Role of a genetic variation in the microRNA-4421 binding site of ERP29 regarding risk of oropharynx cancer and prognosis

Juliana Carron¹, Ana Paula Dalla Costa², José Augusto Rinck-Junior², Fernanda Viviane Mariano³, Benilton de Sá Carvalho¹, Carmen Silvia Passos Lima² & Gustavo Jacob Lourenço¹

We conducted a two-stage association study on patients with oropharynx (OP) squamous cell carcinoma (SCC) and healthy controls to identify single nucleotide variants (SNVs) located at the microRNA (miR)-binding sites of carcinogenesis genes associated with risk and prognosis of the disease. In stage 1, 49 patients and 49 controls were analyzed using Genome-Wide Human SNV Arrays to identify variants in the 3′-untranslated region (3′-UTR) of carcinogenesis-related genes, and one SNV was selected for data validation in stage 2 by TaqMan assays in 250 OPSCC patients and 250 controls. The ERP29 c.*293A > G (rs7114) SNV located at miR-4421 binding site was selected for data validation among 46 SNVs. The ERp29 and miR-4421 levels were evaluated by quantitative-PCR and Western blotting. Interaction between miR-4421 with 3′-UTR of ERP29 was evaluated by luciferase reporter assay. Event-free survival (EFS) was calculated by Kaplan–Meier and Cox methods. ERP29 GG variant genotype was more common in OPSCC patients than in controls (6.4% vs 3.6%, \( p = 0.02 \); odds ratio: 5.67; 95% confidence interval (CI) 1.27–25.26). Shorter EFS were seen in the base of tongue (BT) SCC patients with GG genotype (0.0% vs 36.2%, \( p = 0.01 \); hazard ratio: 2.31; 95% CI: 1.03–5.15). Individuals with ERP29 AG or GG genotypes featured lower levels of ERP29 mRNA (\( p = 0.005 \)), ERp29 protein (\( p < 0.001 \)) and higher levels of miR-4421 (\( p = 0.02 \)). The miR-4421 showed more efficient binding with 3′-UTR of the variant G allele when compared with wild-type allele A (\( p = 0.001 \)). Our data suggest that ERP29 rs7114 SNV may alter the risk and prognosis of OPSCC due to variation in the ERp29 production possibly modulated by miR-4421.

Smoking habit and alcohol consumption consist in the classical risk factors for developing oropharynx (OP) squamous cell carcinoma (SCC)ⁱ. Sexual behavior is also established as a risk factor for human papillomavirus (HPV)-related OPSCC³. Most of OPSCC patients are diagnosed with measurable locally advanced disease⁴, and only about half of these patients achieve complete or partial responses after 5-years survival⁴.

Endoplasmic reticulum (ER) protein 29 (p29) is a chaperone protein that functions in unfolding and facilitating transport of synthesized secretory proteins from the ER to Golgi². Its function in cancer has been actively addressed, but its role in cancer development and progression is still unclear⁶. The ERp29 expression was found to be inversely associated with tumor development in the lung⁷, breast⁴, and gallbladder⁶; moreover, it inhibited breast tumor formation in mice¹⁰. In contrast, ERp29 overexpression was observed in breast, melanoma, lung, cervical⁴, liver⁶, and metastatic colorectal cancer cell line.

Additionally, ERp29 was found to regulate breast cancer cell growth arrest through p38 activation and upregulation of the ER stress protein p58IPK¹⁴, and cancer cell survival against genotoxic stress induced by doxorubicin¹⁵, cisplatin (CDDP)¹⁷, gemcitabine¹⁸, and radiation¹⁹,²⁰. ERp29 overexpression was associated with unfavorable prognosis of colorectal cancer through activation of chloride intracellular channel 4 and second...
and drinkers or abstemious, as previously reported. The Institutional Research Committee of University of and ethnicity from the same Teaching Hospital. All subjects were classified as either smokers and non-smokers,  between June 2000 and April 2016. The control group comprised 250 blood donors of the same sex and OPSCC risk, and in stage 2, one SNV was selected for data validation in 250 OPSCC patients and 250 controls. SNVs previously identified are located in non-coding regions of the genome, including the 3′-untranslated region (3′-UTR). SNVs located in the 3′-UTR of genes could influence micro-RNAs (miR) binding and direct posttranscriptional repression of genes involved in carcinogenesis, including the OPSCC. Besides, SNVs can affect the ability of a protein to fold and remain stable inside cells, often leading to diseases.

The population of Brazil is highly heterogeneous and admixed, as a result of cross-breeding among native Amerindians, Europeans settlers and immigrants, and sub-Saharan Africans. Since SNVs in genes with importance in OPSCC risk may not have been selected in previous studies, we conducted a two-stage association study on patients with OPSCC and healthy controls from the Southeast region of Brazil identifying a SNV in the ERP29 (c.293A > G, rs71114) associated with risk of tumor and prognosis. Moreover, we found that it is a functional SNV that alters the binding of miR-4421 and ERP29 levels.

Material and methods

Study population. This study was conducted in two stages. In stage 1, 49 patients and 49 controls were analyzed with the purpose of identifying SNVs on miR-binding sites of carcinogenesis genes with importance in OPSCC risk, and in stage 2, one SNV was selected for data validation in 250 OPSCC patients and 250 controls. All OPSCC patients were diagnosed at the Clinical Oncology Service of the University of Campinas Teaching Hospital between June 2000 and April 2016. The control group comprised 250 blood donors of the same sex and ethnicity from the same Teaching Hospital. All subjects were classified as either smokers and non-smokers, and drinkers or abstemious, as previously reported. The Institutional Research Committee of University of Campinas approved the study (numbers: 424/2016 and 1.438.601). All procedures were carried out according to the Helsinki Declaration, and appropriate informed consent form was obtained.

OPSCC was diagnosed according to World Health Organization criteria. Histologically, the OPSCC was classified as well, moderately, poorly differentiated, or undifferentiated. In addition, the OPSCC was staged according to the TNM system of the 7th American Joint Committee of Cancer Staging.

HPV status was performed in available tumor fragments embedded in paraffin of 98 OPSCC patients. We could not obtain the HPV status from 152 patients due to unavailable tumor fragments. P16 immunohistochemistry and in situ hybridization were performed in tumor fragments, aiming to test the presence of human papillomavirus type 16. For survival analysis, we selected 226 patients; 24 out of 250 patients were sent to other services for treatment and follow-up, and no consistent clinical information could be obtained. Patients were treated according to the institutional protocol, based on conventional procedures. Patients with locoregional advanced resectable tumors received neoadjuvant treatment (n = 19) or adjuvant treatment (n = 36) with 35 sessions of radiation, 2 Gy per session, with concurrent intravenous CDDP at a dose of 80–100 mg/m² or carboplatin area under the curve of 5 on days 1, 22, and 43, before or after surgery, respectively. For 168 patients, definite treatment with 35 sessions of radiation, 2 Gy per session, and concurrent intravenous CDDP or carboplatin at aforementioned doses, was administered, to whom surgical treatment was not performed because of locoregional unresectable tumors, low Karnofsky performance scale score, refusal of surgery due to expected functional or anatomic sequel, or an organ preservation protocol. We could not obtain treatment information from 3 patients enrolled in survival analysis due to lack of consistent clinical information. Patients’ follow-up was performed at 3-month intervals. The end of the follow-up period was January 2019.

Stage 1: screening of SNVs, candidate gene choice and SNV selection. Ninety-eight individuals were analyzed in the first stage. Genomic DNA from peripheral blood of 49 base of tongue (BT) SCC patients and 49 controls was genotyped for a total of 500,568 SNVs using the Affymetrix Genome-Wide Human SNV Array 5.0 (AFFYMETRIX, USA) according to the manufacturer’s recommended protocols. The intensities resulting from the arrays scanning process were made available via CEL files, one per DNA sample with total quality control higher than 90% (AFFYMETRIX, USA). Tools from the Bioconductor (www.biocosductorm.org) were used to process the CEL files. The genotyping was performed applying the corrected robust linear mixture model (crllmm) algorithm.

After association analysis (patients vs controls), SNVs located in the 3′-UTR of genes previously reported to be associated with, or known to be involved in carcinogenesis pathways, were selected. The analysis of carcinogenesis pathways was performed using The Database for Annotation, Visualization and Integrated Discovery and Kyoto Encyclopedia of Genes and Genomes pathway maps. SNVs showing significant deviation from Hardy–Weinberg (HW) equilibrium in controls, and those with minor allele frequency less than 10%, were excluded from the selection. SNV selection across each of carcinogenesis-related genes was carried out calculating their sample size based on the genotypic frequencies observed in healthy individuals from different ethnic populations and also using MicroSNiPer and MirSNPscore algorithms to select variations in miRNA binding sites sequences of 3′-UTR. We select the SNVs that presented a sample size less or equal than 250 individuals, which is the total patients’ sample available in our biorepository.
We selected the miRNAs that match six, seven, or eight nucleotides in seed region (position 2–8 of miRNA). The RNAhybrid software was used for finding the energetically most favorable hybridization sites using the minimum free energy (MFE) of hybridization of −20 kcal/mol or less. As a result, in this study we selected the ERP29 rs7114 for further validation.

Stage 2: validation of SNV ERp29 rs7114 in OPSCC risk. Five hundred individuals were analyzed in the second stage, including the 98 individuals analyzed preliminary in stage 1. Genomic DNA from leukocytes of peripheral blood of 250 OPSCC patients and 250 controls was analyzed by real-time polymerase chain reaction (PCR), using TaqMan SNV genotyping assay (assay reference: C_7521976_10, APPLIED BIOSYSTEMS, USA) for ERp29 rs7114 genotyping, according to the manufacturer's instructions. The amount of 20% of genotype determination was carried out twice in independent experiments with 100% agreement.

ERP29 expression by quantitative PCR. Total RNA from leukocytes of peripheral blood of 55 healthy individuals with distinct genotypes of ERp29 (22 individuals with AA, 23 with AG and 10 with GG genotypes) was extracted with TRIzol reagent (LIFE TECHNOLOGIES, USA). Experiments were performed using SYBR Green PCR Master Mix reagents (APPLIED BIOSYSTEMS, USA) and specific primers for the ERp29 gene (forward: 5′-CAGAGGTGGGATCTCAGATTAT-3′, and reverse: 5′-GAAGACTGGGTA GCTCTCTTTTGC-3′), in triplicate per sample, and a negative control without template was included in each plate. The relative expression level of ERp29 was normalized to the reference housekeeping gene actin beta level (forward: 5′-AGGCGAACCGCAGAAG-3′, and reverse: 5′-ACAGGCTGGATGCAAGCTACA-3′) using the 2−ΔΔCt cycle threshold method. Values of 20% of the samples were repeated in separate experiments with 100% agreement. Results were expressed in arbitrary units (AUs).

miR-4421 expression. Specific cDNA for miR-4421 and RNU24 (endogenous control) were generated using TaqMan MicroRNA Reverse Transcription kit (LIFE TECHNOLOGIES, USA), according to the manufacturer's instructions. The cDNA (n = 44) was obtained from previously extracted total RNA (14 individuals with AA genotype of ERp29, 18 individuals with AG and 12 individuals with GG). Reactions of qPCR were performed using TaqMan Universal PCR Master Mix II, no UNG (LIFE TECHNOLOGIES, USA) and TaqMan MicroRNA Assay RT-PCR (miR-4421: assay ID 464860_mat and RNU24: assay ID: 001001; LIFE TECHNOLOGIES, USA), in triplicate per sample, and a negative control without cDNA was included in all plate reactions. The relative expression level of miR-4421 was normalized to miRNA RNU24 using the 2−ΔΔCt cycle threshold method. Results were expressed in arbitrary units (AUs).

Western blotting. Leukocytes of peripheral blood of 28 healthy individuals with distinct ERp29 genotypes (15 individuals with AA genotype, 10 individuals with AG and 3 with GG) were used for extraction of total proteins. Briefly, cells were lysed with RIPA buffer containing protease inhibitors. Total protein concentrations were measured by Lowry protein assay. The cell lysates (40 µg) were subjected to 12% SDS-PAGE, and proteins were transferred to nitrocellulose membranes. The proteins reacted with rabbit anti-ERp29 monoclonal antibody (1:1000, sc-47724; SANTA CRUZ, GBR) overnight at 4 °C. A horseradish peroxidase-conjugated goat anti-rabbit IgG antibody was used as the secondary antibody (1:10,000, ab97051; ABCAM, GBR). Chemiluminescent signals were visualized in the capture imaging system ImageQuant 350 (GE HEALTHCARE, SWE) using SuperSignal West Pico PlusChemiluminescent Substrate (THERMO SCIENTIFIC, USA), and signal intensity was analyzed by the ImageJ software (NIH). The level of GAPDH was used as loading control.

Human pharynx SCC cell line culture. The human pharynx SCC cell line (FaDu) (ATCC HTB-43) was cultured in Dulbecco's modified Eagle's (DMEM) medium (GIBCO, USA) supplemented by 10% fetal bovine serum (FBS) and 100 µg/ml penicillin–streptomycin (SIGMA-ALDRICH, DEU) in an incubator at 37 °C with humidified atmosphere of 5% CO2. FaDu cell line was authenticated using short tandem repeat analysis. All experiments were performed with mycoplasma-free cells.

Construction of plasmids. The 3′-UTR of ERp29 rs7114_A (wild-type allele) and rs7114_G (variant allele) mRNA (266 bps) of individuals with known ERp29 AA and GG genotypes, respectively, were amplified by PCR using 2U of Platinum Taq DNA Polymerase High Fidelity (THERMO SCIENTIFIC, USA) and specific primers with restriction site for SpeI (forward: 5′-GCACTATGCTTGGGATCTCAGATTG-3′, where the SpeI site is underlined) and MluI (reverse: 5′-ATACGCTTAGATCCGCTTAGG-3′, where the MluI site is underlined). Fragments were cloned into the pMIR-REPORT miRNA Expression Reporter Vector (AMBION, USA) immediately downstream of the firefly luciferase gene driven by the CMV promoter, using standard protocols. After procedures, the plasmids pMIR_rs7114_A and pMIR_rs7114_G were obtained.

Dual luciferase reporter assay. FaDu cells were transiently transfected with the plasmids pMIR_rs7114_A, pMIR_rs7114_G, Renilla luciferase control reporter (pRL) (normalizing control) (PRIMEGA, USA), and synthetic sequences of miR-4421 mimics and inhibitor mimics (AMBION, USA), using Lipofectamine 2000 (INVITROGEN, USA), according to the manufacturer's instructions. FaDu cells were seeded in DMEM medium 48 h prior to transfection. In summary, 2×104 cells were transfected in four different groups: (1) 10 ng
of pMIR-ERP29_AG co-transfected with 5 ng of pRL and 50 nM of miR-4421; (2) 10 ng of pMIR-ERP29_AA co-transfected with 5 ng of pRL and 50 nM of miR-4421 inhibitor; (3) pMIR-ERP29-AA co-transfected with miR-4421 mimics; (3) pMIR-ERP29_AG (ERP29 rs7114 AG genotype) co-transfected with miR-4421 inhibitor; and (4) pMIR-ERP29_AG co-transfected with miR-4421 mimics, in pharynx squamous cell carcinoma cell line, FaDu (ATCC). (*) FaDu cells co-transfected with pMIR-ERP29_AG and miR-4421 mimics presented lower luciferase activity when compared with those co-transfected with pMIR-ERP29_AA and miR-4421 mimics (p < 0.001). (**) FaDu cells co-transfected with pMIR-ERP29_AG and miR-4421 inhibitor featured an increase in luciferase activity when compared with those co-transfected with pMIR-ERP29_AG and miR-4421 mimics (p < 0.001). (F) Probability of event-free survival (EFS) of 102 base of tongue squamous cell carcinoma patients stratified by ERP29 rs7114 SNV genotypes. The Kaplan–Meier curve indicates lower EFS in patients with the GG variant genotype (0.0% vs 36.2%, p = 0.01) when compared with patients with AA or AG genotypes. All statistical data analyses were performed using SPSS version 21.0 (www.ibm.com/analytics/spss-statistics-software).

Statistical analysis. Association between disease statuses, BTSCC patients vs controls, and genotypes for stage 1 was performed using the logistic regression model. These analyses were adjusted by age at diagnosis, sex and skin color. SNVs that presented raw p-values below the 0.01 thresholds were selected for further inspection. These analyses were implemented in R software (www.r-project.org).

The HW equilibrium was tested using chi-square (χ²) statistics for the goodness-of-fit test. Differences between groups were analyzed by χ² or Fisher's exact test. Multivariate analysis using logistic regression model served to obtain age- and tobacco status-adjusted crude odds ratios (ORs) with 95% confidence intervals (CI), and to assess associations between genotypes and OPSCC in stage 2. Power of analysis (PA) was used to calculate the minimum effect size that is likely to be detected in a study using a given sample size (DSS Research: https://bit.ly/2Fe79BD). χ² and Fisher's exact tests were used to evaluate possible associations among clinical characteristics, tumor aspects, and the genotypes of the selected SNV. Considering continuous variables, data sets were probed for normality using Shapiro–Wilk's test. For the ERP29 gene expression, the data set assumed normal distribution and t test was used for analysis. For ERP29 protein content, miR-4421 expression and luciferase assay, data sets did not assume normal distribution, thus, we used the Mann–Whitney test to compare the groups.

For survival analysis, the event-free survival (EFS) was calculated from the date of diagnosis until the date of progression of disease, the first relapse, death by disease, or the last follow-up. Overall survival (OS) was calculated from the date of diagnosis until the date of death, resulting from any cause, or the date of last follow-up. EFS and OS times were calculated using Kaplan–Meier (K–M) estimate probabilities, and differences between survival curves were analyzed by the log-rank test. The prognostic impact of age at diagnosis, sex, histological grade, TNM stage and ERP29 genotypes in survival of OPSCC patients was examined using Cox proportional hazard ratio (HR) regression. In a second step, all variables with p < 0.15 were included in a multivariate Cox regression (backward conditional step wise selection). For all statistical tests, significance was two-sided and achieved when p-values were ≤ 0.05. Tests were done using the SPSS 21.0 software (SPSS INCORPORATION, USA).

Results

Study population. We present demographic data and smoking and alcohol habits of 250 OPSCC patients and 250 controls in Table S1 Supplement. Control individuals were younger than OPSCC patients (median age: 44 vs 56 years, p < 0.001), and the number of tobacco and alcohol users was higher among patients than in controls (89.6% vs 13.2%, p < 0.001; 78.8% vs 49.2%, p < 0.001; respectively). Differences in age and pattern of tobacco and alcohol habits of individuals of each group were corrected in all comparisons of genotype frequencies by pertinent statistical analyses.
A

5’- CCAGGUCUUUGACAGAUGUAUUUCU - 3’ ERP29 (wild-type allele A)
3’- AUCGAGGAAAGGGUGUCUGUCCA - 5’ miR-4421

5’- CCAGGUCUUUGACAGGUGUAUUUCU - 3’ ERP29 (variant allele G)
3’- AUCGAGGAAAGGGUGUCUGUCCA - 5’ miR-4421 (6mer site) MFE: -27.9 kcal/mol

B

C

D

E

F

p = 0.005

p < 0.001

p = 0.02

p = 0.001

p < 0.001

p = 0.01

Probability of event-free survival (%)

AA or AG (n = 95)

GG (n = 7)

Time (in months)
We show frequencies of tumor characteristics of OPSCC patients in Table S2 Supplement. Most patients presented tumors located in the BT (46.4%), moderately differentiated tumors (64.8%), and advanced tumor stage (IV) (72.8%). HPV type 16 was positive in only 8 out of 98 OPSCC patients analyzed in the study.

Patients were treated with CDDP, radiotherapy (RT), and surgery. One hundred and sixty eight patients (68.6%) were submitted to chemotherapy (CT) and RT combined treatment; 47 patients (19.2%) received CT, RT, and were submitted to surgery; 12 patients (4.9%) received RT and surgery; 7 patients (2.9%) only received CT; 6 patients (2.4%), only RT; and 5 patients (2.0%) were only submitted to surgery. We could not obtain consistent information about the therapeutics of 5 patients.

Stage 1: analysis of screening of SNVs. We present clinical and pathological characteristics of the 98 participants included in stage 1 in Table S3 Supplement. After screening of SNVs (49 BTSCC patients vs 49 controls), we observed 6,609 SNVs associated with risk of BTSCC; 3461 (52.4%) of them were located in regulatory regions; 3,045 (46.0%), in introns; 52 (0.8%), in coding regions; 46 (0.7%), in 3′-UTR; and 5 (0.1%) SNVs, in 5′-UTR. Data on the genome association were deposited at the Gene Expression Omnibus database (www.ncbi.nlm.nih.gov/geo/) with accession number GSE46812.

Stage 1: SNV selection. After the SNVs screening, 16 of 46 SNVs located in the 3′-UTR were related to genes of carcinogenesis pathways (Table S4 Supplement). Only six of the 16 SNVs presented the appropriate sample size (n = 250): MYO6 rs6914716, SLC7A11 rs2370512, JMJD6 rs2240774, SLC20A1 rs5237012, SLC7A1 rs7674870, MYO6 rs6914716, TUSC1 rs1128957 and ERP29 rs7114 (Table S4 Supplement). Among them, we select the SNV ERP29 rs7114 for further validation in stage 2. We found that miR-4421 (MFE: − 27.9 kcal/mol) matched 6mer site complementary in seed sequence of 3′-UTR of variant allele G of ERP29 (rs7114), while the wild-type allele A disrupted these target sites.

Stage 2: analysis of validation of ERP29 rs7114. Samples of 250 OPSCC patients (χ² = 0.82, p = 0.36) and 250 controls (χ² = 1.10, p = 0.29) were in HW equilibrium at the ERP29 rs7114 locus. ERP29 GG variant genotype (6.4% vs 3.6%, p = 0.02) and G allele (23.2% vs 16.4%, p = 0.02) were more common in OPSCC patients than in controls. Individuals with GG genotype were under 5.67-fold increased risk of OPSCC than those with the remaining genotypes (Table 1). Moreover, individuals carrying at least one variant allele G were under 2.13-fold increased risk of OPSCC than carriers of wild-type A allele (Table 1).

Association between clinical and tumorous aspects with ERP29 genotypes. No associations between ERP29 rs7114 genotypes were seen in OPSCC patients stratified by age, sex, tobacco and alcohol consumption, histological grade, tumor stage, and tumor localization (Table 2).

\[
\text{ERP29 expression. Lower mRNA expression levels were seen in individuals with ERP29 AG (n = 23) (1.17 AUs ± standard deviation (SD): 0.45, p = 0.04) and GG (n = 10) (0.94 AUs ± SD: 0.25, p = 0.002) than in those with the AA genotype (n = 22) (1.44 AUs ± SD: 0.45). In addition, individuals with ERP29 AG or GG genotypes (n = 33) presented lower mRNA levels than those with the AA genotype (n = 22) (1.10 AUs ± 0.41 SD vs 1.44 AUs ± 0.45 SD, p = 0.005) (Fig. 1B).}
\]

\[
\text{ERP29 protein content. Lower ERp29 protein content was observed in individuals with ERP29 AG (n = 16) (0.79 AUs ± SD: 0.48, p < 0.001) and GG (n = 4) (0.98 AUs ± SD: 0.58, p = 0.03) than in those with the AA genotype (n = 11) (2.58 AUs ± SD: 1.44). In addition, individuals with ERP29 AG or GG genotypes (n = 20)}
\]

| Genotypes | OR (95% CI) |
|-----------|-------------|
| AA vs AG  | 1.69 (0.83–3.43) |
| AA vs GG  | 6.7 (1.27–25.26) |
| AG vs GG  | 2.58 (1.04–4.05) |
| AA or AG  | 5.5 (1.9–35.38) |
| AA or AG  | 2.13 (1.20–3.78) |

| Genotypes | OR (95% CI) |
|-----------|-------------|
| AA vs AG  | 1.69 (0.83–3.43) |
| AA vs GG  | 6.7 (1.27–25.26) |
| AG vs GG  | 2.58 (1.04–4.05) |
| AA or AG  | 5.5 (1.9–35.38) |
| AA or AG  | 2.13 (1.20–3.78) |

Table 1. ERP29 rs7114 genotypes in 250 oropharynx squamous cell carcinoma and 250 controls. OPSCC oropharynx squamous cell carcinoma, n number of patients or controls, OR odds ratio adjusted by age, smoking and drinking status, CI confidence interval. The significant values are indicated by bold letters.

*Power of analysis (PA): 30.0%; **PA: 30.0%; ***PA: 72.0%; ****PA: 77.1%.
presented lower ERp29 protein content than those with the AA genotype (n = 11) (0.83 AUs ± 0.48 SD vs 2.58 AUs ± 1.44 SD, \( p < 0.001 \)) (Fig. 1C).

**miR-4421 expression.** Higher miR-4421 expression levels were observed in individuals with ERP29 AG (n = 18) (0.60 AUs ± SD: 0.28, \( p = 0.01 \)) than in those with the AA genotype (n = 14) (0.36 AUs ± SD: 0.16). Similar miR-4421 expression levels were perceived in individuals with ERP29 GG (n = 12) (0.79 AUs ± SD: 0.76, \( p = 0.21 \)) than in those with the AA genotype (n = 14) (0.36 AUs ± SD: 0.16). However, individuals with ERP29 AG or GG genotypes (n = 30) presented higher miR-4421 expression than those with the AA genotype (n = 14) (0.67 AUs ± 0.52 SD vs 0.36 AUs ± 0.16 SD, \( p = 0.02 \)) (Fig. 1D).

**Luciferase reporter assay.** FaDu cells co-transfected with pMIR_rs7114_G (containing variant homozygous GG genotype) and miR-4421 mimics featured lower luciferase activity when compared with those co-transfected with pMIR_rs7114_A and miR-4421 mimics (\( p = 0.001 \)) (Fig. 1E). Besides, FaDu cells co-transfected with pMIR_rs7114_G and miR-4421 inhibitor featured an increase in luciferase activity when compared with those co-transfected with pMIR_rs7114_A and miR-4421 mimics (\( p < 0.001 \)) (Fig. 1E).

**Survival analysis of OPSCC patients.** We obtained consistent survival data from 226 OPSCC patients. The median follow-up of patients enrolled in survival analysis was 33.3 months (range 1.5–162.2 months). The final status of patients was established on January 2019. At this time, 60 patients were alive without disease, 6 patients were alive with disease, 123 patients died due to disease, and 37 patients died due to unrelated causes. The 5-year EFS and OS rates were 42.8% and 38.7%, respectively.

At 60 months of follow-up, lower EFS was observed in patients with large tumors (T3 or T4) (37.1% vs 61.6%, \( p = 0.002 \)), with advanced nodal stage (N2 or N3) (31.2% vs 57.8%, \( p < 0.001 \)), and with tumors in the BT (33.6% vs 50.2%, \( p = 0.01 \)) (K–M estimates) compared with others. The significance of differences between groups remained the same in univariate Cox analysis. After multivariate Cox analysis, large tumors (T3 or T4) (HR 1.84, 95% CI 1.16–2.91, \( p = 0.009 \)) and advanced nodal stage (N2 or N3) (HR 1.88, 95% CI 1.30–2.71, \( p = 0.001 \)) were found to be predictors of poor EFS (Table S5 Supplement).

| Characteristics | n | ERP29 rs7114 |
|-----------------|---|-------------|
|                 |   | AA | AG | GG | AA or AG | GG | AA | AG or GG |
| **Sex**         |   |    |    |    |          |    |    |        |
| Male            | 228 | 135 (59.2%) | 79 (34.6%) | 14 (6.2%) | 214 (94.8%) | 14 (6.2%) | 135 (59.2%) | 93 (40.8%) |
| Female          | 22  | 15 (68.2%)  | 5 (22.7%)  | 2 (9.1%)  | 20 (90.9%)  | 2 (9.1%)  | 15 (68.2%)  | 7 (31.8%)  |
| **p-value**     |   | 0.33 | 0.67 | 0.63 | Reference | Reference | Reference | 0.49 |
| **Tobacco consumption** |   |    |    |    |          |    |    |        |
| Smokers         | 224 | 135 (60.3%) | 76 (33.9%) | 13 (5.8%) | 211 (94.2%) | 13 (5.8%) | 135 (60.3%) | 89 (39.7%) |
| Non-smokers     | 26  | 15 (57.7%)  | 8 (30.7%)  | 3 (11.6%) | 23 (88.5%)  | 3 (11.5%) | 15 (57.7%)  | 7 (31.8%)  |
| **p-value**     |   | Reference | 0.38 | 0.63 | Reference | Reference | Reference | 0.49 |
| **Alcohol consumption** |   |    |    |    |          |    |    |        |
| Drinkers        | 197 | 118 (59.9%) | 69 (35.0%) | 10 (5.1%) | 187 (94.9%) | 10 (5.1%) | 118 (59.9%) | 79 (40.1%) |
| Abstainers      | 53  | 32 (60.4%)  | 15 (28.3%) | 6 (11.3%) | 47 (88.7%)  | 6 (11.3%) | 32 (60.4%)  | 21 (39.6%) |
| **p-value**     |   | Reference | 0.61 | 0.20 | Reference | Reference | Reference | 1.00 |
| **Histological grade** |   |    |    |    |          |    |    |        |
| Well or moderately | 176 | 104 (59.1%) | 59 (33.5%) | 13 (7.4%) | 163 (92.6%) | 13 (7.4%) | 104 (59.1%) | 72 (40.9%) |
| Poorly or undifferentiated | 35  | 19 (54.3%)  | 14 (40.0%) | 2 (5.7%)  | 33 (94.3%)  | 2 (5.7%)  | 19 (54.3%)  | 16 (45.7%) |
| **p-value**     |   | Reference | 0.55 | 1.00 | Reference | Reference | Reference | 0.70 |
| **Tumor stage** |   |    |    |    |          |    |    |        |
| I or II         | 21  | 10 (47.6%)  | 11 (52.4%) | 0 (0.0%)  | 21 (100.0%) | 0 (0.0%)  | 10 (47.6%)  | 11 (52.4%) |
| III or IV       | 224 | 137 (61.5%) | 71 (31.4%) | 16 (7.1%) | 208 (92.9%) | 16 (7.1%) | 137 (61.5%) | 87 (38.5%) |
| **p-value**     |   | Reference | 0.15 | 0.60 | Reference | Reference | Reference | 0.25 |
| **Tumor localization** |   |    |    |    |          |    |    |        |
| Base of tongue  | 116 | 67 (57.8%)  | 41 (35.3%) | 8 (6.9%)  | 108 (93.1%) | 8 (6.9%)  | 67 (57.7%)  | 49 (42.3%) |
| Tonsillar complex | 84  | 52 (61.9%)  | 26 (30.9%) | 6 (7.2%)  | 78 (92.9%)  | 6 (7.1%)  | 52 (61.9%)  | 32 (38.1%) |
| Soft palate     | 45  | 27 (60.0%)  | 16 (35.5%) | 2 (4.5%)  | 43 (95.5%)  | 2 (4.5%)  | 27 (60.0%)  | 18 (40.0%) |
| **p-value**     |   | Reference | 0.80 | 0.83 | Reference | Reference | Reference | 0.83 |

Table 2. Association of ERP29 rs7114 genotypes, clinical and tumor characteristics of 250 oropharynx squamous cell carcinoma patients. *The number of patients differed from the total quoted in the study (n = 250), because it was not possible to obtain consistent information about histological grade and tumor stage in some cases. **The number of patients differed from the total quoted in the study (n = 250), because it was considered for analysis only the most frequent tumor localization.
At the same time of follow-up, a shorter OS was observed in patients with large tumors (T3 or T4) (30.3% vs 64.9%, \( p = 0.001 \)), with advanced nodal stage (N2 or N3) (29.9% vs 50.0%, \( p = 0.008 \)), and with tumors in the BT (26.6% vs 49.1%, \( p = 0.001 \)) (K–M estimates) compared with others. The significance of differences between groups remained the same in univariate Cox analysis. After multivariate Cox analysis, large tumors (T3 or T4) (HR 1.75, 95% CI 1.17–2.61, \( p = 0.006 \)), advanced nodal stage (N2 or N3) (HR 1.43, 95% CI 1.04–1.97, \( p = 0.02 \)), and tumors located on the BT (HR 1.53, 95% CI 1.11–2.10, \( p = 0.008 \)) were found to be predictors of poor OS (Table S5 Supplement).

The SNV ERP29 rs7114 did not influence the EFS and OS of our OPSCC patients (Table S5 Supplement).

Survival analysis of BTSCC patients. Considering only BTSCC patients, the median follow-up was 23.0 months (range 3.4–157.1 months). At the final follow-up, 16 patients were alive without disease, 2 patients were alive with disease, and 65 patients died due to disease, and 19 patients died due to unrelated causes. The five-year EFS and OS rates were 33.6% and 26.6%, respectively.

At 60 months of follow-up, lower EFS was observed in patients with advanced nodal stage (N2 or N3) (26.2% vs 49.1%, \( p = 0.001 \)) and in those with ERP29 GG genotype (0.0% vs 36.2%, \( p = 0.01 \)) (Fig. 1F) (K–M estimates) compared with others. The significance of differences between groups remained the same in univariate Cox analysis. After multivariate Cox analysis, advanced nodal stage (N2 or N3) (HR 1.77, 95% CI 1.04–1.97, \( p = 0.02 \)), and tumors located on the BT (HR 1.53, 95% CI 1.11–2.10, \( p = 0.008 \)) were found to be predictors of poor OS (Table S5 Supplement).

The SNV ERP29 rs7114 did not influence the EFS and OS of our OPSCC patients (Table S5 Supplement).

### Table 3.

| Variables                        | N  | Univariate analysis | Multivariate analysis |
|----------------------------------|----|---------------------|-----------------------|
|                                 |    | Event-free survival | Overall survival      | Event-free survival | Overall survival |
|                                 |    | OR (95% CI)         | OR (95% CI)           | OR (95% CI)         | OR (95% CI)     |
|                                 |    | p value             | p value               | p value             | p value         |
| Age (years)                     | 102| ≤ 57                | 1.12 (0.69–1.82)      | 0.62                | 1.01 (0.65–1.56) | 0.95 | NA |
|                                 |    | > 57                | Reference             | Reference           | Reference       | 0.95 | NA |
| Sex                              | 102| Male                | 1.31 (0.52–3.28)      | 0.55                | 1.38 (0.60–3.19) | 0.44 | NA |
|                                 |    | Female              | Reference             | Reference           | Reference       | 0.44 | NA |
| Histological grade              | 83 | Well or moderately  | 1.03 (0.52–2.00)      | 0.93                | 1.29 (0.70–2.38) | 0.40 | NA |
|                                 |    | Poorly or undifferentiated | 1.91 (0.87–4.19)    | Reference          | Reference       | 0.10 | Reference |
| Tumor size                      | 101| T1 or T2            | 1.91 (0.87–4.19)      | 0.10                | 2.33 (1.12–4.84) | 0.02 | 2.36 (1.13–4.92) |
|                                 |    | T3 or T4            | Reference             | Reference           | Reference       | 0.12 | Reference |
| Nodal stage                     | 102| N0 or N1            | 1.85 (1.10–3.12)      | 0.11                | 1.43 (0.91–2.24) | 0.03 | Reference |
|                                 |    | N2 or N3            | Reference             | Reference           | Reference       | 0.10 | Reference |
| ERP29 rs7114                    | 102| AA                  | 1.03 (0.64–1.68)      | 0.88                | 1.44 (0.73–1.76) | 0.55 | NA |
|                                 |    | AG or GG            | Reference             | Reference           | Reference       | 0.21 | Reference |
|                                 |    | AA or AG            | Reference             | Reference           | Reference       | 2.31 (1.03–5.15) | 0.04 | NA |
|                                 |    | GG                  | 2.60 (1.17–5.76)      | 0.01                | 1.64 (0.74–3.59) | 0.21 | Reference |

At the same time of follow-up, a shorter OS was perceived only in patients with large tumors (T3 or T4) (20.2% vs 71.4%, \( p = 0.02 \)) (K–M estimates) compared with others. The significance of differences between groups remained the same in univariate Cox analysis. After multivariate Cox analysis, large tumors (T3 or T4) (HR 2.36, 95% CI 1.13–4.92, \( p = 0.02 \)) were found to be predictors of poor OS (Table 3). No association was found between ERP29 rs7114 genotypes and OS.

### Discussion

We investigated whether the ERP29 SNV rs7114 (c.293A > G) alters the risk of OPSCC and prognosis of patients with the disease. Besides, we also investigated the role of the distinct alleles (A and G) of the referred SNV in ERP29 expression, ERp29 protein content, miR-4421 expression and its interaction with miR-4421 in pharynx SCC cell line.
ERp29 is an ER chaperone protein and plays a role in protein maturation, secretion, and intercellular communication. Although the controversial role of ERp29 in tumor development and progression, ERp29 is a potential tumor suppressor in cancer. Bambang et al. observed that ERp29 acts upregulating a group of genes with tumour suppressor function such as E-cadherin number 1 (CDH1), cyclin-dependent kinase inhibitor 2B (CDKN2B), and spleen tyrosine kinase (SYK). Moreover, ERp29 downregulated a group of genes involved in cell proliferation such as epidermal growth factor receptor (EGFR), plasminogen activator receptor (uPAR), cyclin D2 (CCND2), and serine/threonine kinase 1 (AKT).

Initially, we observed that the variant ERp29 GG genotype was more common in OPSCC patients than in controls, and that individuals with the referred genotype were under 5.67-fold increased risk of OPSCC than others. There are no studies focusing on the role of the referred SNV in the risk of OPSCC or any disease. The ERp29 SNV rs7114 determines the exchange of adenine (A) by guanine (G) at the 293 positions of the ERp29 3′-UTR, and the variant allele G creates a functional binding site for miR-4421. miRNAs inhibit mRNA translation by directly binding to the 3′-UTR of target mRNAs, often accompanied by mRNA degradation. In fact, it was already described that miRNAs are important regulatory molecules in OPSCC. Overexpression of miR-4421 was associated with esophageal carcinoma development. However, the role of miR-4421 in the regulation of gene expression is still unknown.

Actually, we observed that individuals carrying ERp29 AG or GG genotypes presented lower ERp29 expression and ERp29 protein content than individuals carrying wild-type AA genotype. Besides, individuals with AG or GG genotypes also presented higher levels of miR-4421 than individuals with the AA genotype.

Additionally, we cloned the 3′-UTR of ERp29 into a miRNA expression vector co-transfected with miR-4421 in pharynx SCC cells (FaDu) to verify whether ERp29 is a target gene of miR-4421, and to determine the affinity of the different alleles of the SNV rs7114 with 3′-UTR. We observed that cells co-transfected with the variant GG genotype and miR-4421 presented lower luciferase activity when compared with those co-transfected with wild-type AA genotype and miR-4421. This result indicated that miR-4421 downregulated ERp29 expression by targeting the SNV region and, indeed, miR-4421 had more affinity with the variant G allele than with the wild-type A allele. It’s worth noting that miR-4421 is physiologically expressed in FaDu cells.

All in all, our results support the fact that individuals with ERp29 AG or GG genotypes were under increased risk of OPSCC due to ERp29 downexpression, possibly modulated by miR-4421, and consequent loss of the tumour suppressor function.

We also observed that tumors located in BT and those in advanced stages were found to be predictors of poor survival, as reported in previous studies. In fact, the OPSCC consists of a group of heterogeneous tumors with a variety of clinical characteristics, thus, we performed the survival analysis only in BTSCC patients.

We observed that BTSCC patients with ERp29 GG variant genotype had worst EFS. None of the genotypes of the studied SNV have influenced the prognosis of patients in previous studies.

Besides the increase in cell proliferation, downexpression of ERp29 was associated with higher cancer cell motility and invasion, and worst prognosis of pancreatic ductal adenocarcinoma and gastric cancer patients. In fact, ERp29 can drive MET in breast cancer cells and it may have a critical role in promoting distant metastasis during cancer progression. Furthermore, ERp29 downexpression was associated with decreased apoptosis in curcumin-treated breast cancer cells and in fibroblasts and thyrocytes from null ERp29 mice.

In contrast, ERp29 downexpression was associated with decreased RT resistance in nasopharyngeal carcinoma cells, increased CDDP efficacy in lung cancer cell line with null p53, and better prognosis of colorectal cancer patients. Clearly, understanding the association of ERp29 with disease recurrence and distant metastasis is noteworthy for assessing its prognostic value in clinical applications.

Therefore, BTSCC patients carrying the ERp29 GG variant genotype may present worst EFS due to lower ERp29 levels, leading to activation of cell proliferation, loss of cell adhesion, and MET deregulation. In summary, we identified that inheritable abnormality in ERp29 modulates OPSCC occurrence and acts as an independent prognostic factor for EFS of BTSCC patients. We identified that ERp29 rs7114 SNV is capable of modulating ERp29 levels, possible due to miR-4421 affinity. These findings, once validated by studies with functional protein analyses and large sample sizes, will assist in individualizing the medical care provided to patients, in which high-risk patients should receive a closer follow-up.

Data availability
The authors declare that all data of the present study are available for the corresponding author upon reasonable request.

Received: 13 May 2020; Accepted: 24 August 2020
Published online: 12 October 2020

References
1. Mendenhall, W., Werning, I. & Pfister, D. Treatment of head and neck cancers. In Cancer: Principles & Practice of Oncology 9th edn (eds De Vita, V. T. et al.) 729–780 (Lippincott Williams & Wilkins, Philadelphia, 2011).
2. Elrefaey, S. HPV in oropharyngeal cancer: The basics to know in clinical practice. Acta Otorhinolaryngol. Ital. 34(5), 299–309 (2014).
3. Belcher, R., Hayes, K., Fedewa, S. & Chen, A. Y. Current treatment of head and neck squamous cell cancer. J. Surg. Oncol. 110(5), 551–574 (2014).
4. O’Sullivan, B. et al. Development and validation of a staging system for HPV-related oropharyngeal cancer by the International Collaboration on Oropharyngeal cancer Network for Staging (ICON-S): A multicentre cohort study. Lancet Oncol. 17(4), 440–451 (2016).
5. Sargsyan, E. et al. Identification of ERp29, an endoplasmic reticulum luminal protein, as a new member of the thyroglobulin folding complex. J. Biol. Chem. 277(19), 17009–17015 (2002).
6. Chen, S. & Zhang, D. Friend or foe: Endoplasmic reticulum protein 29 (ERp29) in epithelial cancer. FEBS Open Bio. 5, 91–98 (2015).
7. Shnyder, S. D., Mangum, J. E. & Hubbard, M. J. Triplex profiling of functionally distinct chaperones (ERp29/PDI/BiP) reveals marked heterogeneity of the endoplasmic reticulum proteome in cancer. J. Proteome Res. 7(8), 3364–3372 (2008).
8. Bambang, I. F., Lee, Y. K., Richardson, D. R. & Zhang, D. Endoplasmic reticulum protein 29 regulates epithelial cell integrity during the mesenchymal-epithelial transition in breast cancer cells. Oncogene 32(10), 1240–1251 (2013).
9. Yuan, L. W., Liu, D. C. & Yang, Z. L. Correlation of S1P1 and ERp29 expression with progression, metastasis, and poor prognosis of gallbladder adenocarcinoma. Hepatobiliary Pancreat. Dis. Int. 12(2), 189–195 (2013).
10. Bambang, I. F. et al. Overexpression of endoplasmic reticulum protein 29 regulates epithelial-mesenchymal transition and suppresses xenograft tumor growth of invasive breast cancer cells. Lab. Invest. 89(11), 1229–1242 (2009).
11. Myung, J. K. et al. Expressional patterns of chaperones in ten human tumor cell lines. Proteome Sci. 2(1), 8 (2004).
12. Fung, Y., Tian, Z. M., Wan, M. X. & Zheng, Z. P. Protein profile of human hepatocarcinoma cell line SMCC-7721: Identification and functional analysis. World J. Gastroenterol. 13(18), 2608–2614 (2007).
13. Deng, Y. J. et al. CLIC4, ERp29, and Smac/DIABLO derived from metastatic cancer stem-like cells stratify prognostic risks of colorectal cancer. Clin. Cancer Res. 20(14), 3809–3817 (2014).
14. Gao, D. et al. ERp29 induces breast cancer cell growth arrest and survival through modulation of activation of p38 and upregulation of ER stress protein p88IPK. Lab. Invest. 92(2), 200–213 (2012).
15. Zhang, D. & Putti, T. C. Over-expression of ERp29 attenuates doxorubicin-induced cell apoptosis through up-regulation of Hsp27 in breast cancer cells. Exp. Cell Res. 316(20), 3522–3531 (2010).
16. Farmaki, E. et al. ERp29 regulates response to doxorubicin by a PERK-mediated mechanism. Biochim. Biophys. Acta 1813(6), 1165–1171 (2011).
17. Zhang, Y. et al. Proteomic identification of ERp29 as a key chemoresistant factor activated by the aggregating p53 mutant Arg282Trp. Oncogene 36(39), 5473–5483 (2017).
18. Ye, W. et al. ERp29 downregulation enhances lung adenocarcinoma cell chemosensitivity to gemcitabine by upregulating HSP27 phosphorylation. Exp. Ther. Med. 17(1), 817–823 (2019).
19. Qi, L. et al. Inhibiting ERp29 expression enhances radiosensitivity in human nasopharyngeal carcinoma cell lines. Med. Oncol. 29(2), 721–728 (2012).
20. Wu, P. et al. Identification of ERp29 as a biomarker for predicting nasopharyngeal carcinoma response to radiotherapy. Oncol. Rep. 27(4), 987–994 (2012).
21. Zhang, K. et al. Comparison of ILK and ERp29 expressions in benign and malignant pancreatic lesions and their clinicopathological significances in pancreatic ductal adenocarcinomas. Clin. Transl. Oncol. 18(4), 352–359 (2016).
22. Wu, J. et al. ERp29 inhibits tumorigenicity by suppressing epithelial mesenchymal transition in gastric cancer. Oncotarget 8(45), 87787–87766 (2017).
23. Ye, J. et al. ERp29 controls invasion and metastasis of gastric carcinoma by inhibition of epithelial-mesenchymal transition via PI3K/Akt-signaling pathway. BMC Cancer 17(1), 626 (2017).
24. Xu, S. G., Yan, P. J. & Shao, Z. M. Differential proteomic analysis of a highly metastatic variant of human breast cancer cells using two-dimensional differential gel electrophoresis. J. Cancer Res. Clin. Oncol. 136(10), 1545–1556 (2010).
25. Linga, A. et al. Differential expression of fourteen proteins between uveal melanoma from patients who subsequently developed distant metastases versus those who did Not. Invest. Ophthalmol. Vis. Sci. 53(8), 4634–4643 (2012).
26. Guo, L. et al. ERp29 counteracts the suppression of malignancy mediated by endoplasmic reticulum stress and promotes the metastasis of colorectal cancer. Oncol. Rep. 41(3), 1603–1613 (2019).
27. Lesueur, C. et al. Genome-wide association analyses identify new susceptibility loci for oral cavity and pharyngeal cancer. Nat. Genet. 48(12), 1544–1550 (2016).
28. Preskill, C. & Weidhaas, J. B. SNPs in microRNA binding sites as prognostic and predictive cancer biomarkers. Crit. Rev. Oncog. 18(4), 327–340 (2013).
29. Robert, F. & Pelletier, J. Exploring the impact of single-nucleotide polymorphisms on translation. Front. Genet. 9, 507 (2018).
30. Manta, F. S. N. et al. Revisiting the genetic ancestry of Brazilians using autosomal AIM-Indels. PLoS ONE 8(9), e75145 (2013).
31. Costa, E. F. et al. Association between polymorphisms in genes related to DNA base excision repair with risk and prognosis of oropharyngeal squamous cell carcinoma. J. Cancer Res. Clin. Oncol. 142(9), 1917–1926 (2016).
32. Cardesa, A., Gale, N., Nadal, A. & Zidar, N. Squamous cell carcinoma. In Pathology & Genetics, Head and Neck Tumours. World Health Organization Classification of Tumours, 4th edn (eds Damjanov, I. & Fang, F.) 9–29 (IARC Press, Lyon, 2005).
33. Gale, N. & Zidar, N. Tumours of the head and neck. In Cancer Grading Manual 2nd edn (eds Damjanov, I. & Fang, E.) 9–29 (Springer, Heidelberg, 2013).
34. Edge, S. B. & Compton, C. C. The American Joint Committee on Cancer: The 7th edition of the AJCC cancer staging manual and the future of TNM. Ann. Surg. Oncol. 17(6), 1471–1474 (2010).
35. Singh, A. D. & Westra, W. H. Comparison of human papillomavirus in situ hybridization and p16 immunohistochemistry in the detection of human papillomavirus-associated head and neck cancer based on a prospective clinical experience. Cancer 116(9), 2166–2173 (2010).
36. Carvalho, B. S., Louis, T. A. & Irizarry, R. A. Quantifying uncertainty in genotype calls. Bioinformatics 26(2), 242–249 (2010).
37. Huang, D. W. & Sherman, B. T. & Lempicki, R. A. Bioinformatics enrichment tools: Paths to the comprehensive functional analysis of large gene lists. Nucleic Acids Res. 37(1), 1–13 (2009).
38. Kanehisa, M. & Goto, S. KEGG: Kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 28(1), 27–30 (2000).
39. International Hapmap Consortium et al. Integrating common and rare genetic variation in diverse human populations. Nature 467(7311), 52–58 (2010).
40. Barenboim, M., Zoltick, B., Guo, Y. & Weinberger, D. R. MicroSNiPer: A web tool for prediction of SNP effects on putative microRNA targets. Hum. Mutat. 31(11), 1223–1230 (2010).
41. Thomas, L. F., Saito, T. & Setrom, P. Inferring causative variants in microRNA target sites. Nucleic Acids Res. 39(16), e109 (2011).
42. Rehmsmeier, M., Steffen, P., Hochsmann, M. & Gögerich, R. Fast and effective prediction of microRNA/target duplexes. RNA 10(10), 1507–1517 (2004).
43. Capes-Davis, A. et al. Match criteria for human cell line authentication: Where do we draw the line? Int. J. Cancer 132(11), 2510–2519 (2013).
44. Costa, E. F. et al. Association between polymorphisms in genes related to DNA base excision repair with risk and prognosis of oropharyngeal squamous cell carcinoma. J. Cancer Res. Clin. Oncol. 142, 1917–1926 (2016).
45. Bartel, D. P. MicroRNAs: Genomics, biogenesis, mechanism, and function. Cell 116(2), 281–297 (2004).
46. Gunawardena, I. et al. Micro-ribonucleic acids in head and neck cancer: An introduction. J. Laryngol. Otol. 127, 52–57 (2013).
47. Drahos, J. et al. MicroRNA profiles of Barrett’s esophagus and esophageal adenocarcinoma: Differences in glandular non-native epithelium. Cancer Epidemiol. Biomark. Prev. 25(3), 429–437 (2016).
48. Rouillard, A. D. et al. The harmonizome: A collection of processed datasets gathered to serve and mine knowledge about genes and proteins. Database (Oxford) https://doi.org/10.1093/database/baw100 (2016).
49. Prueksavanakon, K. et al. Survival and prognostic factors of different sites of head and neck cancer: An analysis from Thailand. Asian Pac. J. Cancer Prev. 13(3), 885–890 (2012).
50. Yasumatsu, R., Nakashima, T. & Komune, S. Squamous cell carcinoma of the oropharynx: Single-institution outcome analysis of patients treated with concurrent chemoradiotherapy. *J. Laryngol. Otol.* **129**(Suppl 2), S77–S82 (2015).
51. Fang, H. Y. *et al.* Proteomic identification of differentially expressed proteins in curcumin-treated MCF-7 cells. *Phytomedicine* **18**(6–9), 697–703 (2011).
52. Hirsch, I., Weiwad, M., Prell, E. & Ferrari, D. M. ERp29 deficiency affects sensitivity to apoptosis via impairment of the ATF6-CHOP pathway of stress response. *Apoptosis* **19**(5), 801–815 (2014).

**Acknowledgements**
We would like to thank “Fundação de Amparo à Pesquisa do Estado de São Paulo” (FAPESP) (Grants Numbers 2006/06231-0, 2012/17182-1, and 2015/18039-6) for the financial support.

**Author contributions**
J.C., C.S.P.L. and G.J.L. designed the study and wrote the manuscript. J.C., A.P.D.C., J.A.R.J. and F.V.M. contributed to experimental design and data acquisition. J.C., B.S.C. and G.J.L. performed the data analysis.

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

**Additional information**
Supplementary information is available for this paper at https://doi.org/10.1038/s41598-020-73675-z.

**Correspondence** and requests for materials should be addressed to G.J.L.

**Reprints and permissions information** is available at www.nature.com/reprints.

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

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

© The Author(s) 2020