Polimorfismos de base única localizados na região codificadora de aminoácidos do DNA identificados em 49 pacientes com carcinoma de células escamosas de base de língua e 49 controles por meio do algoritmo crlmm de acordo com o número do cromossomo, a região cromossômica, o símbolo do gene, o número de referência e os alelos distintos .....	94
Tabela 14	Polimorfismos de base única localizados nas regiões 3' e 5' do DNA identificados em 49 pacientes com carcinoma de células escamosas de base de língua e 49 controles por meio do algoritmo brlmm de acordo com o número do cromossomo, a região cromossômica, o símbolo do gene, o número de referência e os alelos distintos .....	96
Tabela 15	Polimorfismos de base única localizados nas regiões 3' e 5' do DNA identificados em 49 pacientes com carcinoma de células escamosas de base de língua e 49 controles por meio do algoritmo crlmm de acordo com o número do cromossomo, a região cromossômica, o símbolo do gene, o número de referência e os alelos distintos .....	98



Tabela 16	Polimorfismos de base única localizados nas regiões regulatórias de expressão dos genes identificados em 49 pacientes com carcinoma de células escamosas de base de língua e 49 controles por meio do algoritmo brlmm de acordo com o número do cromossomo, a região cromossômica, o símbolo do gene, o número de referência e os alelos distintos.....	100
Tabela 17	Polimorfismos de base única localizados nas regiões regulatórias de expressão dos genes identificados em 49 pacientes com carcinoma de células escamosas de base de língua e 49 controles por meio do algoritmo crlmm de acordo com o número do cromossomo, a região cromossômica, o símbolo do gene, o número de referência e os alelos distintos.....	102
Tabela 18	Teste de equilíbrio de Hardy-Weinberg de dez polimorfismos de base de única selecionados e identificados em amostras de 49 pacientes com carcinoma de células escamosas de base de língua e 49 controles.....	104
Tabela 19	Frequências dos genótipos dos polimorfismos de base única nas regiões codificadoras de aminoácidos dos genes <i>COL22A1</i> e <i>IKBKAP</i> selecionados em amostras de sangue periférico dos 49 pacientes com carcinoma de células escamosas de base de língua e 49 controles.....	105
Tabela 20	Frequências dos genótipos dos polimorfismos de base única nas regiões 3' e 5' não traduzidas do DNA dos genes <i>ERP29</i> , <i>MCC</i> e <i>ZNF415</i> selecionados em amostras do sangue periférico dos 49 pacientes com carcinoma de células escamosas de base de língua e 49 controles.....	106
Tabela 21	Frequências dos genótipos dos polimorfismos de base única nas regiões regulatórias da expressão gênica dos genes <i>LEF1</i> e <i>LY6K</i> selecionados em amostras do sangue periférico dos 49 pacientes com carcinoma de células escamosas de base de língua e 49 controles.....	107



Tabela 22	Frequências dos genótipos dos polimorfismos de base única nas regiões regulatórias da expressão gênica do gene PTCH1 selecionados em amostras do sangue periférico dos 49 pacientes com carcinoma de células escamosas de base de língua e 49 controles.....	108
Tabela 23	Regiões de ganhos de material genético e seus respectivos cromossomos, genes e tamanhos das regiões em misturas de 49 amostras de DNA de pacientes caucasóides e negróides com carcinoma de células escamosas de base de língua comparadas a misturas de 49 amostras de DNA de controles caucasóides e negróides.....	116
Tabela 24	Regiões de ganhos de material genético e seus respectivos cromossomos, genes e tamanhos das variações em misturas de 39 amostras de DNA de pacientes caucasóides com carcinoma de células escamosas de base de língua comparadas a misturas de 39 amostras de DNA de controles caucasóides .....	117
Tabela 25	Regiões de ganhos de material genético e seus respectivos cromossomos, genes e tamanhos das variações em misturas de dez amostras de DNA de pacientes negróides com carcinoma de células escamosas de base de língua comparadas a misturas de dez amostras de DNA de controles negróides .....	118
Tabela 26	Regiões de perdas de material genético e seus respectivos cromossomos, genes e tamanhos das variações em misturas de 49 amostras de DNA de pacientes caucasóides e negróides com carcinoma de células escamosas de base de língua comparadas a misturas de 49 amostras de DNA de controles caucasóides e negróides.....	119



Tabela 27	Regiões de perdas de material genético e seus respectivos cromossomos, genes e tamanhos das variações em misturas de 39 amostras de DNA de pacientes caucasóides com carcinoma de células escamosas de base de língua comparadas a misturas de 39 amostras de DNA de controles caucasóides .....	120
Tabela 28	Regiões de perdas de material genético e seus respectivos cromossomos, genes e tamanhos das variações em misturas de dez amostras de DNA de pacientes negróides com carcinoma de células escamosas de base de língua comparadas a misturas de dez amostras de DNA de controles negróides .....	121



## LISTA DE FIGURAS

---

	Página
Figura 1 Método de genotipagem em larga escala em portadores de carcinoma de células escamosas de base de língua com os produtos da amplificação dos fragmentos obtidos da digestão com a enzima <i>StyI</i> de DNA genômico em gel de agarose 2%, para a hibridização em lâminas de microarranjos.....	61
Figura 2 Método de genotipagem em larga escala em portadores de carcinoma de células escamosas de base de língua com os produtos da amplificação dos fragmentos obtidos da digestão com a enzima <i>NspI</i> de DNA genômico em gel de agarose 2%, para a hibridização em lâminas de microarranjos.....	63
Figura 3 Método de genotipagem em larga escala em portadores de carcinoma de células escamosas de base de língua com os produtos da fragmentação da mistura dos produtos das PCRs obtidos da digestão enzimática com o reagente de fragmentação, para a hibridização em lâminas de microarranjos.....	65
Figura 4 Representação esquemática do protocolo Affymetrix® <i>Genome-Wide Human SNP Array 5.0</i> para a genotipagem em larga-escala de pacientes e controles.....	68
Figura 5 Imagem de um paciente com carcinoma de células escamosas de base de língua .....	81
Figura 6 Aspectos histológicos do carcinoma de células escamosas de base de língua de paciente analisado no estudo (Caso nº 13) (H&E, 20x e 40x, respectivamente).....	82
Figura 7 Curva de qualidade da amostra de DNA de um paciente (Caso nº 5) com carcinoma de células escamosas de base de língua gerada por espectrofotômetro .....	85



Figura 8	Imagem padrão gerada pela excitação dos fluoróforos contidos em lâmina de um indivíduo controle capturada por <i>scanner</i> para a genotipagem em larga escala.....	87
Figura 9	Representação esquemática de cariótipo obtido pelo programa Partek <sup>®</sup> após a análise do número de variações de cópias de DNA de 49 pacientes com carcinoma de células escamosas de base de língua em relação aos 49 controles.....	110
Figura 10	Representação esquemática de regiões de ganhos de DNA em cromossomos e genes após a análise de variações no número de cópias de DNA de 49 pacientes com carcinoma de células escamosas de base de língua em relação aos controles .....	112
Figura 11	Representação esquemática de regiões de perdas de DNA em cromossomos e genes após a análise de variações no número de cópias de DNA de 49 pacientes com carcinoma de células escamosas de base de língua em relação aos controles .....	113
Figura 12	Representação esquemática das funções conhecidas ou presumidas das proteínas possivelmente produzidas pelos genes polimórficos identificados e selecionados em 49 pacientes com carcinoma de células escamosas de base de língua e 49 controles por meio da genotipagem em larga escala .....	139



## SUMÁRIO

---

	Página
<b>RESUMO</b> .....	xxxvii
<b>ABSTRACT</b> .....	xxxix
<b>INTRODUÇÃO</b> .....	41
1. Considerações iniciais.....	43
2. Aspectos clínicos.....	44
3. Fatores predisponentes.....	45
3.1. Aspectos ambientais.....	45
3.2. Aspectos genéticos.....	45
4. Genotipagem em larga escala.....	47
<b>OBJETIVOS</b> .....	51
<b>PACIENTES E MÉTODOS</b> .....	55
1. Avaliação clínica.....	57
2. Aspectos do tumor.....	58
3. Determinação dos polimorfismos genéticos.....	58
3.1. Extração do DNA a partir dos leucócitos.....	59
3.2. Digestão com a enzima <i>SlyI</i> , ligação a adaptadores e amplificação.....	59
3.3. Digestão com a enzima <i>NspI</i> , ligação a adaptadores e amplificação.....	62
3.4. Mistura dos produtos das PCRs, purificação e quantificação.....	64
3.5. Fragmentação dos produtos das PCRs.....	64
3.6. Marcação do DNA fragmentado com biotina.....	66
3.7. Hibridização.....	66
3.8. Lavagens das lâminas e marcação com fluoróforos.....	66
3.9. Captação das imagens e genotipagem.....	69
4. Análise estatística.....	70
5. Aspectos éticos.....	71
<b>RESULTADOS</b> .....	73
1. Avaliação clínica.....	75
1.1. Pacientes com CEC de BL e controles.....	75
2. Aspectos do tumor.....	79
2.1. Grau de diferenciação e estágio do tumor.....	79



	Página
3. Determinação dos polimorfismos genéticos.....	80
3.1. Quantidade e qualidade de DNA nas amostras.....	80
3.2. Hibridização das amostras nas lâminas de microarranjos.....	86
3.3. Controle de qualidade das amostras hibridizadas nas lâminas.....	86
3.3.1. Controle de qualidade qualitativo das GELEs individual e coletiva....	86
3.3.2. Controle de qualidade quantitativo da GELE individual.....	87
3.3.3. Controle de qualidade quantitativo da GELE coletiva.....	90
3.4. Análise de SNPs.....	90
3.4.1. GELE individual.....	90
3.4.2. Seleção dos SNPs identificados pela GELE individual.....	103
3.4.3. GELE coletiva.....	109
3.5. Análise de CNVs.....	109
3.5.1. GELE individual.....	109
3.5.2. Análises das CNVs previamente descritas.....	114
3.5.3. GELE coletiva.....	114
<b>DISCUSSÃO.....</b>	<b>123</b>
<b>CONCLUSÕES.....</b>	<b>141</b>
<b>REFERÊNCIAS BIBLIOGRÁFICAS.....</b>	<b>145</b>
<b>ANEXOS.....</b>	<b>167</b>



Alterações genéticas herdadas, como os polimorfismos gênicos de base única (SNPs) e as variações no número de cópias do DNA (CNVs), foram associadas com o risco de carcinoma de células escamosas (CEC) de base de língua (BL) em poucos estudos. O CEC de BL é um tumor que determina altas taxas de morbidade e mortalidade, no entanto, sua associação com polimorfismos genéticos não está estabelecida. Frente ao exposto, o objetivo deste estudo foi o de avaliar os papéis de SNPs e CNVs no CEC de BL. O DNA genômico foi obtido de amostras de sangue periférico de 49 pacientes com CEC de BL e de 49 controles. Cada amostra foi analisada por meio de lâminas com microarranjos de DNA contendo 500.568 SNPs e 420.000 CNVs (Affymetrix®). A digestão enzimática do DNA, a ligação de adaptadores, a amplificação, a fragmentação, a marcação, a hibridização, as lavagens e a leitura das intensidades dos sinais das sondas foram realizadas de acordo com instruções do fabricante. Os dados obtidos foram analisados utilizando o programa Bioconductor e o algoritmo crlmm. Para os SNPs, as diferenças entre os grupos foram analisadas por meio da regressão logística múltipla. Para as CNVs, os dados obtidos foram analisados por meio do programa Partek®. Regiões de ganhos ou perdas significativas de DNA foram determinadas pelo algoritmo cbs. Os genes de interesse foram escolhidos por meio do programa DAVID. Nós observamos que a frequência de 6.609 SNPs foi distinta entre pacientes com CEC de BL e controles ( $P < 0,01$ ). Cinquenta e dois SNPs (0,8%) estiveram localizados nas regiões codificantes do DNA, 51 (0,8%) estiveram nas regiões 3' e 5' não traduzidas, 3.461 estiveram em regiões regulatórias de transcrição e 3.045 em íntrons. Os SNPs considerados de interesse estiveram localizados nos genes relacionados ao ciclo celular (*ERP29*, *MCC* e *PTCHI*), à transcrição (*IKBKAP* e *ZNF415*) e à adesão celular (*COL22A1*, *LEF1* e *LY6K*). Nós identificamos regiões do DNA que apresentaram duplicações em genes relacionados com a proliferação celular (*ADAM3A*, *ADAM5P* e *DDT*), apoptose (*FAM90A*), mecanismo de defesa (*DEFB*) e metabolismo de carcinógenos (*GSTs*). Nós também observamos genes deletados relacionados à apoptose (*BLC2*) e aos receptores do olfato (*ORs*). Nossos resultados sugerem que SNPs e CNVs em genes relacionados com a origem e a progressão de tumores podem predispor indivíduos ao CEC de BL. No entanto, esses resultados devem ser validados por genotipagens de número maior de indivíduos e por análises funcionais de proteínas codificadas por alelos distintos de genes polimórficos.



## ABSTRACT

---

Inherited genetic alterations, such as single nucleotide polymorphisms (SNPs) and copy number variations (CNVs), were described in association with base of tongue (BT) squamous cell carcinoma (SCC) risk in only few reports. BTSCC are tumours with high morbidity and mortality rates; however, the association of SNPs and CNVs and BTSCC risk is still not clarified and, therefore, this was the aim of the present study. DNA was extracted of the peripheral blood samples of 49 BTSCC patients and 49 controls. Each sample was genotyped using DNA high-resolution microarrays containing 500.568 SNPs and 420.000 CNVs (Affymetrix<sup>®</sup>). Further sample processing, including digestion, adaptor ligation, amplification, fragmentation, labelling, hybridization, washing and scanning was assayed according to the standard protocol. Genotype data were acquired by genotyping calling of samples using the crlmm algorithm provided by Bioconductor software, as per the recommended guidelines. For SNPs, the differences between groups were analysed by the logistic regression model. For CNVs, the patients' and controls' data files were imported into the Partek<sup>®</sup> Genomic Suite. Common aberration analysis was performed on all samples to identify genomic intervals that had statistically significant aberrations. Significantly different regions were determined using the segmentation algorithm. For SNPs, we observed 6.609 SNPs with distinct frequencies between BTSCC patients and controls ( $P < 0.01$ ). Fifty two SNPs (0.8%) were located in coding sequence of amino acids, 51 (0.8%) in 3' and 5' untranslated regions, 3.461 (52.4%) in up or downstream regions and 3.045 (46.0%) in introns. The SNPs were clustered to their main function, evidencing those localized in genes related to cell cycle and apoptosis (*ERP29*, *MCC* and *PTCH1*), transcriptional process (*IKBKAP* and *ZNF415*) and cell adhesion and metastasis (*COL22A1*, *LEF1* and *LY6K*). We also identified a consistent number of altered regions including duplicated genes, such as involved in cell proliferation and angiogenesis (*ADAM3A*, *ADAM5P* and *DDT*), apoptosis (*FAM90A*), defensins proteins (*DEFB*) and metabolism of carcinogens (*GSTs*); and deleted genes, such as in olfactory receptors (*ORs*) and apoptosis (*BCL2*). Our preliminary results suggest that SNPs and CNVs in genes involved in tumour origin and progression may predispose individuals to BTSCC. However, these results should be confirmed by functional studies of coded proteins and validated by genotyping in larger epidemiological studies.



# INTRODUÇÃO

---



## **1. Considerações iniciais**

O termo carcinoma de células escamosas (CEC) de cabeça e pescoço engloba um conjunto de tumores do referido tipo histológico identificado na cavidade oral, faringe, fossa nasal, seios paranasais e laringe (MENDENHALL et al., 2005).

O CEC de cabeça e pescoço é um tumor de alta incidência e de alta mortalidade em várias regiões do mundo, incluindo o Brasil (LAMBERT et al., 2011). Cerca de 645.000 novos casos de portadores desses tumores são identificados a cada ano em todo o mundo e, aproximadamente 65% deles apresentam a doença em estágios avançados (MARCUS & TISHLER, 2010). Além disso, esse tumor determina mais de 350.000 mortes ao ano (KAMANGAR et al., 2006; JEMAL et al., 2007).

Os dados epidemiológicos do CEC de cabeça e pescoço não são totalmente conhecidos no Brasil. Segundo as Estimativas de Incidência por Câncer no país, publicadas pelo Instituto Nacional do Câncer, foram estimados 14.120 novos casos do câncer de cavidade oral para o ano de 2010, sendo 7.590 casos somente na região sudeste (INSTITUTO NACIONAL DO CÂNCER, 2011).

O CEC de faringe constitui cerca de 15 a 25% desses tumores e inclui os tumores de nasofaringe, hipofaringe e orofaringe (MENDENHALL et al., 2005).

Com os avanços no conhecimento de aspectos clínicos e moleculares dos diferentes subtipos, esses tumores, antes considerados doenças homogêneas, agora, são reconhecidos como entidades distintas (RISCHIN, 2010).

O CEC de orofaringe merece destaque entre os tumores de faringe devido a seu comportamento agressivo (LICITRA et al., 2002; LAMBERT et al., 2011). É classificado em tumores da região amigdaliana, da base de língua (BL), da parede posterior da faringe, do palato mole e da úvula (LAMBERT et al., 2011). Entre os tumores de orofaringe, o CEC de BL é um dos mais importantes devido a sua alta incidência, morbidade e mortalidade (SESSIONS et al., 2003; KOWALSKI et al., 2005; PALERI et al., 2010).

## 2. Aspectos clínicos

As queixas mais frequentes referidas por pacientes com o CEC de BL são o aumento de volume na região, a dificuldade na ingestão de alimentos, o sangramento e a dor local (COHAN et al., 2009).

A incidência do tumor aumenta com a idade, à semelhança das demais doenças neoplásicas (MENDENHALL et al., 2005). O CEC de BL acomete, em geral, indivíduos com idade maior do que 50 anos e do sexo masculino (MENDENHALL et al., 2005; SECRETARIA MUNICIPAL DE SAÚDE DE SÃO PAULO, 2011).

O CEC pode ser classificado de acordo com o grau de diferenciação celular em indiferenciado, pouco e moderadamente diferenciado (MORI et al., 1998).

O principal sistema de estadiamento do tumor utilizado para a determinação do prognóstico e da terapêutica adequada a ser ministrada a cada caso é o TNM, realizado segundo as normas da União Internacional Contra o Câncer e do Comitê Americano do Câncer, que foram unificadas no ano de 2002 (GREENE et al., 2002). O sistema TNM considera a extensão local do acometimento do tumor em tecidos adjacentes (T), a identificação de metástases para linfonodos regionais (N) e a identificação de metástases à distância (M) (INTERNATIONAL UNION AGAINST CANCER, 2004).

A remoção cirúrgica do tumor e a radioterapia constituem os principais tipos de tratamento para o CEC de BL (COHAN et al., 2009; VARTANIAN et al., 2010). Os procedimentos isolados ou combinados possibilitaram probabilidades de sobrevida global em cinco anos para cerca de 40 a 80% dos pacientes portadores do tumor (SESSIONS et al., 2003). Assim, um contingente considerável de pacientes evolui para o óbito em decorrência de complicações da doença em período menor do que cinco anos do diagnóstico. Recentemente, a quimioterapia, em conjunto com a radioterapia e a cirurgia, foi associada com o aumento de 16% na sobrevida global de pacientes com CEC de orofaringe (FURNESS et al., 2010).

A alta incidência da doença e o prognóstico reservado observado na maioria dos casos justificam a busca de maior conhecimento sobre a sua etiologia e fisiopatologia.

### **3. Fatores predisponentes**

#### **3.1. Aspectos ambientais**

Já é bem conhecido que indivíduos com os hábitos de fumar e ingerir bebidas alcoólicas por longos anos têm maior probabilidade de desenvolver o CEC de cabeça e pescoço, incluindo o de BL (MENDENHALL et al., 2005; LAMBERT et al., 2011). Os indivíduos que consomem tabaco e álcool estão sob risco cerca de 20 vezes maior de apresentarem o CEC de cabeça e pescoço, incluindo o de BL, do que indivíduos sem os referidos hábitos (LAZARUS & PARK, 2000). O benzopireno do tabaco e o acetaldeído do metabolismo do álcool determinam mutações em células do epitélio da faringe e, conseqüentemente, o tumor (HAYES et al., 2005).

O padrão da dieta também parece contribuir para a ocorrência da doença: a carência de verduras e frutas e o excesso de alimentos gordurosos foram associadas ao CEC de BL em alguns estudos (GARAVELLO et al., 2009; LUCENTEFORTE et al., 2009). Os produtos do metabolismo do ácido fólico de frutas e verduras participam da síntese adequada de DNA e do silenciamento de genes envolvidos com a proliferação celular (GARAVELLO et al., 2009; LUCENTEFORTE et al., 2009).

Recentemente, o papiloma vírus humano (HPV), especialmente o tipo 16, foi associado com o surgimento de tumor de BL em indivíduos jovens (com menos de 50 anos) sem o histórico dos hábitos do tabagismo e do etilismo (NGUYEN et al., 2010; PSYRRI & COHEN, 2011). A inserção de DNA viral no DNA das células epiteliais determina a inativação das proteínas p53 e retinoblastoma (Rb) e, possivelmente, o surgimento do tumor (PSYRRI & COHEN, 2011).

#### **3.2. Aspectos genéticos**

Múltiplos eventos genéticos e epigenéticos, como a perda ou a aquisição de funções de moléculas que regulam a sinalização, o crescimento, a sobrevivência e a motilidade celular, a angiogênese, o controle do ciclo celular e o reparo do DNA, são responsáveis pela ocorrência e agressividade dos tumores de cabeça e pescoço (MOLINOLO et al., 2009; LEEMANS et al., 2011).

Já é conhecido que mutações em oncogenes, como os da família *RAS*, e em genes supressores tumorais, como os *TP16* e *P53*, parecem estar associadas ao surgimento desses

tumores (SIDRANSKY, 2005; GOLD & KIM, 2009). Os resultados de estudos recentes sugerem que as vias de sinalização que incluem as proteínas do receptor do fator de crescimento epidermal (EGFR), do fator de transcrição nuclear *kappa* de células B (NfκB), do sinal de transdução e transcrição (STAT), das proteínas transmembranas *wingless* (Wnt) e β-catenina, e do fator de crescimento transformante beta (TGF-β) estão alteradas em tumores de cabeça e pescoço, devido a hipo ou hiper expressão dos genes que as codificam (MOLINOLO et al., 2009). Entretanto, não foram ainda identificadas anormalidades gênicas recorrentes em pacientes com o CEC de BL.

Diferenças individuais geradas por polimorfismos genéticos relacionados à capacidade de metabolizar carcinógenos, de controlar o ciclo celular e de reparar danos ao DNA indicam que a constituição genética pode predispor indivíduos saudáveis ao câncer, inclusive ao CEC de BL (DUTT & BEROUKHIM, 2007; ASSIE et al., 2008; CHEN & CHEN, 2008; GOLD & KIM, 2009). Polimorfismos de base única (SNPs) em genes diversos, cujos alelos selvagem e variante codificam quantidades distintas de proteínas ou proteínas com distintas atividades funcionais, foram associados com a ocorrência de tumores de orofaringe, incluindo o de BL (HIYAMA et al., 2008).

SNPs em genes codificadores de enzimas envolvidas na detoxificação de carcinógenos, como a glucoronosil transferase (ZHENG et al., 2001; LACKO et al., 2009) e no metabolismo do álcool, como a álcool desidrogenase (NISHIMOTO et al., 2004) pareceram influenciar a susceptibilidade ao CEC de BL. SNPs em genes que codificam moléculas de fatores de crescimento, como os fatores de crescimento transformante (STREIT et al., 2004) e de fibroblasto (GUAN et al., 2010); de controle do ciclo celular e da apoptose, como a ciclinas (WAGNER et al., 2001) e a p53 (PERRONE et al., 2007; CHEN et al., 2008; JI et al., 2008); e do reparo a danos do DNA, como o xeroderma pigmentoso (CUI et al., 2006), também foram associados ao risco de ocorrência do tumor.

Entretanto, apenas pequenos números de SNPs de uma única ou duas vias metabólicas foram avaliados na maioria dos estudos (WAGNER et al., 2001; ZHENG et al., 2001; NISHIMOTO et al., 2004; STREIT et al., 2004; CUI et al., 2006; PERRONE et al., 2007; CHEN et al., 2008; JI et al., 2008; LACKO et al., 2009; GUAN et al., 2010; LOURENÇO et al., 2011), o que impediu a avaliação da interação entre eles no metabolismo de carcinógenos ou no reparo de lesões no DNA celular de um indivíduo. Ainda, é possível que outros SNPs envolvidos no metabolismo de carcinógenos ou demais

processos da carcinogênese, como a angiogênese (CHEN et al., 2009; RUIZ et al., 2010; XU et al., 2010), isolados ou combinados, possam ter papéis tão ou mais importantes na origem ou manifestações clínicas da doença.

A variação no número de cópias (CNVs) do DNA também ocorre de maneira polimórfica em indivíduos saudáveis da população em geral (CARTER, 2007; ASSIE et al., 2008; BACOLOD et al., 2009; ZHANG et al., 2009) e em portadores de tumores diversos (CHEN & CHEN, 2008), indicando que possam predispor indivíduos ao câncer. De fato, CNVs de diversos genes, como o *CCND1* e o *CDKA*, incluindo os oncogenes *EGF* e *c-MYC*, e os gene supressor tumoral *P53*, também foram associadas ao risco de ocorrência de CEC de cabeça e pescoço, incluindo o de BL (CHEN & CHEN, 2008)

As regiões dos cromossomos que apresentaram maiores alterações foram as 3p, 4q, 5q21-22, 8p21-23, 9p21-22, 11q13, 11q23, 13q, 14q, 17p, 18q e 22q (CHEN & CHEN, 2008).

Ainda, é fundamental comentar que as frequências dos genótipos de genes polimórficos e CNVs variam substancialmente em diferentes populações étnicas (CHUNG & CHANOCK, 2011) e, desta forma, suas relações com a origem ou com manifestações clínicas de doenças não podem ser generalizadas para todas as populações.

#### **4. Genotipagem em larga escala**

A avaliação de SNPs e CNVs em genes de portadores de tumores e controles por métodos baseados na reação em cadeia da polimerase (PCR), digestão enzimática (RFLP) e na hibridização genômica comparativa (CGH) são consideradas lentas e trabalhosas. Além disso, esses métodos possibilitaram a avaliação de pequeno número de genes na maioria dos estudos sobre a predisposição hereditária ao CEC de cabeça e pescoço (WAGNER et al., 2001; ZHENG et al., 2001; NISHIMOTO et al., 2004; STREIT et al., 2004; CUI et al., 2006; PERRONE et al., 2007; CHEN et al., 2008; CHEN & CHEN, 2008; JI et al., 2008; GUAN et al., 2010; LOURENÇO et al., 2011).

Recentemente, foi descrito um método de genotipagem individual, tanto para pacientes como para controles, que permitiu analisar centenas ou milhares de polimorfismos genéticos com rapidez e confiabilidade em um único estudo. O método tem como base a análise de microarranjos de DNA de cada indivíduo separadamente e recebeu a denominação de genotipagem em larga escala (GELE) individual (SYVANEN, 2001; MAO

et al., 2007; XING et al., 2008). A alternativa para reduzir custos e o tempo de análise de grande número de amostras, necessárias para a busca sistemática de influências genéticas em doenças, foi a de realizar a GELE coletiva (mistura de DNA), para o conjunto de amostras de pacientes e para o conjunto de amostras de controles (SHAM et al., 2002; CRAIG et al., 2005; KIROV et al., 2006; DOCHERTY et al., 2007; LIN et al., 2008; MACGREGOR et al., 2008).

No método de GELE coletiva, quantidades idênticas de DNA de pacientes e controles são misturadas em dois conjuntos distintos: um conjunto para pacientes e um conjunto para controles. A seguir, os genótipos são analisados em ambas as amostras por meio do método de microarranjos e a frequência alélica é estimada por métodos estatísticos específicos (SHAM et al., 2002; CRAIG et al., 2005; KIROV et al., 2006; DOCHERTY et al., 2007; LIN et al., 2008; MACGREGOR et al., 2008).

Até o momento, a GELE individual por meio de microarranjos de DNA somente foi realizada em tumores de cabeça e pescoço localizados na região da nasofaringe. Ng et al. (2009) utilizaram a GELE individual para a identificação de SNPs em 168 amostras de DNA do sangue periférico de pacientes com CEC de nasofaringe e 260 controles. Os autores identificaram oito SNPs localizados nos introns 1 e 3 do gene da integrina  $\alpha$ -9 (*ITGA9*). Indivíduos que apresentaram os genótipos variantes desses polimorfismos, isolados ou agrupados, apresentaram por volta de três vezes mais chances de desenvolver o carcinoma de nasofaringe do que aqueles com os genótipos selvagens (NG et al., 2009). Tse et al. (2009) também utilizaram a GELE individual para identificar SNPs em amostras de 277 pacientes com CEC de nasofaringe e 285 controles. Eles identificaram 12 SNPs localizados na região do cromossomo 6p21.3. Após a validação dos resultados em 339 pacientes e 696 controles, os pesquisadores identificaram dois SNPs localizados no gene *HLA-A2* e dois SNPs no gene *HLA-F*, ambos codificadores de proteínas do complexo de histocompatibilidade e um SNP no gene do receptor do ácido gama aminobutírico B (*GABBR1*). Recentemente, Bei et al. (2010) também utilizaram a GELE individual para avaliar SNPs em 1.583 amostras de DNA de pacientes com CEC de nasofaringe e 1.894 controles. Eles identificaram 49 SNPs que foram validados em 3.507 pacientes e 3.063 controles. Três novos SNPs localizados nos genes *TNFRSF19*, *MDS1-EVII* e *CDKN2A-CDKN2B* foram identificados após a validação dos resultados.

Até o momento, a GELE coletiva não foi realizada em tumores de cabeça e pescoço.

Entre os outros tumores analisados por meio deste método, encontram-se os de pulmão e pâncreas (SPINOLA et al., 2007; DIEGAARDE et al., 2010). Spinola et al. (2007) utilizaram a GELE coletiva para a identificação de SNPs em 338 amostras de DNA do sangue periférico de pacientes com adenocarcinoma de pulmão e 335 controles. Os pesquisadores realizaram a mistura de DNA de 50 pacientes e 50 controles em cada microarranjo de DNA. Eles identificaram 38 SNPs candidatos ao risco de ocorrência do tumor. Um único SNP, localizado no gene Kruppel fator 6 (*KLF6*), foi associado com o aumento do risco para o câncer de pulmão, após a validação dos resultados em amostras individuais. Recentemente, Diergaarde et al. (2010) também utilizaram a GELE coletiva para a identificação de SNPs em amostras de DNA do sangue periférico de 103 pacientes com câncer de pâncreas e 182 controles. Os autores realizaram a mistura de DNA de aproximadamente 20 pacientes e 20 controles em cada microarranjo de DNA. Eles identificaram 26 SNPs candidatos e após a validação em amostras individuais, apenas um SNP localizado no gene da gama glutamiltransferase 1 (*GGT1*) foi associado com o aumento do risco para o câncer de pâncreas.

Até o momento, somente a GELE individual foi utilizada para a análise de CNVs. Para a identificação de CNVs em CEC de cavidade oral, Tong et al. (2004) utilizaram a GELE individual em 35 amostras de tumores e identificaram sítios de perda alélica, ainda não descritos na doença, no braço longo do cromossomo 6. Já Zhou et al. (2004a) utilizaram a mesma plataforma para caracterizar a perda de heterozigosidade (LOH) em 11 linhagens celulares de câncer de cavidade oral e identificaram a frequente LOH no braço curto do cromossomo 8 em oito linhagens celulares tumorais. O mesmo grupo de pesquisadores (ZHOU et al., 2004b) também utilizou a GELE individual em linhagens celulares normais e de câncer de cavidade oral. Eles identificaram as regiões cromossômicas dos braços curtos dos cromossomos 3 e 11 associadas com a progressão maligna da doença. Ainda, Zhou et al. (2005) também observaram a deleção da região 3p11-p12 em tumores de cinco pacientes com CEC da cavidade oral em relação a dois controles. Além disso, em 2007, Ye et al. identificaram regiões de LOH nos braços dos cromossomos 3p, 4p, 4q, 5q, 8p, 9p, 10p, 11q e 17p no DNA de linhagens celulares de câncer de cavidade oral. A perda da região genômica 8p21.3~p22 foi a mais frequente entre as descritas. Assim, o grupo de pesquisadores sugeriu como candidato para a tumorigênese, o gene supressor tumoral mitocondrial 1 (*MTUS1*), que está localizado nessa região.

A eficácia da GELE coletiva foi descrita como similar à da GELE individual (diferença alélica: 0,5-2,0%) para identificar SNPs em indivíduos caucasianos saudáveis (CRAIG et al., 2005; MEABURN et al., 2005; DOCHERTY et al., 2006; KIROV et al., 2006; MEABURN et al., 2006). Entretanto, Chiang et al. (2010) observaram que a GELE coletiva foi menos eficaz que a GELE individual para identificar SNPs em indivíduos de outras etnias.

Em nosso conhecimento, não há descrições sobre o uso da GELE para a identificação de SNPs e CNVs em amostras de DNA de pacientes com CEC de BL e indivíduos saudáveis. Além disso, também não há estudos comparativos entre os métodos da GELE individual e da GELE coletiva para identificar SNPs e CNVs em indivíduos saudáveis e portadores do CEC de BL.

O câncer é a segunda causa de óbito na região sudeste do Brasil (DUNCAN et al., 1992). Por outro lado, o tabagismo, o etilismo e o HPV constituem problemas importantes na saúde de indivíduos da nossa população (INSTITUTO NACIONAL DO CÂNCER, 2011). Além disso, a população brasileira é altamente heterogênea, composta por indígenas e imigrantes da Europa, Ásia e África (ALVES-SILVA, 2000; CARVALHO-SILVA, 2001).

Frente ao exposto, nos pareceu de interesse identificar os papéis de grande número de SNPs e CNVs em uma população de alta heterogeneidade racial e de uma área com alta incidência de câncer de cabeça e pescoço, como a de Campinas em São Paulo.

# **OBJETIVOS**

---



Foram avaliados pacientes com CEC de BL atendidos nos ambulatórios de Oncologia Clínica e de Otorrinolaringologia do Hospital de Clínicas da UNICAMP e indivíduos saudáveis do HEMOCENTRO da UNICAMP, tendo como objetivos:

- Identificar SNPs e CNVs envolvidos com a alteração do risco para o CEC de BL por meio de microarranjos de DNA, utilizando a GELE individual e coletiva,
- Verificar a eficácia da GELE coletiva na identificação de SNPs e CNVs utilizando o padrão dos resultados da GELE individual de pacientes e controles da nossa população.



# **PACIENTES E MÉTODOS**

---



Foram avaliados 55 pacientes com CEC de BL atendidos ao diagnóstico nos ambulatórios de Oncologia Clínica e de Otorrinolaringologia do Hospital de Clínicas da UNICAMP durante o período de fevereiro de 2007 a março de 2010 . Foi também avaliado um grupo controle, durante o mesmo período, constituído por 55 doadores de sangue atendidos no HEMOCENTRO da UNICAMP, que não referiram antecedente familiar de câncer e que foram pareados aos pacientes por sexo e raça.

## **1. Avaliação clínica**

Os dados relativos à identificação, à idade, ao sexo, à raça, ao padrão da dieta, ao hábito de fumar e ao etilismo, foram coletados de pacientes e controles por meio de questionário específico (Anexo 1), aplicado pelo pesquisador responsável pelo estudo.

A definição de raça foi autoreferida pelos indivíduos inseridos no estudo e realizada por meio do biótipo de cada um.

A ingestão dos diferentes grupos de alimentos foi considerada adequada ou inadequada de acordo com o número de porções preconizadas pelo Ministério da Saúde para a população brasileira (Anexo 1) (GOUVEIA & BRUNO, 2001).

Foram considerados usuários de bebidas alcoólicas aqueles indivíduos que ingeriam algum tipo de bebida alcoólica pelo menos uma vez na semana, de forma regular. Foram classificados em etilistas discretos, aqueles que ingeriram bebida alcoólica na proporção de uma até sete doses por semana, etilistas moderados, aqueles que ingeriram de oito doses até 42 doses por semana, e como etilistas acentuados, aqueles que ingeriram 43 doses ou mais por semana. Nos demais casos, foram classificados como não usuários de bebidas alcoólicas, de acordo com critério estabelecido por Huang et al. (2003). Vale comentar que a quantidade considerada para um drinque foi de cerca de 355ml de cerveja ou 148ml de vinho ou ainda 44ml de bebida destilada de acordo com padrão estabelecido pelo Departamento de Agricultura dos Estados Unidos da América e pelo Departamento de Saúde e Serviços Humanos (NATIONAL INSTITUTE ON ALCOHOL ABUSE AND ALCOHOLISM-NIAAA, 1995).

Foram considerados tabagistas aqueles que mantiveram o hábito de fumar até o momento do diagnóstico do CEC de BL (pacientes) ou da entrevista para inclusão no estudo (controles). Foram considerados não tabagistas aqueles que nunca fumaram e ex-tabagistas aqueles que mantiveram o hábito de fumar até, no máximo, cinco anos antes do

diagnóstico do CEC de BL (pacientes) ou da entrevista para inclusão no estudo (controles), de acordo com o proposto por Freedman et al. (1996). Os tabagistas por sua vez foram classificados como tabagistas discretos, aqueles que fumaram menos de dez cigarros por dia por menos de dez anos, tabagistas acentuados, aqueles que fumaram 30 ou mais cigarros por dia por 30 anos ou mais, ou que fumaram charuto ou cachimbo. Foram considerados tabagistas moderados aqueles indivíduos classificados em qualquer outro tipo de situação (HUANG et al., 2003).

## **2. Aspectos do tumor**

Os resultados dos exames laboratoriais necessários ao diagnóstico, à determinação do grau de diferenciação e do estágio do tumor foram obtidos dos prontuários de cada paciente pelo pesquisador responsável pelo estudo.

O diagnóstico do CEC de BL foi realizado em cortes histológicos de fragmentos do tumor incluídos em parafina e corados por hematoxilina e eosina. Eles foram processados e analisados no Laboratório de Anatomia Patológica do Hospital de Clínicas da UNICAMP. Cada tumor foi classificado de acordo com o grau de diferenciação, como bem diferenciado, moderadamente diferenciado, pouco diferenciado e indiferenciado (MORI et al., 1998).

O estadiamento do tumor foi realizado com base nos resultados obtidos do exame clínico realizado por médico oncologista ou otorrinolaringologista e os resultados dos exames de rastreamento de metástases, que incluíram a tomografia computadorizada do pescoço e a radiografia do tórax, de acordo com critérios convencionais (Anexos 2 e 3) (INTERNATIONAL UNION AGAINST CANCER, 2004).

## **3. Determinação dos polimorfismos genéticos**

Para a análise dos SNPs e das CNVs foi utilizada a plataforma de GELE da empresa Affymetrix<sup>®</sup>, disponível para utilização no Laboratório Nacional de Luz Síncrotron (Barão Geraldo, Campinas, São Paulo).

Foram utilizados microarranjos de DNA contendo 500.568 SNPs e 420.000 sondas de DNA para CNVs (Genome-Wide Human SNP Array 5.0). O sistema incluiu oligonucleotídeos pré-arranjados sobre lâminas de vidro, instrumentos, conjunto de

reagentes e programas computacionais necessários para o desenvolvimento do protocolo otimizado na análise simultânea dos SNPs e das CNVs.

Essas lâminas de microarranjos foram utilizadas para a GELE individual de todos os pacientes e controles incluídos no estudo (uma lâmina por indivíduo). Para a GELE coletiva, em uma lâmina foi colocada a mistura do DNA de todos os pacientes caucasóides e em outra a mistura do DNA dos pacientes negróides. O mesmo esquema foi utilizado para a GELE coletiva dos controles. As amostras de DNA foram cuidadosamente quantificadas antes de serem agrupadas para se certificar que a frequência alélica obtida seria real e não resultado de variações de concentração de DNA.

Os procedimentos para a identificação dos SNPs e das CNVs foram realizados de acordo com o protocolo do fabricante, descrito sucintamente abaixo.

### **3.1. Extração do DNA a partir dos leucócitos**

O DNA genômico foi obtido de amostras de sangue periférico dos pacientes com CEC de BL e dos indivíduos controle, utilizando o kit de extração de DNA *QIAamp DNA Blood Kit* (QIAGEN, Hilden, Germany). Sua concentração foi determinada por meio de espectrofotômetro (Nanodrop ND100, Wilmington, Delaware, USA). O DNA foi diluído em tampão TE (Tris EDTA) (0,1mM ácido etilenodiamino tetracético, 10,0mM Tris HCl, pH 8,0) para 50ng/μl e foi armazenado a -20°C até o próximo passo.

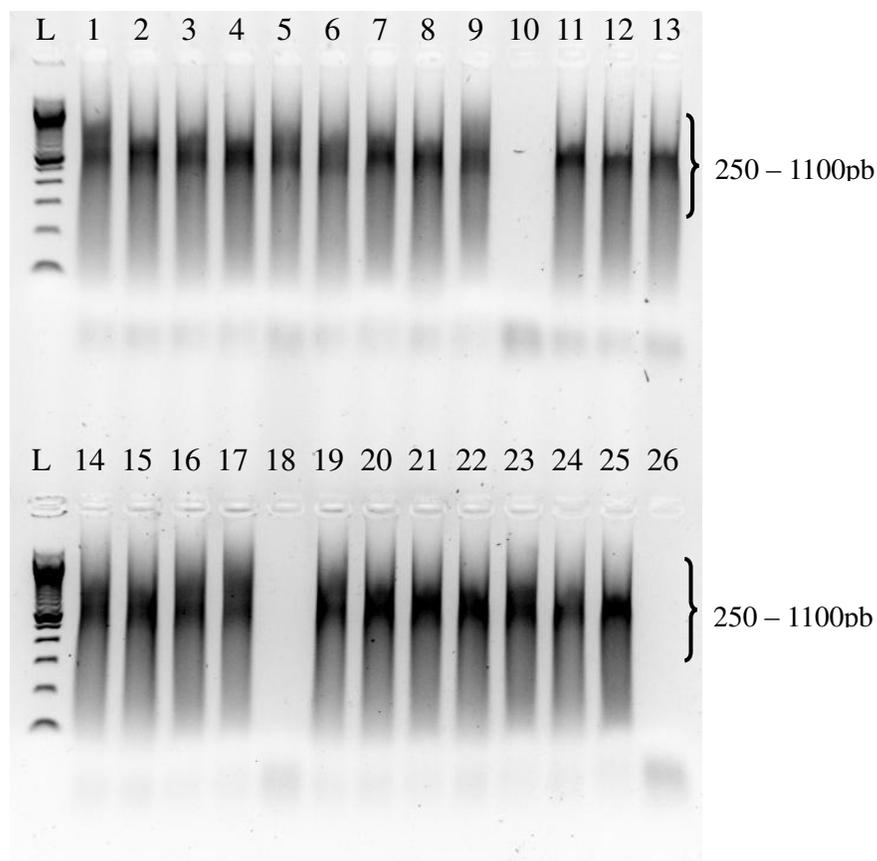
### **3.2. Digestão com a enzima *StyI*, ligação a adaptadores e amplificação**

Os DNAs dos pacientes e controles foram submetidos à digestão com a enzima *StyI* para a obtenção de fragmentos de DNA de tamanhos variados. Foi utilizada uma mistura de 5μl de DNA (50ng/μl), 11,5μl de água estéril, 2,0μl de tampão (10X), 0,2μl de albumina bovina sérica (BSA) (100X, 10mg/ml) e 1,0μl da enzima (10U/μl). A seguir, a mistura foi colocada em termociclador (Eppendorf, Hamburg, Germany) por duas horas a 37°C e 20 minutos a 65°C. Os produtos da digestão foram armazenados a -20°C até o próximo passo.

A seguir, foi realizada a ligação dos fragmentos de DNA a adaptadores específicos, com a mistura de 0,75μl de adaptador *StyI* (50μM), 2,5μl de tampão (10X) e 2,0μl da enzima T4 DNA ligase (400U/μl). A mistura foi colocada em termociclador por três horas a 16°C e 20 minutos a 70°C. Os produtos da reação de ligação foram diluídos com 75μl de

água estéril e foram armazenados a -20°C até o próximo passo.

Os fragmentos de DNA ligados a adaptadores *StyI* foram submetidos a três PCRs realizadas com a utilização de 39,5µl de água estéril, 10µl de tampão (10X), 20µl de potencializador para PCR (5M), 14µl de dNTP (desoxirribonucleosídeo trifosfato) (2,5mM cada), 4,5µl do iniciador específico (100µM) e 2µl da enzima TITANIUM™ *Taq* DNA polimerase (50X). As reações compreenderam 30 ciclos de incubação a 94°C (30 segundos), a 60°C (45 segundos) e a 68°C (15 segundos). Foram obtidos fragmentos de 250 a 1100 pares de bases (pb). A presença dos fragmentos foi analisada por eletroforese em gel de agarose 2% (Figura 1) e os produtos das PCRs foram armazenados a -20°C até o próximo passo.



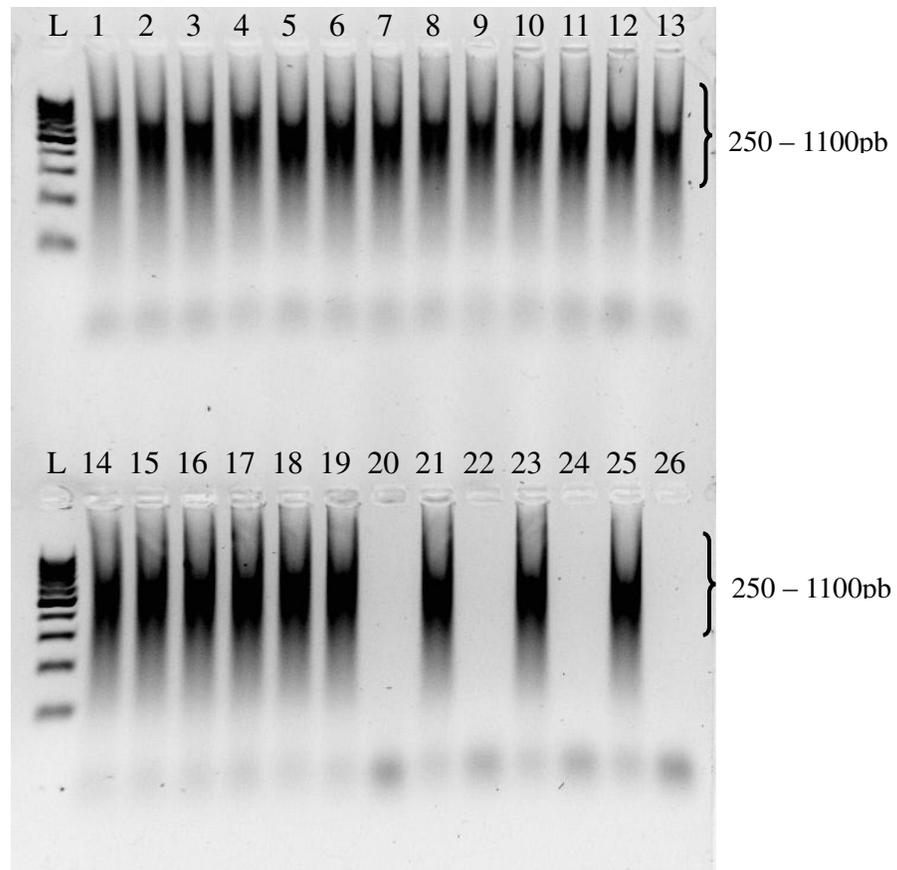
**Figura 1. Método de genotipagem em larga escala em portadores de carcinoma de células escamosas de base de língua.** Produtos da amplificação dos fragmentos obtidos da digestão com a enzima *StyI* de DNA genômico em gel de agarose 2%, para a hibridização em lâminas de microarranjos. Os marcadores do tamanho do DNA, *ladder* 100 pares de bases (pb), estão representados nas colunas L. Os fragmentos de 250 a 1100 pb, que corresponderam à amplificação dos fragmentos da digestão enzimática, estão representados nas colunas 1 a 9, 11 a 17 e 19 a 25. Os resultados obtidos de controles negativos (água) estão representados nas colunas 10, 18 e 26

### 3.3. Digestão com a enzima *NspI*, ligação a adaptadores e amplificação

Os DNAs dos pacientes e controles também foram submetidos à digestão com a enzima *NspI* para a obtenção de fragmentos de DNA de tamanhos variados. Foi utilizada uma mistura de 5µl de DNA (50ng/µl), 11,5µl de água estéril, 2,0µl de tampão (10X), 0,2µl de BSA (100X, 10mg/ml) e 1,0µl da enzima (10U/µl). A seguir, a mistura foi colocada no termociclador por duas horas a 37°C e 20 minutos a 65°C. Os produtos da digestão foram armazenados a -20°C até o próximo passo.

A seguir, foi realizada a ligação dos fragmentos de DNA a adaptadores específicos com a mistura de 0,75µl de adaptador *NspI* (50µM), 2,5µl de tampão (10X) e 2,0µl da enzima T4 DNA ligase (400U/µl). A mistura foi colocada em termociclador por três horas a 16°C e 20 minutos a 70°C. Os produtos da reação de ligação foram diluídos com 75µl de água estéril e foram armazenados a -20°C até o próximo passo.

Os fragmentos de DNA ligado a adaptadores *NspI* foram submetidos a quatro PCRs com a utilização de 39,5µl de água estéril, 10µl de tampão (10X), 20µl de potencializador para PCR (5M), 14µl de dNTP (2,5mM cada), 4,5µl do iniciador específico (100µM) e 2µl da enzima TITANIUM™ *Taq* DNA polimerase (50X). As reações compreenderam 30 ciclos de incubação a 94°C (30 segundos), a 60°C (45 segundos) e a 68°C (15 segundos). Foram obtidos fragmentos de 250 a 1100 pb. A presença dos fragmentos foi analisada por eletroforese em gel de agarose 2% (Figura 2). Os produtos das PCRs foram armazenados a -20°C até o próximo passo.



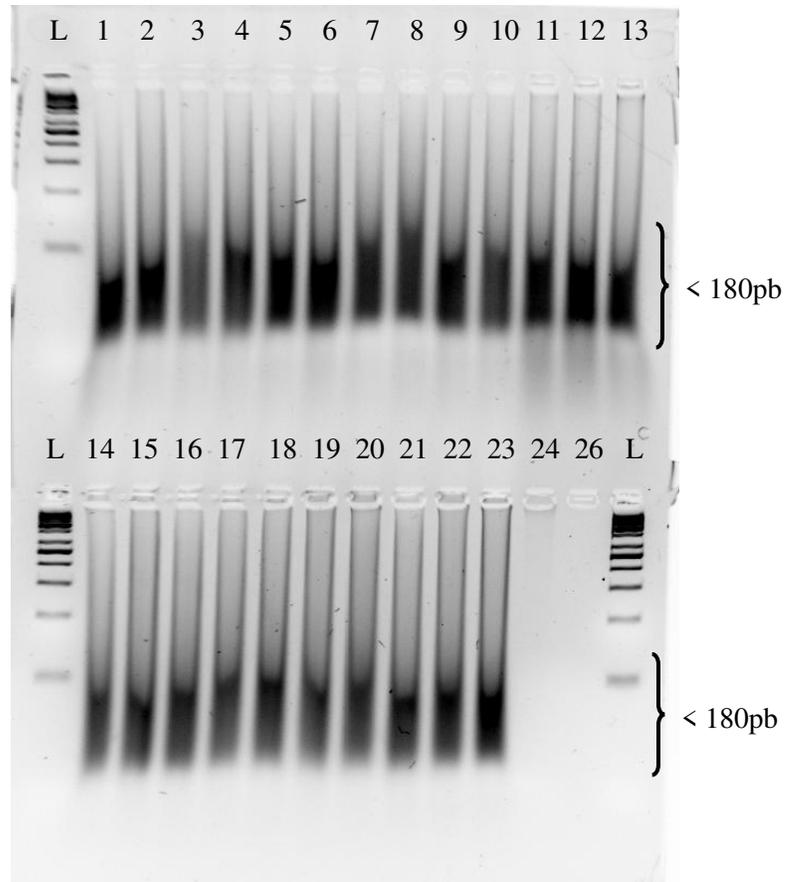
**Figura 2. Método de genotipagem em larga escala em portadores de carcinoma de células escamosas de base de língua.** Produtos da amplificação dos fragmentos obtidos da digestão com a enzima *NspI* de DNA genômico em gel de agarose 2%, para a hibridização em lâminas de microarranjos. Os marcadores do tamanho do DNA, *ladder* 100 pares de bases (pb), estão representados nas colunas L. Os fragmentos de 250 a 1100 pb, que corresponderam à amplificação dos fragmentos da digestão enzimática, estão representados nas colunas 1 a 19, 21, 23 e 25. Os resultados obtidos de controles negativos estão representados nas colunas 20, 22, 24 e 26

### **3.4. Mistura dos produtos das PCRs, purificação e quantificação**

Os produtos das PCRs das enzimas *StyI* e *NspI* foram misturados. A mistura compreendeu 100µl do produto de cada uma das três PCRs da enzima *StyI* e 100µl do produto de cada uma das quatro PCRs da *NspI*. Após a mistura (volume final de 700µl), os produtos foram purificados por meio de pérolas magnéticas e colunas magnéticas, utilizando o sistema de vácuo da empresa Millipore® (Billerica, MA, EUA) e, a seguir, quantificados em espectrofotômetro, com leitura de absorbância a 260nm. A concentração dos produtos purificados das PCRs foi ajustada para 50µg. Os produtos das PCRs purificadas foram armazenados a -20°C até o próximo passo.

### **3.5. Fragmentação dos produtos das PCRs**

Os produtos das PCRs inicialmente foram submetidos à diluição com a utilização de 45µl do produto da PCR purificada e 5µl de tampão de fragmentação (mistura de fragmentação). Em seguida a mistura foi submetida à fragmentação com a adição de 5µl de reagente de fragmentação (0,1U/µl). A mistura foi colocada em um termociclador por 35 minutos a 37°C e 15 minutos a 95°C. A fragmentação foi observada pela presença de fragmentos menores que 180 pb em eletroforese em gel de agarose 4% (Figura 3). Os fragmentos obtidos foram utilizados imediatamente.



**Figura 3. Método de genotipagem em larga escala em portadores de carcinoma de células escamosas de base de língua.** Produtos da fragmentação da mistura dos produtos das PCRs obtidos da digestão enzimática com o reagente de fragmentação, para a hibridização em lâminas de microarranjos, em gel de agarose 4%. Os marcadores do tamanho do DNA, *ladder* 100 pares de bases (pb), estão representados nas colunas L. Os fragmentos menores que 180 pb, que corresponderam à fragmentação das amostras, estão representados nas colunas 1 a 23. Os resultados obtidos de controles negativos estão representados nas colunas 24 e 26

### **3.6. Marcação do DNA fragmentado com biotina**

Inicialmente foi preparada a mistura de marcação com 14µl de tampão *terminal deoxynucleotidyl transferase* (TdT) (5X), 2µl de reagente de marcação de DNA (30mM) (biotina) e 3,5µl da enzima TdT (30U/µl). A marcação foi realizada com a mistura de 53,5µl de DNA fragmentado e 19,5µl da mistura de marcação. A mistura foi colocada em um termociclador por quatro horas a 37°C e 15 minutos a 95°C. O DNA marcado com biotina foi armazenado a -20°C até o próximo passo.

### **3.7. Hibridização**

Primeiramente foi preparado a mistura de hibridização com 12µl de ácido *4-morpholineethanesulfonic* (MES) (1,25M), 13µl de dimetil sulfóxido (DMSO) (100%), 13µl de solução *Denhardt's* (50X), 3µl de EDTA (0,5M), 3µl de DNA de esperma de salmão (HSDNA) (10mg/ml), 2µl *oligo control reagent*, 0,100 (OCR), 3µl da enzima Cot-1 humana (1mg/ml), 1µl de tween-20 (3%) e 140µl de cloreto de amônio tetra metil (TMACL) (5M). Em seguida foi realizada a mistura de 73µl de DNA biotilado e 190µl da mistura de hibridização. Esta mistura foi aquecida em banho-maria a 95°C por 10 minutos e resfriada a 49°C até o próximo passo. Após estes procedimentos, 200µl da mistura desnaturada foi depositada em cada uma das lâminas de vidro com os oligonucleotídeos pré-arranjados. As lâminas e o material com o DNA foram hibridizados a 50°C durante 16 a 18 horas a 60 rpm.

### **3.8. Lavagens das lâminas e marcação com fluoróforos**

Inicialmente foram preparados os tampões de lavagens e de marcações. O tampão de lavagem “A” consistiu da mistura de 300ml de solução de hibridização (SSPE) (3M NaCl, 0,2M NaH<sub>2</sub>PO<sub>4</sub>, 0,02M EDTA) (20X), 1ml de tween-20 (10%) e 699ml de água estéril. O tampão de lavagem “B” consistiu da mistura de 30ml de SSPE (20X), 1ml de tween-20 (10%) e 969ml de água estéril. O tampão de marcação resultou da mistura de 800,04µl de água estéril, 360µl de SSPE (6X), 3,96µl de tween-20 (3%) e 24µl de *Denhardt's* (50X).

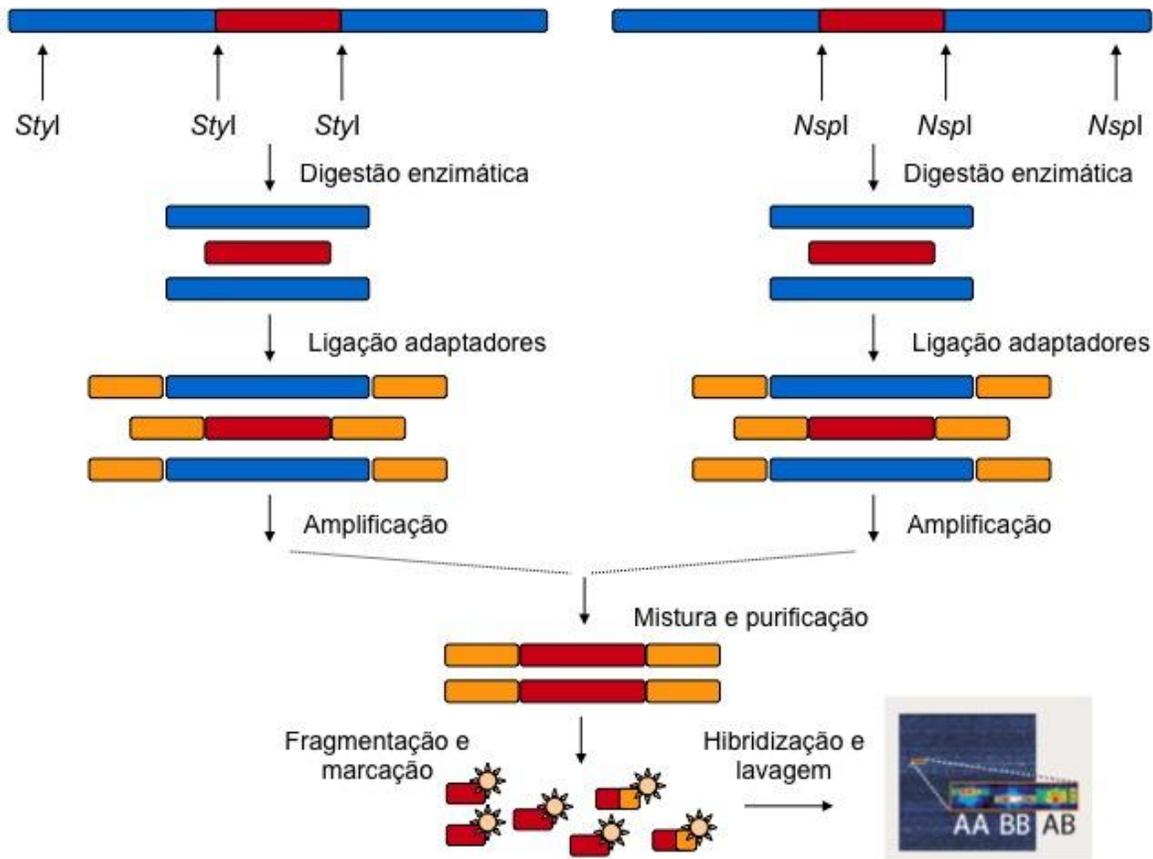
Em seguida, foi preparada a solução de marcação com estreptavidina e ficoeritrina (SAPE), que consistiu da mistura de 594µl de tampão de marcação, preparado anteriormente, e 6µl de SAPE (1mg/ml); e a solução de marcação com o anticorpo biotilado antiestreptavidina, que consistiu da mistura de 594µl de tampão de marcação, preparado anteriormente, e 6µl de anticorpo biotilado antiestreptavidina (5µg/ml).

O tampão de microarranjos foi preparado com a mistura de 8,3ml de MES (12X), 18,5ml de cloreto de sódio (5M), 0,1ml de tween-20 (10%) e 73,1ml de água estéril. Após a hibridização, as lavagens as marcações foram realizadas com a utilização da estação *Fluidics Station 450*. As sequências de lavagens e marcações dos microarranjos contidos nas lâminas por fluoróforos estão descritas na Tabela 1.

**Tabela 1.** Sequências de lavagens e marcações dos microarranjos contidos em lâminas por fluoróforos para a genotipagem em larga escala de polimorfismos gênicos de base única e identificação do número de cópias do DNA em pacientes com tumores de células escamosas de base de língua e controles

<b>Etapas</b>	<b>Descrição</b>
Primeira lavagem pós-hibridização	Seis ciclos (cinco misturas por ciclo) com tampão de lavagem “A” a 25°C
Segunda lavagem pós-hibridização	24 ciclos (cinco misturas por ciclo) com tampão de lavagem “B” a 45°C
Marcação com estreptavidina ligada a ficoeritrina	Dez minutos em solução SAPE a 25°C
Lavagem pós-marcação	Seis ciclos (cinco misturas por ciclo) com tampão de lavagem “A” a 25°C
Segunda marcação (com o anticorpo biotilado antiestreptavidina)	Dez minutos em solução de marcação com o anticorpo antiestreptavidina a 25°C
Terceira marcação (com estreptavidina ligada a ficoeritrina)	Dez minutos em solução SAPE a 25°C
Lavagem final	Dez ciclos (seis misturas por ciclo) com tampão de lavagem “A” a 25°C

A representação esquemática da sequência de passos técnicos envolvidos no procedimento descrito acima está apresentada na Figura 4.



**Figura 4.** Representação esquemática do protocolo Affymetrix® *Genome-Wide Human SNP Array 5.0* para a genotipagem em larga-escala de pacientes e controles. O DNA genômico dos indivíduos é digerido com as enzimas *StyI* e *NspI*, em seguida, é realizada a reação de ligação de adaptadores específicos para os fragmentos obtidos após a digestão enzimática. As reações de amplificação, mistura e purificação desses fragmentos são realizadas, seguidas da fragmentação e marcação dos fragmentos de DNA com a biotina. Após, o DNA biotilado é hibridizado na lâmina de microarray seguido das etapas de lavagens, marcações com estreptavidina ligada a ficoeritrina, e leitura dos resultados por meio da captação de fluorescência pelo *scanner*

### 3.9. Captação de imagens e genotipagem

As lâminas com microarranjos foram colocadas no *GeneChip Scanner 3000 7G*, gerenciado pelo programa *Command Console*. As imagens geradas pela excitação dos fluoróforos (ficoeritrina) captadas pelo *scanner* foram submetidas primeiramente à análise de controle de qualidade geral (CQ) por meio do programa *Genotyping Console* (versão 3.0.2, Affymetrix<sup>®</sup>, Santa Clara, CA, USA). Só foram considerados adequados para análise da GELE individual e coletiva as lâminas que obtiveram taxas iguais ou maiores que 86% de CQ.

Os dados obtidos foram analisados para identificar genótipos de SNPs e CNVs pelos programas *Genotyping Console* (versão 3.0.2, Affymetrix<sup>®</sup>, Santa Clara, CA, USA), com o algoritmo computacional *bayesian robust linear model with mahalanobis* (brlmm) e o programa Bioconductor (versão 2.7, linguagem R), com o algoritmo *corrected robust linear model with maximum likelihood classification* (crlmm). Os algoritmos utilizaram as etapas de normalização e sumarização das intensidades das sondas, seguido da genotipagem. Os resultados obtidos por meio do algoritmo crlmm (programa Bioconductor) foram considerados “padrão ouro” para a genotipagem dos SNPs.

Vale comentar que todas as análises realizadas por meio do programa Bioconductor foram desenvolvidas pelo pesquisador responsável pelo estudo junto com o Prof. Dr. Benilton Carvalho, do Grupo de Biologia Computacional, no Departamento de Oncologia da Universidade de Cambridge, em Cambridge, Inglaterra.

Cada amostra também foi analisada por meio do programa Partek<sup>®</sup> (St. Louis, MO, USA) para a identificação das CNVs por meio dos algoritmos *circular binary segmentation* (cbs) e *hidden Markov model* (hmm). Os resultados obtidos por meio da GELE individual foram considerados “padrão ouro” para a identificação de CNVs.

Vale comentar que todas as análises realizadas por meio do programa Partek<sup>®</sup> foram realizadas pelo pesquisador responsável pelo estudo com o auxílio da pesquisadora Renata Pellegrino, doutoranda do Departamento de Psicobiologia da Universidade Federal de São Paulo.

Os SNPs e CNVs de genes de interesses foram escolhidos entre os diferencialmente frequentes em pacientes e controles ( $P < 0,01$ ) com base nas funções de cada um por meio do programa *Database for Annotation, Visualization and Integrated Discovery (DAVID)* e do banco de dados de genes do *National Center for Biotechnology Information (NCBI)*, todos disponíveis para a utilização em *websites* públicos. Optamos por selecionar genes relacionados com os processos de carcinogênese.

#### **4. Análise estatística**

O teste de verificação do equilíbrio de Hardy-Weinberg foi realizado com o intuito de verificar se ocorreu distribuição preferencial de algum dos genótipos dos SNPs avaliados no grupo de pacientes e controles utilizados no estudo.

O significado estatístico das diferenças dos SNPs entre os grupos foi calculado por meio do teste de Fisher ou  $\chi^2$  utilizando o programa Partek<sup>®</sup> para a genotipagem obtida pelo algoritmo brlmm e por meio da regressão logística múltipla utilizando o programa Bioconductor para a genotipagem obtida por meio do algoritmo crlmm.

Para as diferenças entre as CNVs, o significado estatístico foi calculado por meio do programa Partek<sup>®</sup>. O algoritmo hmm foi utilizado para a determinação dos estados das CNVs: 0, 1, 2, 3 e 4, onde 0 = deleção em homozigose, 1 = deleção em heterozigose, 2 = diplóide normal, 3 = ganho de uma única cópia e 4 = ganho de duas cópias. Os controles saudáveis foram utilizados como base de referência para as análises. Os ganhos e as perdas de material genômico foram determinados com base das razões  $\log_2$  das intensidades das sondas obtidas, e as sequências foram consideradas amplificadas ou deletadas quando ultrapassaram a variação de três desvios padrão.

As análises das CNVs em GELE coletiva foram realizadas considerando pacientes e controles caucasianos, pacientes e controles negróides e pacientes em geral (caucasianos e negróides) comparados a controles em geral (caucasianos e negróides).

Todos os resultados foram selecionados utilizando-se um nível de significância de 1% ( $P < 0,01$ ). As determinações dos riscos de ocorrência do CEC de BL, a que as pacientes e controles foram submetidos, foram obtidas por meio das razões das chances (ORs) e calculadas considerando-se um intervalo de confiança de 95%.

## **5. Aspectos éticos**

O estudo molecular foi realizado em amostras de sangue periférico obtidas por ocasião da punção venosa realizada para a coleta de exames necessários para os exames pré-operatórios e de estagiamento no momento do diagnóstico ou acompanhamento clínico dos pacientes e de uma única punção venosa em controles. Nenhum material adicional foi coletado dos pacientes ou controles.

O raio-X de tórax e a tomografia computadorizada do pescoço foram exames realizados de rotina para o estagiamento da doença e para as determinações do prognóstico e da terapêutica a ser administrada.

Todos os procedimentos foram realizados após a assinatura dos termos de consentimento livre e esclarecido pelos pacientes e controles (Anexos 4 e 5) que aceitaram participar do estudo proposto. O projeto foi aprovado pelo Comitê de Ética em Pesquisa da Faculdade de Ciências Médicas da UNICAMP (Processo no. 446/2005) (Anexo 6).



# **RESULTADOS**

---



## **1. Avaliação clínica**

### **1.1. Pacientes e controles**

As distribuições individualizadas dos 55 pacientes com CEC de BL, de acordo com a idade, o sexo, a raça, o padrão da dieta e o hábito do tabagismo e do etilismo estão apresentadas no Anexo 7.

As frequências das distribuições dos 55 pacientes, de acordo com a idade, o sexo, a raça e o hábito de fumar e de ingerir bebidas alcoólicas estão apresentadas na Tabela 2. As frequências das distribuições desses pacientes de acordo com o padrão alimentar estão apresentadas na Tabela 3.

Observamos que a maioria dos pacientes foi do sexo masculino, caucasianos, com hábitos acentuados de tabagismo e etilismo e que ingeriam quantidade preconizada ou maior que a preconizada de carnes.

As distribuições individualizadas dos 55 controles de acordo com a idade, o sexo e a raça, o padrão da dieta e os hábitos do tabagismo e do etilismo estão apresentadas no Anexo 8.

As frequências das distribuições dos 55 controles de acordo com a idade, o sexo, a raça, e os hábitos de fumar e de ingerir bebidas alcoólicas estão apresentadas na Tabela 4. As características desses controles de acordo com o padrão da dieta estão apresentadas na Tabela 5.

Observamos que os indivíduos controles foram, em sua maioria, do sexo masculino, caucasianos, não tabagistas ou tabagistas discretos e que ingeriram quantidade preconizada ou maior que a preconizada de carnes, vegetais, frutas e cereais.

**Tabela 2.** Frequências das distribuições dos 55 pacientes com carcinoma de células escamosas de base de língua de acordo com a idade, o sexo, a raça e o hábito de fumar e de ingerir bebidas alcoólicas

<b>Variáveis</b>	<b>Número de pacientes (%)</b>
<b>Idade (anos)*</b>	
< 56	25 (45,4)
≥ 56	30 (54,6)
<b>Sexo</b>	
Masculino	52 (94,6)
Feminino	3 (5,4)
<b>Raça</b>	
Caucasóide	44 (80,0)
Negróide	11 (20,0)
<b>Hábito de fumar</b>	
Não tabagista	1 (1,8)
Discreto	0 (0,0)
Moderado	11 (20,0)
Acentuado	39 (70,9)
Não obtido	4 (7,3)
<b>Etilismo</b>	
Não ingere	1 (1,8)
Discreto	1 (1,8)
Moderado	9 (16,4)
Acentuado	40 (72,7)
Não obtido	4 (7,3)

\*idade mediana

**Tabela 3.** Frequências da distribuição dos 55 pacientes com carcinoma de células escamosas de base de língua de acordo com o padrão alimentar

<b>Variáveis</b>	<b>Número de pacientes (%)</b>
<b>Ingestão de carnes, aves e peixes</b>	
Preconizada	30 (54,6)
Menor que a preconizada	14 (25,4)
Maior que a preconizada	7 (12,7)
Não obtido	4 (7,3)
<b>Ingestão de vegetais</b>	
Preconizada	23 (41,8)
Menor que a preconizada	23 (41,8)
Maior que a preconizada	5 (9,1)
Não obtido	4 (7,3)
<b>Ingestão de frutas</b>	
Preconizada	22 (40,0)
Menor que a preconizada	26 (47,3)
Maior que a preconizada	3 (5,4)
Não obtido	4 (7,3)
<b>Ingestão de cereais</b>	
Preconizada	27 (49,1)
Menor que a preconizada	22 (40,0)
Maior que a preconizada	2 (3,6)
Não obtido	4 (7,3)

**Tabela 4.** Frequências das distribuições dos 55 controles de acordo com a idade, o sexo, a raça e os hábitos de fumar e ingerir bebida alcoólica

<b>Variáveis</b>	<b>Número de controles (%)</b>
<b>Idade (anos)*</b>	
< 48	24 (43,6)
≥ 48	31 (56,4)
<b>Sexo</b>	
Masculino	52 (94,6)
Feminino	3 (5,4)
<b>Raça</b>	
Caucasóide	44 (80,0)
Negróide	11 (20,0)
<b>Hábito de fumar</b>	
Não tabagista	42 (76,4)
Discreto	5 (9,1)
Moderado	6 (10,9)
Acentuado	1 (1,8)
Não obtido	1 (1,8)
<b>Etilismo</b>	
Não ingere	23 (41,8)
Discreto	26 (47,3)
Moderado	5 (9,1)
Acentuado	0 (0,0)
Não obtido	1 (1,8)

\*idade mediana

**Tabela 5.** Frequências da distribuição dos 55 controles de acordo com o padrão alimentar

<b>Variáveis</b>	<b>Número de controles (%)</b>
<b>Ingestão de carnes, aves e peixes</b>	
Preconizada	33 (68,3)
Menor que a preconizada	14 (20,2)
Maior que a preconizada	8 (6,6)
<b>Ingestão de vegetais</b>	
Preconizada	36 (60,9)
Menor que a preconizada	15 (29,6)
Maior que a preconizada	4 (4,6)
<b>Ingestão de frutas</b>	
Preconizada	26 (51,8)
Menor que a preconizada	23 (38,7)
Maior que a preconizada	6 (4,6)
<b>Ingestão de cereais</b>	
Preconizada	41 (74,6)
Menor que a preconizada	14 (25,4)
Maior que a preconizada	0 (0,0)

## **2. Aspectos do tumor**

### **2.1. Grau de diferenciação e estágio do tumor**

As distribuições individualizadas dos 55 pacientes de acordo com o grau de diferenciação e o estágio do tumor estão apresentadas no Anexo 9.

As distribuições dos 55 pacientes com CEC de BL, de acordo com as características do tumor estão apresentadas na Tabela 6.

**Tabela 6.** Frequências das distribuições dos 55 pacientes com carcinoma de células escamosas de base de língua de acordo com a diferenciação e o estágio do tumor

<b>Variáveis</b>	<b>Número de pacientes (%)</b>
<b>Diferenciação do tumor</b>	
Bem diferenciado	3 (5,4)
Moderadamente diferenciado	37 (67,3)
Pouco diferenciado	9 (16,4)
Indiferenciado	1 (1,8)
Não definido	5 (9,1)
<b>Estágio do tumor</b>	
I	1 (1,8)
II	6 (10,9)
III	8 (14,6)
IV	39 (70,9)
Ainda não definido	1 (1,8)

Observamos que a maioria dos nossos pacientes apresentou tumor moderadamente diferenciado e em estágio avançado.

Um paciente com CEC de BL está apresentado nas Figuras 5 e 6.

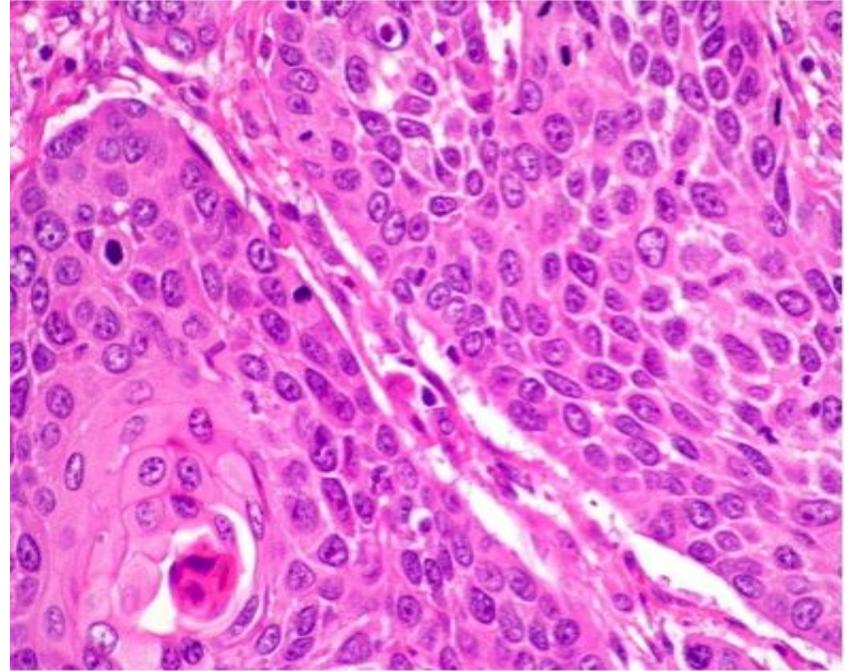
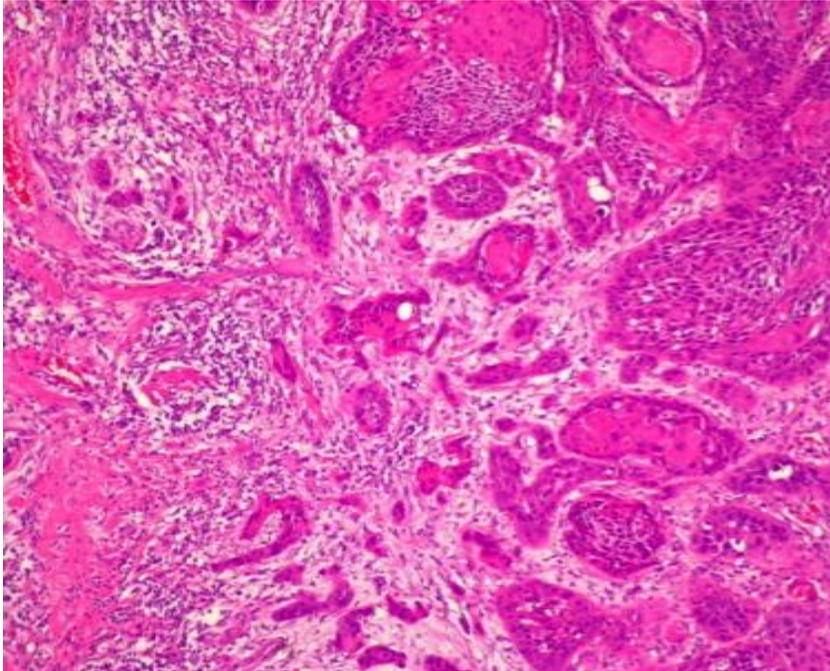
### **3. Determinação dos polimorfismos genéticos**

#### **3.1. Quantidade e qualidade de DNA nas amostras**

Cada amostra de DNA obtida do sangue periférico dos indivíduos incluídos no estudo foi cuidadosamente quantificada por meio do espectrofotômetro *NanoDrop<sup>TM</sup> 2000* (Thermo Scientific, Wilmington, DE, USA). As concentrações de DNAs de pacientes e controles estão apresentadas nas Tabelas 7 e 8, respectivamente.



**Figura 5.** Imagem de um paciente com carcinoma de células escamosas de base de língua  
(fonte: <http://www.meded.ucsd.edu>)



**Figura 6.** Aspectos histológicos do carcinoma de células escamosas de base de língua de paciente analisado no estudo (Caso nº 13)  
(H&E, 20x e 40x, respectivamente)

**Tabela 7.** Concentrações de DNA em ng/μl de amostras de sangue periférico dos 55 pacientes com carcinoma de células escamosas de base de língua

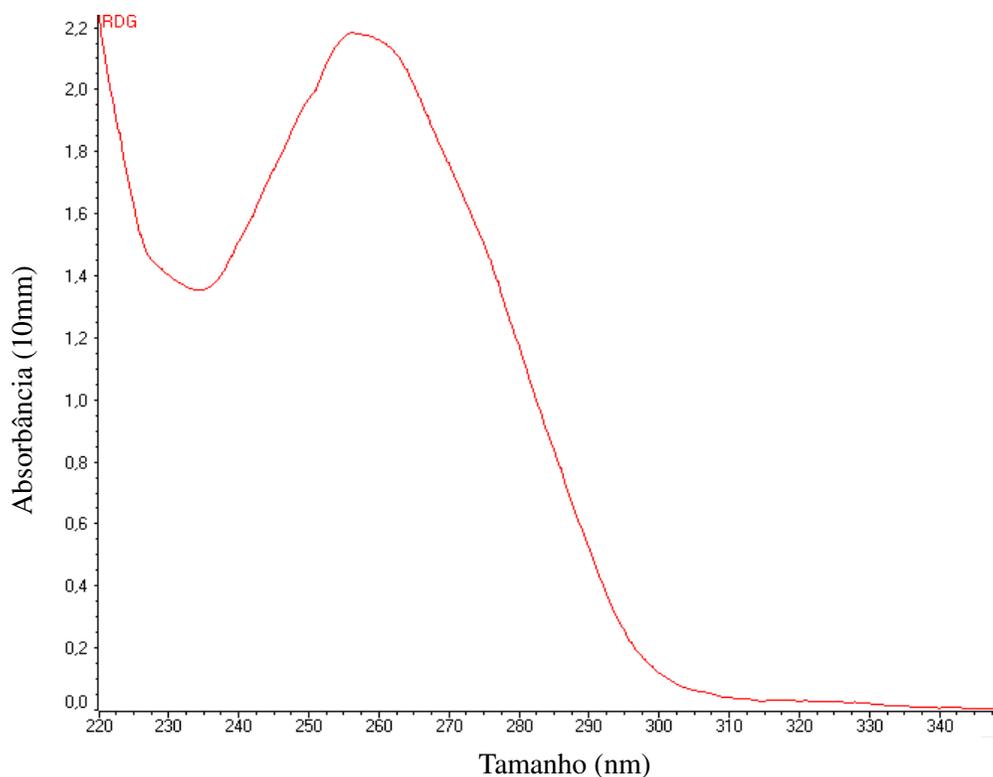
Nº	Iniciais	DNA (ng/μl)	Nº	Iniciais	DNA (ng/μl)
1	AAP	121,1	29	MM	1067,5
2	EAC	347,0	30	MPP	143,0
3	JF	531,7	31	MR	274,1
4	PLB	58,6	32	PA	379,2
5	RDG	129,1	33	DB	473,9
6	SAL	166,3	34	JF	52,9
7	SAC	95,5	35	OD	174,9
8	SRJL	185,3	36	JLT	441,8
9	WS	81,1	37	OBF	100,7
10	JAOD	194,6	38	AV	189,1
11	JS	6,4	39	JJ	62,0
12	JAR	104,6	40	BS	63,6
13	JBF	94,6	41	AC	111,2
14	LCO	5,3	42	JBR	53,5
15	JRSF	27,5	43	JBSO	99,5
16	LCM	1,9	44	FSRJ	128,0
17	EEF	6,9	45	JLGRJ	588,0
18	JRF	79,1	46	JGSRJ	128,8
19	OG	108,5	47	BA	188,6
20	ACS	2,2	48	EDA	119,4
21	VRS	79,6	49	JFS	133,4
22	NGR	384,2	50	JMB	101,0
23	GAC	165,6	51	ACS	241,0
24	ANS	540,7	52	JGS2	188,0
25	MAMS	99,1	53	GBS	236,4
26	LCC	292,2	54	ESB	182,6
27	CDM	93,5	55	JRSF2	127,3
28	JRF	197,8			

Somente as amostras dos pacientes com CEC de BL que apresentaram a concentração de DNA  $\geq 50$ ng/μl foram utilizadas. Assim, as amostras de 49 pacientes foram utilizadas para hibridização em lâminas de microarranjos.

**Tabela 8.** Concentrações de DNA em ng/μl de amostras de sangue periférico dos 55 controles

Nº	Iniciais	DNA (ng/μl)	Nº	Iniciais	DNA (ng/μl)
1	JBC	165,2	29	DRC	103,6
2	JAS	62,0	30	SVC	71,6
3	HAN	320,1	31	ELB	101,5
4	SAIS	54,6	32	IJS	157,7
5	LN	164,3	33	NL	88,7
6	PGS	54,9	34	JBAC	87,4
7	AHS	168,7	35	RB	85,7
8	MJF	61,2	36	MAV	127,0
9	JPR	107,4	37	JSS	66,6
10	MCGP	209,1	38	JRA	54,8
11	LS	165,8	39	DV	139,9
12	MP	621,7	40	RT	199,8
13	AMA	142,1	41	JPS	146,0
14	VF	60,0	42	ERR	122,5
15	APS	89,0	43	AOM	104,7
16	ABP	63,9	44	AMM	217,5
17	JVDF	97,2	45	JAF	180,2
18	COM	83,8	46	IFS	567,4
19	IC	117,1	47	SLS	312,0
20	FCM	103,3	48	DRL	94,7
21	CRSF	96,1	49	PSM	177,5
22	RLMJ	140,0	50	WLO	173,0
23	IPN	95,5	51	MAV	172,0
24	MSCO	66,0	52	MDN	98,1
25	GIO	255,0	53	SF	63,0
26	JM	58,4	54	JAS2	150,8
27	JPS	133,6	55	COM	108,0
28	AS	189,0			

Todas as amostras dos controles apresentaram a concentração do DNA  $\geq 50\text{ng}/\mu\text{l}$  e DNA com boa qualidade. Entretanto, as amostras 3, 18, 21, 24, 27 e 36 foram retiradas dos procedimentos devido ao pareamento das amostras dos controles com os pacientes. Assim, as amostras de 49 controles foram hibridizadas nas lâminas de microarranjos. A qualidade destas amostras foi observada por meio das curvas geradas pelo espectrofotômetro. Todas as curvas apresentaram um padrão adequado de qualidade (Figura 7).



**Figura 7.** Curva de qualidade da amostra de DNA de um paciente (Caso nº 5) com carcinoma de células escamosas de base de língua gerada por espectrofotômetro. O ápice da curva com tamanho de 260nm indica a qualidade ideal da amostra

### **3.2. Hibridização das amostras nas lâminas de microarranjos**

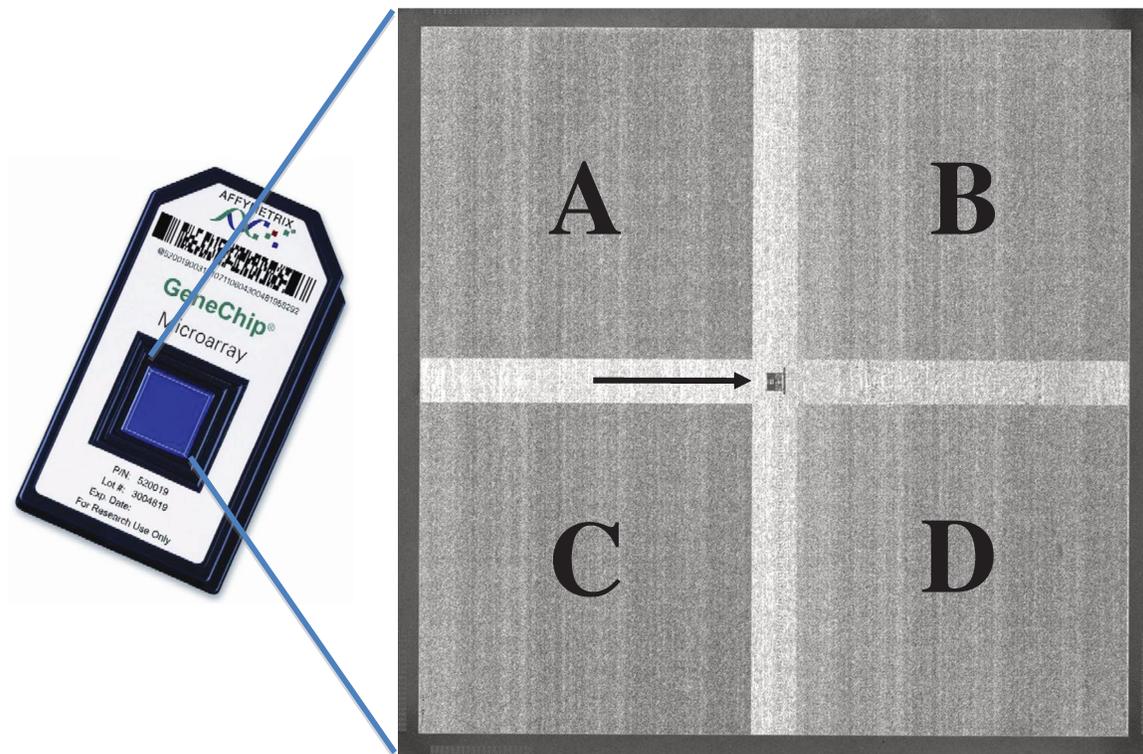
Foram utilizadas 98 lâminas para a GELE individual. Quarenta e nove delas foram utilizadas para a GELE individual de 49 pacientes com CEC de BL e 49 para os 49 indivíduos saudáveis.

Foram utilizadas quatro lâminas para a GELE coletiva. Duas delas para os pacientes com CEC de BL e duas para os controles. Em uma lâmina da GELE coletiva dos pacientes, foi utilizada a mistura de amostras de DNA de 39 pacientes caucasóides com CEC de BL e em outra lâmina, a mistura de DNA de dez pacientes negróides. O mesmo esquema foi utilizado para a GELE coletiva dos controles.

### **3.3. Controle de qualidade das amostras hibridizadas nas lâminas**

#### **3.3.1. Controle de qualidade qualitativo das GELEs individual e coletiva**

A imagem padrão gerada pela excitação dos fluoróforos de uma lâmina de um indivíduo controle capturada pelo *scanner* está apresentada na Figura 8.



**Figura 8. Imagem padrão gerada pela excitação dos fluoróforos contidos em lâmina de um indivíduo controle capturada por *scanner* para a genotipagem em larga escala. Os quatro quadrados (A, B, C e D) representam a hibridização adequada da amostra com as sondas que contêm os polimorfismos de base única. A imagem mais clara em forma de cruz representa a hibridização adequada da amostra em sondas para a variação do número de cópias do DNA. A imagem no centro (→) representa um controle interno da lâmina**

O controle de qualidade qualitativo adequado foi observado em todas as lâminas de microarranjos utilizadas para a análise de SNPS e CNVs.

### **. 3.3.2. Controle de qualidade quantitativo da GELE individual**

As distribuições individualizadas das amostras de DNA de 49 pacientes com CEC de BL e 49 amostras de controles hibridizadas nas lâminas de acordo com as taxas de CQ, CQ *StyI* e CQ *NspI* estão apresentadas nas Tabelas 9 e 10 respectivamente

**Tabela 9.** Taxas dos controles de qualidade das amostras de DNA dos 49 pacientes com carcinoma de células escamosas de base de língua hibridizadas nas lâminas com microarranjos de DNA

Nº	Iniciais	CQ	CQ <i>StyI</i>	CQ <i>NspI</i>	Nº	Iniciais	CQ	CQ <i>StyI</i>	CQ <i>NspI</i>
1	AAP	93,75	90,66	93,32	29	OD	96,23	95,81	94,60
2	EAC	88,02	85,99	86,25	30	JLT	98,44	98,07	98,33
3	JF	99,07	98,07	99,36	31	OBF	97,55	97,17	96,46
4	PLB	98,48	97,42	98,33	32	AV	96,43	95,33	94,73
5	RDG	87,36	81,96	84,83	33	JJ	87,56	82,45	94,45
6	SAL	92,59	86,96	91,13	34	BS	97,25	96,62	95,89
7	SAC	98,08	96,62	97,43	35	AC	88,85	84,54	84,83
8	SRJL	98,28	97,58	97,30	36	JBR	98,21	98,55	96,40
9	WS	98,71	97,91	98,20	37	JBSO	87,36	82,93	83,42
10	JAOD	95,23	92,75	94,47	38	FSRJ	91,33	85,35	90,10
11	JAR	96,39	95,33	95,37	39	JLGRJ	95,60	96,46	92,03
12	JBF	97,72	97,58	96,79	40	JGSRJ	98,51	98,71	97,81
13	JRF	95,80	93,24	94,09	41	BA	96,39	95,33	95,63
14	OG	97,52	95,17	97,04	42	EDA	97,68	95,81	97,43
15	VRS	98,21	96,46	98,33	43	JFS	92,79	89,69	90,62
16	NGR	96,49	94,69	95,12	44	JMB	91,86	87,44	88,30
17	GAC	98,25	97,42	98,20	45	ACS	90,93	85,35	88,95
18	ANS	98,61	98,23	98,33	46	JGS2	92,75	88,89	91,26
19	MAMS	94,74	94,20	91,13	47	GBS	95,43	90,82	95,63
20	LCC	98,28	97,10	97,69	48	ESB	92,55	86,96	92,29
21	CDM	96,72	94,52	95,89	49	JRSF2	83,49	80,03	82,01
22	JRF	98,28	96,94	98,33					
23	MM	98,08	97,10	97,04					
24	MPP	98,91	98,71	98,97					
25	MR	96,46	93,56	95,37					
26	PA	97,98	97,10	97,81					
27	DB	98,64	98,23	98,71					
28	JF	97,72	95,81	97,17					

CQ: taxa do controle de qualidade geral em porcentagem da amostra de DNA total hibridizada na lâmina com microarranjos, CQ *StyI*: taxa do controle de qualidade em porcentagem da amostra de DNA clivada pela enzima *StyI*, CQ *NspI*: taxa do controle de qualidade em porcentagem da amostra de DNA clivada pela enzima *NspI*

**Tabela 10.** Taxas dos controles de qualidade das amostras de DNA dos 49 controles hibridizadas nas lâminas com microarranjos de DNA

Nº	Iniciais	CQ (%)	CQ <i>StyI</i>	CQ <i>NspI</i>	Nº	Iniciais	CQ	CQ <i>StyI</i>	CQ <i>NspI</i>
1	JBC	97,62	94,36	97,56	29	JBAC	91,93	85,99	90,10
2	JAS	99,07	99,19	98,97	30	RB	97,39	94,85	97,69
3	SAIS	90,83	84,70	91,52	31	JSS	95,86	92,43	95,12
4	LN	98,54	97,75	98,33	32	JRA	86,33	82,29	83,42
5	PGS	97,65	94,69	97,17	33	DV	97,49	95,65	97,17
6	AHS	97,68	96,78	97,17	34	RT	98,61	96,94	98,33
7	MJF	98,64	96,78	98,33	35	JPS	91,63	86,80	92,54
8	JPR	93,55	91,30	90,87	36	ERR	93,98	91,47	91,90
9	MCGP	97,85	95,17	97,81	37	AOM	86,57	81,80	83,16
10	LS	97,92	95,65	98,20	38	AMM	97,82	95,49	97,56
11	MP	95,23	91,63	94,22	39	JAF	97,68	95,33	97,56
12	AMA	92,95	90,82	91,00	40	IFS	93,15	89,53	93,06
13	VF	96,92	95,01	95,76	41	SLS	94,94	94,36	92,03
14	APS	94,64	90,18	93,70	42	DRL	90,04	87,60	85,73
15	ABP	91,13	85,67	88,43	43	PSM	96,72	94,20	96,14
16	JVDF	98,51	97,75	98,20	44	WLO	97,62	96,30	96,92
17	IC	98,31	95,65	97,81	45	MAV	92,06	88,41	90,49
18	FCM	96,79	94,20	95,37	46	MDN	98,21	96,46	98,59
19	RLMJ	89,81	85,35	89,85	47	SF	92,55	91,63	89,33
20	IPN	97,62	95,33	97,81	48	JAS2	98,11	96,46	98,84
21	GIO	95,47	92,59	93,70	49	COM	96,62	93,40	96,14
22	JM	94,47	89,53	94,22					
23	AS	97,58	95,65	98,20					
24	DRC	95,76	90,50	97,17					
25	SVC	97,75	94,85	97,81					
26	ELB	96,76	90,82	97,69					
27	IJS	98,11	97,10	97,30					
28	NL	93,85	92,11	91,77					

CQ: taxa do controle de qualidade geral em porcentagem da amostra de DNA total hibridizada na lâmina com microarranjos, CQ *StyI*: taxa do controle de qualidade em porcentagem da amostra de DNA clivada pela enzima *StyI*, CQ *NspI*: taxa do controle de qualidade em porcentagem da amostra de DNA clivada pela enzima

Apenas uma lâmina hibridizada com a amostra de um paciente com CEC de BL (caso 49) não atingiu o CQ geral determinado pelo fabricante (83,49%). Todas as outras lâminas hibridizadas com as amostras dos pacientes e dos controles obtiveram a taxa de CQ igual ou maior do que 86,0%. Entretanto, todas as lâminas foram incluídas nas análises por julgarmos que a única lâmina com menor CQ geral apresentou valor próximo do valor de corte padronizado pelo fabricante e que não influenciaria nos resultados finais.

### **3.3.3. Controle de qualidade quantitativo da GELE coletiva**

Duas lâminas foram utilizadas para a GELE coletiva de amostras de DNA de pacientes com CEC de BL (uma lâmina com 39 pacientes caucasóides e a outra com dez pacientes negróides) e duas para as amostras de controles (uma lâmina com 39 controles caucasóides e a outra com dez controles negróides).

As taxas de CQ, CQ *StyI* e CQ *NspI* das lâminas com 39 pacientes caucasóides e dez pacientes negróides foram 85,2%, 78,6% e 86,0%; e 81,0%, 72,6% e 81,2%, respectivamente.

As taxas de CQ, CQ *StyI* e CQ *NspI* das lâminas com 39 controles caucasóides e dez controles negróides foram 84,5%, 76,2% e 84,6%; e 83,8%, 77,9% e 82,2%, respectivamente.

Por convenção, todas as lâminas foram incluídas nas análises por julgarmos que elas apresentaram valores próximos dos de corte padronizado pelo fabricante.

## **3.4. Análise de polimorfismos gênicos de base única**

### **3.4.1. GELE individual**

Observamos que 5.402 e 6.609 SNPs apresentaram frequências distintas (maiores ou menores) entre os pacientes com CEC de BL e controles, por meio dos algoritmos brlmm e crlmm, respectivamente. As regiões do DNA onde estiveram localizados os SNPs estão apresentadas na Tabela 11.

**Tabela 11.** Regiões do DNA que possuem os polimorfismos gênicos de base única identificados em 49 pacientes com carcinoma de células escamosas de base de língua e 49 controles por meio dos algoritmos brlmm e crlmm

Região do DNA	Número de SNPs (%)	
	Algoritmo brlmm	Algoritmo crlmm
Total	5.402 (100,0%)	6.609 (100,0%)
Codificadora de aminoácidos	53 (1,0%)	52 (0,8%)
3' e 5' não traduzidas	58 (1,1%)	51 (0,8%)
Regulatória de transcrição	3.379 (62,5%)	3.461 (52,4%)
Íntron	1.896 (35,1%)	3.045 (46,0%)
Não identificada	16 (0,3%)	0 (0,0%)

SNPs: polimorfismos gênicos de base única

As regiões de regulação de transcrição e intrônica do DNA apresentaram os maiores números de SNPs identificados por ambos os algoritmos de análise ( $P < 0,01$ ).

As distribuições de todos os SNPs diferencialmente frequentes entre pacientes e controles, identificados pelos algoritmos brlmm ou crlmm, localizados nas regiões codificadora de aminoácidos e nas regiões 3' e 5' não traduzidas do DNA, de acordo com o número do cromossomo, a região do cromossomo, o símbolo do gene, o número de referência (rs) do SNP, o alelo selvagem e o alelo variante estão apresentadas nas Tabelas de 12 a 15, respectivamente.

**Tabela 12.** Polimorfismos de base única localizados na região codificadora de aminoácidos do DNA identificados em 49 pacientes com carcinoma de células escamosas de base de língua e 49 controles por meio do algoritmo brlmm de acordo com o número do cromossomo, a região cromossômica, o símbolo do gene, o número de referência e os alelos distintos

<b>Cromossomo</b>	<b>Região</b>	<b>Gene</b>	<b>rs do SNP</b>	<b>Alelo selvagem</b>	<b>Alelo variante</b>	<b>Valor de P</b>
3	p26.1	<i>GRM7</i>	rs3749380	C	T	0,0002
19	q13.31	<i>ZNF227</i>	rs922063	C	G	0,0003
16	p13.3	<i>TBL3</i>	rs17602	G	T	0,0004
22	q11.21	<i>LOC150185</i>	rs8138364	C	G	0,0005
10	q25.2	<i>SMC3</i>	rs2419565	A	G	0,0006
12	p13.32	<i>C12orf5</i>	rs7133330	A	G	0,0007
14	q11.2	<i>C14orf167</i>	rs8003491	G	T	0,0008
5	q33.1	<i>NDST1</i>	rs2273234	A	G	0,0009
10	q23.33	<i>SORBS1</i>	rs726176	C	T	0,001
2	q12.3	<i>ST6GAL2</i>	rs12615112	C	T	0,001
2	p24.2	<i>NT5C1B</i>	rs16985301	A	G	0,001
16	q22.1	<i>CDH3</i>	rs2296408	A	C	0,001
3	p21.31	<i>TGM4</i>	rs1995641	A	G	0,001
2	q23.3	<i>RIF1</i>	rs2444263	A	G	0,002
6	p22.1	<i>OR2B6</i>	rs9380030	A	G	0,002
11	q22.3	<i>CARD16</i>	rs542571	A	T	0,002
17	q11.1	<i>NOS2</i>	rs16966563	A	G	0,002
9	p21.3	<i>CDKN2BAS</i>	rs564398	A	G	0,003
2	q21.2	<i>ZNF285A</i>	rs17394868	C	T	0,003
6	q22.31	<i>RNF217</i>	rs475076	C	T	0,003
9	q22.33	<i>XPA</i>	rs3176753	C	T	0,003
11	q22.3	<i>CASP5</i>	rs3181179	C	T	0,003
X	q22.1	<i>NXF4</i>	rs35532477	C	T	0,003
15	q24.1	<i>ULK3</i>	rs936227	A	G	0,003
1	q24.2	<i>SCYL3</i>	rs17602701	A	G	0,003
3	q27.3	<i>RTP4</i>	rs1533594	C	T	0,003
5	q22.1	<i>WDR36</i>	rs17624563	C	G	0,004
20	q13.12	<i>JPH2</i>	rs6093935	A	G	0,004
12	p13.32	<i>EFCAB4B</i>	rs7138453	C	T	0,004
1	q24.2	<i>SELE</i>	rs5355	A	G	0,004
3	q26.31	<i>PLD1</i>	rs6766336	A	G	0,004
12	q24.31	<i>SBNO1</i>	rs3825141	G	T	0,005

Continuação						
3	p21.31	<i>LARS2</i>	rs2128361	C	T	0,005
7	p12.3	<i>ABCA13</i>	rs6583546	A	G	0,005
1	p32.2	<i>C8A</i>	rs706479	C	T	0,005
15	q15.1	<i>RPAP1</i>	rs1200349	C	T	0,006
10	q22.1	<i>CALM2</i>	rs4746024	C	T	0,006
1	p31.3	<i>KANK4</i>	rs2260581	A	G	0,006
11	p11.2	<i>PTPRJ</i>	rs16905649	C	T	0,007
11	p15.5	<i>LOC338651</i>	rs1809667	G	T	0,007
1	p13.2	<i>RP11</i>	rs11102233	A	C	0,007
6	q23.2	<i>ENPP3</i>	rs9321309	C	T	0,007
15	q21.3	<i>ZNF280D</i>	rs28620278	A	G	0,008
15	q22.1	<i>HSP90AB4P</i>	rs7161889	C	G	0,008
7	q21.3	<i>PON3</i>	rs1053275	A	G	0,008
1	p31.1	<i>DLST</i>	rs12123977	A	G	0,008
1	p13.2	<i>RP11</i>	rs10776722	C	G	0,009
1	p36.13	<i>PAX7</i>	rs2076020	A	G	0,009
1	p34.2	<i>MACF1</i>	rs11206102	A	C	0,009
3	p14.3	<i>CCDC66</i>	rs6807338	A	G	0,009
19	q13.2	<i>LGALS14</i>	rs10755	C	G	0,009
6	q25.3	<i>RSPH3</i>	rs3756987	A	G	0,009
19	q13.12	<i>GAPDHS</i>	rs2239945	A	G	0,009

rs: número de referência do SNP, SNP: polimorfismo de base única

**Tabela 13.** Polimorfismos de base única localizados na região codificadora de aminoácidos do DNA identificados em 49 pacientes com carcinoma de células escamosas de base de língua e 49 controles por meio do algoritmo crlmm de acordo com o número do cromossomo, a região cromossômica, o símbolo do gene, o número de referência e os alelos distintos

<b>Cromossomo</b>	<b>Região</b>	<b>Gene</b>	<b>rs do SNP</b>	<b>Alelo selvagem</b>	<b>Alelo variante</b>	<b>Valor de P</b>
17	q23.3	<i>SCN4A</i>	rs7218917	T	C	0,006
9	q22.31	<i>C9orf89</i>	rs17593496	G	A	0,008
8	q24.23	<i>COL22A1</i>	rs2292927	G	A	0,003
6	p12.3	<i>GLYATL3</i>	rs1480615	T	A	0,002
20	p12.3	<i>FERMT1</i>	rs10373	T	C	0,001
19	q13.2	<i>LGALS14</i>	rs10755	G	C	0,002
1	q23.1	<i>APOA1BP</i>	rs6668178	G	A	0,003
2	p24.2	<i>NT5C1B</i>	rs16985301	G	A	0,006
2	p24.2	<i>GEN1</i>	rs300175	G	A	0,009
2	q12.2	<i>ST6GAL2</i>	rs12615112	T	C	0,008
6	p21.1	<i>TREML2</i>	rs3747742	T	C	0,003
6	q25.3	<i>RSPH3</i>	rs3756988	G	A	0,006
6	q25.3	<i>RSPH3</i>	rs3756987	G	A	0,005
2	q37.3	<i>COL6A3</i>	rs4433949	T	C	0,002
7	q36.1	<i>MLL3</i>	rs6464211	G	A	0,002
11	q22.3	<i>CASP5</i>	rs3181320	G	C	0,002
4	p16.1	<i>WFS1</i>	rs1805070	T	C	0,003
10	q26.13	<i>FAM24B</i>	rs1891110	G	A	0,008
22	q13.33	<i>CPT1B</i>	rs2269383	G	A	0,005
5	q31.1	<i>DDX46</i>	rs17167672	G	A	0,00005
6	p12.1	<i>GFRAL</i>	rs9370418	T	C	0,008
6	q23.2	<i>ENPP3</i>	rs9321309	T	C	0,005
19	q13.2	<i>LGALS14</i>	rs4830	G	A	0,003
4	q34.2	<i>SPATA4</i>	rs6832177	G	A	0,008
1	q23.1	<i>SPTA1</i>	rs863931	G	A	0,007
11	q23.3	<i>USP2</i>	rs2241646	T	C	0,004
6	p22.1	<i>OR2B6</i>	rs9380030	G	A	0,008
7	q22.1	<i>ZAN</i>	rs17147735	T	C	0,009
15	q26.3	<i>LYSMD4</i>	rs6598264	G	A	0,008
5	q33.1	<i>NDST1</i>	rs2273234	G	A	0,009
6	q25.3	<i>SYNJ2</i>	rs1744177	T	C	0,001
3	q13.33	<i>POLQ</i>	rs3218637	G	A	0,005

Continuação						
19	p13.11	<i>KIAA0892</i>	rs769267	G	A	0,005
9	q31.3	<i>IKBKAP</i>	rs3204145	T	A	0,0003
1	p31.1	<i>CRYZ</i>	rs17552114	G	A	0,005
11	q22.3	<i>CASP5</i>	rs7935132	G	A	0,002
15	q15.1	<i>EIF2AK4</i>	rs566792	G	A	0,002
19	p13.3	<i>SHC2</i>	rs2287960	T	G	0,009
1	p31.3	<i>KANK4</i>	rs2260581	G	A	0,0006
8	q21.3	<i>SLC7A13</i>	rs7814198	G	A	0,009
5	q32	<i>LARS</i>	rs2895647	T	C	0,007
15	q15.1	<i>RPAP1</i>	rs1200349	T	C	0,007
5	q32	<i>PLAC8L1</i>	rs12187913	T	A	0,007
14	q32.32	<i>C14orf73</i>	rs2297067	G	A	0,004
20	p11.21	<i>C20orf56</i>	rs1203887	T	C	0,005
22	q11.21	<i>LOC150185</i>	rs8138364	G	C	0,0001
6	p22.2	<i>BTN2A3</i>	rs12201546	G	A	0,007
22	q13.33	<i>CPT1B</i>	rs2269383	G	A	0,005
16	q24.1	<i>LOC732275</i>	rs2696845	C	A	0,009
5	q32	<i>LOC728264</i>	rs17110069	G	A	0,0007
9	q22.33	<i>LOC441455</i>	rs16911482	G	A	0,003
19	q13.42	<i>ZNF415</i>	rs3814	T	C	0,008

rs: número de referência do SNP, SNP: polimorfismo de base única

**Tabela 14.** Polimorfismos de base única localizados nas regiões 3' e 5' do DNA identificados em 49 pacientes com carcinoma de células escamosas de base de língua e 49 controles por meio do algoritmo brlmm de acordo com o número do cromossomo, a região cromossômica, o símbolo do gene, o número de referência e os alelos distintos

<b>Cromossomo</b>	<b>Região</b>	<b>Gene</b>	<b>rs do SNP</b>	<b>Alelo selvagem</b>	<b>Alelo variante</b>	<b>Valor de P</b>
14	q32.2	<i>BDKRB2</i>	rs2069590	A	T	0,0001
16	p12.1	<i>PRKCB</i>	rs3380	A	T	0,0002
4	q28.3	<i>SLC7A11</i>	rs7674870	C	T	0,0004
15	q13.3	<i>MTMR10</i>	rs2873	C	T	0,0005
5	q31.1	<i>C5orf20</i>	rs4976298	C	T	0,0008
5	q31.1	<i>C5orf20</i>	rs744247	A	G	0,0008
2	p23.1	<i>CAPN13</i>	rs2280404	A	G	0,001
2	q14.2	<i>STEAP3</i>	rs7581339	C	T	0,001
17	p13.2	<i>UBE2G1</i>	rs41408848	A	G	0,001
11	q23.1	<i>C11orf92</i>	rs11213822	A	G	0,001
11	q13.2	<i>TPCN2</i>	rs1060435	C	T	0,001
9	q21.13	<i>GCNT1</i>	rs4532668	A	C	0,001
9	q21.13	<i>GCNT1</i>	rs4532668	A	C	0,001
9	q21.33	<i>NTRK2</i>	rs3654	A	G	0,002
14	q12	<i>PCK2</i>	rs1951634	G	T	0,002
13	q14.12	<i>CPB2</i>	rs940	C	G	0,002
1	p12	<i>MANIA2</i>	rs1146297	A	G	0,002
8	q24.12	<i>SAMD12</i>	rs4467962	A	G	0,003
12	p12.1	<i>ST8SIA1</i>	rs16924743	C	T	0,003
1	q24.2	<i>SCYL3</i>	rs1046058	A	G	0,003
17	q25.2	<i>JMJD6</i>	rs2240774	A	G	0,003
5	q22.2	<i>MCC</i>	rs7033	A	G	0,003
3	p14.1	<i>PRICKLE2</i>	rs26938	G	T	0,003
10	q26.13	<i>ACADSB</i>	rs6838	A	T	0,003
6	p12.1	<i>KLHL31</i>	rs9474613	C	T	0,003
12	q24.13	<i>ERP29</i>	rs7114	C	T	0,004
2	q21.1	<i>RAB6A</i>	rs6719989	C	G	0,004
19	q13.2	<i>CEACAM6</i>	rs2701	C	T	0,005
11	q13.5	<i>ACER3</i>	rs6592695	A	T	0,005
11	q13.5	<i>ACER3</i>	rs1062170	A	G	0,005
15	q24.3	<i>SGK269</i>	rs14726	C	T	0,005
6	p12.1	<i>KLHL31</i>	rs997604	C	T	0,006

Continuação						
9	p21.2	<i>TUSC1</i>	rs1128957	C	G	0,006
2	q32.2	<i>INPP1</i>	rs1882892	A	G	0,006
19	q13.33	<i>ZNF766</i>	rs1048399	A	G	0,006
2	q33.1	<i>MPP4</i>	rs3754932	C	T	0,006
X	q11.1	<i>ZC4H2</i>	rs41326649	C	G	0,006
1	q42.3	<i>TOMM20</i>	rs7930	A	G	0,006
1	p36.12	<i>WNT4</i>	rs2072920	A	G	0,007
7	q11.23	<i>FGL2</i>	rs3093279	C	T	0,007
11	p11.2	<i>ACP2</i>	rs4640	A	G	0,007
9	p21.2	<i>IFNK</i>	rs700782	C	T	0,007
4	p16.1	<i>PPP2R2C</i>	rs6446489	A	G	0,008
3	q26.1	<i>C3orf57</i>	rs163327	A	G	0,008
1	q21.2	<i>GOLPH3L</i>	rs9733	A	C	0,008
5	q32	<i>RBM27</i>	rs11954297	C	T	0,008
3	q28	<i>CLDN1</i>	rs17500920	A	T	0,008
1	q21.2	<i>ENSA</i>	rs1971378	C	T	0,008
17	q21.31	<i>EFTUD2</i>	rs2304987	A	G	0,008
1	q42.12	<i>C1orf95</i>	rs878150	C	G	0,008
12	q24.31	<i>CAMKK2</i>	rs1140886	C	T	0,008
7	q31.1	<i>LAMB4</i>	rs1614188	A	C	0,009
4	p16.1	<i>AFAP1</i>	rs2285766	A	G	0,009
9	p13.2	<i>GRHPR</i>	rs1057507	C	T	0,009
14	q24.3	<i>HEATR4</i>	rs1074501	A	G	0,009
10	q23.33	<i>TBC1D12</i>	rs2281791	C	T	0,009
1	p35.2	<i>NKAIN1</i>	rs4949356	A	G	0,009
8	q21.13	<i>ZNF704</i>	rs10504721	A	G	0,009

rs: número de referência do SNP, SNP: polimorfismo de base única

**Tabela 15.** Polimorfismos de base única localizados nas regiões 3' e 5' do DNA identificados em 49 pacientes com carcinoma de células escamosas de base de língua e 49 controles por meio do algoritmo crlmm de acordo com o número do cromossomo, a região cromossômica, o símbolo do gene, o número de referência e os alelos distintos

<b>Cromossomo</b>	<b>Região</b>	<b>Gene</b>	<b>rs do SNP</b>	<b>Alelo selvagem</b>	<b>Alelo variante</b>	<b>Valor de P</b>
16	q23.1	<i>ZNRF1</i>	rs1043509	T	A	0,002
17	p13.2	<i>UBE2G1</i>	rs41408848	G	A	0,001
19	p13.3	<i>ARRDC5</i>	rs2656938	G	A	0,004
6	q24.3	<i>SAMD5</i>	rs702346	G	A	0,002
17	q25.1	<i>JMJD6</i>	rs2240774	G	A	0,008
5	q32	<i>RBM27</i>	rs11954297	T	C	0,006
9	q22.33	<i>CORO2A</i>	rs942164	G	A	0,004
9	q22.33	<i>CORO2A</i>	rs1891323	T	C	0,004
7	q21.11	<i>PCLO</i>	rs2715148	C	A	0,006
12	q24.13	<i>ERP29</i>	rs7114	T	C	0,002
5	q22.2	<i>MCC</i>	rs7033	G	A	0,003
1	q44	<i>CNST</i>	rs4335342	G	A	0,005
10	q21.2	<i>CCDC6</i>	rs7917618	G	A	0,004
7	q11.23	<i>FGL2</i>	rs3093279	T	C	0,001
8	q22.1	<i>RBM12B</i>	rs11549490	T	C	0,007
6	p12.1	<i>KLHL31</i>	rs9474613	T	C	0,006
15	q13.3	<i>MTMR10</i>	rs2873	T	C	0,00009
6	q14.1	<i>MYO6</i>	rs6914716	G	A	0,008
3	q22.1	<i>SLCO2A1</i>	rs2370512	T	A	0,007
14	q13.3	<i>MBIP</i>	rs4900833	T	C	0,0008
7	p22.3	<i>IQCE</i>	rs1045753	T	C	0,007
4	q28.3	<i>SLC7A11</i>	rs7674870	T	C	0,004
3	p14.3	<i>FLNB</i>	rs1131312	T	C	0,006
1	p31.3	<i>LITD1</i>	rs11207936	T	C	0,004
11	q13.3	<i>TPCN2</i>	rs4930265	G	A	0,006
19	p13.2	<i>ARHGEF18</i>	rs7245775	T	C	0,005
14	q23.1	<i>TRMT5</i>	rs1958281	G	A	0,004
11	p14.1	<i>MPPED2</i>	rs16925538	G	A	0,004
13	q34	<i>IRS2</i>	rs2289046	G	A	0,004
8	p23.1	<i>CTSB</i>	rs1736078	C	A	0,009
14	q23.1	<i>TRMT5</i>	rs1887742	C	A	0,004
3	p14.1	<i>PRICKLE2</i>	rs26938	T	G	0,0004

Continuação						
3	p21.1	<i>DCPIA</i>	rs1108506	G	C	0,007
3	p14.3	<i>FLNB</i>	rs11412	G	A	0,006
11	q13.3	<i>TPCN2</i>	rs1060435	T	C	0,00009
14	q11.2	<i>PCK2</i>	rs1951634	T	G	0,001
15	q21.2	<i>AP4E1</i>	rs2619685	T	C	0,009
12	q24.33	<i>CHFR</i>	rs4758909	G	A	0,005
14	q24.3	<i>ANGEL1</i>	rs1076584	T	C	0,009
3	p21.31	<i>LARS2</i>	rs267266	G	A	0,005
19	q13.42	<i>ZNF415</i>	rs3814	T	C	0,008
1	p36.12	<i>PLA2G5</i>	rs534397	G	C	0,007
9	p21.2	<i>TUSC1</i>	rs1128957	G	C	0,003
6	p12.3	<i>GLYATL3</i>	rs9395504	T	C	0,007
1	q42.3	<i>TOMM20</i>	rs7930	G	A	0,0009
16	p13.3	<i>UBN1</i>	rs1045643	T	C	0,0001
9	q21.13	<i>GCNT1</i>	rs4532668	C	A	0,005
16	p11.2	<i>ZNF668</i>	rs4889615	T	A	0,002
14	q24.3	<i>HEATR4</i>	rs1074501	G	A	0,005
19	q13.33	<i>PIH1D1</i>	rs2293011	C	A	0,004
2	q21.1	<i>RAB6C</i>	rs6719989	G	C	0,0006

rs: número de referência do SNP, SNP: polimorfismo de base única

As distribuições dos 50 SNPs mais frequentes (menores valores de *P*) entre os pacientes com CEC de BL e os controles, identificados pelos dois algoritmos de análises, localizados nas regiões regulatórias da expressão gênica, de acordo com o número do cromossomo, a região cromossômica, o símbolo do gene, o rs do SNP, o alelo selvagem e o alelo variante estão apresentadas nas Tabelas 16 e 17, respectivamente.

**Tabela 16.** Polimorfismos de base única localizados nas regiões regulatórias de expressão dos genes identificados em 49 pacientes com carcinoma de células escamosas de base de língua e 49 controles por meio do algoritmo brlmm de acordo com o número do cromossomo, a região cromossômica, o símbolo do gene, o número de referência e os alelos distintos

<b>Cromossomo</b>	<b>Região</b>	<b>Gene</b>	<b>rs do SNP</b>	<b>Alelo selvagem</b>	<b>Alelo variante</b>	<b>Valor de P</b>
14	q21.3	<i>RPS29</i>	rs1769526	C	T	0,0001
6	p22.1	<i>ZNF165</i>	rs149946	A	C	0,0001
1	p33	<i>SKINTL</i>	rs1508604	C	T	0,0001
2	p13.3	<i>CYP26B1</i>	rs1559979	C	T	0,0001
15	q25.2	<i>MEX3B</i>	rs7176101	A	T	0,0001
11	p15.4	<i>OR52K2</i>	rs10742295	C	G	0,0001
6	p12.3	<i>DEFB114</i>	rs4715153	A	G	0,0001
19	p13.13	<i>CCDC130</i>	rs10409463	C	T	0,0001
4	q23	<i>PPP3CA</i>	rs13121245	C	T	0,0001
5	q22.3	<i>CDO1</i>	rs415395	C	G	0,0001
5	q21.2	<i>RAB9PI</i>	rs12514595	A	G	0,0001
12	q21.31	<i>SLC6A15</i>	rs11116465	A	G	0,0001
7	q21.11	<i>PCLO</i>	rs10281234	A	G	0,0001
1	q25.2	<i>CEP350</i>	rs474258	C	T	0,0001
1	q25.2	<i>CEP350</i>	rs475233	A	G	0,0001
9	q31.1	<i>SMC2</i>	rs10990522	C	T	0,0001
12	q22	<i>NDUFA12</i>	rs2008023	C	T	0,0001
4	q34.3	<i>MGC45800</i>	rs6811144	C	T	0,0001
8	p21.3	<i>GFRA2</i>	rs1441799	C	T	0,0001
1	p31.3	<i>TM2D1</i>	rs1933293	A	G	0,0001
2	q31.1	<i>CDCA7</i>	rs1093977	C	T	0,0001
1	p31.3	<i>DIRAS3</i>	rs6588287	A	G	0,0001
1	q25.2	<i>CEP350</i>	rs1298479	C	T	0,0001
15	q26.2	<i>MCTP2</i>	rs8041151	A	G	0,0001
14	q31.3	<i>GALC</i>	rs1003205	A	G	0,0001
3	q27.1	<i>EPHB3</i>	rs4912577	A	G	0,0001
5	p14.1	<i>ZCCHC10</i>	rs11948763	A	G	0,0001
15	q25.1	<i>RPL21</i>	rs7176187	A	G	0,0001
10	p15.1	<i>PRKCQ</i>	rs11258213	C	T	0,0001
6	p22.1	<i>RSL24D1</i>	rs9461405	G	T	0,0001
4	q13.2	<i>CENPC1</i>	rs17763413	G	T	0,0001
10	q24.2	<i>PYROXD2</i>	rs1925636	A	G	0,0001

Continuação						
5	q14.3	<i>RPL10A</i>	rs2112457	C	T	0,0001
5	q14.3	<i>RPL10A</i>	rs1422249	C	T	0,0001
1	p36.22	<i>Clorf127</i>	rs12755921	C	T	0,0001
8	q21.11	<i>ZFHX4</i>	rs4007357	C	T	0,0001
10	p15.1	<i>PRKCQ</i>	rs744254	A	G	0,0001
7	q21.11	<i>GNAI1</i>	rs10244538	A	G	0,0001
15	q12	<i>GABRB3</i>	rs2042753	A	C	0,0001
2	p25.1	<i>TRIB2</i>	rs7590007	A	C	0,0001
3	q11.2	<i>LOC255025</i>	rs10936849	A	G	0,0001
10	p15.1	<i>PRKCQ</i>	rs744253	A	G	0,0002
4	p15.2	<i>SOD3</i>	rs809607	C	T	0,0002
1	q32.1	<i>NR5A2</i>	rs2821353	G	T	0,0002
6	p22.1	<i>GPR89A</i>	rs9393848	C	T	0,0002
9	q21.33	<i>CTSL3</i>	rs1927981	C	G	0,0002
11	q22.3	<i>ZC3H12C</i>	rs949751	C	T	0,0002
11	q22.3	<i>ZC3H12C</i>	rs1509728	C	T	0,0002
8	p23.2	<i>MCPHI</i>	rs7464987	A	C	0,0002
14	q21.3	<i>RPL10L</i>	rs1869099	C	T	0,0002

rs: número de referência do SNP, SNP: polimorfismo de base única

**Tabela 17.** Polimorfismos de base única localizados nas regiões regulatórias de expressão dos genes identificados em 49 pacientes com carcinoma de células escamosas de base de língua e 49 controles por meio do algoritmo crlmm de acordo com o número do cromossomo, a região cromossômica, o símbolo do gene, o número de referência e os alelos distintos

<b>Cromossomo</b>	<b>Região</b>	<b>Gene</b>	<b>rs do SNP</b>	<b>Alelo selvagem</b>	<b>Alelo variante</b>	<b>Valor de P</b>
22	q11.1	<i>CECR5</i>	rs5994231	C	T	0,007
3	q11.2	<i>OR5AC2</i>	rs10935236	C	T	0,002
15	q14	<i>MEIS2</i>	rs12101578	C	T	0,009
15	q26.2	<i>MCTP2</i>	rs2161936	C	T	0,003
11	q25	<i>B3GAT1</i>	rs1892993	A	G	0,006
1	q21.3	<i>S100A7</i>	rs2879553	A	G	0,001
1	q23.2	<i>AIM2</i>	rs863035	A	G	0,006
13	q21.33	<i>KLHL1</i>	rs7337751	G	T	0,008
16	q23.3	<i>HSBP1</i>	rs11862993	C	T	0,003
2	q35	<i>MARCH4</i>	rs998029	C	T	0,006
17	q12	<i>FBXO47</i>	rs9912445	A	T	0,005
9	p24.2	<i>RFX3</i>	rs10814587	A	G	0,006
15	q26.2	<i>RGMA</i>	rs7181082	A	T	0,008
3	p12.2	<i>GBE1</i>	rs6548843	A	G	0,007
5	q34	<i>MAT2B</i>	rs12189272	C	G	0,002
15	q21.1	<i>SQRDL</i>	rs2463383	C	T	0,003
11	q25	<i>OPCML</i>	rs2723589	C	T	0,002
21	q22.2	<i>B3GALT5</i>	rs12233349	A	G	0,006
17	q22	<i>PCTP</i>	rs8080394	C	G	0,007
7	p12.1	<i>POM121L12</i>	rs6944190	A	G	0,004
3	q22.2	<i>EPHB1</i>	rs7615982	C	T	0,008
1	p21.3	<i>SLC44A3</i>	rs2933917	A	G	0,004
19	p13.3	<i>PTBP1</i>	rs8104762	C	T	0,004
6	q25.1	<i>PLEKHG1</i>	rs9384539	C	T	0,008
11	p13	<i>WT1</i>	rs5030118	C	T	0,00004
2	p24.3	<i>FAM49A</i>	rs17383609	A	G	0,001
11	p11.2	<i>CD82</i>	rs41358444	A	G	0,008
2	q36.3	<i>DOCK10</i>	rs4674981	A	G	0,009
9	q22.1	<i>NXNL2</i>	rs10120738	C	T	0,002
15	q26.3	<i>ALDH1A3</i>	rs304103	A	G	0,001
3	p26.3	<i>CHL1</i>	rs11709432	A	C	0,0006
5	p15.32	<i>MED10</i>	rs2617545	A	G	0,003

Continuação						
11	p14.3	<i>SVIP</i>	rs12279918	A	G	0,007
8	p23.1	<i>SOX7</i>	rs17152626	C	T	0,009
1	p31.3	<i>GNG12</i>	rs6588286	A	G	0,003
19	q13.2	<i>LGALS14</i>	rs10409279	C	T	0,002
5	p14.2	<i>PRDM9</i>	rs2139793	A	C	0,002
11	p13	<i>RCN1</i>	rs7924866	A	T	0,0002
6	p22.3	<i>HDGFL1</i>	rs6905924	C	T	0,009
7	q31.33	<i>POT1</i>	rs2107097	A	C	0,001
8	p21.3	<i>LZTS1</i>	rs2631880	C	T	0,009
22	q12.3	<i>TST</i>	rs16997638	A	C	0,004
11	q25	<i>B3GAT1</i>	rs4439527	G	T	0,006
5	q32	<i>HTR4</i>	rs1363543	C	G	0,009
3	p26.1	<i>EDEM1</i>	rs4635693	A	G	0,007
6	q22.31	<i>GJA1</i>	rs2389542	A	G	0,007
7	q31.32	<i>SLC13A1</i>	rs2215576	A	C	0,007
1	p35.2	<i>PTPRU</i>	rs909624	A	T	0,009
7	q34	<i>MRPS33</i>	rs491554	C	T	0,006
6	p22.3	<i>HDGFL1</i>	rs9379626	A	G	0,003

rs: número de referência do SNP, SNP: polimorfismo de base única

### 3.4.2. Seleção dos SNPs identificados pela GELE individual

Escolhemos dez SNPs entre os obtidos por meio do algoritmo crlmm (“padrão ouro”), considerando a função de cada gene nos quais eles estiveram localizados, por meio dos programas *DAVID* e do banco de dados do *NCBI*. Dois deles estiveram localizados em regiões codificadoras de aminoácidos (rs2292927 e rs3204145), três deles em regiões 3' e 5' não traduzidas do DNA (rs7114, rs7033 e rs3814) e cinco deles em regiões regulatórias da expressão gênica (rs2107028, rs4245926, rs2082801, rs16909856 e rs16909859).

As amostras dos pacientes e dos controles avaliadas estiveram em equilíbrio de Hardy-Weinberg para oito SNPs selecionados, sugerindo similaridade de frequências dos genótipos entre os grupos. Já as amostras dos pacientes não confirmaram a expectativa de Hardy-Weinberg no SNP rs4245926 do gene *LEF1* e rs2082801 do gene *LY6K*, sugerindo frequências genóticas diferentes das encontradas em controles nestes loci (Tabela 18).

**Tabela 18.** Teste de equilíbrio de Hardy-Weinberg de dez polimorfismos de base de única selecionados e identificados em amostras de 49 pacientes com carcinoma de células escamosas de base de língua e 49 controles

Símbolo do gene	rs do SNP	Pacientes	Controles
<i>COL22A1</i>	2292927	$\chi^2 = 0,00$ ; $P = 1,00$	$\chi^2 = 1,29$ ; $P = 0,26$
<i>IKBKAP</i>	3204145	$\chi^2 = 0,63$ ; $P = 0,43$	$\chi^2 = 1,29$ ; $P = 0,26$
<i>ERP29</i>	7114	$\chi^2 = 0,04$ ; $P = 0,84$	$\chi^2 = 0,78$ ; $P = 0,38$
<i>MCC</i>	7033	$\chi^2 = 1,60$ ; $P = 0,21$	$\chi^2 = 1,32$ ; $P = 0,57$
<i>ZNF415</i>	3814	$\chi^2 = 0,24$ ; $P = 0,62$	$\chi^2 = 0,24$ ; $P = 0,62$
<i>LEF1</i>	2107028	$\chi^2 = 0,29$ ; $P = 0,59$	$\chi^2 = 0,06$ ; $P = 0,81$
<i>LEF1</i>	4245926	$\chi^2 = 11,23$ ; $P < 0,01$	$\chi^2 = 2,83$ ; $P = 0,09$
<i>LY6K</i>	2082801	$\chi^2 = 5,87$ ; $P = 0,01$	$\chi^2 = 0,01$ ; $P = 0,92$
<i>PTCH1</i>	16909856	$\chi^2 = 1,01$ ; $P = 0,31$	$\chi^2 = 0,04$ ; $P = 0,84$
<i>PTCH1</i>	16909859	$\chi^2 = 0,58$ ; $P = 0,45$	$\chi^2 = 0,16$ ; $P = 0,69$

rs: número de referência do SNP, SNP: polimorfismo de base única

As frequências dos genótipos dos dez SNPs selecionados em pacientes com CEC de BL e controles estão apresentadas nas Tabelas 19, 20, 21 e 22.

**Tabela 19.** Frequências dos genótipos dos polimorfismos de base única nas regiões codificadoras de aminoácidos dos genes *COL22A1* e *IKBKAP* selecionados em amostras de sangue periférico dos 49 pacientes com carcinoma de células escamosas de base de língua e 49 controles

<b>Genótipos</b>	<b>Pacientes (%)</b>	<b>Controles (%)</b>	<b>OR (IC 95%)*</b>	<b>P valor</b>
<b><i>COL22A1</i> A1405G</b>				
AA	1 (2,0)	5 (10,2)	1,00 (referência)	
AG	12 (24,5)	16 (32,6)	0,01 (0,01-3,06)	0,12
GG	36 (73,5)	28 (57,2)	38,94 (0,67-2238,01)	0,07
AA+AG	13 (26,5)	21 (42,8)	1,00 (referência)	
GG	36 (73,5)	28 (57,2)	1,77 (0,65-4,78)	0,25
AA	1 (2,0)	5 (10,2)	1,00 (referência)	
AG+GG	48 (98,0)	44 (89,8)	3,23 (1,75-2,34)	0,04
<b><i>IKBKAP</i> T3734A</b>				
TT	39 (79,6)	28 (57,2)	-	0,99
TA	10 (20,4)	16 (32,6)	-	0,99
AA	0 (0,0)	5 (10,2)	1,00 (referência)	
TT+TA	49 (100,0)	44 (89,8)	-	0,99
AA	0 (0,0)	5 (10,2)	1,00 (referência)	
TT	39 (79,6)	28 (57,2)	3,82 (1,28-11,36)	0,01
TA+AA	10 (20,4)	21 (42,8)	1,00 (referência)	

OR: razão das chances, IC: intervalo de confiança, \*OR ajustada por idade, sexo e raça

**Tabela 20.** Frequências dos genótipos dos polimorfismos de base única nas regiões 3' e 5' não traduzidas do DNA dos genes *ERP29*, *MCC* e *ZNF415* selecionados em amostras do sangue periférico dos 49 pacientes com carcinoma de células escamosas de base de língua e 49 controles

<b>Genótipos</b>	<b>Pacientes (%)</b>	<b>Controles (%)</b>	<b>OR (IC 95%)*</b>	<b>P valor</b>
<b><i>ERP29</i> A1058G</b>				
AA	4 (8,2)	0 (0,0)	-	0,99
AG	21 (42,8)	11 (22,4)	-	0,99
GG	24 (49,0)	38 (77,6)	1,00 (referência)	
AA+AG	25 (51,0)	11 (22,4)	2,94 (1,06-8,13)	0,03
GG	24 (49,0)	38 (77,6)	1,00 (referência)	
AA	4 (8,2)	0 (0,0)	-	0,99
AG+GG	45 (91,8)	49 (100,0)	1,00 (referência)	
<b><i>MCC</i> A8553G</b>				
AA	34 (69,4)	20 (40,8)	-	0,99
AG	15 (30,6)	24 (49,0)	-	0,99
GG	0 (0,0)	5 (10,2)	1,00 (referência)	
AA+AG	49 (100,0)	44 (89,8)	-	0,99
GG	0 (0,0)	5 (10,2)	1,00 (referência)	
AA	34 (69,4)	20 (40,8)	3,81 (1,37-10,52)	0,009
AG+GG	15 (30,6)	29 (59,2)	1,00 (referência)	
<b><i>ZNF415</i> C2475T</b>				
CC	9 (18,4)	18 (36,7)	1,00 (referência)	
CT	22 (44,9)	22 (44,9)	0,37 (0,09-1,43)	0,15
TT	18 (36,7)	9 (18,4)	8,69 (1,75-43,07)	0,008
CC+CT	31 (63,3)	40 (81,6)	1,00 (referência)	
TT	18 (36,7)	9 (18,4)	3,64 (1,16-11,42)	0,02
CC	9 (18,4)	18 (36,7)	1,00 (referência)	
CT+TT	40 (81,6)	31 (63,3)	3,87 (1,18-12,63)	0,02

OR: razão das chances, IC: intervalo de confiança, \*OR ajustada por idade, sexo e raça

**Tabela 21.** Frequências dos genótipos dos polimorfismos de base única nas regiões regulatórias da expressão gênica dos genes *LEF1* e *LY6K* selecionados em amostras do sangue periférico dos 49 pacientes com carcinoma de células escamosas de base de língua e 49 controles

<b>Genótipos</b>	<b>Pacientes (%)</b>	<b>Controles (%)</b>	<b>OR (IC 95%)*</b>	<b>P valor</b>
<b><i>LEF1 C126892T</i></b>				
CC	0 (0,0)	3 (6,1)	1,00 (referência)	
CT	7 (14,3)	17 (34,7)	-	0,99
TT	42 (85,6)	29 (59,2)	-	0,99
CC+CT	7 (14,3)	20 (40,8)	1,00 (referência)	
TT	42 (85,6)	29 (59,2)	3,88 (1,26-11,95)	0,01
CC	0 (0,0)	3 (6,1)	1,00 (referência)	
CT+TT	49 (100,0)	46 (93,9)	-	0,99
<b><i>LEF1 C127267T</i></b>				
CC	41 (83,7)	30 (61,2)	-	0,99
CT	8 (16,3)	16 (32,7)	-	0,99
TT	0 (0,0)	3 (6,1)	1,00 (referência)	
CC+CT	49 (100,0)	46 (93,9)	-	0,99
TT	0 (0,0)	3 (6,1)	1,00 (referência)	
CC	41 (83,7)	30 (61,2)	2,97 (1,01-8,92)	0,04
CT+TT	8 (16,3)	19 (38,8)	1,00 (referência)	
<b><i>LY6K C2893T</i></b>				
CC	15 (30,6)	11 (22,4)	4,08 (0,82-20,09)	0,08
CT	31 (63,3)	24 (49,0)	3,58 (0,84-15,19)	0,08
TT	3 (6,1)	14 (28,6)	1,00 (referência)	
CC+CT	46 (93,9)	35 (71,4)	3,67 (1,01-14,93)	0,04
TT	3 (6,1)	14 (28,6)	1,00 (referência)	
CC	15 (30,6)	11 (22,4)	1,52 (0,52-4,47)	0,44
CT+TT	34 (69,4)	38 (77,6)	1,00 (referência)	

OR: razão das chances, IC: intervalo de confiança, \*OR ajustada por idade, sexo e raça

**Tabela 22.** Frequências dos genótipos dos polimorfismos de base única nas regiões regulatórias da expressão gênica do gene *PTCH1* selecionados em amostras do sangue periférico dos 49 pacientes com carcinoma de células escamosas de base de língua e 49 controles

Genótipos	Pacientes (%)	Controles (%)	OR (IC 95%)*	P valor
<b><i>PTCH1 C79755T</i></b>				
CC	41 (83,7)	26 (53,0)	20,76 (1,18-364,06)	0,03
CT	7 (14,3)	19 (38,8)	0,17 (0,01-4,45)	0,29
TT	1 (2,0)	4 (8,2)	1,00 (referência)	
CC+CT	48 (98,0)	45 (91,8)	12,94 (0,84-198,21)	0,06
TT	1 (2,0)	4 (8,2)	1,00 (referência)	
CC	41 (83,7)	26 (53,0)	4,87 (1,57-15,10)	0,006
CT+TT	8 (16,3)	23 (47,0)	1,00 (referência)	
<b><i>PTCH1 C79456T</i></b>				
CC	40 (81,6)	27 (55,1)	0,05 (0,01-1,17)	0,06
CT	8 (16,4)	18 (36,7)	0,19 (0,01-5,71)	0,34
TT	1 (2,0)	4 (8,2)	1,00 (referência)	
CC+CT	48 (98,0)	45 (91,8)	10,48 (0,60-182,58)	0,10
TT	1 (2,0)	4 (8,2)	1,00 (referência)	
CC	40 (81,6)	27 (55,1)	4,54 (1,46-14,06)	0,008
CT+TT	9 (18,4)	22 (44,9)	1,00 (referência)	

OR: razão das chances, IC: intervalo de confiança, \*OR ajustada por idade, sexo e raça

Entre os SNPs localizados nas regiões codificadores de aminoácidos dos genes *COL22A1* e *IKBKAP*, nós observamos frequências maiores dos genótipos *COL22A1* 1405AG+GG e *IKBKAP* 3734TT em pacientes quando comparados aos controles. Os portadores dos genótipos estiveram sob riscos 3,23 e 3,82 vezes maiores de apresentarem o CEC de BL do que os demais, respectivamente.

Entre os SNPs localizados nas regiões 3' e 5' não traduzidas dos genes *ERP29*, *MCC* e *ZNF415*, nós observamos frequências maiores dos genótipos *ERP29* 1058AA+AG, *MCC* 8553AA e *ZNF415* 2475CT+TT em pacientes quando comparados aos controles. Os portadores dos genótipos estiveram sob riscos 2,94, 3,81 e 3,87 vezes

maiores de apresentarem o CEC de BL do que os demais, respectivamente.

Já entre os SNPs localizados nas regiões regulatórias da expressão dos genes *LEF1*, *LY6K* e *PTCH1*, nós observamos frequências maiores dos genótipos *LEF1* 126892TT e 127267CC, *LY6K* 2893CC+CT, *PTCH1* 79755CC e 79456CC em pacientes quando comparados aos controles. Portadores dos genótipos estiveram sob riscos 3,88, 2,97, 3,67, 4,87 e 4,54 vezes maiores de apresentarem o CEC de BL do que os demais, respectivamente.

É interessante comentar que os dez SNPs selecionados por meio do algoritmo crlmm (“padrão ouro”) também foram identificados por meio do algoritmo brlmm, entretanto, considerando um nível de significância de 5%.

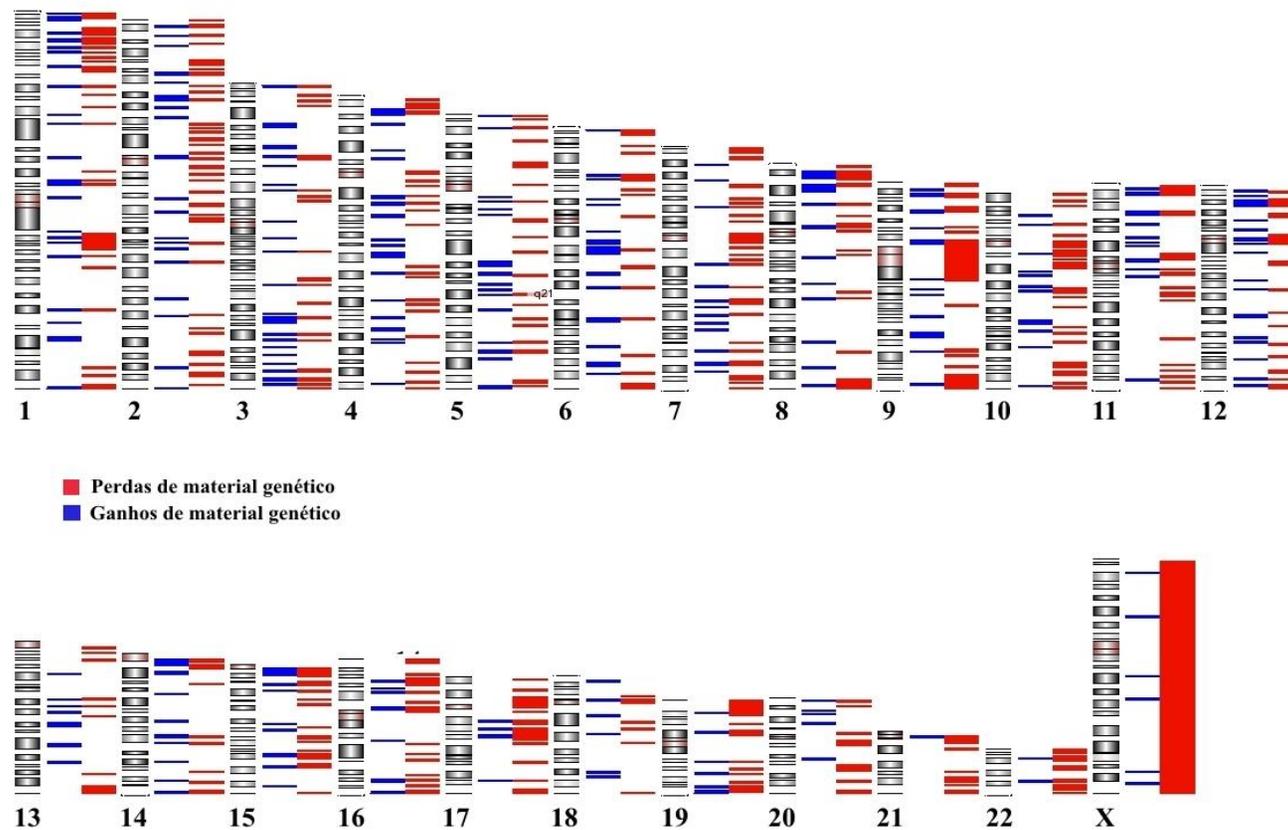
### **3.4.3. GELE coletiva**

Não foi possível realizar as análises dos SNPs contidos nas lâminas com mistura das amostras de DNA de pacientes com CEC de BL e controles por meio dos algoritmos brlmm e crlmm. Os parâmetros de controle de qualidade para a inserção dos dados nos programas não foram suficientes para a realização das análises.

## **3.5. Análise de CNVs**

### **3.5.1. GELE individual**

Identificamos, no total, 1.947 variações no número de cópias de regiões cromossômicas de DNA genômico de pacientes com CEC de BL quando comparados com as amostras de DNA genômico dos controles. Dessas variações, 1.017 (52,2%) foram ganhos de regiões cromossômicas e 930 (47,8%) foram perdas de regiões dos cromossomos (Figura 9).

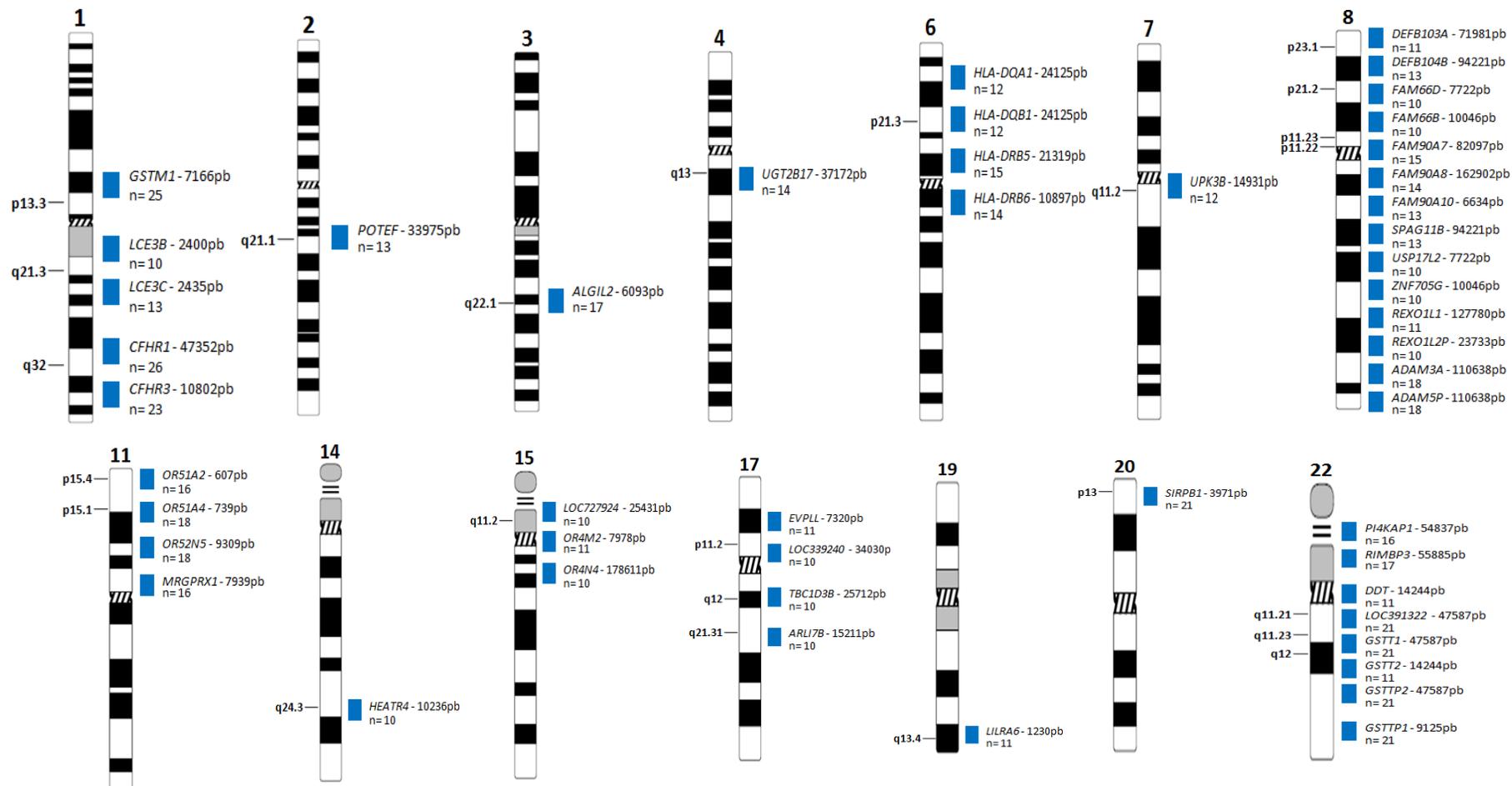


**Figura 9.** Representação esquemática de cariótipo obtido pelo programa Partek<sup>®</sup> após a análise do número de variações de cópias de DNA de 49 pacientes com carcinoma de células escamosas de base de língua em relação aos 49 controles. As regiões azuis e vermelhas ao lado dos cromossomos representam perdas e ganhos de materiais genéticos dos pacientes, respectivamente

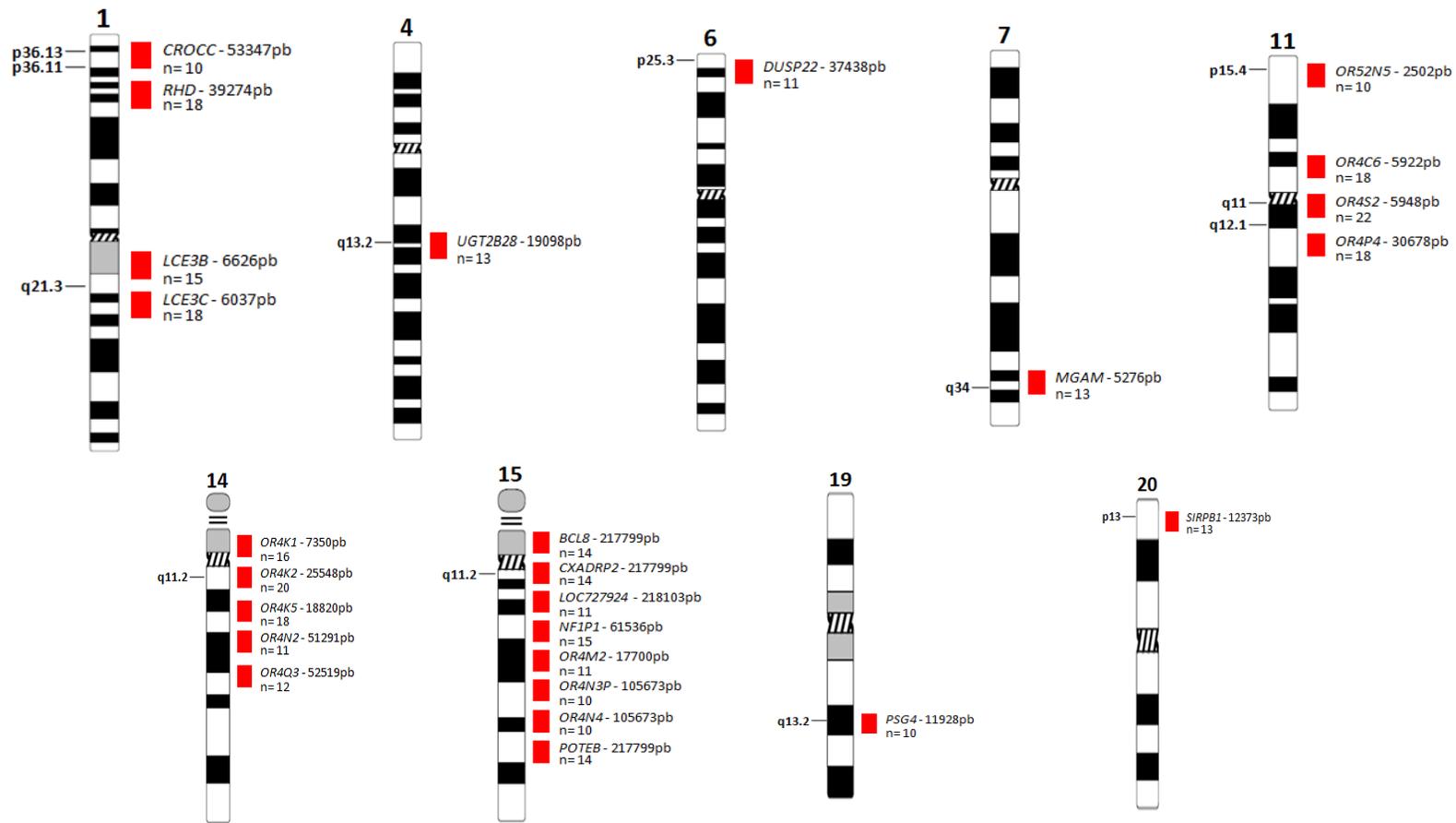
Todos os 49 pacientes com CEC de BL apresentaram ganhos em regiões dos cromossomos com tamanhos menores ou iguais a 99kb. Cerca de 96% deles (47 pacientes) apresentaram ganhos com tamanhos entre 100kb e 499kb. Ainda, cerca de 67% deles (33 pacientes) apresentaram ganhos de tamanhos de regiões cromossômicas maiores ou iguais a 500kb.

Todos os 49 pacientes com CEC de BL também apresentaram perdas de fragmentos de cromossomos menores ou iguais a 99kb. Quarenta e seis pacientes (94%) tiveram perdas de fragmentos de tamanho entre 100kb e 499kb deletados e 28 pacientes (57%) apresentaram perdas maiores ou iguais a 500kb.

Para identificar possíveis regiões cromossômicas e genes envolvidos com o risco de CEC de BL, nós selecionamos e definimos aqueles descritos em regiões em excesso ou deletadas que foram encontradas em pelo menos dez pacientes. Assim, identificamos as regiões de CNVs em 26 (53%) dos 49 pacientes estudados. As regiões de ganhos e perdas de DNA de genes e seus respectivos cromossomos, tamanho da variação do DNA e o número de pacientes que apresentaram as variações estão apresentados nos Anexos 10 e 11, respectivamente. As regiões de ganhos e perdas de DNA de genes e seus respectivos cromossomos e genes estão apresentadas nas Figuras 10 e 11, respectivamente.



**Figura 10.** Representação esquemática de regiões de ganhos de DNA em cromossomos e genes após a análise de variações no número de cópias de DNA de 49 pacientes com carcinoma de células escamosas de base de língua em relação aos controles, onde n é o número de pacientes que apresentou ganhos cromossômicos



**Figura 11.** Representação esquemática de regiões de perdas de DNA em cromossomos e genes após a análise de variações no número de cópias de DNA de 49 pacientes com carcinoma de células escamosas de base de língua em relação aos controles, onde n é o número de pacientes que apresentaram perdas cromossômicas

Identificamos expressivos ganhos de material genético em 49 genes de amostras de DNA individuais dos pacientes com CEC de BL. Observamos 85 diferentes tamanhos de ganhos de material genético dentre os 49 genes identificados. Ganhos de DNA menores que 100kb foram identificados na maioria dos casos (94%). Cerca de 6% dos ganhos de DNA estiveram entre 100 e 180kb.

Dentre os 49 genes com ganhos expressivos de DNA, 14 genes (28,6%) estiveram localizados no cromossomo 8 e oito genes (16,3%) no cromossomo 22

Identificamos expressivas perdas de material genético em 26 genes de amostras de DNA individuais dos pacientes com CEC de BL.

Observamos 41 diferentes tamanhos de perdas de material genético dentre os 26 genes identificados. A maioria das perdas de DNA foi menor que 100kb (85,4%). Cerca de 14,6% dos tamanhos das perdas de DNA estiveram entre 100 e 218kb.

Dentre os 26 genes identificados com perdas expressivas de DNA, oito genes (30,8%) estiveram localizados no cromossomo 15, cinco (19,2%) no cromossomo 14, e quatro (15,4%) no cromossomo 11.

### **3.5.2. Seleção das CNVs identificadas pela GELE individual**

Entre as CNVs que apresentaram ganhos de regiões cromossômicas, nós escolhemos as regiões do braço curto do cromossomo 8 (p11 e p23) e as regiões do braço longo do cromossomo 22 (q11.23) e entre as que apresentaram perdas de regiões cromossômicas, nós escolhemos as regiões dos braços longos dos cromossomos 14 (q11.2) e 15 (q11.2) para futuro processo de validação por meio da PCR em tempo real com sondas específicas para cada região de interesse, considerando a função dos genes localizados nestas regiões, por meio dos programas *DAVID* e do banco de dados do *NCBI*.

### **3.5.3. GELE coletiva**

Identificamos, no total, 26 variações no número de cópias de regiões cromossômicas nas misturas do DNA genômico de 39 pacientes caucasóides e de dez pacientes negróides com CEC de BL quando comparados com as misturas do DNA genômico de 39 controles caucasóides e dez controles negróides, respectivamente. Destas variações, 18 (69,2%) foram ganhos de regiões cromossômicas e dez (38,5%) foram perdas de regiões

cromossômicas. A maioria das variações (88,5%) foi de ganhos ou perdas de material genético menores que 100kb. Cerca de 19% dos tamanhos das variações foi entre 100 e 240kb.

Quando consideramos apenas a mistura de DNA de 39 pacientes caucasóides com CEC de BL, identificamos, no total, 34 variações no número de cópias de regiões cromossômicas quando comparados com as misturas dos respectivos controles caucasóides. Destas variações, quatro (11,8%) foram ganhos e 30 (88,2%) foram perdas de regiões cromossômicas. A maioria das variações (94,0%), ganhos ou perdas de material genético, foi menor que 100kb. Cerca de 6% das variações de tamanho de cromossomos esteve entre 100 e 200kb.

Quando consideramos apenas a mistura de DNA de dez pacientes negróides com CEC de BL, identificamos 15 variações no número de cópias de regiões cromossômicas quando comparados com as misturas dos DNAs dos respectivos controles negróides. Destas variações, 13 (86,7%) foram ganhos e duas (13,3%) foram perdas de regiões cromossômicas. A maioria das variações (86,7%), ganhos ou perdas de material genético, foi menor que 100kb. Cerca de 13% das variações esteve entre 100 e 200kb.

Quando avaliamos os resultados da GELE coletiva dos pacientes em geral (caucasóides e negróides) (2 lâminas) comparados aos da GELE coletiva dos controles em geral (caucasóides e negróides) (2 lâminas), observamos ganhos de material genético apresentados na Tabela 23.

Já quando avaliamos os resultados das GELEs coletivas de pacientes (1 lâmina) e controles (1 lâmina) caucasóides e as GELEs coletivas de pacientes (1 lâmina) e controles (1 lâmina) negróides, observamos ganhos de material genético apresentados nas Tabelas 24 e 25, respectivamente.

**Tabela 23.** Regiões de ganhos de material genético e seus respectivos cromossomos, genes e tamanhos das regiões em misturas de 49 amostras de DNA de pacientes caucasóides e negróides com carcinoma de células escamosas de base de língua comparadas a misturas de 49 amostras de DNA de controles caucasóides e negróides

Cromossomo	Símbolo do gene	Região de início	Região de término	Tamanho (pb)
1	<i>SPEN</i>	16089102	16136067	46965
2	<i>DUSP11</i>	73850639	73892140	41501
3	<i>FANCD2</i>	10047122	10098021	50899
4	<i>UGT2B17</i>	68899654	69143076	243422
8	<i>ADAM3A</i>	39374430	39489516	115086
9	<i>EXD3</i>	139315069	139346860	31791
11	<i>MUC2</i>	1064543	1076571	12028
11	<i>TNNT3</i>	1883376	1899646	16270
17	<i>GRAP</i>	18854462	19097832	243370
20	<i>RTEL1</i>	61787628	61794808	7180
22	<i>GSTT1</i>	22674016	22730726	56710

pb: pares de bases

Identificamos expressivos ganhos de material genético em 11 genes em amostras de DNA de 39 pacientes caucasóides e dez pacientes negróides com CEC de BL.

O tamanho da maioria dos ganhos de DNA foi menor que 100kb (72,7%). Cerca de 27% dos ganhos de DNA estiveram entre 100 e 240kb.

Vale comentar que os genes *UGT2B17*, *ADAM3A* e *GSTT1* também foram identificados nas análises das amostras individuais de DNA dos pacientes com CEC de BL.

**Tabela 24.** Regiões de ganhos de material genético e seus respectivos cromossomos, genes e tamanhos das variações em misturas de 39 amostras de DNA de pacientes caucasóides com carcinoma de células escamosas de base de língua comparadas a misturas de 39 amostras de DNA de controles caucasóides

<b>Cromossomo</b>	<b>Símbolo do gene</b>	<b>Região de início</b>	<b>Região de término</b>	<b>Tamanho (pb)</b>
3	<i>VHL</i>	10149453	10188726	39273
10	<i>MGMT</i>	131214465	131270997	56532

pb: pares de bases

Identificamos expressivos ganhos de material genético em apenas dois genes em amostras de mistura de DNA de 39 pacientes caucasóides com CEC de BL.

Os tamanhos das variações foram menores que 100kb.

Nenhum dos dois genes identificados nessas misturas de DNA foram encontrados nas amostras individuais.

**Tabela 25.** Regiões de ganhos de material genético e seus respectivos cromossomos, genes e tamanhos das variações em misturas de dez amostras de DNA de pacientes negróides com carcinoma de células escamosas de base de língua comparadas a misturas de dez amostras de DNA de controles negróides

Cromossomo	Símbolo do gene	Região de início	Região de término	Tamanho (pb)
1	<i>CELA3A</i>	22188651	22217957	29306
2	<i>DUSP11</i>	73850639	73892140	41501
2	<i>ASB18</i>	236795577	236829818	34241
4	<i>UGT2B17</i>	69000102	69169546	169444
7	<i>C7orf28B</i>	6825413	6853010	27597
8	<i>ADAM3A</i>	39371150	39489516	118366
9	<i>EXD3</i>	139315069	139346860	31791
11	<i>KDM2A</i>	66592082	66656213	64131
11	<i>TNNT3</i>	1883376	1899646	16270
19	<i>PSG4</i>	48400861	48445571	44710
20	<i>RTEL1</i>	61779368	61794808	15440
22	<i>GSTT1</i>	22674016	22730726	56710
X	<i>ARMCX3</i>	100749443	100764805	15362

pb: pares de bases

Identificamos expressivos ganhos de material genético em 13 genes em amostras de mistura de DNA de dez pacientes negróides com CEC de BL.

A maioria dos ganhos de DNA foi menor que 100kb (84,6%). Ganhos de material genético maiores que 100kb foram observados em apenas duas CNVs.

Vale comentar que os genes *UGT2B17*, *ADAM3A*, *PSG4* e *GSTT1* também foram identificados nas análises das amostras individuais de DNA dos pacientes com CEC de BL.

Quando avaliamos os resultados da GELE coletiva dos pacientes em geral (caucasóides e negróides) (2 lâminas) comparados aos da GELE coletiva dos controles em geral (caucasóides e negróides) (2 lâminas), observamos perdas de material genético apresentados na Tabela 26. Já quando avaliamos os resultados das GELEs coletivas de pacientes (1 lâmina) e controles (1 lâmina) caucasóides e as GELEs coletivas de pacientes (1 lâmina) e controles (1 lâmina) negróides, observamos perdas de material genético apresentados nas Tabelas 27 e 28, respectivamente.

**Tabela 26.** Regiões de perdas de material genético e seus respectivos cromossomos, genes e tamanhos das variações em misturas de 49 amostras de DNA de pacientes caucasóides e negróides com carcinoma de células escamosas de base de língua comparadas a misturas de 49 amostras de DNA de controles caucasóides e negróides

Cromossomo	Símbolo do gene	Região de início	Região de término	Tamanho (pb)
1	<i>ESPNP</i>	16909632	16919371	9739
1	<i>LCE3B</i>	150818718	150855084	36366
8	<i>FAM66E</i>	7832058	7975607	143549
11	<i>ASRGL1</i>	61877851	61907411	29560
X	<i>ARMCX3</i>	100755403	100766680	11277

pb: pares de bases

Identificamos expressivas perdas de material genético em cinco genes em amostras de mistura de DNA de 39 pacientes caucasóides e dez pacientes negróides com CEC de BL.

O tamanho da maioria dos ganhos de DNA foi menor que 100kb. Perda de material genético maior que 100kb foi identificada em apenas uma CNV.

Vale comentar que apenas o gene *LCE3B* também foi identificado nas análises das amostras individuais de DNA dos pacientes com CEC de BL.

**Tabela 27.** Regiões de perdas de material genético e seus respectivos cromossomos, genes e tamanhos das variações em misturas de 39 amostras de DNA de pacientes caucasóides com carcinoma de células escamosas de base de língua comparadas a misturas de 39 amostras de DNA de controles caucasóides

Cromossomo	Símbolo do gene	Região de início	Região de término	Tamanho (pb)
1	<i>LCE3C</i>	150818194	150848459	30265
2	<i>OR7E91P</i>	71090136	71121687	31551
4	<i>MIR548I2</i>	9067492	9185881	118389
7	<i>MGAM</i>	141406414	141444658	38244
7	<i>MUC12</i>	100421037	100447345	26308
9	<i>FAM27A</i>	45416679	45620362	203683
9	<i>KANK1</i>	668255	679311	11056
9	<i>RAPGEF1</i>	133410144	133446226	36082
10	<i>ADK</i>	75916067	75948220	32153
11	<i>ASRGL1</i>	61894027	61908217	14190
12	<i>OR6C1</i>	54000759	54020050	19291
19	<i>RPS28</i>	8241954	8306588	64634
22	<i>RRP7A</i>	41238113	41279614	41501
X	<i>ARMCX3</i>	100755403	100778731	23328

pb: pares de bases

Identificamos expressivas perdas de material genético em 14 genes em amostras de mistura de DNA de 39 pacientes caucasóides com CEC de BL.

O tamanho da maioria das perdas de DNA foi menor que 100kb (85,7%). Perdas maiores que 100kb foram observadas em apenas dois genes (14,3%).

Vale comentar que os genes *LCE3C*, *MGAM* também foram identificados nas análises das amostras individuais de DNA dos pacientes com CEC de BL.

**Tabela 28.** Regiões de perdas de material genético e seus respectivos cromossomos, genes e tamanhos das variações em misturas de dez amostras de DNA de pacientes negróides com carcinoma de células escamosas de base de língua comparadas a misturas de dez amostras de DNA de controles negróides

<b>Cromossomo</b>	<b>Símbolo do gene</b>	<b>Região de início</b>	<b>Região de término</b>	<b>Tamanho (pb)</b>
10	<i>CCDC109A</i>	74258144	74304119	45975
11	<i>MRGPRX1</i>	18905894	18918208	12314

pb: pares de bases

Identificamos expressivas perdas de material genético em apenas dois genes de amostras de mistura de DNA de 10 pacientes negróides com CEC de BL.

Os tamanhos das variações foram menores que 100kb.

Nenhum dos dois genes identificados nessas misturas de DNA foram encontrados nas amostras individuais.



# **DISCUSSÃO**

---



Já é conhecido que o CEC de BL é um tumor de alta incidência e que ocasiona altas taxas de morbidade e mortalidade (SESSIONS et al., 2003; KOWALSKI et al., 2005; MENDENHALL et al., 2005; PALERI et al., 2010). A probabilidade mediana de sobrevivência é cerca de quatro anos para os pacientes do sexo masculino com CEC de BL que realizam a remoção cirúrgica do tumor e a radioterapia (SESSIONS et al., 2003). Recentemente, a quimioterapia associada à radioterapia e a remoção do tumor, aumentou em 16% a sobrevida destes pacientes (FURNESS et al., 2010).

Com os avanços no conhecimento da biologia molecular do câncer, os tumores de cabeça e pescoço, antes considerados doenças homogêneas, agora, são reconhecidos como entidades distintas (RISCHIN, 2010). É importante identificar fatores prognósticos para determinar o curso clínico de pacientes com doenças em sítios distintos desses tumores, incluindo o de BL. O conhecimento de tais fatores pode contribuir também para desenvolver estratégias terapêuticas, principalmente aquelas baseadas em riscos e, assim, possibilitar a identificação de subgrupos de indivíduos que mereçam receber tratamento mais ou menos agressivo (CHEN & CHEN, 2008; VAN DEN EYNDE et al., 2011).

Quando os nossos pacientes com CEC de BL foram distribuídos por idade, sexo e estágio do tumor, notamos que foram semelhantes a pacientes do hemisfério norte (MENDENHALL et al., 2005; MARCUS & TISHLER, 2010) e, desta forma, parecem constituir um grupo representativo da doença em nosso meio.

Quando consideramos a distribuição dos nossos casos por raça, verificamos que a doença ocorreu predominantemente em caucasianos, de forma semelhante às descrições obtidas em outros países (GOURIN & PODOLSKY, 2006). Entretanto, esses resultados devem ser vistos com reserva, uma vez que a população brasileira é altamente heterogênea e composta por indivíduos miscigenados (ALVES et al., 2000; CARVALHO et al., 2001) e que avaliamos os indivíduos por raça apenas pelo biótipo dos mesmos.

Observamos que a maioria dos pacientes inseridos neste estudo eram tabagistas e etilistas. O tabagismo e a exposição ao álcool atuam de forma sinérgica no risco de desenvolvimento ao CEC de BL (LAMBERT et al., 2011). É difícil avaliar o efeito individual destes agentes na etiologia do CEC de BL, pois os indivíduos fumantes tendem a ser os mesmos a consumir álcool e vice-versa (ABBAS et al. 2004; DURAZZO et al., 2005).

O benzopireno é um composto presente no tabaco cujos efeitos carcinogênicos são capazes de induzir mutações em células e o CEC de BL. O benzopireno transforma-se em produtos epóxidos, que são substâncias reativas capazes de induzir danos ao DNA celular (HAYES et al., 2005).

O álcool atua como um fator irritante direto da mucosa. Pode também atuar indiretamente, promovendo má nutrição e imunodeficiência. Esses fatores estimulam e promovem mudanças na reatividade da mucosa à vários estímulos carcinogênicos (KOWALSKI, 2006) e podem, assim, determinar o tumor. O etanol das bebidas alcoólicas é metabolizado a acetaldeído, que é também um carcinógeno (BRENNAN et al., 2004).

A distribuição dos pacientes deste estudo em relação ao padrão alimentar mostrou que a doença ocorreu predominantemente em indivíduos com ingestão preconizada ou alta ingestão de carnes, sugerindo que este tipo de alimento pode contribuir para a origem da doença. O alto consumo de carnes salgadas e defumadas representam fatores de risco para a doença, por serem fontes de hidrocarbonetos aromáticos policíclicos (BRENNAN et al., 2004).

Metade da nossa casuística referiu ingestão insuficiente de frutas e vegetais, como previamente descrito, sugerindo que a carência de ácido fólico contido nesses alimentos possam contribuir para a origem do tumor (GARAVELLO et al., 2009; LUCENTEFORTE et al., 2009). Brennan et al. (2004) observaram um aumento para o risco de CEC em indivíduos com consumo insuficiente de frutas e vegetais (BRENNAN et al., 2004).

Observou-se que cerca de 80% dos nossos casos apresentaram tumores bem ou moderadamente diferenciados e apenas 20% deles, como pouco ou indiferenciados. Frequências similares de graus de diferenciação tumoral foram descritos por Mori et al. (1998) em pacientes com CEC de cabeça e pescoço.

Quanto à distribuição de acordo com o estágio da doença, mais da metade da amostra apresentou tumores avançados, de acordo com descrições prévias (SESSIONS et al., 2003). A explicação para o fato de a maioria dos pacientes com CEC de BL chegarem aos hospitais de referência com a doença em estágios avançados, pode resultar da dificuldade de acesso de pacientes a serviços públicos de saúde e a falta de profissionais de saúde preparados para diagnosticar precocemente as lesões e encaminhar os pacientes em estágios mais iniciais da doença para centros especializados.

Em seu conjunto os resultados deste estudo indicam que os pacientes avaliados com

CEC de BL apresentaram características clínicas e anatomopatológicas do tumor semelhantes às de pacientes de outros continentes e, portanto, constituíram um grupo representativo da doença em nosso meio, próprios para avaliações moleculares à predisposição do tumor.

Os CEC de cabeça e pescoço foram extensivamente estudados nos últimos anos, particularmente em suas características genéticas, por meio de métodos de identificação de mutações adquiridas em possíveis genes envolvidos em sua fisiopatologia (MOLINOLO et al., 2009; LEEMANS et al., 2011). Mutações em oncogenes, como os da família *RAS*, em genes supressores de tumor, como o *TP16* e o *P53*, e em genes codificadores de fatores de crescimento, como os *EGFR* e *TGF- $\beta$* , parecem estar associadas com a origem destes tumores (SIDRANSKY, 2005; GOLD & KIM, 2009; MOLINOLO et al., 2009).

Atualmente, os polimorfismos genéticos são considerados fatores importantes para a predisposição herdada de um indivíduo a tumores, incluindo o CEC de BL (HIYAMA et al., 2008). Essas variações genéticas determinam a capacidade distinta de cada indivíduo metabolizar carcinógenos, controlar o mecanismo do ciclo celular, reparar danos ao DNA e outros mecanismos fisiológicos. Essas variações indicam que a constituição genética pode predispor um indivíduo ao câncer (DUTT & BEROUKHIM, 2007; ASSIE et al., 2008; CHEN & CHEN, 2008; GOLD & KIM, 2009).

SNPs e CNVs em genes diversos foram associados com a ocorrência de CEC de cabeça e pescoço, incluindo o de BL (WAGNER et al., 2001; ZHENG et al., 2001; NISHIMOTO et al., 2004; STREIT et al., 2004; CUI et al., 2006; PERRONE et al., 2007; CHEN et al., 2008; CHEN & CHEN, 2008; JI et al., 2008; GUAN et al., 2010; LOURENÇO et al., 2011). Entretanto, apenas pequenos números destes polimorfismos genéticos foram avaliados na maioria dos estudos (CHEN & CHEN, 2008; HIYAMA et al., 2008).

Os SNPs mais estudados foram aqueles localizados em genes que codificam enzimas que metabolizam carcinógenos originados da combustão do tabaco e do metabolismo de bebidas alcoólicas, como as do citocromo P450 (CYPs) e as glutathione S-transferases (GSTs) (HIYAMA et al., 2008; LACKO et al., 2009). Recentemente, o nosso grupo de pesquisadores observou que os indivíduos portadores dos genótipos variantes dos polimorfismos A4889G e T6235C do gene *CYP1A1* e da deleção homocigótica dos genes *GSTM1* e *GSTT1* estiveram sob risco cerca de 16 vezes maior para o CEC de faringe

quando comparados com aqueles com os genótipos selvagens (LOURENÇO et al., 2011) (Anexo 12). Ainda, é possível que outros SNPs envolvidos nesta via de metabolização possam ter papéis importantes na origem do CEC de BL.

SNPs em genes que codificam proteínas relacionadas com fatores de crescimento celular (STREIT et al., 2004; GUAN et al., 2010), ao controle do ciclo celular (WAGNER et al., 2001; PERRONE et al., 2007; CHEN et al., 2008; JI et al., 2008) e ao reparo a danos no DNA (CUI et al., 2006), também foram associados ao risco de ocorrência do tumor.

As CNVs encontradas frequentemente em CEC de BL foram as 3p, 4q, 5q21-22, 8p21-23, 9p21-22, 11q13, 11q23, 13q, 14q, 17p, 18q e 22q (CHEN & CHEN, 2008). Entretanto, a maioria dos estudos avaliou as CNVs encontradas no DNA tumoral (CHEN & CHEN, 2008). Até onde sabemos, estes autores não verificaram se as CNVs estavam presentes no sangue periférico, assim, não é possível saber se foram anormalidades adquiridas no tumor ou se constituíram anormalidades hereditárias associadas com a predisposição do tumor (ASSIE et al., 2008; CHEN & CHEN, 2008).

Recentemente, a GELE, com base em microarranjos de DNA, possibilitou analisar milhares de polimorfismos genéticos em um único estudo. A GELE proporciona a varredura de diversos marcadores genéticos, localizados em uma única lâmina com microarranjos de DNA, que podem estar relacionados com o surgimento e a agressividade de tumores. Além disso, ela proporciona maior rapidez e confiabilidade para a avaliação de SNPs e CNVs quando comparada com métodos tradicionais, como a PCR-RFLP e o CGH, respectivamente (SYVANEN, 2001; MAO et al., 2007; CHEN & CHEN, 2008; HIYAMA et al., 2008; XING et al., 2008).

Nos dias atuais, o procedimento de GELE só é possível devido aos avanços de pesquisas na área de bioinformática. As sequências ordenadas de regras e cálculos determinados por algoritmos computacionais possibilitam avaliar lâminas com microarranjos de DNA que possuem alta densidade de marcadores genéticos em um pequeno espaço físico (CARVALHO & IRIZARRY, 2010; CARVALHO et al., 2010).

O algoritmo de análise da GELE de SNPs utilizado como padrão “ouro” em nosso estudo foi o *crlmm*, disponível no programa computacional Bioconductor, cujo código fonte está disponível para qualquer pesquisador, sem custo. A sequência básica de regras deste algoritmo que possibilita a genotipagem de SNPs de microarranjos de DNA consiste na normalização e sumarização das intensidades das sondas contidas nas lâminas de

microarranjos seguido da genotipagem. O algoritmo crlmm foi descrito por possuir uma melhor acurácia quando comparado com o algoritmo brlmm, descrito previamente (CARVALHO et al., 2007).

Por meio da GELE individual, nós observamos que as frequências de 6.609 SNPs foram distintas entre os pacientes com CEC de BL e controles. Cerca da metade do número destes SNPs esteve localizada em regiões do DNA que controlam a expressão gênica e a outra metade, esteve localizada nas regiões de íntrons do DNA. Aproximadamente 1% deles estiveram localizados em regiões do DNA que codificam proteínas e 1% em regiões 3' e 5' não traduzidas do DNA. Estes números foram considerados adequados, uma vez que a maioria dos estudos com CEC de cabeça e pescoço, que utilizou a GELE individual, observou que cerca da metade dos SNPs estão localizados em regiões regulatórias da expressão de genes (NG et al., 2009; TSE et al., 2009; BEI et al., 2010).

Identificamos inúmeros SNPs em genes relacionados com o metabolismo de carcinógenos, como o *CRYZ* e o *ALDH1A3*; regulação do ciclo celular, como o *USP2*, o *SYNJ2*, o *RAB6C*, o *S100A7*, o *AIM2* e o *POT1*; reparo de DNA, como o *GEN1*; apoptose, como o *CASP5*, o *SPATA4* e o *SLC7A11*; angiogênese, como o *FERMT1*, o *JMJD6*, o *CTSB*, o *GCNT1*, o *CD82* e o *CHL1* e processo transcricional, como os genes *MLL3*, *DDX46*, *POLQ*, *RPAP1*, *CCDC6*, *ARHGEF18*, *TRMT5*, *HSBP1*, *RFX3*, *PTBP1*, *WT1* e *SOX7*. E, ainda, SNPs localizados em genes candidatos a supressores de tumores, como os *CHFR*, *TUSC1*, *OPCML* e o *LZTS1*. A identificação das funções dos genes foi realizada por meio dos programas *DAVID* e do banco de dados do *NCBI*.

A seguir, escolhemos dez SNPs para validar os nossos resultados com a GELE. A validação será realizada em breve por pesquisadores do nosso grupo por meio da PCR em tempo real, utilizando sondas *TaqMan* específicas para cada lócus dos genes polimórficos. Os SNPs selecionados apresentaram frequências dos alelos variantes maiores que 1%. Os SNPs estiveram localizados em genes cujas funções estão relacionadas principalmente ao controle do processo transcricional de genes (*IKBKAP* e *ZNF415*) e do controle do ciclo celular (*ERP29*, *MCC*, *LEF1*, *LY6K* e *PTCH1*), mecanismos que podem estar relacionados com o surgimento de tumores (LEEMANS et al., 2011).

Vale comentar que não há descrições das funções e associações dos dez SNPs selecionados com o surgimento de tumores, incluindo o CEC de BL. As descrições observadas são apenas das funções das proteínas codificadas pelos alelos selvagens dos

genes distintos.

Entre os SNPs selecionados e localizados na região do DNA codificadora de cadeias de aminoácidos, encontram-se os localizados nos genes *COL22A1* e *IKBKAP*, descritos aqui, pela primeira vez em associação a tumores de BL.

A proteína codificada pelo gene *COL22A1*, o colágeno XXII, foi descrita como um novo marcador de adesão celular e junção de tecidos por Koch et al. (2004). Proteínas da família do colágeno, por exemplo os colágenos I, IV e XVIII, foram relacionados com o mecanismo de metástase e angiogênese tumoral (SUND & KALLURI, 2009; EGEBLAD et al., 2010). Entretanto, o papel do colágeno XXII no mecanismo de carcinogênese é desconhecido. Frente ao exposto, nos pareceu interessante selecionar e avaliar um SNP que codifica proteína do colágeno. O SNP A1405G (rs2292927) ocasiona a troca do aminoácido serina pela glicina no alelo variante. Nós observamos que os indivíduos que apresentaram o alelo variante (G) estiveram sob risco cerca de três vezes maior de ocorrência do CEC de BL do que aqueles que apresentaram o alelo selvagem (A).

Já o gene *IKBKAP* codifica uma proteína do complexo de alongação do processo da transcrição do DNA pela enzima RNA polimerase II (KIM et al., 2002). Além da função de modular o processo transcricional, ela pode influenciar a sensibilidade de indivíduos a agentes quimioterápicos, como a hidroxiuréia (LI et al., 2009). Apesar de mutações neste gene estarem associadas à doenças do sistema nervoso autônomo (HOLMBERG et al., 2002) e ao risco da asma brônquica em crianças (TAKEOKA et al., 2001), os papéis de mutações e de SNPs em genes que regulam a transcrição do DNA não estão totalmente conhecidos (KIM et al., 2002). O SNP T3734A (rs3204145), localizado neste gene, ocasiona a troca do aminoácido cisteína por uma serina na cadeia de aminoácidos do alelo variante. O papel deste SNP é desconhecido. Nós observamos que indivíduos que apresentaram o genótipo selvagem (TT) estiveram sob risco cerca de quatro vezes maior de ocorrência do CEC de BL em comparação àqueles que apresentaram os demais genótipos.

Até onde atinge o nosso conhecimento, não se sabe se os referidos polimorfismos determinam variações quantitativas ou qualitativas nas proteínas produzidas.

As regiões 3' e 5' não traduzidas do DNA controlam a expressão gênica por meio da capacidade de manter a estabilidade do RNA mensageiro (RNAm), além de possuírem informações de posicionamento do RNAm nos ribossomos para o início da tradução da sequência de nucleotídeos para aminoácidos e proteínas (HOLLAMS et al., 2002; CHEN et

al., 2006). Os papéis de SNPs encontrados nessas regiões não foram ainda estabelecidos (HOLLAMS et al., 2002). Os SNPs nos genes *ERP29*, *MCC* e *ZNF415* relacionados com o CEC de BL foram identificados pela primeira vez em nosso estudo.

A hiperexpressão da proteína do retículo endoplasmático 29 (Erp29), codificada pelo gene *ERP29*, está inversamente relacionada com a progressão de tumores (BAMBANG et al., 2009; ZHANG & RICHARDSON, 2011). Bambang et al. (2009) sugerem que o gene *ERP29* é um novo candidato a gene supressor tumoral. Esta proteína regulou a parada do ciclo celular de um estado proliferativo para outro de quiescência em linhagem celular de tumor invasivo de mama (BAMBANG et al., 2009). Até o momento, o SNP A1058G (rs7114), localizado neste gene, não possui a sua função estabelecida. Nós observamos que indivíduos que apresentaram alelo selvagem (A) estiveram sob risco cerca de três vezes maior de ocorrência do CEC de BL do que os indivíduos com o genótipo variante (GG).

O gene *MCC* também codifica uma proteína candidata a supressora de tumor, devido à regulação negativa da proliferação celular em linhagem de células de tumor colorretal (FUKUYAMA et al., 2008). A proteína MCC interage com a proteína  $\beta$ -catenina e inibem a cascata da proliferação celular iniciada pela glicoproteína Wnt (FUKUYAMA et al., 2008). É importante comentar que a sinalização da proliferação celular iniciada pelas proteínas  $\beta$ -catenina e Wnt também são encontradas em diversos tipos de cânceres, entre eles o de cabeça e pescoço (IWAI et al., 2010). O SNP A8553G (rs7033) neste gene ainda não tem a sua função conhecida. Nós observamos que os indivíduos que apresentaram o genótipo selvagem (AA) estiveram sob risco de ocorrência do CEC de BL cerca de quatro vezes maior do que indivíduos que apresentaram o alelo variante (G).

Frente ao exposto, julgamos importante definir os papéis das proteínas codificadas pelos alelos selvagens e variantes dos referidos SNPs para identificar a função dos mesmos na supressão de CEC de BL.

As proteínas denominadas dedos de zinco (*zinc finger*) têm papel importante no controle da especificidade de ligação dos fatores de transcrição na sequência do DNA, determinando, assim, o controle da transcrição do DNA (CHENG et al., 2006). O gene *ZNF415* codifica uma proteína da família dedo de zinco número 415 que pode estar envolvida na atividade transcricional de genes importantes relacionados com a carcinogênese, como o *P53* (CHENG et al., 2006). Até o momento, não há descrições da

relação deste gene e do SNP C2475T (rs3814) com o surgimento de tumores, incluindo o de BL. Indivíduos que apresentaram o alelo variante (T) estiveram sob risco de ocorrência do CEC de BL cerca de quatro vezes maior do que os com o genótipo selvagem (CC).

A validação desse resultado, seguido da avaliação do papel da proteína codificada pelo alelo selvagem 2475CC do gene será importante para definir a relação desta proteína com o gene supressor tumoral *P53* e o surgimento do CEC de BL.

A posição e a direção em que enzimas específicas se ligam na fita do DNA para realizar o processo de transcrição e tradução do DNA e do RNAm, respectivamente, é determinada pela ligação nas moléculas no carbono 5' (ribose) e no carbono 3' (desoxirribose) na cadeia de nucleotídeos. A região *downstream* possui a sequência de nucleotídeos de ligação destas enzimas que reconhece a sequência 5' – 3' da fita do DNA (HOLMES & ERIE, 2003). Os papéis de SNPs localizados nestas regiões não estão estabelecidos (PALANGAT et al., 2004). Os SNPs selecionados neste estudo e localizados nesta região estão inseridos nos genes *LEF1*, *LY6K* e *PTCH1* e ainda não foram associados com o surgimento do CEC de BL.

O gene *LEF1* codifica a proteína do fator de potencialização linfóide 1 (Lef1). Esta proteína participa, juntamente com as proteínas  $\beta$ -catenina e Wnt, de um complexo de ativação ou repressão da proliferação celular, por meio do controle da adesão celular, metástase e angiogênese (CONACCI-SORRELL et al., 2002; HOLNTHONER et al., 2002; TORRES et al., 2007; NGUYEN et al., 2009; ZHANG & WAXMAN, 2010). A Lef1 também foi associada com o mecanismo de apoptose via sinalização da proteína Notch1 (SPAUDING et al., 2007). A hiperexpressão da Lef1 foi associada com a origem de tumores do cólon (WATERMAN et al., 2004; LI et al., 2006; LIN et al., 2011), tumores epiteliais (WANG et al., 2005), hematológicos (WANG et al., 2005), de mama (NGUYEN et al., 2005) e da próstata (LI et al., 2009). Os SNPs C126892T (rs2107028) e C127267T (rs4245926) localizados na região *downstream* desse gene não possuem as funções estabelecidas. Nós observamos que os indivíduos que apresentaram o genótipo variante 126892TT estiveram sob risco do CEC de BL cerca de quatro vezes maior que os portadores do alelo selvagem. Em contraste, os indivíduos que apresentaram o genótipo selvagem 127267CC estiveram sob risco três vezes maior de ocorrência do tumor que os demais. A análise funcional e da formação de um possível haplótipo é necessária para definir os papéis destes dois SNPs no gene *LEF1*.

A proteína do complexo 6 de antígenos de linfócitos do locus K (Ly6K) é codificada pelo gene *LY6K*. A família de proteínas linfocitárias possui funções relacionadas com a sinalização e ou adesão celular (BAMEZAI & ROCK, 1995). A hiperexpressão desta proteína foi observada em linhagens celulares de tumores de cabeça e pescoço (DE NOOIJ-VAN DALEN et al., 2003) e bexiga (MATSUDA et al., 2011). Recentemente, a Ly6K foi descrita como uma possível proteína alvo para o desenvolvimento de novas terapias para o tratamento do câncer de esôfago e pulmão (SUDA et al., 2007; ISHIKAWA et al., 2011). De fato, uma vacina derivada da Ly6K foi testada em pesquisas clínicas com pacientes com CEC de esôfago (KONO et al., 2009; IWAHASHI et al., 2010). Em uma delas, a terapia com essa vacina induziu uma melhor resposta clínica dos pacientes ao tratamento convencional (KONO et al., 2009). Os resultados da pesquisa clínica de fase I sugeriram o aumento da imunidade inata nestes pacientes e a recomendação de pesquisas de fase II (IWAHASHI et al., 2010). O SNP C2893T localizado neste gene também não possui a sua função estabelecida. Entretanto, indivíduos com o genótipo selvagem (CC) estiveram sob risco de cerca de quatro vezes maior de ocorrência do CEC de BL quando comparados com àqueles com o alelo variante (T).

O gene *PTCH1* codifica uma proteína transmembrana da família *patched* (NAGAO et al. 2005). Ela é uma proteína receptora de sinalização das proteínas da família *hedgehog*. As proteínas dessa família controlam a proliferação celular por meio de ativação de genes que codificam as proteínas ciclinas D e E, e Myc, além de genes envolvidos no processo de angiogênese tumoral (SHIMOKAWA et al., 2004; LINDSTROM et al., 2006). Já é conhecido que mutações no gene *PTCH1* estão associadas com o surgimento de tumores de células basais da pele (de ZWAAN & HAASS, 2010), o que o caracterizou como um gene supressor de tumores (LINDSTROM et al., 2006; de ZWAAN & HAASS, 2010). Além da associação com tumores de pele, a alteração da expressão da *PTCH1* também foi associada com o surgimento de tumores de nasofaringe (XIAO et al., 2010), da tireóide (RHODEN et al., 2004), de mama (WOLF et al., 2007; SINHA et al., 2008), de ovário (CRETNIK et al., 2007), gástrico (MA et al., 2005; DU et al., 2009), pancreático (SHAO et al., 2006; QUINT et al., 2009), hepatocelular (FU et al., 2008) e coloretal (BIAN et al., 2007; STEFANIUS et al., 2011). Além disso, o SNP C3944T (rs357564) localizado na região codificadora de aminoácidos do *PTCH1* foi associado com o aumento do risco de câncer de mama em mulheres que utilizam contraceptivos orais (CHANG-CLAUDE et al., 2003). Outro

polimorfismo gênico, caracterizado por oito repetições dos nucleotídeos CGG na região 5' não traduzida, também foi associado com o aumento do risco de ameloblastomas (KAWABATA et al., 2005). Recentemente, a hiperexpressão do PTCH1 foi relacionada com a resistência à radioterapia e quimioterapia em pacientes com CEC de esôfago (ZHU et al., 2011). Os SNPs C79755T (rs16909856) e C794567T (rs16909859) localizados na região *downstream* desse gene não possuem as funções estabelecidas. Nós observamos que os indivíduos que apresentaram o genótipo selvagem 79755CC estiveram sob risco do CEC de BL cerca de cinco vezes maior que àqueles com o alelo variante. Em contraste, os indivíduos que apresentaram o genótipo selvagem 79456CC também estiveram sob risco cinco vezes maior de ocorrência do tumor. A análise da formação de um possível haplótipo também é necessária para definir os papéis destes dois SNPs na origem do CEC de BL e na resposta dos pacientes com os distintos genótipos à radioterapia e quimioterapia.

Adicionalmente, será importante a realização das análises destes SNPs em relação aos dados clínicos dos pacientes e características do tumor. Com isso, será possível determinar os papéis destes SNPs em grupos distintos de pacientes, além da influência dos mesmos na agressividade do tumor. Além disso, as combinações dos genótipos dos diferentes SNPs também deverão ser realizadas para definir o papel dessas possíveis combinações no surgimento do CEC de BL.

Infelizmente, não foi possível avaliar e selecionar os SNPs por meio da GELE coletiva. Os parâmetros de controle de qualidade para a inserção dos dados nos programas de análise não foram suficientes. A eficácia da GELE com mistura de DNA de indivíduos distintos não está totalmente estabelecida (CHIANG et al., 2010) e assim, novos estudos devem ser realizados.

Os algoritmos de análise da GELE de CNVs utilizados em nosso estudo foi o cbs e o hmm, disponíveis no programa computacional Partek<sup>®</sup>, cujo código fonte pertence à empresa com o mesmo nome. Por meio desses algoritmos, foi possível determinar o número de cópias de cada segmento de DNA localizado nos microarranjos de DNA das lâminas estudadas. Em breve, nós também analisaremos as CNVs das lâminas de microarranjos de DNA com o programa Bioconductor para confirmar os resultados deste estudo.

Por meio da GELE individual, nós identificamos 1.017 CNVs relacionadas a ganhos de DNA genômico de pacientes com CEC de BL quando comparados com as amostras de

DNA genômico dos controles. Entre essas regiões que apresentaram ganhos de regiões cromossômicas, nós selecionamos as regiões do braço curto do cromossomo 8 (p11 e p23) e as regiões do braço longo do cromossomo 22 (q11.23) para futuro processo de validação por meio da PCR em tempo real com sondas específicas para cada região de interesse. Essas regiões foram descritas previamente relacionadas com o surgimento de tumores de cabeça e pescoço (YE et al., 2007; CHEN & CHEN et al., 2008).

As proteínas da família de metaloprotease (ADAMs) são proteínas transmembranas que estão relacionadas com o crescimento tumoral e metástase em diversos tipos de tumores, incluindo os de cabeça e pescoço (DEMIRCAN et al., 2009; ECKSCHLAGER et al., 2009). A frequência da duplicação da região 8p11 foi maior em pacientes com CEC de BL do que em controles. Os genes *ADAM3A* e *ADAM5P* estão localizados nesta região. Entretanto, não há descrições da relação das proteínas codificadas por esses genes com o CEC de BL. Recentemente, a deleção do gene *ADAM3A* foi relacionada com tumores do sistema nervoso central (BARROW et al., 2011). No entanto, os papéis de ganhos e perdas desta região cromossômica com o surgimento e agressividade de tumores não estão definidos.

A frequência da duplicação da região 8p23 também foi maior nos pacientes quando comparados aos controles. Os genes de interesse *DEFB* e *FAM90A* estão localizados nesta região. A expressão da proteína humana defensina  $\beta$ -3 (hBD-3), codificada pelo gene *DEFB*, é encontrada em tecidos saudáveis da cavidade oral e orofaringe e estão relacionadas com processos inflamatórios localizados nessas regiões (DUNSCHE et al., 2002; YOSHIMOTO et al., 2003). Entretanto, a expressão desta proteína não está estabelecida em CEC de cavidade oral (JOLY et al., 2009; KESTING et al., 2009). Recentemente, Mburu et al. (2011) observaram que a hBD-3 inibe a apoptose em linhagem celular de CEC de cabeça e pescoço tratadas com cisplatina. Frente ao exposto, é provável que a duplicação deste gene possa influenciar no mecanismo de morte celular e resistência ao tratamento quimioterápico em indivíduos com tumores de BL. As proteínas da família FAM também estão relacionadas com o controle do mecanismo da apoptose (SCHWICKART et al., 2010; SUN et al., 2011). No entanto, não há descrições da proteína FAM do tipo 90A, codificada pelo gene *FAM90A*, com mecanismos da carcinogênese. Assim, estudos funcionais dessas proteínas são necessários para avaliar o seu papel no mecanismo de morte celular.

Outras duplicações que nos parecem de interesse são as genes *GSTT1*, *GSTT2*, *GSTTP1* e *DDT*, localizados na região 22q11.23. As frequências dessas duplicações foram maiores em pacientes com CEC de BL do que em controles. Já é amplamente conhecido a importância das enzimas da glutathione S-transferase (GSTs), codificada pelos genes *GSTT1*, *GSTT2* e *GSTTP1*, na detoxificação de carcinógenos e risco de tumores, incluindo o CEC de BL. Diversos estudos relacionaram a ausência desses genes com o aumento do risco para o CEC de orofaringe (LACKO et al., 2009). Surpreendentemente, nós observamos um maior risco de ocorrência do CEC de BL naqueles indivíduos que apresentaram duplicações nesses genes. Entretanto, é necessário avaliar, além da quantidade, a qualidade das proteínas codificadas pelos genes distintos. A proteína D-DT, codificada pelo gene *DDT*, está relacionada com o aumento da angiogênese em linhagens de células de carcinoma de pulmão (COLEMAN et al., 2008) e do adenocarcinoma de cólon (XIN et al., 2010). Entretanto, o seu papel na origem do CEC de BL não está estabelecido. Frente ao exposto, a duplicação deste gene pode influenciar a angiogênese em tumores de BL. Entretanto, estudos funcionais são necessários para comprovar essa hipótese.

Por meio da GELE individual, nós também identificamos 930 CNVs relacionadas a perdas de DNA genômico de pacientes com CEC de BL quando comparados com as amostras de DNA genômico dos controles. Entre essas regiões que apresentaram perdas de regiões cromossômicas, nós selecionamos as regiões do braços longos dos cromossomos 14 (q11.2) e 15 (q11.2) para futuro processo de validação desses resultados.

A frequência da perda da região 14q11.2 foi maior em pacientes com CEC de BL do que em controles. Os genes dos receptores de olfato (*OR4K1*, *OR4K2*, *OR4K5*, *OR4N2* e *OR4Q3*) estão localizados nesta região. Interessantemente, a proteína OR do tipo 51E2, codificada pelo gene *OR51E2*, foi relacionada com a diminuição da proliferação de células tumorais da próstata (NEUHAUS et al., 2009). No entanto, estudos funcionais com essas proteínas são necessários para definir o real papel das mesmas no surgimento de tumores.

Já é conhecido o papel da proteína Bcl-2 na regulação da apoptose em tumores de cabeça e pescoço (DOS SANTOS & CARVALHO, 2011). Nós observamos que a frequência da deleção da região 15q11.2 foi maior nos nossos pacientes quando comparados com os controles. O gene *BCL8* está localizado nesta região e o seu papel na regulação da apoptose ainda é desconhecido. Em adição, a proteína Bcl-8 foi relacionada, anteriormente, com o surgimento de linfomas de células B (DYOMIN et al., 1997). O seu

papel no surgimento do CEC de BL deve ser avaliado.

Por meio da GELE coletiva, nós observamos que os genes *UGT2B17*, *ADAM3A* e *GSTT1* apresentaram CNVs relacionadas com ganho de DNA e também foram identificados anteriormente por meio da GELE individual. A proteína UDP-glucuronosil-transferase (UGT), codificada pelo gene *UGT2B17*, está relacionada com o metabolismo da nicotina presente nos cigarros (CHEN et al., 2010). Além disso, o papel da presença ou ausência deste gene não está estabelecido (GALLAGHER et al., 2007; OLSSON et al., 2008; MENARD et al., 2009). Os papéis dos genes *ADAM3A* e *GSTT1*, selecionados pela GELE individual, já foram discutidos anteriormente.

Também por meio da GELE coletiva, nós observamos que os genes *MGAM* e *LCE3B* apresentaram CNVs relacionadas com a perda de DNA e também foram identificados anteriormente por meio da GELE individual. A proteína maltase glucoamilase humana é responsável pelo processo de digestão de alimentos e foi relacionada com o surgimento de tumores gástricos (YANG, 2007; SIM et al., 2010). No entanto, não há relação da função do gene *LCE3B* com mecanismos de surgimento de tumores (DOCAMPO et al., 2011).

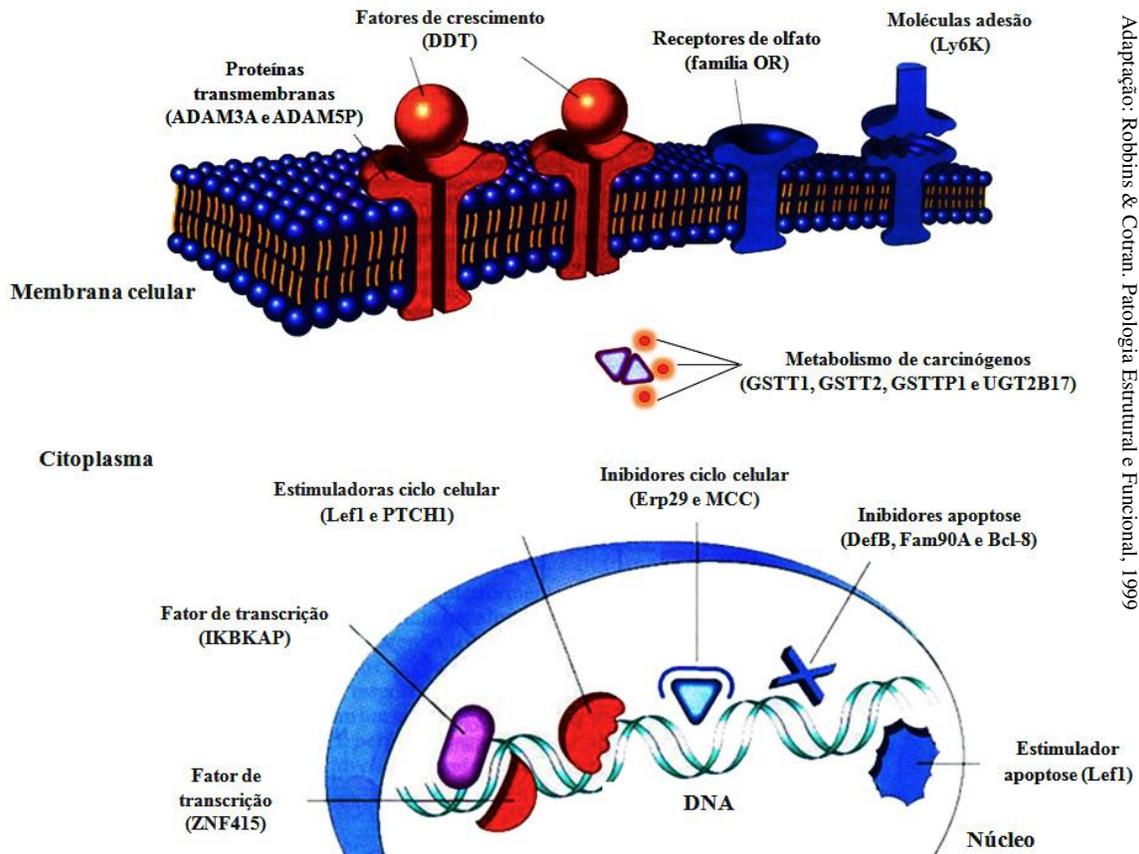
A eficácia da GELE coletiva em relação à GELE individual não foi ainda determinada (CRAIG et al., 2005; MEABURN et al., 2005; DOCHERTY et al., 2006; KIROV et al., 2006; MEABURN et al., 2006; CHIANG et al., 2010). Novos estudos são necessários para determinar a eficácia da GELE em mistura de amostras de DNA de uma população altamente miscigenada e heterogênea, como a população brasileira (ALVES-SILVA, 2000; CARVALHO-SILVA, 2001).

Adicionalmente, também será importante relacionar as CNVs selecionadas com os dados clínicos dos pacientes e as características do tumor. Além disso, a validação dos nossos resultados será realizada em breve pelos pesquisadores do nosso grupo por meio da PCR em tempo real, utilizando sondas específicas para cada gene.

Ainda, após a validação dos SNPs e CNVs selecionados em nossa própria casuística, será necessário avaliar um grupo maior de pacientes e de indivíduos saudáveis de regiões distintas do Brasil e de outros continentes para estabelecer o papel desses polimorfismos genéticos no CEC de BL.

De maneira geral, nós identificamos SNPs localizados em genes envolvidos em vias metabólicas de controle do processo transcricional de genes (*IKBKAP* e *ZNF415*) e do ciclo

celular (*ERP29*, *MCC*, *LEF1*, *LY6K* e *PTCH1*) de células tumorais, e CNVs localizadas em genes envolvidos com o metabolismo de carcinógenos (*GSTT1*, *GSTT2*, *GSTTP1* e *UGT2B17*), com o processo de angiogênese e metástase (*ADAM3A*, *ADAM5P* e *DDT*) e apoptose (*BCL8* e *DEFB*). Estes resultados, em seu conjunto, certamente contribuem para o entendimento da fisiopatologia do CEC de BL e, ainda, podem funcionar como possíveis alvos para o desenvolvimento de terapias específicas para a doença (Figura 12).



**Figura 12.** Representação esquemática das funções conhecidas ou presumidas das proteínas possivelmente produzidas pelos genes polimórficos identificados e selecionados em 49 pacientes com carcinoma de células escamosas de base de língua e 49 controles por meio da genotipagem em larga escala. Os genes envolvidos em vias metabólicas de controle do processo transcrricional (*IKBKAP* e *ZNF415*), de estimulação (*LEF1* e *PTCH1*) e inibição (*ERP29* e *MCC*) do ciclo celular, de estimulação da apoptose (*LEF1*) e do mecanismo de adesão celular (*LY6K*), possuem polimorfismos gênicos de base única que podem influenciar nas referidas funções possivelmente por meio da alteração da quantidade ou da qualidade das proteínas produzidas pelos alelos distintos. A variação no número de cópias de DNA de genes envolvidos com o metabolismo de carcinógenos (*GSTT1*, *GSTT2*, *GSTTP1* e *UGT2B17*), com fatores de crescimento celular (*ADAM3A*, *ADAM5P*, *ORs* e *DDT*), e com a inibição da apoptose (*BCL8*, *FAM90A* e *DEFB*) podem alterar a quantidade ou a qualidade das proteínas produzidas pelos genes



# CONCLUSÕES

---



- O método de GELE individual possibilitou a identificação de SNPs localizados em genes envolvidos com o controle do processo transcricional (*IKBKAP* e *ZNF415*) e do ciclo celular (*ERP29*, *MCC*, *LEF1*, *LY6K* e *PTCH1*), e de CNVs em genes envolvidos no metabolismo de carcinógenos (*GSTT1*, *GSTT2*, *GSTTP1*, *UGT2B17*), no processo de angiogênese e metástase (*ADAM3A*, *ADAM5P* e *DDT*), e apoptose (*BCL8* e *DEFB*), que podem predispor indivíduos para o CEC de BL;
- O método de GELE individual foi mais eficaz que o método de GELE coletiva para identificação de SNPs e CNVs em nossas amostras, possivelmente pela heterogeneidade racial da nossa população;
- Em seu conjunto, esses resultados contribuem para a compreensão da fisiopatologia do CEC de BL, bem como para o possível desenvolvimento de novas terapias alvo-específicas para a doença.



## **REFERÊNCIAS BIBLIOGRÁFICAS**

---



Abbas A, Delvinqui re K, Lechevrel M, Lebailey P, Gauduchon P, Launoy G, et al. GSTM1, GSTT1, GSTP1 and CYP1A1 genetic polymorphisms and susceptibility to esophageal cancer in a french population: different pattern of squamous cell carcinoma and adenocarcinoma. *World J Gastroenterol.* 2004;10(23):3389-93.

Alves-Silva J, Silva-Santos M, Guimaraes PE. The ancestry of Brazilian mtDNA lineages. *Am J Hum Genet.* 2000;67:444-61.

Assie G, Laframboise T, Platzer P, Eng C. Frequency of germline genomic homozygosity associated with cancer cases. *JAMA.* 2008;299(12):1437-45.

Bacolod MD, Schemmann GS, Giardina SF, Paty P, Notterman DA, Barany F. Emerging paradigms in cancer genetics: some important findings from high-density single nucleotide polymorphism array studies. *Cancer Res.* 2009;69(3):723-7.

Bambang IF, Xu S, Zhou J, Salto-Tellez M, Sethi SK, Zhang D. Overexpression of endoplasmic reticulum protein 29 regulates mesenchymal-epithelial transition and suppresses xenograft tumor growth of invasive breast cancer cells. *Lab Invest.* 2009;89(11):1229-42.

Bamezai A, Rock KL. Overexpressed Ly-6A.2 mediates cell-cell adhesion by binding a ligand expressed on lymphoid cells. *Proc Natl Acad Sci U S A.* 1995;92(10):4294-8.

Barrow J, Adamowicz-Brice M, Cartmill M, MacArthur D, Lowe J, Robson K, et al. Homozygous loss of ADAM3A revealed by genome-wide analysis of pediatric high-grade glioma and diffuse intrinsic pontine gliomas. *Neuro Oncol.* 2011;13(2):212-22.

Bei JX, Li Y, Jia WH, Feng BJ, Zhou G, Chen LZ, et al. A genome-wide association study of nasopharyngeal carcinoma identifies three new susceptibility loci. *Nat Genet.* 2010;42(7):599-603.

Bian YH, Huang SH, Yang L, Ma XL, Xie JW, Zhang HW. Sonic hedgehog-Gli1 pathway in colorectal adenocarcinomas. *World J Gastroenterol*. 2007;13(11):1659-65.

Brennan P, Lewis S, Hashibe M, Bell DA, Boffetta P, Bouchardy C, et al. Pooled analysis of alcohol dehydrogenase genotypes and head and neck cancer: a huGE review. *Am J Epidemiol*. 2004;159(1):1-16.

Carter N.P. Methods and strategies for analyzing copy number variation using DNA microarrays. *Nat Genet*. 2007;39(7): S16-21.

Carvalho B, Bengtsson H, Speed TP, Irizarry RA. Exploration, normalization, and genotype calls of high-density oligonucleotide SNP array data. *Biostatistics*. 2007;8(2):485-99.

Carvalho BS, Irizarry RA. A framework for oligonucleotide microarray preprocessing. *Bioinformatics*. 2010;26(19):2363-7.

Carvalho BS, Louis TA, Irizarry RA. Quantifying uncertainty in genotype calls. *Bioinformatics*. 2010;26(2):242-9.

Carvalho-Silva DR, Santos FR, Rocha J, Pena SD. The phylogeography of Brazilian Y-chromosome lineages. *Am J Hum Genet*. 2001;68:281-6.

Chang-Claude J, Dunning A, Schnitzbauer U, Galmbacher P, Tee L, Wjst M, et al. The patched polymorphism Pro1315Leu (C3944T) may modulate the association between use of oral contraceptives and breast cancer risk. *Int J Cancer*. 2003;103(6):779-83.

Chen G, Giambone NE Jr, Dluzen DF, Muscat JE, Berg A, Gallagher CJ, et al. Glucuronidation genotypes and nicotine metabolic phenotypes: importance of functional UGT2B10 and UGT2B17 polymorphisms. *Cancer Res*. 2010;70(19):7543-52.

Chen JM, Férec C, Cooper DN. A systematic analysis of disease-associated variants in the 3' regulatory regions of human protein-coding genes II: the importance of mRNA secondary structure in assessing the functionality of 3' UTR variants. *Hum Genet.* 2006;120(3):301-33.

Chen K, Zhao H, Hu Z, Wang LE, Zhang W, Sturgis EM, et al. CASP3 polymorphisms and risk of squamous cell carcinoma of the head and neck. *Clin Cancer Res.* 2008;14(19):6343-9.

Chen MK, Chiou HL, Su SC, Chung TT, Tseng HC, Tsai HT, et al. The association between hypoxia inducible factor-1alpha gene polymorphisms and increased susceptibility to oral cancer. *Oral Oncol.* 2009;45(12):e222-6.

Chen Y., Chen C. DNA copy number variation and loss of heterozygosity in relation to recurrence of and survival from head and neck squamous cell carcinoma: a review. *Head Neck.* 2008;30(10):1361-83.

Cheng Y, Wang Y, Li Y, Deng Y, Hu J, Mo X, et al. A novel human gene ZNF415 with five isoforms inhibits AP-1- and p53-mediated transcriptional activity. *Biochem Biophys Res Commun.* 2006;351(1):33-9.

Chiang CW, Gajdos ZK, Korn JM, Kuruvilla FG, Butler JL, Hackett R, et al. Rapid assessment of genetic ancestry in populations of unknown origin by genome-wide genotyping of pooled samples. *PLoS Genet.* 2010;6(3):e1000866.

Chung CC, Chanock SJ. Current status of genome-wide association studies in cancer. *Hum Genet.* 2011;130(1):59-78.

Cohan DM, Popat S, Kaplan SE, Rigual N, Loree T, Hicks WL Jr. Oropharyngeal cancer: current understanding and management. *Curr Opin Otolaryngol Head Neck Surg.* 2009;17(2):88-94.

Coleman AM, Rendon BE, Zhao M, Qian MW, Bucala R, Xin D, et al. Cooperative regulation of non-small cell lung carcinoma angiogenic potential by macrophage migration inhibitory factor and its homolog, D-dopachrome tautomerase. *J Immunol.* 2008;181(4):2330-7.

Conacci-Sorrell ME, Ben-Yedidia T, Shtutman M, Feinstein E, Einat P, Ben-Ze'ev A. Nr-CAM is a target gene of the beta-catenin/LEF-1 pathway in melanoma and colon cancer and its expression enhances motility and confers tumorigenesis. *Genes Dev.* 2002;16(16):2058-72.

Costa LLA, Júnior FAR, Ramos CFC. Correlação entre a classificação clínica TNM e as características histológicas de malignidade do carcinoma epidermóide oral. *Rev Bras Otorrinolaringol.* 2005;71(2):181-7.

Craig DW, Huentelman MJ, Hu-Lince D, Zismann VL, Kruer MC, Lee AM, et al. Identification of disease causing loci using an array-based genotyping approach on pooled DNA. *BMC Genomics.* 2005;30(6):138.

Cretnik M, Musani V, Oreskovic S, Leovic D, Levanat S. The Patched gene is epigenetically regulated in ovarian dermoids and fibromas, but not in basocellular carcinomas. *Int J Mol Med.* 2007;19(6):875-83.

Cui Y, Morgenstern H, Greenland S, Tashkin DP, Mao J, Cao W, et al. Polymorphism of Xeroderma Pigmentosum group G and the risk of lung cancer and squamous cell carcinomas of the oropharynx, larynx and esophagus. *Int J Cancer.* 2006;118(3):714-20.

De Nooij-van Dalen AG, van Dongen GA, Smeets SJ, Nieuwenhuis EJ, Stigter-van Walsum M, et al. Characterization of the human Ly-6 antigens, the newly annotated member Ly-6K included, as molecular markers for head-and-neck squamous cell carcinoma. *Int J Cancer.* 2003;103(6):768-74.

De Zwaan SE, Haass NK. Genetics of basal cell carcinoma. *Australas J Dermatol*. 2010;51(2):81-92.

Demircan K, Gunduz E, Gunduz M, Beder LB, Hirohata S, Nagatsuka H, et al. Increased mRNA expression of ADAMTS metalloproteinases in metastatic foci of head and neck cancer. *Head Neck*. 2009;31(6):793-801.

Diergaarde B, Brand R, Lamb J, Cheong SY, Stello K, Barmada MM, et al. Pooling-based genome-wide association study implicates gamma-glutamyltransferase 1 (GGT1) gene in pancreatic carcinogenesis. *Pancreatology*. 2010;10(2-3):194-200.

Docampo E, Giardina E, Riveira-Muñoz E, de Cid R, Escaramís G, Perricone C, et al. Deletion of LCE3C and LCE3B is a susceptibility factor for psoriatic arthritis: a study in Spanish and Italian populations and meta-analysis. *Arthritis Rheum*. 2011;63(7):1860-5.

Docherty SJ, Butcher LM, Schalkwyk LC, Plomin R. Applicability of DNA pools on 500K SNP microarrays for cost-effective initial screens in genome-wide association studies. *BMC Genomics*. 2007;8:214.

Dos Santos LV, Carvalho AL. Bcl-2 targeted-therapy for the treatment of head and neck squamous cell carcinoma. *Recent Pat Anticancer Drug Discov*. 2011;6(1):45-57.

Du P, Ye HR, Gao J, Chen W, Wang ZC, Jiang HH, et al. Methylation of PTCH1a gene in a subset of gastric cancers. *World J Gastroenterol*. 2009;15(30):3799-806.

Duncan BB, Schmidt MI, Polanczyk CA. Altos coeficientes de mortalidade em populações adultas brasileiras: uma comparação internacional. *Rev Ass Med Brasil*. 1992;38:138-44.

Dunsche A, Açil Y, Dommisch H, Siebert R, Schröder JM, Jepsen S. The novel human beta-defensin-3 is widely expressed in oral tissues. *Eur J Oral Sci*. 2002;110(2):121-4.

Durazzo MD, Araújo CEN, Brandão Neto JS, Potenza AS, Costa P, Takeda F, et al. Clinical and epidemiological features of oral cancer in a medical school teaching hospital from 1994 to 2002: incidence in women, predominance of advanced local disease, and low incidence of neck metastase. *Clinics*. 2005;60 (4):293-8.

Dutt A, Beroukhim R. Single nucleotide polymorphism array analysis of cancer. *Curr Opin Oncol*. 2007;19(1):43-9.

Dyomin VG, Rao PH, Dalla-Favera R, Chaganti RS. BCL8, a novel gene involved in translocations affecting band 15q11-13 in diffuse large-cell lymphoma. *Proc Natl Acad Sci U S A*. 1997;94(11):5728-32.

Eckschlager T, Adam V, Hrabeta J, Figova K, Kizek R. Metallothioneins and cancer. *Curr Protein Pept Sci*. 2009;10(4):360-75.

Egeblad M, Rasch MG, Weaver VM. Dynamic interplay between the collagen scaffold and tumor evolution. *Curr Opin Cell Biol*. 2010;22(5):697-706.

Freedman AN, Michalek AM, Marshall JR, Mettlin CJ, Petrelli NJ, Zhang ZF, et al. The relationship between smoking exposure and p53 overexpression in colorectal cancer. *Br J Cancer*. 1996;73:902-8.

Fu X, Wang Q, Chen X, Huang X, Cao L, Tan H, et al. Expression patterns and polymorphisms of PTCH in Chinese hepatocellular carcinoma patients. *Exp Mol Pathol*. 2008;84(3):195-9.

Fukuyama R, Nicolaita R, Ng KP, Obusez E, Sanchez J, Kalady M, et al. Mutated in colorectal cancer, a putative tumor suppressor for serrated colorectal cancer, selectively represses beta-catenin-dependent transcription. *Oncogene*. 2008;27(46):6044-55.

Furness S, Glenny AM, Worthington HV, Pavitt S, Oliver R, Clarkson JE, et al. Interventions for the treatment of oral cavity and oropharyngeal cancer: chemotherapy. *Cochrane Database Syst Rev.* 2010;(4):CD006386.

Gallagher CJ, Muscat JE, Hicks AN, Zheng Y, Dyer AM, Chase GA, et al. The UDP-glucuronosyltransferase 2B17 gene deletion polymorphism: sex-specific association with urinary 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanol glucuronidation phenotype and risk for lung cancer. *Cancer Epidemiol Biomarkers Prev.* 2007;16(4):823-8.

Garavello W, Lucenteforte E, Bosetti C, La Vecchia C. The role of foods and nutrients on oral and pharyngeal cancer risk. *Minerva Stomatol.* 2009;58(1-2):25-34.

Gold KA, Kim ES. Role of molecular markers and gene profiling in head and neck cancers. *Curr Opin Oncol.* 2009;21(3):206-11.

Gouveia GR, Bruno LPC. *Manual de contagem de carboidratos*, 2001.

Gourin CG, Podolsky RH. Racial disparities in patients with head and neck squamous cell carcinoma. *Laryngoscope.* 2006;116(7):1093-106.

Greene FL, Page D, Morrow M, Balch C, Haller D, Fritz A, et al. *AJCC Cancer Staging Manual 6th*. New York: Springer; 2002.

Guan X, Sturgis EM, Lei D, Liu Z, Dahlstrom KR, Wei Q, et al. Association of TGF-beta1 genetic variants with HPV16-positive oropharyngeal cancer. *Clin Cancer Res.* 2010;16(5):1416-22.

Hayes JD, Flanagan JU, Jowsey IR. Glutathione Transferases. *Annu Rev Pharmacol Toxicol.* 2005; 45:51-88.

Hiyama T, Yoshihara M, Tanaka S, Chayama K. Genetic polymorphisms and head and neck cancer risk (Review). *Int J Oncol.* 2008;32(5):945-73.

Hollams EM, Giles KM, Thomson AM, Leedman PJ. mRNA stability and the control of gene expression: implications for human disease. *Neurochem Res.* 2002;27(10):957-80.

Holmberg C, Katz S, Lerdrup M, Herdegen T, Jäättelä M, Aronheim A, et al. A novel specific role for I kappa B kinase complex-associated protein in cytosolic stress signaling. *J Biol Chem.* 2002;277(35):31918-28.

Holmes SF, Erie DA. Downstream DNA sequence effects on transcription elongation. Allosteric binding of nucleoside triphosphates facilitates translocation via a ratchet motion. *J Biol Chem.* 2003;278(37):35597-608.

Holnthoner W, Pillinger M, Groger M, Wolff K, Ashton AW, Albanese C, et al. Fibroblast growth factor-2 induces Lef/Tcf-dependent transcription in human endothelial cells. *J Biol Chem.* 2002;277(48):45847-53.

Huang WY, Winn DM. Alcohol concentration and risk of oral cancer in Puerto Rico. *American J Epidemiol.* 2003;157(10):881-7.

Instituto Nacional do Câncer (INCA). Estimativas de câncer 2010. Ministério da Saúde, Brasil, 2011. Disponível em: <<http://www.inca.gov.br>> Acesso: 7 set 2011.

International Union Against Cancer (UICC). TNM classification of malignant tumors. Instituto Nacional do Câncer (INCA), 6.ed., Rio de Janeiro, RJ, 2004.

Ishikawa N, Takano A, Yasui W, Inai K, Nishimura H, Ito H, et al. Cancer-testis antigen lymphocyte antigen 6 complex locus K is a serologic biomarker and a therapeutic target for lung and esophageal carcinomas. *Cancer Res.* 2007;67(24):11601-11.

Iwahashi M, Katsuda M, Nakamori M, Nakamura M, Naka T, Ojima T, et al. Vaccination with peptides derived from cancer-testis antigens in combination with CpG-7909 elicits strong specific CD8+ T cell response in patients with metastatic esophageal squamous cell carcinoma. *Cancer Sci.* 2010;101(12):2510-7.

Iwai S, Yonekawa A, Harada C, Hamada M, Katagiri W, Nakazawa M, et al. Involvement of the Wnt- $\beta$ -catenin pathway in invasion and migration of oral squamous carcinoma cells. *Int J Oncol.* 2010;37(5):1095-103.

Jemal A, Siegel R, Ward E, Murray T, Xu J, Thun MJ. Cancer statistics, 2007. *CA Cancer J Clin.* 2007;57:43-66.

Ji X, Neumann AS, Sturgis EM, Adler-Storthz K, Dahlstrom KR, Schiller JT, et al. p53 codon 72 polymorphism associated with risk of human papillomavirus-associated squamous cell carcinoma of the oropharynx in never-smokers. *Carcinogenesis.* 2008;29(4):875-9.

Joly S, Compton LM, Pujol C, Kurago ZB, Guthmiller JM. Loss of human beta-defensin 1, 2, and 3 expression in oral squamous cell carcinoma. *Oral Microbiol Immunol.* 2009;24(5):353-60.

Kamangar F, Dores GM, Anderson WF. Patterns of cancer incidence, mortality, and prevalence across five continents: defining priorities to reduce cancer disparities in different geographic regions of the world. *J Clin Oncol.* 2006;24:2137-50.

Kawabata T, Takahashi K, Sugai M, Murashima-Suginami A, Ando S, Shimizu A, et al. Polymorphisms in PTCH1 affect the risk of ameloblastoma. *J Dent Res.* 2005;84(9):812-6.

Kesting MR, Loeffelbein DJ, Hasler RJ, Wolff KD, Rittig A, Schulte M, et al. Expression profile of human beta-defensin 3 in oral squamous cell carcinoma. *Cancer Invest.* 2009;27(5):575-81.

Kim JH, Lane WS, Reinberg D. Human Elongator facilitates RNA polymerase II transcription through chromatin. *Proc Natl Acad Sci U S A*. 2002;99(3):1241-6.

Kirov G, Nikolov I, Georgieva L, Moskina V, Owen MJ, Donovan MCO. Pooled DNA genotyping on Affymetrix SNP genotyping arrays. *BMC Genomics*. 2006;7(27):1-10.

Koch M, Schulze J, Hansen U, Ashwodt T, Keene DR, Brunken WJ, et al. A novel marker of tissue junctions, collagen XXII. *J Biol Chem*. 2004 279(21):22514-21.

Kono K, Mizukami Y, Daigo Y, Takano A, Masuda K, Yoshida K, et al. Vaccination with multiple peptides derived from novel cancer-testis antigens can induce specific T-cell responses and clinical responses in advanced esophageal cancer. *Cancer Sci*. 2009;100(8):1502-9.

Kowalski LP, Carvalho AL, Martins Priante AV, Magrin J. Predictive factors for distant metastasis from oral and oropharyngeal squamous cell carcinoma. *Oral Oncol*. 2005;41(5):534-41.

Kowalski LP. *Manual de Condutas Diagnósticas e Terapêuticas em Oncologia*. AC Camargo, 3. ed. Âmbito. 2006.

Lacko M, Roelofs HM, te Morsche RH, Voogd AC, Ophuis MB, Peters WH, et al. Genetic polymorphisms in the tobacco smoke carcinogens detoxifying enzyme UGT1A7 and the risk of head and neck cancer. *Head Neck*. 2009;31(10):1274-81.

Lambert R, Sauvaget C, de Camargo Cancela M, Sankaranarayanan R. Epidemiology of cancer from the oral cavity and oropharynx. *Eur J Gastroenterol Hepatol*. 2011;(8):633-41.

Lazarus P, Park JY. Metabolizing enzyme genotype and risk for upper aerodigestive tract cancer. *Oral Oncol*. 2000;36:421-31.

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

Li Q, Fazly AM, Zhou H, Huang S, Zhang Z, Stillman B. The elongator complex interacts with PCNA and modulates transcriptional silencing and sensitivity to DNA damage agents. *PLoS Genet*. 2009;5(10):e1000684.

Li TW, Ting JH, Yokoyama NN, Bernstein A, van de Wetering M, Waterman ML. Wnt activation and alternative promoter repression of LEF1 in colon cancer. *Mol Cell Biol*. 2006;26(14):5284-99.

Li Y, Wang L, Zhang M, Melamed J, Liu X, Reiter R, et al. LEF1 in androgen-independent prostate cancer: regulation of androgen receptor expression, prostate cancer growth, and invasion. *Cancer Res*. 2009;69(8):3332-8.

Licitra L, Bernier J, Grandi C, Merlano M, Bruzzi P, Lefebvre JL. Cancer of the oropharynx. *Crit Rev Oncol Hematol*. 2002;41(1):107-22.

Lin AY, Chua MS, Choi YL, Yeh W, Kim YH, Azzi R, et al. Comparative profiling of primary colorectal carcinomas and liver metastases identifies LEF1 as a prognostic biomarker. *PLoS One*. 2011;24;6(2):e16636.

Lin CH, Huang MC, Li LH, Wu JY, Chen YT, Fann CS. Genome-wide copy number analysis using copy number inferring tool (CNIT) and DNA pooling. *Hum Mutat*. 2008;29(8):1055-62.

Lindström E, Shimokawa T, Toftgård R, Zaphiropoulos PG. PTCH mutations: distribution and analyses. *Hum Mutat*. 2006;27(3):215-9.

Lopes MA, Nikitakis NG, Reynolds MA, Ord RA, Sauk J. Biomarkers predictive of lymphonode metastases in oral squamous cell carcinoma. *J Oral Maxill Surg*. 2002;60:142-7.

Lourenço GJ, Silva EF, Rinck-Junior JA, Chone CT, Lima CS. CYP1A1, GSTM1 and GSTT1 polymorphisms, tobacco and alcohol status and risk of head and neck squamous cell carcinoma. *Tumor Biol* 2011; 32(6):1209-15.

Lucenteforte E, Garavello W, Bosetti C, La Vecchia C. Dietary factors and oral and pharyngeal cancer risk. *Oral Oncol*. 2009;45(6):461-7.

Ma X, Chen K, Huang S, Zhang X, Adegboyega PA, Evers BM, et al. Frequent activation of the hedgehog pathway in advanced gastric adenocarcinomas. *Carcinogenesis*. 2005;26(10):1698-705.

Macgregor S, Zhao ZZ, Henders A, Nicholas MG, Montgomery GW, Visscher PM. Highly cost-efficient genome-wide association studies using DNA pools and dense SNP arrays. *Nucleic Acids Res*. 2008;36(6):e35.

Mao X, Young BD, Lu YJ. The application of single nucleotide polymorphism microarrays in cancer research. *Curr Genomics*. 2007;8(4):219-28.

Marcus KJ, Tishler RB. Head and neck carcinomas across the age spectrum: epidemiology, therapy, and late effects. *Semin Radiat Oncol*. 2010;20(1):52-7.

Matsuda R, Enokida H, Chiyomaru T, Kikkawa N, Sugimoto T, Kawakami K, et al. LY6K is a novel molecular target in bladder cancer on basis of integrate genome-wide profiling. *Br J Cancer*. 2011;104(2):376-86.

Mburu YK, Abe K, Ferris LK, Sarkar SN, Ferris RL. Human  $\beta$ -defensin 3 promotes NF- $\kappa$ B-mediated CCR7 expression and anti-apoptotic signals in squamous cell carcinoma of the head and neck. *Carcinogenesis*. 2011;32(2):168-74.

Meaburn E, Butcher LM, Liu L, Fernandes C, Hansen V, Al-Chalabi A, et al. Genotyping DNA pools on microarrays: Tackling the QTL problem of large samples and large numbers of SNPs. *BMC Genomics*. 2005;6(52):1-8.

Meaburn E, Butcher LM, Schalkwyk LC, Plomin R. Genotyping pooled DNA using 100K SNP microarrays: a step towards genomewide association scans. *Nucl Ac Res*. 2006;34(4):1-8.

Menard V, Eap O, Harvey M, Guillemette C, Lévesque E. Copy-number variations (CNVs) of the human sex steroid metabolizing genes UGT2B17 and UGT2B28 and their associations with a UGT2B15 functional polymorphism. *Hum Mutat*. 2009;30(9):1310-9.

Mendenhall WM, Riggs-Jr CE, Cassisi NJ. Cancer of the head and neck: treatment of head and neck cancers. In: DeVita-Jr VT, Hellman S, Rosenberg SA. *Cancer Principles & Practice of Oncology*. 7.ed., Lippincott Williams & Wilks, 2005.

Molinolo AA, Amornphimoltham P, Squarize CH, Castilho RM, Patel V, Gutkind JS. Dysregulated molecular networks in head and neck carcinogenesis. *Oral Oncol*. 2009;45(4-5):324-34.

Mori S, Nose M, Morikawa H, Sato A, Saito T, Song ST, et al. A novel evaluation system of metastatic potential of oral squamous cell carcinoma according to the histopathological and histochemical grading. *Oral Oncol*. 1998;34(6):549-57.

Nagao K, Toyoda M, Takeuchi-Inoue K, Fujii K, Yamada M, Miyashita T. Identification and characterization of multiple isoforms of a murine and human tumor suppressor, patched, having distinct first exons. *Genomics*. 2005;85(4):462-71.

National Institute on Alcohol Abuse and Alcoholism (NIAAA) n° 30, PH 359, 1995. *Alcohol Res Health*. 2000;24(1):1-11.

Neuhaus EM, Zhang W, Gelis L, Deng Y, Noldus J, Hatt H. Activation of an olfactory receptor inhibits proliferation of prostate cancer cells. *J Biol Chem.* 2009;284(24):16218-25.

Ng CC, Yew PY, Pua SM, Krishnan G, Yap LF, Teo SH. A genome-wide association study identifies ITGA9 conferring risk of nasopharyngeal carcinoma. *J Hum Genet.* 2009;54(7):392-97.

Nguyen A, Rosner A, Milovanovic T, Hope C, Planutis K, Saha B, et al. Wnt pathway component LEF1 mediates tumor cell invasion and is expressed in human and murine breast cancers lacking ErbB2 (her-2/neu) overexpression. *Int J Oncol.* 2005;27(4):949-56.

Nguyen DX, Chiang AC, Zhang XH, Kim JY, Kris MG, Ladanyi M, et al. WNT/TCF signaling through LEF1 and HOXB9 mediates lung adenocarcinoma metastasis. *Cell.* 2009;138(1):51-62.

Nguyen NP, Chi A, Nguyen LM, Ly BH, Karlsson U, Vinh-Hung V. Human papillomavirus-associated oropharyngeal cancer: a new clinical entity. *QJM.* 2010;103(4):229-36.

Nishimoto IN, Pinheiro NA, Rogatto SR, Carvalho AL, de Moura RP, Caballero OL, et al. Alcohol dehydrogenase 3 genotype as a risk factor for upper aerodigestive tract cancers. *Arch Otolaryngol Head Neck Surg.* 2004;130(1):78-82.

Olsson M, Lindström S, Häggkvist B, Adami HO, Bälter K, Stattin P, et al. The UGT2B17 gene deletion is not associated with prostate cancer risk. *Prostate.* 2008;68(5):571-5.

Palangat M, Hittinger CT, Landick R. Downstream DNA selectively affects a paused conformation of human RNA polymerase II. *J Mol Biol.* 2004;341(2):429-42.

Paleri V, Wight RG, Silver CE, Haigentz M Jr, Takes RP, Bradley PJ, et al. Comorbidity in head and neck cancer: a critical appraisal and recommendations for practice. *Oral Oncol.* 2010;46(10):712-9.

Perrone F, Mariani L, Pastore E, Orsenigo M, Suardi S, Marcomini B, et al. p53 codon 72 polymorphisms in human papillomavirus-negative and human papillomavirus-positive squamous cell carcinomas of the oropharynx. *Cancer.* 2007;109(12):2461-5.

Psyrri A, Cohen E. Oropharyngeal cancer: clinical implications of the HPV connection. *Ann Oncol.* 2011;22(5):997-9.

Quint K, Stintzing S, Alinger B, Hauser-Kronberger C, Dietze O, Gahr S, et al. The expression pattern of PDX-1, SHH, Patched and Gli-1 is associated with pathological and clinical features in human pancreatic cancer. *Pancreatology.* 2009;9(1-2):116-26.

Rhoden KJ, Johnson C, Brandao G, Howe JG, Smith BR, Tallini G. Real-time quantitative RT-PCR identifies distinct c-RET, RET/PTC1 and RET/PTC3 expression patterns in papillary thyroid carcinoma. *Lab Invest.* 2004;84(12):1557-70.

Rischin D. Oropharyngeal cancer, human papilloma virus, and clinical trials. *J Clin Oncol.* 2010;28(1):1-3.

Ruiz MT, Biselli PM, Maniglia JV, Pavarino-Bertelli EC, Goloni-Bertollo EM. Genetic variability of vascular endothelial growth factor and prognosis of head and neck cancer in a Brazilian population. *Braz J Med Biol Res.* 2010;43(2):127-33.

Schwickart M, Huang X, Lill JR, Liu J, Ferrando R, French DM, et al. Deubiquitinase USP9X stabilizes MCL1 and promotes tumour cell survival. *Nature.* 2010;463(7277):103-7.

Secretaria Municipal de Saúde de São Paulo. Tabnet. Prefeitura da cidade de São Paulo, 2009. Disponível em: < <http://portal.prefeitura.sp.gov.br/secretarias/saude>> Acesso: 16 ago 2011.

Sessions DG, Lenox J, Spector GJ, Chao C, Chaudry OA. Analysis of treatment results for base of tongue cancer. *Laryngoscope*. 2003;113(7):1252-61.

Sham PC, Bader JS, Craig I, O'Donovan M, Owen M. DNA pooling: A tool for large-scale association studies. *Nat Rev Genet*. 2002;3:862-71.

Shao J, Zhang L, Gao J, Li Z, Chen Z. Aberrant expression of PTCH (patched gene) and Smo (smoothed gene) in human pancreatic cancerous tissues and its association with hyperglycemia. *Pancreas*. 2006;33(1):38-44.

Shimokawa T, Rahnama F, Zaphiropoulos PG. A novel first exon of the Patched1 gene is upregulated by Hedgehog signaling resulting in a protein with pathway inhibitory functions. *FEBS Lett*. 2004;578(1-2):157-62.

Sidransky D. Cancer of the head and neck: molecular biology of head and neck tumors. In: DeVita-JR VT, Hellman S, Rosenberg SA. *Cancer Principles & Practice of Oncology*. 7.ed., Lippincott Williams & Wilks, 2005.

Sim L, Willemsma C, Mohan S, Naim HY, Pinto BM, Rose DR. Structural basis for substrate selectivity in human maltase-glucoamylase and sucrase-isomaltase N-terminal domains. *J Biol Chem*. 2010;285(23):17763-70.

Sinha S, Singh RK, Alam N, Roy A, Roychoudhury S, Panda CK. Alterations in candidate genes PHF2, FANCC, PTCH1 and XPA at chromosomal 9q22.3 region: pathological significance in early- and late-onset breast carcinoma. *Mol Cancer*. 2008;7:84.

Spaulding C, Reschly EJ, Zagort DE, Yashiro-Ohtani Y, Beverly LJ, Capobianco A, et al. Notch1 co-opts lymphoid enhancer factor 1 for survival of murine T-cell lymphomas. *Blood*. 2007;110(7):2650-8.

Spinola M, Leoni VP, Galvan A, Korsching E, Conti B, Pastorino U, et al. Genome-wide single nucleotide polymorphism analysis of lung cancer risk detects the KLF6 gene. *Cancer Lett*. 2007;251(2):311-16.

Stefanius K, Kantola T, Tuomisto A, Vahteristo P, Karttunen TJ, Aaltonen LA, et al. Downregulation of the hedgehog receptor PTCH1 in colorectal serrated adenocarcinomas is not caused by PTCH1 mutations. *Virchows Arch*. 2011;458(2):213-9.

Streit S, Bange J, Fichtner A, Ihrler S, Issing W, Ullrich A. Involvement of the FGFR4 Arg388 allele in head and neck squamous cell carcinoma. *Int J Cancer*. 2004;111(2):213-7.

Suda T, Tsunoda T, Daigo Y, Nakamura Y, Tahara H. Identification of human leukocyte antigen-A24-restricted epitope peptides derived from gene products upregulated in lung and esophageal cancers as novel targets for immunotherapy. *Cancer Sci*. 2007 [Epub ahead of print].

Sun H, Kapuria V, Peterson LF, Fang D, Bornmann WG, Bartholomeusz G, et al. Bcr-Abl ubiquitination and Usp9x inhibition block kinase signaling and promote CML cell apoptosis. *Blood*. 2011;117(11):3151-62.

Sund M, Kalluri R. Tumor stroma derived biomarkers in cancer. *Cancer Metastasis Rev*. 2009 Jun;28(1-2):177-83.

Syvanen A.C. Accessing genetic variation: Genotyping single nucleotide polymorphisms. *Nat Rev Genet*. 2001;2:930-42.

Takeoka S, Unoki M, Onouchi Y, Doi S, Fujiwara H, Miyatake A, et al. Amino-acid substitutions in the IKAP gene product significantly increase risk for bronchial asthma in children. *J Hum Genet.* 2001;46(2):57-63.

Tong BT, Dhir K, Ha PK, Westra WH, Alter BP, Sidransky D, et al. Use of single nucleotide polymorphism arrays to identify a novel region of loss on chromosome 6q in squamous cell carcinomas of the oral cavity. *Head Neck.* 2004;26(4):345-52.

Torres VA, Tapia JC, Rodriguez DA, Lladser A, Arredondo C, Leyton L, et al. E-cadherin is required for caveolin-1-mediated down-regulation of the inhibitor of apoptosis protein survivin via reduced beta-catenin-Tcf/Lef-dependent transcription. *Mol Cell Biol.* 2007;27(21):7703-17.

Tse KP, Su WH, Chang KP, Tsang NM, Yu CJ, Tang P, et al. Genome-wide association study reveals multiple nasopharyngeal carcinoma-associated loci within the HLA region at chromosome 6p21.3. *Am J Hum Genet.* 2009;85(2):194-203.

Van den Eynde M, Baurain JF, Mazzeo F, Machiels JP. Epidermal growth factor receptor targeted therapies for solid tumours. *Acta Clin Belg.* 2011;66(1):10-7.

Vartanian JG, Magrin J, Kowalski LP. Total glossectomy in the organ preservation era. *Curr Opin Otolaryngol Head Neck Surg.* 2010;18(2):95-100.

Wagner M, Klusmann JP, Fangmann R, Linder R, Elewa ME, Eidt S, et al. Cyclin-dependent kinase-inhibitor 1 (CDKN1A) in the squamous epithelium of the oropharynx: possible implications of molecular biology and compartmentation. *Anticancer Res.* 2001;21(1A):333-45.

Wang W, Ji P, Steffen B, Metzger R, Schneider PM, Halfter H, et al. Alterations of lymphoid enhancer factor-1 isoform expression in solid tumors and acute leukemias. *Acta Biochim Biophys Sin (Shanghai).* 2005;37(3):173-80.

Waterman ML. Lymphoid enhancer factor/T cell factor expression in colorectal cancer. *Cancer Metastasis Rev.* 2004;23(1-2):41-52.

Wolf I, Bose S, Desmond JC, Lin BT, Williamson EA, Karlan BY, et al. Unmasking of epigenetically silenced genes reveals DNA promoter methylation and reduced expression of PTCH in breast cancer. *Breast Cancer Res Treat.* 2007;105(2):139-55.

Xiao BY, Dang H, Gan JY, Cai Q, Zhang GP, Chang H. Expression of PTCH-1 and SMO mRNA in nasopharyngeal carcinoma. *Xi Bao Yu Fen Zi Mian Yi Xue Za Zhi.* 2010;26(10):955-8.

Xin D, Rendon BE, Zhao M, Winner M, McGhee Coleman A, et al. The MIF homologue D-dopachrome tautomerase promotes COX-2 expression through  $\beta$ -catenin-dependent and -independent mechanisms. *Mol Cancer Res.* 2010;8(12):1601-9.

Xing J, Watkins WS, Zhang Y, Witherspoon DJ, Jorde LB. High fidelity of whole-genome amplified DNA on high-density single nucleotide polymorphism arrays. *Genomics.* 2008;92(6):452-6.

Xu B, Li JM, Tong N, Tao J, Li PC, Song NH, et al. VEGFA +936C>T polymorphism and cancer risk: a meta-analysis. *Cancer Genet Cytogenet.* 2010;198(1):7-14.

Yang S. Gene amplifications at chromosome 7 of the human gastric cancer genome. *Int J Mol Med.* 2007;20(2):225-31.

Ye H, Pungpravat N, Huang BL, Muzio LL, Marigliò MA, Chen Z, et al. Genomic assessments of the frequent loss of heterozygosity region on 8p21.3-p22 in head and neck squamous cell carcinoma. *Cancer Genet Cytogenet.* 2007;176(2):100-6.

Yoshimoto T, Yamaai T, Mizukawa N, Sawaki K, Nakano M, Yamachika E, et al. Different expression patterns of beta-defensins in human squamous cell carcinomas. *Anticancer Res.* 2003;23(6C):4629-33.

Zhang D, Richardson DR. Endoplasmic reticulum protein 29 (ERp29): An emerging role in cancer. *Int J Biochem Cell Biol.* 2011;43(1):33-6.

Zhang F, Gu W, Hurles ME, Lupski JR. Copy Number Variation in Human Health, Disease, and Evolution. *Annu Rev Genomics Hum Genet.* 2009;10:451-81.

Zhang K, Waxman DJ. PC3 prostate tumor-initiating cells with molecular profile FAM65Bhigh/MFI2low/LEF1low increase tumor angiogenesis. *Mol Cancer.* 2010;9:319.

Zheng Z, Park JY, Guillemette C, Schantz SP, Lazarus P. Tobacco carcinogen-detoxifying enzyme UGT1A7 and its association with orolaryngeal cancer risk. *J Natl Cancer Inst.* 2001;93(18):1411-8.

Zhou X, Li C, Mok SC, Chen Z, Wong DTW. Whole genome loss of heterozygosity profiling on oral squamous cell carcinoma by high-density single nucleotide polymorphic allele (SNP) array. *Cancer Genet Cytogenet.* 2004b;151:82-4.

Zhou X, Mok SC, Chen Z, Li Y, Wong DT. Concurrent analysis of loss of heterozygosity (LOH) and copy number abnormality (CNA) for oral premalignancy progression using the Affymetrix 10K SNP mapping array. *Hum Genet.* 2004a;115(4):327-30.

Zhou X, Temam S, Chen Z, Ye H, Mao L, Wong DT. Allelic imbalance analysis of oral tongue squamous cell carcinoma by high-density single nucleotide polymorphism arrays using whole-genome amplified DNA. *Hum Genet.* 2005;118(3-4):504-7.

Zhu W, You Z, Li T, Yu C, Tao G, Hu M, Chen X. Correlation of hedgehog signal activation with chemoradiotherapy sensitivity and survival in esophageal squamous cell carcinomas. *Jpn J Clin Oncol.* 2011;41(3):386-93.

# **ANEXOS**

---



## ANEXO 1

### Aspectos clínicos de pacientes e controles participantes do estudo “Identificação de Genes de Susceptibilidade Herdada para o Carcinoma de Células Escamosas de Base de Língua por Genotipagem em Larga Escala”

**Nome do paciente:** .....

**HC:** ..... **Idade:** .....anos **Sexo:** ..... **Raça:** .....

**Data da entrevista:** ...../ ...../ .....

#### **Padrão da dieta:**

(A = maior que preconizada; B = menor que preconizada; P = preconizada)

Carnes (2 a 3 porções diárias): ..... Vegetais (3 a 5 porções diárias): .....

Frutas (2 a 4 porções diárias): ..... Cereais (6 a 11 porções diárias): .....

Observações:

.....  
.....

#### **Etilismo:**

( ) Não                      ( ) Sim

Tipo(s) de bebida alcoólica e quantidade diária: .....

Tempo: ..... Observações: .....

#### **Tabagismo:**

( ) Não                      ( ) Sim

Tipo(s) de fumo e quantidade diária: .....

Tempo: ..... Observações: .....



## ANEXO 2

### Critérios do sistema de estadiamento TNM para o câncer de orofaringe

TNM – classificação clínica	
<b>T – tumor primário</b>	
<b>TX</b>	O tumor primário não pode ser avaliado
<b>T0</b>	Não há evidência de tumor primário
<b>Tis</b>	Carcinoma <i>in situ</i>
<b>T1</b>	Tumor com 2cm ou menos em sua maior dimensão
<b>T2</b>	Tumor com mais de 2cm até 4cm em sua maior dimensão
<b>T3</b>	Tumor com mais de 4cm em sua maior dimensão
<b>T4a</b>	Tumor que invade qualquer das seguintes estruturas: laringe, músculos profundos/extrínsecos da língua (genioglosso, hioglosso, palatoglosso e estiloglosso), pterigóide medial, palato duro e mandíbula)
<b>T4b</b>	Tumor que invade qualquer das seguintes estruturas: músculo pterigóide lateral, lâminas pterigóides, nasofaringe lateral, base do crânio ou adjacentes a artéria carótida
<b>N – linfonodos regionais</b>	
<b>NX</b>	Os linfonodos regionais não podem ser avaliados
<b>N0</b>	Ausência de metástase em linfonodos regionais
<b>N1</b>	Metástase em um único linfonodo homolateral, com 3cm ou menos em sua maior dimensão
<b>N2</b>	Metástase em um único linfonodo homolateral, com mais de 3cm, porém, não mais de 6cm em sua maior dimensão, ou em linfonodos homolaterais múltiplos, nenhum deles com mais de 6cm em sua maior extensão, ou em linfonodos bilaterais ou contralaterais, nenhum deles com mais de 6cm em sua maior dimensão
<b>N2a</b>	Metástase em um único linfonodo homolateral, com mais de 3cm até 6cm em sua maior dimensão
<b>N2b</b>	Metástase em linfonodos homolaterais múltiplos, nenhum deles com mais de 6cm em sua maior dimensão
<b>N2c</b>	Metástase em linfonodos bilaterais ou contralaterais, nenhum deles com mais de 6cm em sua maior dimensão
<b>N3</b>	Metástase em linfonodo com mais de 6cm em sua maior dimensão
<b>M – metástase à distância</b>	
<b>MX</b>	A presença de metástase à distância não pode ser avaliada
<b>M0</b>	Ausência de metástase à distância
<b>M1</b>	Metástase à distância



### ANEXO 3

#### Grupos de estagiamento e características TNM do câncer de orofaringe

	<b>T</b>	<b>N</b>	<b>M</b>
<b>Estágio 0</b>	T0	N0	M0
<b>Estágio I</b>	T1	N0	M0
<b>Estágio II</b>	T2	N0	M0
<b>Estágio III</b>	T1, T2	N1	M0
	T3	N0, N1	M0
<b>Estágio IVA</b>	T1, T2, T3	N2	M0
	T4a	N0, N1, N2	M0
<b>Estágio IVB</b>	T4b	Qualquer N	M0
	Qualquer T	N3	M0
<b>Estágio IVC</b>	Qualquer T	Qualquer N	M1



## ANEXO 4

### **Carta de consentimento pós-informação dos pacientes, para participação no estudo intitulado “Identificação de Genes de Susceptibilidade Herdada para o Carcinoma de Células Escamosas de Base de Língua por Genotipagem em Larga Escala”**

**Nome do paciente:** .....

**Idade:** .....anos **Sexo:** ..... **Raça:** .....

**RG:** ..... **HC:** .....

**Endereço:**

.....  
.....  
.....

**Nome do responsável legal (se paciente incapacitado):** .....

.....

**RG:** ..... **Grau de parentesco:** .....

**Endereço:**

.....  
.....  
.....

Aceito participar do estudo proposto, no qual fornecerei uma amostra de sangue (volume: 10ml) a ser colhida em veia de um dos braços. Estou ciente de que este sangue será utilizado para a avaliação de uma predisposição familiar para o desenvolvimento de câncer de cabeça e pescoço. Estou ciente que um fragmento do tumor retirado para estabelecer o diagnóstico da minha doença será utilizado para a avaliação de características próprias deste tumor. Estou ciente de que não terei prejuízos com a realização destes exames. Sei que posso sentir dor de pequena intensidade e curta duração no local de punção da veia. Sei que terei direito de saber o resultado do meu exame, caso esteja interessado. Neste caso, poderei marcar uma consulta no ambulatório de Onco-Genética, que é realizado no mesmo local do ambulatório de Oncologia Clínica, no terceiro andar do hospital de Clínicas da UNICAMP, no telefone 3521 7363, com a Profa. Dra. Carmen Silvia Bertuzzo ou a Profa. Dra Fátima Botcher, que darão esclarecimentos necessários para o meu caso. Sei que posso sair do estudo a qualquer momento e que isto não vai prejudicar o meu tratamento na UNICAMP.

Sei que meus dados pessoais serão mantidos em sigilo pelo pesquisador. Estou ainda ciente, que nenhum outro estudo usando o mesmo material será realizado sem a aprovação do Comitê de Ética da Faculdade de Ciências Médicas da UNICAMP. Se tiver qualquer dúvida sobre o estudo poderei procurar a Profa. Dra. Carmen Silvia Passos Lima, no ambulatório de Oncologia Clínica do Hospital das Clínicas da UNICAMP, telefone: (19) 3521 7496 ou 3521 7363. Se tiver reclamações sobre qualquer procedimento do estudo, poderei procurar a secretaria do Comitê de Ética do Hospital das Clínicas UNICAMP, telefone:(19) 3521 8936 ou 3521 7187. Tenho a opção ou não de aceitar o armazenamento do material genético por mim fornecido (Resolução 340/04), abaixo relacionado.

(    ) aceito o armazenamento                      (    ) não aceito o armazenamento

Eu li/ouvi o conteúdo deste termo e recebi esclarecimentos sobre as minhas dúvidas oralmente.

.....

Assinatura do paciente

.....

Assinatura do responsável legal

.....

Assinatura do pesquisador legal

Campinas, ..... / ..... / .....

## ANEXO 5

### **Carta de consentimento pós-informação dos controles, para participação no estudo intitulado “Identificação de Genes de Susceptibilidade Herdada para o Carcinoma de Células Escamosas de Base de Língua por Genotipagem em Larga Escala”**

**Nome do controle:** .....

**Idade:** .....anos **Sexo:** ..... **Raça:** .....

**RG:** .....

**Endereço:**

.....  
.....  
.....

Aceito participar do estudo proposto, no qual fornecerei uma amostra de sangue (volume: 10ml) a ser colhida em veia de um dos braços. Estou ciente de que este sangue será utilizado para a avaliação de uma predisposição familiar para o desenvolvimento de câncer de cabeça e pescoço. Estou ciente que um fragmento do tumor retirado para estabelecer o diagnóstico da minha doença será utilizado para a avaliação de características próprias deste tumor. Estou ciente de que não terei prejuízos com a realização destes exames. Sei que posso sentir dor de pequena intensidade e curta duração no local de punção da veia. Sei que terei direito de saber o resultado do meu exame, caso esteja interessado. Neste caso, poderei marcar uma consulta no ambulatório de Onco-Genética, que é realizado no mesmo local do ambulatório de Oncologia Clínica, no terceiro andar do hospital de Clínicas da UNICAMP, no telefone 3521 7363, com a Profa. Dra. Carmen Silvia Bertuzzo ou a Profa. Dra Fátima Botcher, que darão esclarecimentos necessários para o meu caso. Sei que posso sair do estudo a qualquer momento e que isto não vai prejudicar o meu tratamento na UNICAMP. Sei que meus dados pessoais serão mantidos em sigilo pelo pesquisador. Estou ainda ciente, que nenhum outro estudo usando o mesmo material será realizado sem a aprovação do Comitê de Ética da Faculdade de Ciências Médicas da UNICAMP. Se tiver qualquer dúvida sobre o estudo poderei procurar a Profa. Dra. Carmen Silvia Passos Lima, no ambulatório de Oncologia Clínica do Hospital das Clínicas da UNICAMP, telefone: (19) 3521 7496 ou 3521 7363. Se tiver reclamações sobre qualquer procedimento do estudo, poderei procurar a secretaria do Comitê de Ética do Hospital das Clínicas UNICAMP, telefone:(19) 3521 8936 ou 3521 7187. Tenho a opção ou não de aceitar o armazenamento do material genético por mim fornecido (Resolução 340/04), abaixo relacionado.

( ) aceito o armazenamento      ( ) não aceito o armazenamento

Eu li/ouvi o conteúdo deste termo e recebi esclarecimentos sobre as minhas dúvidas oralmente.

.....

Assinatura do paciente

.....

Assinatura do responsável legal

.....

Assinatura do pesquisador legal

Campinas, ..... / ..... / .....

## ANEXO 6

### Parecer do Comitê de Ética em Pesquisa da Faculdade de Ciências Médicas da UNICAMP para o projeto intitulado “Identificação de Genes de Susceptibilidade Herdada para o Carcinoma de Células Escamosas de Base de Língua por Genotipagem em Larga Escala”

FACULDADE DE CIÊNCIAS MÉDICAS  
COMITÊ DE ÉTICA EM PESQUISA

[www.fcm.unicamp.br/pesquisa/etica/index.html](http://www.fcm.unicamp.br/pesquisa/etica/index.html)

CEP. 28/11/06.  
(PARECER PROJETO: N° 446/2005)

### PARECER

#### I-IDENTIFICAÇÃO:

PROJETO: “INFLUÊNCIA DOS POLIMORFISMOS DE ALELOS DO SISTEMA DA GLUTATIONA S-TRANSFERASE 1 (GSTM1) E THETA (GSTT1) E CYP1A1\*2A E CYP1A1\*2B DO GENE *CYP1A1* DO SISTEMA DO CITOCROMO P450 NA SUSCETIBILIDADE AO CARCINOMA DE CÉLULAS ESCAMOSAS DE CABEÇA E PESCOÇO”

PESQUISADOR RESPONSÁVEL: Érica Furquim Soledade Neves Silva

#### II - PARECER DO CEP

O Comitê de Ética em Pesquisa da Faculdade de Ciências Médicas da UNICAMP tomou ciência e aprovou o adendo que inclui o projeto de pesquisa intitulado “IDENTIFICAÇÃO DE GENES DE SUSCEPTIBILIDADE PARA O CARCINOMA DE CÉLULAS ESCAMOSAS DE FARINGE”, com finalidade de tese de doutorado do aluno Gustavo Jacob Lourenço, bem como a ampliação do número de amostras a ser coletada no Ambulatório de Otorrinolaringologia e a nova versão do Termo do Consentimento Livre e Esclarecido, referente ao protocolo de pesquisa supracitado.

O conteúdo e as conclusões aqui apresentados são de responsabilidade exclusiva do CEP/FCM/UNICAMP e não representam a opinião da Universidade Estadual de Campinas nem a comprometem.

Recomendação: Retificar o nome do Comitê de Ética em Pesquisa da FCM/UNICAMP no Termo de Consentimento Livre e Esclarecido.

Homologado na XI Reunião Ordinária do CEP/FCM, em 28 de novembro de 2006.

  
**Prof. Dra. Carmen Silvia Bertuzzo**  
PRESIDENTE DO COMITÊ DE ÉTICA EM PESQUISA  
FCM / UNICAMP

Comitê de Ética em Pesquisa - UNICAMP  
Rua: Tessália Vieira de Camargo, 126  
Caixa Postal 6111  
13084-971. Campinas – SP

FONE (019) 3521-8936  
FAX (019) 3521-7187  
cep@fcm.unicamp.br



## ANEXO 7

**Distribuição dos 55 pacientes com carcinoma de células escamosas de base de língua,  
de acordo com a idade, o sexo, a raça, os hábitos de ingerir bebidas alcoólicas e de  
fumar e ao padrão da dieta**

Nº	Iniciais	Idade	Sexo	Raça	Etilismo	Tabagismo	Padrão da dieta			
							Carnes	Vegetais	Frutas	Cereais
1	AAP	45	M	C	NO	M	B	B	B	B
2	EAC	50	M	C	A	A	P	P	P	P
3	JF	66	M	N	A	A	B	P	P	B
4	PLB	75	M	C	A	M	P	A	P	P
5	RDG	59	M	C	A	NO	A	A	P	P
6	SAL	49	M	C	A	M	P	B	B	B
7	SAC	58	M	N	D	A	B	B	B	B
8	SRJL	43	M	C	A	A	P	P	P	P
9	WS	46	M	N	A	M	P	P	B	B
10	JAOD	39	M	C	A	M	P	B	B	B
11	JS	56	M	C	M	A	P	P	P	B
12	JAR	50	M	C	A	A	A	P	P	B
13	JBF	58	M	C	A	A	A	A	A	B
14	LCO	73	M	C	M	A	A	A	A	P
15	JRSF	66	M	C	A	A	A	P	P	B
16	LCM	52	M	C	A	A	P	B	P	B
17	EEF	47	F	C	M	M	A	A	A	B
18	JRF	50	M	N	M	A	P	B	B	A
19	OG	61	M	C	A	A	P	P	B	B
20	ACS	70	M	N	A	A	B	B	P	P
21	VRS	52	M	C	A	A	P	B	B	P
22	NGR	59	M	C	NO	NO	NO	NO	NO	NO
23	GAC	51	M	C	A	A	P	B	B	A
24	ANS	55	M	C	M	A	B	P	P	B
25	MAMS	55	F	C	M	N	NO	NO	NO	NO
26	LCC	49	M	C	A	A	P	P	P	P
27	CDM	41	M	C	NO	NO	NO	NO	NO	NO
28	JRF	51	M	C	A	A	P	P	B	P

Continuação

29	MM	71	M	C	A	A	P	P	P	P
30	MPP	56	M	N	A	A	A	P	P	P
31	MR	53	M	C	A	A	B	B	B	B
32	PA	53	M	C	A	M	B	B	B	B
33	DB	61	M	C	A	A	P	P	P	P
34	JF	53	M	N	A	A	P	P	P	P
35	OD	65	M	C	A	A	P	P	P	P
36	JLT	55	M	C	A	A	P	P	P	P
37	OBF	77	M	C	A	A	P	B	P	B
38	AV	77	M	C	M	A	P	P	B	P
39	JJ	84	F	C	N	A	P	P	B	P
40	BS	65	M	N	A	A	P	B	P	P
41	AC	60	M	N	A	A	P	P	B	P
42	JBR	67	M	C	A	A	B	B	B	P
43	JBSO	60	M	C	A	A	P	B	P	B
44	FSRJ	65	M	N	A	A	B	B	B	B
45	JLGRJ	44	M	C	A	A	P	B	B	P
46	JGSRJ	67	M	C	A	A	B	B	B	P
47	BA	68	M	C	A	A	B	B	B	P
48	EDA	48	M	C	A	A	P	B	B	P
49	JFS	74	M	C	M	M	B	B	B	B
50	JMB	78	M	C	A	M	B	B	B	B
51	ACS	44	M	C	A	A	B	P	P	P
52	JGS2	59	M	C	M	M	P	P	B	P
53	GBS	56	M	N	A	M	P	B	B	B
54	ESB	49	M	C	A	A	P	P	B	P
55	JRSF2	66	M	C	NO	NO	NO	NO	NO	NO

M: masculino, F: feminino, C: caucasóide, N: negróide, Etilismo e Tabagismo: A: ingestão ou consumo acentuado, D: ingestão ou consumo discreto, M: ingestão ou consumo moderado, N: não etilista ou tabagista, NO: informação não obtida, Padrão da dieta: A: acima da ingestão preconizada, B: abaixo da ingestão preconizada, P: ingestão preconizada, NO: informação não obtida

## ANEXO 8

### Distribuição dos 55 controles, de acordo com a idade, o sexo, a raça, os hábitos de ingerir bebidas alcoólicas e de fumar e ao padrão da dieta

Nº	Iniciais	Idade	Sexo	Raça	Etilismo	Tabagismo	Padrão da dieta			
							Carnes	Vegetais	Frutas	Cereais
1	JBC	54	M	C	M	D	B	B	B	P
2	JAS	37	M	N	N	N	P	P	P	P
3	HAN	45	M	N	M	N	P	B	B	B
4	SAIS	49	M	C	D	N	P	P	B	P
5	LN	50	M	C	D	M	P	P	B	P
6	PGS	58	M	C	N	N	B	P	A	B
7	AHS	47	M	N	N	N	B	P	P	P
8	MJF	48	M	N	D	N	P	P	P	P
9	JPR	48	M	C	D	M	P	B	B	P
10	MCGP	52	M	C	D	M	P	P	B	P
11	LS	49	M	C	N	N	P	B	A	B
12	MP	48	F	C	D	N	B	A	A	B
13	AMA	46	M	C	N	N	A	P	P	P
14	VF	52	M	C	N	N	B	P	A	B
15	APS	47	M	C	D	N	A	P	B	P
16	ABP	50	M	C	D	N	P	P	B	P
17	JVDF	51	M	C	D	N	B	B	B	B
18	COM	41	M	C	D	N	B	B	B	B
19	IC	42	M	C	D	D	P	P	P	P
20	FCM	49	M	C	D	N	B	P	P	P
21	CRSF	49	M	C	D	D	P	P	P	P
22	RLMJ	39	M	C	M	D	P	P	P	P
23	IPN	46	M	C	D	N	A	P	B	P
24	MSCO	59	F	C	N	N	B	B	P	B
25	GIO	51	M	C	D	N	P	P	P	P
26	JM	45	M	C	D	A	P	P	B	P
27	JPS	52	M	C	D	M	P	B	P	B
28	AS	49	M	C	N	N	P	P	B	P
29	DRC	50	M	C	N	M	P	P	P	P

Continuação

30	SVC	51	M	C	M	N	P	P	P	P
31	ELB	56	F	C	N	N	P	P	B	P
32	IJS	43	M	C	D	N	P	P	P	P
33	NL	54	M	C	N	N	P	P	P	P
34	JBAC	56	M	N	D	N	P	P	B	P
35	RB	50	M	C	D	N	A	P	P	P
36	MAV	47	M	C	D	N	A	A	A	P
37	JSS	29	M	N	N	N	P	P	P	P
38	JRA	47	M	N	D	N	P	P	P	P
39	DV	49	M	N	N	N	P	P	P	P
40	RT	46	M	N	N	N	P	B	B	P
41	JPS	46	M	N	M	M	P	P	P	P
42	ERR	50	M	C	N	N	B	B	B	P
43	AOM	50	M	C	N	N	P	B	B	B
44	AMM	46	M	C	D	N	B	P	P	P
45	JAF	47	M	C	N	N	P	B	P	B
46	IFS	55	M	C	N	N	B	P	P	P
47	SLS	57	M	C	NO	NO	A	P	P	P
48	DRL	44	M	C	N	N	B	P	B	B
49	PSM	44	M	C	D	N	A	A	P	P
50	WLO	53	M	C	N	D	P	B	B	P
51	MAV	47	M	C	D	N	A	A	A	P
52	MDN	41	M	C	N	N	P	P	B	P
53	SF	43	M	C	N	N	P	B	B	B
54	JAS2	48	M	N	N	N	P	P	P	P
55	COM	41	M	C	D	N	B	B	B	B

M: masculino, F: feminino, C: caucasóide, N: negróide, Etilismo e Tabagismo: A: ingestão ou consumo acentuado, D: ingestão ou consumo discreto, M: ingestão ou consumo moderado, N: não etilista ou tabagista, NO: informação não obtida, Padrão da dieta: A: acima da ingestão preconizada, B: abaixo da ingestão preconizada, P: ingestão preconizada, NO: informação não obtida

## ANEXO 9

### Distribuição dos 55 pacientes com carcinoma de células escamosas de base língua com o grau de diferenciação do tumor e o estágio TNM

Nº	Iniciais	Grau de diferenciação	Estadiamento clínico
1	AAP	moderadamente	IV
2	EAC	bem	IV
3	JF	bem	IV
4	PLB	moderadamente	IV
5	RDG	moderadamente	IV
6	SAL	moderadamente	II
7	SAC	moderadamente	IV
8	SRJL	moderadamente	IV
9	WS	pouco	III
10	JAOD	moderadamente	IV
11	JS	pouco	IV
12	JAR	moderadamente	IV
13	JBF	moderadamente	IV
14	LCO	moderadamente	IV
15	JRSF	moderadamente	IV
16	LCM	pouco	III
17	EEF	moderadamente	IV
18	JRF	moderadamente	II
19	OG	moderadamente	IV
20	ACS	pouco	II
21	VRS	moderadamente	IV
22	NGR	moderadamente	I
23	GAC	moderadamente	III
24	ANS	moderadamente	IV
25	MAMS	pouco	IV
26	LCC	moderadamente	IV
27	CDM	bem	IV
28	JRF	moderadamente	II
29	MM	moderadamente	IV
30	MPP	ND	III

Continuação			
31	MR	ND	ND
32	PA	ND	IV
33	DB	indiferenciado	IV
34	JF	moderadamente	III
35	OD	moderadamente	IV
36	JLT	moderadamente	IV
37	OBF	pouco	III
38	AV	moderadamente	IV
39	JJ	moderadamente	III
40	BS	moderadamente	IV
41	AC	pouco	IV
42	JBR	ND	II
43	JBSO	moderadamente	IV
44	FSRJ	ND	IV
45	JLGRJ	pouco	IV
46	JGSRJ	moderadamente	IV
47	BA	moderadamente	IV
48	EDA	moderadamente	IV
49	JFS	moderadamente	IV
50	JMB	moderadamente	IV
51	ACS	pouco	II
52	JGS2	moderadamente	IV
53	GBS	moderadamente	III
54	ESB	moderadamente	IV
55	JRSF2	moderadamente	IV

ND: informação não definida

## ANEXO 10

**Distribuição das regiões de ganho de material genético de acordo com os cromossomos, genes identificados, tamanhos das variações e número de pacientes com carcinoma de células escamosas de base língua**

<b>Cromossomo</b>	<b>Gene</b>	<b>Início</b>	<b>Término</b>	<b>Tamanho (pb)</b>	<b>Nº pacientes</b>
1	<i>CFHR1</i>	195014203	195061554	47352	26
1	<i>CFHR1</i>	195061554	195066803	5250	25
1	<i>CFHR1</i>	195066803	195068287	1485	24
1	<i>CFHR3</i>	195003402	195014203	10802	23
1	<i>GSTM1</i>	110032541	110039706	7166	25
1	<i>GSTM1</i>	110030457	110032541	2085	17
1	<i>LCE3B</i>	150852685	150855084	2400	10
1	<i>LCE3C</i>	150838581	150841015	2435	13
2	<i>POTEF</i>	130541281	130575255	33975	13
2	<i>POTEF</i>	130575255	130578875	3621	10
3	<i>ALG1L2</i>	131285743	131291835	6093	17
3	<i>ALG1L2</i>	131280072	131285743	5672	18
4	<i>UGT2B17</i>	69078374	69115545	37172	14
4	<i>UGT2B17</i>	69115545	69145079	29535	20
6	<i>HLA-DQA1</i>	32719235	32743359	24125	12
6	<i>HLA-DQA1</i>	32713894	32719235	5342	10
6	<i>HLA-DQB1</i>	32719235	32743359	24125	12
6	<i>HLA-DRB5</i>	32581631	32602949	21319	15
6	<i>HLA-DRB5</i>	32602949	32605110	2162	14
6	<i>HLA-DRB5</i>	32605110	32606436	1327	13
6	<i>HLA-DRB6</i>	32625797	32636693	10897	14
7	<i>UPK3B</i>	75984673	75999603	14931	12
7	<i>UPK3B</i>	75983132	75984673	1542	13
8	<i>ADAM3A</i>	39374430	39485067	110638	18
8	<i>ADAM3A</i>	39489516	39505950	16435	16
8	<i>ADAM3A</i>	39485067	39489516	4450	17
8	<i>ADAM5P</i>	39374430	39485067	110638	18
8	<i>ADAM5P</i>	39360384	39374430	14047	17
8	<i>ADAM5P</i>	39349983	39360384	10402	16

Continuação

8	<i>DEFB103A</i>	7226817	7298797	71981	11
8	<i>DEFB104B</i>	7298797	7393017	94221	13
8	<i>FAM66D</i>	12028478	12036199	7722	10
8	<i>FAM66D</i>	12024502	12028478	3977	11
8	<i>FAM66B</i>	7193832	7203877	10046	10
8	<i>FAM90A7</i>	7418101	7500197	82097	15
8	<i>FAM90A7</i>	7393017	7418101	25085	13
8	<i>FAM90A8</i>	7500197	7663098	162902	14
8	<i>FAM90A10</i>	7663098	7669731	6634	13
8	<i>REXO1L1</i>	86748106	86875885	127780	11
8	<i>REXO1L2P</i>	86875885	86899617	23733	10
8	<i>SPAG11B</i>	7298797	7393017	94221	13
8	<i>SPAG11B</i>	7226817	7298797	71981	11
8	<i>USP17L2</i>	12028478	12036199	7722	10
8	<i>ZNF705G</i>	7193832	7203877	10046	10
11	<i>MRGPRX1</i>	18905673	18913611	7939	16
11	<i>OR51A2</i>	4932398	4933004	607	16
11	<i>OR51A2</i>	4933004	4933404	401	13
11	<i>OR51A4</i>	4924647	4925385	739	18
11	<i>OR52N5</i>	5756214	5765522	9309	18
11	<i>OR52N5</i>	5746907	5755808	8902	20
11	<i>OR52N5</i>	5755808	5756214	407	19
14	<i>HEATR4</i>	73089369	73099604	10236	10
15	<i>LOC727924</i>	19822216	19847646	25431	10
15	<i>LOC727924</i>	19847646	19869335	21690	12
15	<i>OR4M2</i>	19869335	19877312	7978	11
15	<i>OR4N4</i>	19877312	20055922	178611	10
17	<i>ARL17B</i>	41721036	41736246	15211	10
17	<i>EVPLL</i>	18230518	18237837	7320	11
17	<i>LOC339240</i>	18237837	18271866	34030	10
17	<i>TBC1D3B</i>	31501268	31526979	25712	10
19	<i>LILRA6</i>	59433868	59435097	1230	11
19	<i>LILRA6</i>	59432360	59433489	1130	13
19	<i>LILRA6</i>	59431897	59432360	464	12
19	<i>LILRA6</i>	59433489	59433868	380	12
20	<i>SIRPB1</i>	1531305	1535275	3971	21

Continuação					
20	<i>SIRPB1</i>	1516816	1520058	3243	24
20	<i>SIRPB1</i>	1520058	1523157	3100	23
20	<i>SIRPB1</i>	1528688	1531305	2618	25
20	<i>SIRPB1</i>	1512197	1514456	2260	13
20	<i>SIRPB1</i>	1525793	1527796	2004	25
20	<i>SIRPB1</i>	1515173	1516816	1644	22
20	<i>SIRPB1</i>	1524569	1525793	1225	24
20	<i>SIRPB1</i>	1527796	1528688	893	25
20	<i>SIRPB1</i>	1523157	1523891	735	23
20	<i>SIRPB1</i>	1514456	1515173	718	20
20	<i>SIRPB1</i>	1523891	1524569	679	24
22	<i>DDT</i>	22650793	22665036	14244	11
22	<i>LOC391322</i>	22683140	22730726	47587	21
22	<i>GSTT1</i>	22683140	22730726	47587	21
22	<i>GSTT2</i>	22650793	22665036	14244	11
22	<i>GSTTP1</i>	22674016	22683140	9125	21
22	<i>GSTTP1</i>	22665036	22674016	8981	16
22	<i>GSTTP2</i>	22683140	22730726	47587	21
22	<i>PI4KAP1</i>	18751735	18806571	54837	16
22	<i>RIMBP3</i>	18806571	18862455	55885	17

Gene: símbolo do gene, Início: região de início, Término: região de término, N°: número, pb: pares de bases



## ANEXO 11

**Distribuição das regiões de perda de material genético de acordo com os cromossomos, genes identificados, tamanhos das variações e número de pacientes com carcinoma de células escamosas de base língua**

<b>Cromossomo</b>	<b>Gene</b>	<b>Início</b>	<b>Término</b>	<b>Tamanho (pb)</b>	<b>Nº pacientes</b>
1	<i>CROCC</i>	17079137	17132483	53347	10
1	<i>LCE3B</i>	150848459	150855084	6626	15
1	<i>LCE3C</i>	150838581	150844617	6037	18
1	<i>RHD</i>	25467144	25506417	39274	18
1	<i>RHD</i>	25506417	25526969	20553	17
4	<i>UGT2B28</i>	70170720	70189817	19098	13
4	<i>UGT2B28</i>	70189817	70205355	15539	15
6	<i>DUSP22</i>	200298	237735	37438	11
6	<i>DUSP22</i>	239611	243966	4356	10
6	<i>DUSP22</i>	237735	239611	1877	12
7	<i>MGAM</i>	141424419	141429694	5276	13
7	<i>MGAM</i>	141430697	141434106	3410	16
7	<i>MGAM</i>	141435388	141437522	2135	13
7	<i>MGAM</i>	141434106	141435388	1283	14
7	<i>MGAM</i>	141429694	141430697	1004	13
11	<i>OR4C6</i>	55186827	55192748	5922	18
11	<i>OR4P4</i>	55137350	55168027	30678	18
11	<i>OR4S2</i>	55170448	55176395	5948	22
11	<i>OR52N5</i>	5753713	5756214	2502	10
14	<i>OR4K1</i>	19467539	19474888	7350	16
14	<i>OR4K2</i>	19412605	19438152	25548	20
14	<i>OR4K5</i>	19442649	19461468	18820	18
14	<i>OR4N2</i>	19323008	19374298	51291	11
14	<i>OR4Q3</i>	19270490	19323008	52519	12
15	<i>BCL8</i>	19137740	19355538	217799	14
15	<i>BCL8</i>	19128929	19136095	7167	11
15	<i>BCL8</i>	19136095	19137740	1646	15
15	<i>CXADRP2</i>	19137740	19355538	217799	14
15	<i>LOC727924</i>	19573448	19791550	218103	11

Continuação					
15	<i>LOC727924</i>	19813501	19847646	34146	11
15	<i>LOC727924</i>	19847646	19862284	14639	10
15	<i>LOC727924</i>	19791550	19798947	7398	10
15	<i>NF1P1</i>	19355538	19417073	61536	15
15	<i>OR4M2</i>	19862284	19879983	17700	11
15	<i>OR4N3P</i>	19879983	19985655	105673	10
15	<i>OR4N4</i>	19879983	19985655	105673	10
15	<i>POTEB</i>	19137740	19355538	217799	14
19	<i>PSG4</i>	48394729	48406656	11928	10
20	<i>SIRPB1</i>	1512197	1524569	12373	13
20	<i>SIRPB1</i>	1525793	1528688	2896	11
20	<i>SIRPB1</i>	1524569	1525793	1225	12

Gene: símbolo do gene, Início: região de início, Término: região de término, N°: número, pb: pares de bases

## ANEXO 12

### ARTIGO CIENTÍFICO

#### ***CYP1A1, GSTM1 and GSTT1 Polymorphisms, Tobacco and Alcohol Status, and Risk of Head and Neck Squamous Cell Carcinoma***

Gustavo Jacob Lourenço<sup>a</sup>, Erika Furquim Soledade Neves Silva<sup>a</sup>, José Augusto  
Rinck-Junior<sup>a</sup>, Carlos Takahiro Chone<sup>b</sup>, Carmen Silvia Passos Lima<sup>a\*</sup>

<sup>a</sup>Department of Internal Medicine, <sup>b</sup>Department of Ophthalmology and  
Otorhinolaryngology; Faculty of Medical Sciences, State University of Campinas,  
Campinas, São Paulo, Brazi



# *CYP1A1*, *GSTM1* and *GSTT1* polymorphisms, tobacco and alcohol status and risk of head and neck squamous cell carcinoma

Gustavo Jacob Lourenço · Erika Furquim Soledade Neves Silva · José Augusto Rinck-Junior · Carlos Takahiro Chone · Carmen Silvia Passos Lima

Received: 9 June 2011 / Accepted: 10 August 2011 / Published online: 26 August 2011  
© International Society of Oncology and BioMarkers (ISOBM) 2011

**Abstract** We examined the influence of the *CYP1A1* A4889G and T6235C, *GSTM1* and *GSTT1* polymorphisms, involved in carcinogen metabolism, on the head and neck (HN) squamous cell carcinoma (SCC) risk. DNA from 142 HNSCC patients and 142 controls was analysed by polymerase chain reaction (PCR)–restriction fragment length polymorphism or multiplex-PCR for the polymorphisms analyses. Excesses of the *CYP1A1* 4889AG+GG and 4889AG+GG plus *GSTT1* null genotype were seen in patients with heavy tobacco habit compared with controls (41.9% versus 26.8%,  $P=0.03$ ; 26.2% versus 10.3%,  $P=0.04$ , respectively). Carriers of the referred genotypes and heavy tobacco consumption were under a 2.0-fold and 2.8-fold increased risks for HNSCC than others, respectively. The *CYP1A1* 6235TC+CC plus *GSTM1* and *GSTT1* null genotypes were more common in pharyngeal SCC patients than in controls (5.3% versus 0.7%,  $P=0.04$ ). Carriers of the combined genotype had 16.0-fold increased risk for the disease than others. The frequency of one null genotype of the *GSTM1* or *GSTT1* gene was higher in patients with pharyngeal SCC and heavy smoking status than in

controls (76.3% versus 57.7%,  $P=0.04$ ). Carriers of the referred genotype and heavy tobacco status had a 2.4-fold increased risk for pharyngeal SCC than others. In contrast, the *CYP1A1* 6235TC+CC genotype was more common in controls than in laryngeal SCC patients (35.9% versus 21.6%,  $P=0.01$ ). Carriers of the genotype had a 0.2-fold decreased risk for the disease than others. Our data present preliminary evidence that inherited combined *CYP1A1* A4889G and T6235C abnormalities and *GSTM1* and *GSTT1* pathways are important determinants of HNSCC, particularly pharyngeal SCC in heavy smoking individuals from south-eastern Brazil.

**Keywords** Head and neck tumour · Risk · Tobacco · Alcohol · *CYP1A1* · *GSTM1* · *GSTT1*

## Introduction

The exposure to tobacco smoke and the consumption of alcohol are considered the most important aetiological factors for the development of head and neck (HN) squamous cell carcinomas (SCC) [1]. Polycyclic aromatic hydrocarbons (PAHs), such as benzo(a)pyrene of the tobacco [2] and ethanol and acetaldehyde of the alcoholic drinks [3], are important chemical agents related to the tumour origin.

The cytochrome P450 (CYP) enzymes are the first line of defence against PAHs, participating in their activation in order to facilitate their inactivation by the glutathione S-transferase (GST) family enzymes [4]. In addition, the GST enzymes also participate in the alcohol metabolism [5].

The ability to metabolise PAHs and alcohol varies among human beings and, therefore, healthy individuals may be at

G. J. Lourenço · E. F. S. N. Silva · J. A. Rinck-Junior · C. S. P. Lima  
Department of Internal Medicine, Faculty of Medical Sciences, State University of Campinas, Campinas, São Paulo, Brazil

C. T. Chone  
Department of Ophthalmology and Otorhinolaryngology, Faculty of Medical Sciences, State University of Campinas, Campinas, São Paulo, Brazil

C. S. P. Lima (✉)  
Clinical Oncology Service, Department of Internal Medicine, Faculty of Medical Sciences, Cidade Universitária “Zeferino Vaz”, Rua Alexander Fleming 181, Barão Geraldo, 13083-970, Campinas, São Paulo, Brazil  
e-mail: carmenl@fcm.unicamp.br

distinct risks for HNSCC [4, 5]. The variant alleles of the *CYP1A1* A4889G and T6235C polymorphisms give rise to enzymes with increased activity in PAHs activation [6], and individuals with the absence of the *GSTM1* and *GSTT1* genes have no enzymatic functional activity [3, 7]. The variant GG and CC genotypes of *CYP1A1* gene [6] and the homozygous *GSTM1* and *GSTT1* null genotypes [7] were found in 1–20%, 4–40% and 10–60% of healthy individuals from distinct ethnic populations, respectively. However, their roles for the risk of HNSCC are controversial [4, 8, 9].

The environmental diseases resulting from tobacco habit [10] and consumption of alcohol [11] were described as serious health problems in south-eastern Brazil. In addition, the Brazilian population is highly heterogeneous, consisting of indigenous Amerindians and immigrants from Europe, Africa and Asia [12]. Thus, the identification of the *CYP1A1* and the *GSTs* genotypes, and the evaluation of the smoking and alcohol habits in healthy and HNSCC individuals from this region were considered necessary to test their influences for the HNSCC risk in our population.

## Materials and methods

### Study subjects

The case group was comprised of all 142 consecutive histopathologically diagnosed HNSCC patients at diagnosis and 142 gender and ethnic origin-matched blood donors seen at our service during the period from June 2007 to May 2010. The patients with a previous family history of HNSCC and those who did not accept to participate of the study were excluded from the analyses.

The HNSCC was diagnosed according to standard criteria [13] and staged according to American Joint Committee of Cancer Staging [14]. The study was approved by the institutional review board guidelines, according to the Declaration of Helsinki.

### Environmental risk factors

All the subjects were interviewed using standardized questionnaire regarding the lifestyle information such as use of tobacco and alcohol consumption. Patients and controls were classified as smokers or non-smokers and drinkers or non-drinkers [15]. The lifetime exposure to smoking and alcohol in patients was also specifically measured. On the basis of smoking status, the patients were classified into the following groups: non-smoking, subjects who never smoked in their lifetime; ex-smokers, those who stopped smoking for more than 10 years; light smokers, subjects who smoked 10 cigarettes per day for 10 years; moderate smokers, those who smoked >10 and

<30 cigarettes per day between 10 to 30 years and heavy smokers, individuals who smoked 30 cigarettes per day for 30 years or more. The patients were divided into four groups on the basis of their alcohol status as follows: non-drinkers, subjects who never consumed any type of alcohol in their lifetime; ex-drinkers, those who stopped drinking for more than 10 years; light drinkers, those who consumed <14 g alcohol/day; moderate drinkers, individuals who consumed 14–84 g alcohol/day and heavy drinkers, those who had consumed >84 g alcohol/day.

### Genotyping

Genomic DNA was obtained from peripheral blood samples from all subjects. The *CYP1A1* A4889G (rs 1048943) and T6235C (rs 4646903) genotypes were amplified by the polymerase chain reaction (PCR) followed by the enzymatic digestion [16] and the *GSTM1* and *GSTT1* genes were amplified by the multiplex-PCR [17]. All the molecular reactions had positive and negative controls sequenced previously.

### Statistical analysis

The Hardy–Weinberg (HW) equilibrium was tested with chi square ( $\chi^2$ ) statistic for the goodness-to-fit. The differences between groups were analysed by  $\chi^2$  or Fisher exact test. Multivariate analysis was performed using the logistic regression model and served to obtain age, gender and ethnic origin adjusted crude odds ratios (ORs), and assess associations among genotypes and HNSCC. For all statistical tests, significance is two-sided and achieved when *P* values were <0.05. The analyses of the data were done using the SPSS 15.0 software (SPSS Incorporation, Chicago, IL, USA).

## Results

Demographic, smoking and drinking characteristics of HNSCC patients and controls are presented in Table 1. Individuals with tobacco and alcohol history consumption were at a 54.4-fold (95% CI, 19.4–152.0) and a 121.5-fold (95% CI, 36.5–404.6) increased risks for HNSCC, respectively. Among the patients, 29 (20.4%) presented SCC oral cavity, 76 (53.5%) of pharynx and 37 (26.1%) of larynx.

The HNSCC patients' and controls' samples confirmed the HW expectations at the *CYP1A1* A4889G ( $\chi^2=0.81$ , *P*=0.37 and  $\chi^2=0.001$ , *P*=0.97, respectively) and T6235C loci ( $\chi^2=0.002$ , *P*=0.96 and  $\chi^2=0.02$ , *P*=0.88, respectively). The samples of patients with SCC of cavity oral and pharynx also confirmed the HW expectations at the *CYP1A1* A4889G ( $\chi^2=0.18$ , *P*=0.91 and  $\chi^2=0.48$ , *P*=0.79, respectively) and

**Table 1** Demographic, smoking and alcohol characteristics of 142 head and neck squamous cells carcinoma patients and 142 controls

Characteristics	Patients (%)	Controls (%)	P value	OR (95% CI)
Mean age at diagnosis (range)	58.3 (31–93)	53.7 (46–60)	<0.01	
Gender				
Male	134 (94.4)	134 (94.4)	0.87	0.9 (0.3–2.6)
Female	8 (5.6)	8 (5.6)		
Ethnic origin				
Caucasian	111 (78.2)	111 (78.2)	0.73	0.9 (0.5–1.6)
Black	31 (21.8)	31 (21.8)		
Smoking consumption <sup>a</sup>				
Yes	137 (96.5)	31 (21.8)	<0.001	54.4 (19.4–152.0)
No	5 (3.5)	64 (45.1)		
Alcohol consumption <sup>a</sup>				
Yes	128 (90.1)	8 (5.6)	<0.001	121.5 (36.5–404.6)
No	13 (9.1)	48 (33.8)		

OR odds ratio adjusted by age, gender and ethnic origin, CI confidence interval

<sup>a</sup> The total numbers of individuals with smoking and alcohol consumption differed from the total quoted ( $n=142$ ) because it was not possible to obtain consistent information about tobacco exposure and alcohol ingestion in some individuals

T6235C loci ( $\chi^2=4.21$ ,  $P=0.12$  and  $\chi^2=0.28$ ,  $P=0.87$ , respectively). The laryngeal SCC patients' samples confirmed the HW expectations at the *CYP1A1* A4889G locus ( $\chi^2=4.85$ ,  $P=0.09$ ) but not at the *CYP1A1* T6235C locus ( $\chi^2=8.04$ ,  $P=0.02$ ).

The frequencies of genotypes in patients with HNSCC in general and with SCC of cavity oral, pharynx and larynx, and in controls are presented in Table 2. The *CYP1A1* 6235TC+CC plus *GSTM1* plus *GSTT1* null genotypes were more common in patients with SCC of pharynx than in controls; carriers of the variant alleles were under a 16.0-fold (95% CI, 1.1–221.8) increased risk for the pharyngeal SCC than others (power of analyses, 73%). Higher frequency of the *CYP1A1* 6235TC+CC genotype was seen in controls compared with laryngeal SCC patients; carriers of the variant allele were under a 0.2-fold (95% CI, 0.1–0.7) decreased risk for the SCC of the larynx than carriers of the remaining genotypes (power of analysis, 75%). Similar frequencies of the *CYP1A1* A4889G plus T6235C and the *GSTM1* plus *GSTT1* combined genotypes were seen in patients with HNSCC, patients with tumours of oral cavity, pharynx and larynx and controls (data not shown).

The frequencies of genotypes in HNSCC patients stratified by age, tobacco consumption and subtypes of the tumour are presented in Table 3. The *CYP1A1* 6235TC+CC, 4889AG+GG plus 6235TC+CC and the 6235TC+CC plus *GSTM1* null genotypes were more common in HNSCC patients aged 57 years or more than in younger patients. Excesses of the *CYP1A1* 4889AG+GG and the 4889AG+GG plus the *GSTT1* null genotype were seen in HNSCC patients with a heavy tobacco habit compared with those with moderate smoking status. The frequencies of these genotypes in HNSCC patients were also higher than that

found in controls (41.9% versus 26.8%,  $P=0.03$  and 26.2% versus 10.3%,  $P=0.04$ , respectively). Carriers of the variant alleles of the *CYP1A1* A4889G and the A4889G plus *GSTM1* polymorphisms, who had a heavy tobacco consumption, were under a 2.0-fold (95% CI, 1.1–3.6) and a 2.8-fold (95% CI, 1.1–7.2) increased risk for HNSCC than others, respectively. Moreover, the *CYP1A1* 6235TC+CC genotypes and the 6235TC+CC plus *GSTM1* null genotype were less common in patients with laryngeal SCC than in patients with other types of HNSCC. Similar frequencies of the distinct genotypes were seen in HNSCC patients stratified by ethnic origin and alcohol consumption (data not shown).

When the pharyngeal SCC patients stratified by clinical (age, gender, ethnic origin, tobacco and drinking status) and tumour (stage) variables were considered, the frequency of one null genotype of the *GSTM1* or *GSTT1* gene was higher in heavy smoker patients compared with those with moderate smoking status (76.3% versus 51.4%,  $P=0.02$ ), and also compared with controls (76.3% versus 57.7%,  $P=0.04$ ). Carriers of one null *GST* genotype and heavy tobacco habit had a 2.4-fold (95% CI, 1.0–5.3) increased risk for SCC of pharynx than others. Similar frequencies of the distinct genotypes of genes of interest were seen in HNSCC patients in general and in patients with SCC of oral cavity and of larynx stratified by the referred variables (data not shown).

## Discussion

We investigated herein whether the tobacco and alcohol habits, the *CYP1A1* A4889G and T6235C and the *GSTM1*

**Table 2** *CYP1A1*, *GSTM1* and *GSTT1* genotypes in head and neck carcinoma in patients and controls

Genotypes	Head and neck Number (%)	Oral cavity	Pharynx	Larynx	Controls
<i>CYP1A1</i> A4889G					
AA	94 (66.2)	21 (72.4)	48 (63.2)	25 (67.6)	104 (73.2)
AG+GG	48 (33.8)	8 (27.6)	28 (36.8)	12 (32.4)	38 (26.8)
OR (95% CI)	1.4 (0.8–2.4)	1.0 (0.4–2.8)	1.6 (0.9–3.0)	1.4 (0.5–3.5)	1.0 (ref)
<i>CYP1A1</i> T6235C					
TT	90 (63.4)	13 (44.8)	45 (59.2)	29 (78.4)	91 (64.1)
TC+CC	52 (36.6)	16 (55.2)	31 (40.8)	8 (21.6)	51 (35.9)
OR (95% CI)	0.9 (0.5–1.5)	1.1 (0.4–3.0)	1.2 (0.7–2.2)	0.2 (0.1–0.7)**	1.0 (ref)
<i>GSTM1</i>					
Present	84 (59.2)	19 (65.5)	42 (55.3)	23 (62.2)	76 (53.5)
Null	58 (40.8)	10 (34.5)	34 (44.7)	14 (37.8)	66 (46.5)
OR (95% CI)	0.8 (0.5–1.3)	0.4 (0.1–1.1)	1.0 (0.5–1.7)	0.8 (0.3–1.8)	1.0 (ref)
<i>GSTT1</i>					
Present	106 (74.7)	22 (75.9)	56 (73.7)	28 (75.7)	115 (81.0)
Null	36 (25.3)	7 (24.1)	20 (26.3)	9 (24.3)	27 (19.0)
OR (95% CI)	1.3 (0.7–2.4)	0.8 (0.2–2.6)	1.5 (0.8–3.0)	1.1 (0.4–3.3)	1.0 (ref)
<i>CYP1A1</i> A4889G+ <i>GSTT1</i>					
AA+Present	72 (50.7)	16 (55.2)	37 (48.7)	19 (51.3)	87 (61.3)
AG+GG+Null	14 (9.9)	2 (6.9)	9 (11.8)	3 (8.1)	10 (7.0)
OR (95% CI)	1.6 (0.6–4.0)	0.6 (0.1–4.0)	2.1 (0.8–5.7)	1.6 (0.3–7.9)	1.0 (ref)
<i>CYP1A1</i> T6235C+ <i>GSTM1</i>					
TT+Present	53 (37.3)	12 (41.4)	24 (31.6)	17 (46.0)	48 (33.8)
TC+CC+Null	21 (14.8)	6 (20.7)	13 (17.1)	2 (5.4)	23 (16.2)
OR (95% CI)	0.7 (0.3–1.6)	0.6 (0.1–2.2)	1.2 (0.5–2.8)	0.2 (0.1–1.1)	1.0 (ref)
<i>CYP1A1</i> T6235C+Both <i>GST</i>					
TT+Present+Present	35 (24.6)	10 (34.5)	13 (17.1)	12 (32.4)	35 (24.6)
TC+CC+Null+Null	7 (4.9)	3 (10.3)	4 (5.3)	0 (0.0)	1 (0.7)
OR (95% CI)	5.4 (0.6–52.7)	4.7 (0.3–80.0)	16.0 (1.1–221.8)*	1.0	1.0 (ref)

OR odds ratio adjusted by age at diagnosis, gender and origin ethnic, CI confidence interval

\* $P=0.04$ ; \*\* $P=0.01$

and the *GSTT1* genotypes alter the risk, demographic characteristics and biological features of tumour in Brazilian HNSCC patients.

As it would be expected, we found that individuals with history of tobacco and alcohol consumption were at a 54.4-fold and a 121.5-fold increased risks for HNSCC, respectively, as reported [1–3].

When the genotypes were included in analyses, we initially found similar frequencies of the *CYP1A1* A4889G and T6235C and *GSTM1* and *GSTT1* genotypes in patients and controls, suggesting that the polymorphisms did not alter the risk for HNSCC, according to previous evaluations [8, 9, 18–23]. In contrast, increased risks for disease in carriers of the variant alleles of the *CYP1A1* A4889G [22, 24] and T6235CC [22, 25–28] and the *GSTM1* [8, 23, 25, 26, 28–32] and *GSTT1* null [26, 33] genotypes were also

previously reported. The divergent results obtained in our study and two previous Brazilian studies [25, 30] could be attributed to the different composition of HNSCC subtypes and selection of controls. In the studies of HNSCC, tumours of the oral cavity, pharynx and larynx have traditionally been combined on the basis of similar aetiology (tobacco and alcohol) and sensitivity to radiotherapy and systemic therapy [34]. However, with advances in understanding of biology of HNSCC, tumours that were once considered to be relatively homogeneous diseases are now being recognized as distinct subtypes [34]. In fact, the SCC of oral cavity was more common in the patients analysed by Gattás et al. [30] than in our group (40.0% versus 20.4%,  $P=0.001$ ). In addition, the frequency of the *GSTM1* null genotype was lower in the Olivieri's [25] controls, including hospital-based patients without cancer,

**Table 3** The frequencies of the genotypes in 142 head and neck squamous cell carcinoma patients stratified by clinical variables

Genotypes	Age at diagnosis (years)		Tobacco consumption <sup>a</sup>		Squamous cell carcinoma subtype		
	≤ 57	> 57	Moderate	Heavy	Oral cavity	Pharynx	Larynx
<i>CYP1A1</i> A4889G							
AA	52 (69.3)	42 (62.7)	47 (77.0)	43 (58.1)	21 (72.4)	48 (63.2)	25 (67.6)
AG+GG	23 (30.7)	25 (37.3)	14 (23.0)	31 (41.9)***	8 (27.6)	28 (36.8)	12 (32.4)
<i>CYP1A1</i> T6235C							
TT	54 (72.0)	36 (53.7)	42 (68.8)	44 (59.5)	16 (55.2)	45 (59.2)	29 (78.4)
TC+CC	21 (28.0)	31 (46.3)*	19 (31.2)	30 (40.5)	13 (44.8)	31 (40.8)	8 (21.6)*****
Both <i>CYP1A1</i>							
AA+TT	38 (84.4)	26 (63.4)	34 (85.0)	27 (65.9)	13 (72.2)	32 (68.1)	19 (90.5)
AG+GG+TC+CC	7 (15.6)	15 (36.6)*	6 (15.0)	14 (34.1)	5 (27.8)	15 (31.9)	2 (9.5)
<i>GSTM1</i>							
Present	46 (61.3)	38 (56.7)	38 (62.3)	41 (55.4)	19 (65.5)	42 (55.3)	23 (62.2)
Null	29 (38.7)	29 (43.3)	23 (37.7)	33 (44.6)	10 (34.5)	34 (44.7)	14 (37.8)
<i>GSTT1</i>							
Present	57 (76.0)	49 (73.1)	49 (80.3)	51 (68.9)	22 (75.9)	56 (73.7)	28 (75.7)
Null	18 (24.0)	18 (26.9)	12 (19.7)	23 (31.1)	7 (24.1)	20 (26.3)	9 (24.3)
<i>CYP1A1</i> A4889G+ <i>GSTT1</i>							
AA+Present	41 (85.4)	31 (81.6)	38 (92.7)	31 (73.8)	16 (88.9)	37 (80.4)	19 (86.4)
AG+GG+Null	7 (14.6)	7 (18.4)	3 (7.3)	11 (26.2)****	2 (11.1)	9 (19.6)	3 (13.6)
<i>CYP1A1</i> T6235C+ <i>GSTM1</i>							
TT+Present	31 (83.8)	22 (59.5)	24 (82.8)	26 (63.4)	12 (66.7)	24 (64.9)	17 (89.5)
TC+CC+Null	6 (16.2)	15 (40.5)**	5 (17.2)	15 (36.6)	6 (33.3)	13 (35.1)	2 (10.5)*****
<i>CYP1A1</i> T6235C+ <i>GSTM1</i> + <i>GSTT1</i>							
TT+Present+Present	20 (90.9)	15 (75.0)	18 (94.7)	15 (71.4)	10 (76.9)	13 (76.5)	12 (100.0)
TC+CC+Null+Null	2 (9.1)	5 (25.0)	1 (5.3)	6 (28.6)	3 (23.1)	4 (23.5)	0 (0.0)

<sup>a</sup> The total number ( $n=135$ ) differed from the total quoted in Table 1 ( $n=142$ ) because it was not possible to obtain consistent information about tobacco exposure in some individuals

\* $P=0.03$ ; \*\* $P=0.02$ ; \*\*\* $P=0.02$ , when compared with controls,  $P=0.03$  (OR=2.0); \*\*\*\* $P=0.04$ , when compared with controls,  $P=0.04$  (OR=2.8); \*\*\*\*\* $P=0.02$ , larynx versus others subtypes; \*\*\*\*\* $P=0.04$ , larynx versus others subtypes.  $P$  values are presented herein after adjustment by the multivariate analysis

than in our blood donor's controls (3.4% versus 46.5%,  $P<0.001$ ). The reason for the differences obtained in our study and Indian studies [22, 24, 26, 27] may be attributed to the ethnic origin of individuals, since the *CYP1A1* 4889AG+GG genotype was over-represented in Sabitha's [24] controls (57.8% versus 26.8%,  $P<0.0001$ ) and the *CYP1A1* 6235TC+TT genotype was underrepresented in Sam's [22, 26, 27] and Sabitha's [24] controls compared with our control group (54.2% versus 35.9%,  $P=0.03$ ; 54.8% versus 35.9%,  $P=0.01$ ; respectively). In addition, excesses of the *GSTM1* plus *GSTT1* null genotypes were seen in our controls compared with Indian controls [26, 28] (7.8% versus 1.4%,  $P=0.0002$ ). However, the divergent results obtained among our study and others around the world [22, 23, 26–29, 31–33] in HNSCC risk could not be attributed to the differences in the ethnic origin, since similar frequencies of the genotypes were seen in our and

their controls. They could be attributed, again, to the different composition of HNSCC subtypes in studies. In fact, the SCC of oral cavity was more common in the Hamel et al. [33] (40.0% versus 20.4%,  $P=0.001$ ), Peters et al. [31] (51.5% versus 20.4%,  $P<0.0001$ ), Sam et al. [22, 26, 27] (53.5% versus 34.6%;  $P<0.0001$ ) and Sharma et al. [28] (36.0% versus 20.4%,  $P=0.002$ ) patients than in our group. In contrast, the pharyngeal SCC was more common in our patients group than in Hamel et al. [33], Peters et al. [31] and Capoluongo et al. [29] (53.5% versus 25.5%, versus 28.7% and versus 15.0%, respectively;  $P<0.0001$ ) and Sam et al. [22, 26, 27] (53.5% versus 41.4%;  $P=0.03$ ) groups. In addition, the laryngeal SCC was also over-represented in the Capoluongo's [29] patients than that found in our patients (65.0% versus 26.1%,  $P<0.0001$ ). Unfortunately, the different SCC subtypes were not computed in the two studies [23, 32].

Secondly, we found a higher frequency of the combined *CYP1A1* 6235TC+CC plus *GSTM1* plus *GSTT1* null genotype in patients with SCC of pharynx than in controls. To the best of our knowledge, there are no studies focusing on the influence of the combined *CYP1A1* T6235C, *GSTM1* and *GSTT1* genotype in pharyngeal SCC, suggesting for the first time that the variant alleles of referred genes act together increasing risk for this type of tumour.

Surprisingly, the *CYP1A1* 6235TC+CC genotype was less common in laryngeal SCC patients than in our controls, suggesting a role for the genotype in protection of tumour occurrence in our region. In contrast, increased risks for laryngeal SCC associated with the heterozygous *CYP1A1* A4889G and T6235C genotypes were seen in Portugal [35]. Moreover, no association of the *CYP1A1* T6235C polymorphism was also previously reported [9, 25, 30, 36–38].

Thirdly, we found associations of the *CYP1A1* 6235TC+CC genotype isolated and combined with the *CYP1A1* 4889AG+GG and the *GSTM1* null genotype with an older age of HNSCC onset, and the *CYP1A1* 6235TC+CC genotype isolated and combined with *GSTM1* null genotype with laryngeal tumours, suggesting that the genotypes may alter the clinical characteristics of patients and tumour location of the HNSCC in our sample. No additional studies focusing the roles of the referred genotypes on the age of HNSCC onset were previously reported. We also observed an excess of the *CYP1A1* 4889AG+GG plus *GSTT1* null combined genotype in heavy smoker's subgroup of HNSCC patients compared with controls, suggesting that genotypes are also related to increased risk of HNSCC in individuals who smoke heavily. Similar results were previously reported in India [24, 28] and USA [39]. In contrast, no association of these combined genotypes and smoking habits in HNSCC risk were found in the studies conducted in Germany [38, 40]. These differences may be attributed to the ethnic origin of individuals, since the frequencies of the *CYP1A1* 4889AG+GG genotypes were over-represented in our controls compared with their controls [38, 40] (26.8% versus 13.6%,  $P=0.007$  and 26.8% versus 8.3%,  $P<0.001$ , respectively).

Finally, the frequency of the *GSTM1* or *GSTT1* null genotype in heavy smoker patients with pharyngeal SCC was higher than that found in controls, suggesting that they alter risks for this subtype of tumour in individuals who smoke heavily. An association between oral cavity plus pharyngeal cancer and cumulative cigarette dose exposure among people with *GSTT1* null genotype was previously observed [41]. In fact, a proposal has been made for a mechanism whereby the *CYP1A1* variant genotypes and *GSTM1* null gene–gene interactions result in a greater-than-additive risk for DNA damage and cancer [6, 42]. Our data support the hypothesis that the *CYP1A1* and *GST* combined variant genotypes induces tobacco-associated DNA damage and HNSCC.

In conclusion, our results present preliminary evidence that the variant combined genotypes of the *CYP1A1* A4889G and T6235C, *GSTM1* and *GSTT1* polymorphisms are associated with increased risks for HNSCC and pharyngeal SCC, particularly in heavy smokers, and with the age of disease onset and location of the tumour in south-eastern Brazil. Facing these results, we believe that carriers of the *CYP1A1* and the *GST* variant genotypes, especially smokers, should be warned about the effects of the tobacco for the risk of HNSCC and should be submitted to periodic otolaryngologist evaluations with the purpose of performing an early diagnosis of a pharyngeal SCC in case of disease occurrence. However, we recognise that our conclusions are based on relatively small numbers of individuals and will require verification from an additional larger study enrolling age, gender, race, tobacco and alcohol habits matched patients and controls.

**Financial support** This study was funded by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and Financiadora de Estudos e Projetos do Ministério da Ciência e Tecnologia do Brasil (FINEP).

**Conflicts of interest** None

## References

1. Sturgis EM. The International Head and Neck Cancer Epidemiology consortium. *Head Neck*. 2011;33(1):141.
2. Hecht SS. Tobacco carcinogens, their biomarkers and tobacco-induced cancer. *Nat Rev Cancer*. 2003;3(10):733–44.
3. Hayes PC, Bouchier IAD, Beckett GJ. Glutathione S-transferase in humans in healthy and disease. *Gut*. 1991;32:813–8.
4. Lacko M, Oude Ophuis MB, Peters WH, Manni JJ. Genetic polymorphisms of smoking-related carcinogen detoxifying enzymes and head and neck cancer susceptibility. *Anticancer Res*. 2009;29(2):753–61.
5. Hayes JD, Flanagan JU, Jowsey IR. Glutathione transferases. *Annu Rev Pharmacol Toxicol*. 2005;45:51–88.
6. Bartsch H, Nair U, Risch A, Rojas M, Wikman H, Alexandrov K. Genetic polymorphism of CYP genes, alone or in combination, as a risk modifier of tobacco-related cancers. *Cancer Epidemiol Biomarkers Prev*. 2000;9(1):3–28.
7. Geisler SA, Olshan AF. *GSTM1*, *GSTT1*, and the risk of squamous cell carcinoma of the head and neck: a mini-HuGE review. *Am J Epidemiol*. 2001;154(2):95–105.
8. Hashibe M, Brennan P, Strange RC, Bhisey R, Cascorbi I, Lazarus P, et al. Meta- and pooled analyses of *GSTM1*, *GSTT1*, *GSTP1*, and *CYP1A1* genotypes and risk of head and neck cancer. *Cancer Epidemiol Biomarkers Prev*. 2003;12(12):1509–17.
9. Hiyama T, Yoshihara M, Tanaka S, Chayama K. Genetic polymorphisms and head and neck cancer risk (Review). *Int J Oncol*. 2008;32(5):945–73.
10. Marcopito LF, Coutinho AP, Valencich DM, Moraes MA, Brumini R, Ribeiro SA. Exposition to smoking and attitudes: comparison

- between inquiries applied to adult population in São Paulo city, 1987 and 2002. *Arq Bras Cardiol.* 2007;89:333–40.
11. Galduróz JCF, Carlini EA. Use of alcohol among the inhabitants of the 107 largest cities in Brazil—2001. *Braz J Med Biol Res.* 2007;40:367–75.
  12. Carvalho-Silva DR, Santos FR, Rocha J, Pena SD. The phylogeography of Brazilian Y-chromosome lineages. *Am J Hum Genet.* 2001;68(1):281–6.
  13. Mori S, Nose M, Morikawa H, Sato A, Saito T, Song ST, et al. A novel evaluation system of metastatic potential of oral squamous cell carcinoma according to the histopathological and histochemical grading. *Oral Oncol.* 1998;34(6):549–57.
  14. American Joint Committee on Cancer. In: *Cancer staging handbook from the AJCC cancer staging annual.* 6th ed. Philadelphia: Springer; 2002. p. 27–98.
  15. Huang WY, Winn DM. Alcohol concentration and risk of oral cancer in Puerto Rico. *American J Epidemiol.* 2003;157(10):881–7.
  16. Cascorbi I, Brockmüller J, Roots I. A C4887A polymorphism in exon 7 of human CYP1A1: population frequency, mutation linkages, and impact on lung cancer susceptibility. *Cancer Res.* 1996;56(21):4965–9.
  17. Arruda VR, Lima CS, Grignoli CR, de Melo MB, Lorand-Metze I, Alberto FL, et al. Increased risk for acute myeloid leukaemia in individuals with glutathione S-transferase mu 1 (GSTM1) and tetha 1 (GSTT1) gene defects. *Eur J Haematol.* 2001;66:383–8.
  18. Biselli JM, de Angelo Calsaverini Leal RC, Ruiz MT, Goloni-Bertollo EM, Maniglia JV, Rossit AR, et al. GSTT1 and GSTM1 polymorphism in cigarette smokers with head and neck squamous cell carcinoma. *Braz J Otorhinolaryngol.* 2006;72(5):654–8.
  19. Goloni-Bertollo EM, Biselli JM, Corrêa LC, Maniglia JV, Rossit AR, Ruiz MT, et al. Evaluation of the influence of GSTT1 and GSTM1 null genotypes in head and neck carcinogenesis. *Rev Assoc Med Bras.* 2006;52(5):365–8.
  20. Leme CV, Raposo LS, Ruiz MT, Biselli JM, Galbiatti AL, Maniglia JV, et al. GSTM1 and GSTT1 genes analysis in head and neck cancer patients. *Rev Assoc Med Bras.* 2010;56(3):299–303.
  21. Reszka E, Czekaj P, Adamska J, Wasowicz W. Relevance of glutathione S-transferase M1 and cytochrome P450 1A1 genetic polymorphisms to the development of head and neck cancers. *Clin Chem Lab Med.* 2008;46(8):1090–6.
  22. Sam SS, Thomas V, Reddy SK, Surianarayanan G, Chandrasekaran A. CYP1A1 polymorphisms and the risk of upper aerodigestive tract cancers in an Indian population. *Head Neck.* 2008;30(12):1566–74.
  23. Singh AP, Shah PP, Ruwali M, Mathur N, Pant MC, Parmar D. Polymorphism in cytochrome P4501A1 is significantly associated with head and neck cancer risk. *Cancer Invest.* 2009;27(8):869–76.
  24. Sabitha K, Reddy MV, Jamil K. Smoking related risk involved in individuals carrying genetic variants of CYP1A1 gene in head and neck cancer. *Cancer Epidemiol.* 2010;34(5):587–92.
  25. Olivieri EH, da Silva SD, Mendonça FF, Urata YN, Vidal DO, Faria MA, et al. CYP1A2\*1C, CYP2E1\*5B, and GSTM1 polymorphisms are predictors of risk and poor outcome in head and neck squamous cell carcinoma patients. *Oral Oncol.* 2009;45(9):e73–9.
  26. Sam SS, Thomas V, Reddy KS, Surianarayanan G, Chandrasekaran A. Gene–gene interactions of drug metabolizing enzymes and transporter protein in the risk of upper aerodigestive tract cancers among Indians. *Cancer Epidemiol.* 2010;34(5):626–33.
  27. Sam SS, Thomas V, Reddy KS, Surianarayanan G, Chandrasekaran A. Gene–environment interactions associated with CYP1A1 MspI and GST polymorphisms and the risk of upper aerodigestive tract cancers in an Indian population. *J Cancer Res Clin Oncol.* 2010;136(6):945–51.
  28. Sharma R, Ahuja M, Panda NK, Khullar M. Combined effect of smoking and polymorphisms in tobacco carcinogen-metabolizing enzymes CYP1A1 and GSTM1 on the head and neck cancer risk in North Indians. *DNA Cell Biol.* 2010;29(8):441–8.
  29. Capoluongo E, Almadori G, Concolino P, Bussu F, Santonocito C, Vendittelli F, et al. GSTT1 and GSTM1 allelic polymorphisms in head and neck cancer patients from Italian Lazio Region. *Clin Chim Acta.* 2007;376(1–2):174–8.
  30. Gattás GJ, de Carvalho MB, Siraque MS, Curioni OA, Kohler P, Eluf-Neto J, et al. Genetic polymorphisms of CYP1A1, CYP2E1, GSTM1, and GSTT1 associated with head and neck cancer. *Head Neck.* 2006;28(9):819–26.
  31. Peters ES, McClean MD, Marsit CJ, Lockett B, Kelsey KT. Glutathione S-transferase polymorphisms and the synergy of alcohol and tobacco in oral, pharyngeal, and laryngeal carcinoma. *Cancer Epidemiol Biomarkers Prev.* 2006;15(11):2196–202.
  32. Trizna Z, Clayman GL, Spitz MR, Briggs KL, Goepfert H. Glutathione S-transferase genotypes as risk factors for head and neck cancer. *Am J Surg.* 1995;170(5):499–501.
  33. Hamel N, Karimi S, Hébert-Blouin MN, Brunet JS, Gilfix B, Ghadirian P, et al. Increased risk of head and neck cancer in association with GSTT1 nullizygosity for individuals with low exposure to tobacco. *Int J Cancer.* 2000;87(3):452–4.
  34. Rischin D. Oropharyngeal cancer, human papilloma virus, and clinical trials. *J Clin Oncol.* 2010;28(1):1–3.
  35. Varzim G, Monteiro E, Silva RA, Fernandes J, Lopes C. CYP1A1 and XRCC1 gene polymorphisms in SCC of the larynx. *Eur J Cancer Prev.* 2003;12(6):495–9.
  36. Chatzimichalis M, Xenellis J, Tzagaroulakis A, Sarof P, Banis K, Gazouli M, et al. GSTT1, GSTM1, GSTM3 and NAT2 polymorphisms in laryngeal squamous cell carcinoma in a Greek population. *J Laryngol Otol.* 2010;124(3):318–23.
  37. Gajecka M, Rydzanicz M, Jaskula-Sztul R, Kujawski M, Szyfyer W, Szyfyer K. CYP1A1, CYP2D6, CYP2E1, NAT2, GSTM1 and GSTT1 polymorphisms or their combinations are associated with the increased risk of the laryngeal squamous cell carcinoma. *Mutat Res.* 2005;574(1–2):112–23.
  38. Gronau S, Koenig-Greger D, Jerg M, Riechelmann H. Gene polymorphisms in detoxification enzymes as susceptibility factor for head and neck cancer? *Otolaryngol Head Neck Surg.* 2003;128(5):674–80.
  39. Olshan AF, Weissler MC, Watson MA, Bell DA. GSTM1, GSTT1, GSTP1, CYP1A1, and NAT1 polymorphisms, tobacco use, and the risk of head and neck cancer. *Cancer Epidemiol Biomarkers Prev.* 2000;9(2):185–91.
  40. Harth V, Schafer M, Abel J, Maintz L, Neuhaus T, Besuden M, et al. Head and neck squamous-cell cancer and its association with polymorphic enzymes of xenobiotic metabolism and repair. *J Toxicol Environ Health A.* 2008;71(13–14):887–97.
  41. Jourenkova-Mironova N, Voho A, Bouchardy C, Wikman H, Dayer P, Benhamou S, et al. Glutathione S-transferase GSTM1, GSTM3, GSTP1 and GSTT1 genotypes and the risk of smoking-related oral and pharyngeal cancers. *Int J Cancer.* 1999;81(1):44–8.
  42. Matthias C, Bockmühl U, Jahnke V, Petersen I, Diemel M, Fryer A, et al. Effect of gene polymorphism on detoxifying glutathione-S-transferase enzymes on chromosomal stability of squamous epithelial carcinomas in the area of the head-neck. *Laryngorhinootologie.* 1998;77(4):201–6.