Identification of polymorphisms associated with head and neck squamous cell carcinoma susceptibility in the Spanish population

Javier Fernández-Mateos, Raquel Seijas-Tamayo, Juan Carlos Adansa Klain, Miguel Pastor Borgoñón, Elisabeth Pérez-Ruiz, Ricard Mesía, Elvira del Barco-Morillo, Carmen Salvador Coloma, Antonio Rueda Dominguez, Javier Caballero Daroqui, Encarnación Fernández Ruiz, Rogelio González-Sarmiento, Juan Jesús Cruz-Hernández

International Journal of Cancer

Sometido a publicación

Factor de impacto 2015: 5.531

-Journal Citation Reports Science Edition (Thomson Reuters, 2015)

Oncology: 29/213 Q1

2

Artículo 2: "Identification of polymorphisms associated with head and neck squamous cell carcinoma susceptibility in the Spanish population"

A pesar del efecto de los factores de riesgo ambientales, la presencia de factores genéticos de predisposición desempeña también un papel importante en el desarrollo de CECC. La carcinogénesis del tracto aerodigestivo implica la desregulación de múltiples vías de señalización y variantes en los genes que regulan estas vías dan lugar a diferencias en la susceptibilidad a desarrollar ciertos tumores, entre ellos el CECC. Sin embargo, estos resultados son contradictorios debido a sesgos en el diseño de los estudios en cuanto a heterogeneidad de la población a estudio o en los factores de riesgo.

No todos los individuos expuestos a los carcinógenos clásicos desarrollan CECC, por lo que la identificación de variantes genéticas en vías de señalización que regulan la carcinogénesis podría definir el riesgo diferencial a desarrollar la enfermedad. Por ello, se planteó un estudio multicéntrico en el cual se seleccionaron 296 pacientes con CECC (126 laringes, 100 faringes y 70 cánceres orales) y se compararon con una población control totalmente pareada en sexo, consumo de tabaco y alcohol y edad. Se estudiaron un total de 29 polimorfismos en genes implicados en reparación de DNA, inflamación, apoptosis o metabolismo de carcinógenos previamente asociados con riesgo a desarrollar CECC, así como variaciones en la respuesta y/o toxicidad en el tratamiento. EL DNA se extrajo de sangre periférica mediante el método de fenol-cloroformo. La discriminación alélica se realizó mediante sondas TaqMan® o PCR y posterior digestión con enzimas específicas (RFLP).

El análisis descriptivo mostró una población control totalmente pareada en factores de riesgo y edad, excepto en el grupo de cáncer laríngeo donde la edad fue usada como factor de ajuste. Teniendo en cuenta la distribución entre localizaciones observamos diferencias estadísticamente significativas en la distribución de varios SNPs.

En cáncer laríngeo, los genotipos con el alelo menos común de los polimorfismos *ERCC1* rs11615 y *ERCC2* rs13181 fueron asociados con un menor riesgo a desarrollar este tumor. Ambos SNPs se localizan en genes que forman parte de la vía de reparación por escisión de nucleótidos (NER) y se han relacionado con una menor expresión génica y una menor reparación del DNA, así como una mejor respuesta y supervivencia tras el tratamiento con platino. La diferencia de nuestro resultado podría estar en la naturaleza retrospectiva del estudio, donde la mayoría de pacientes incluidos tendrían el alelo común asociado a una mayor reparación del DNA, proceso altamente activo en pacientes fumadores y bebedores.

La inflamación es otra de las vías esenciales en la carcinogénesis. Ésta es mediada por las citocinas, donde destacan IL-2 y la proinflamatoria IL-6. Las variantes rs1800975 en *IL6* y rs2069762 en *IL2* se ha asociado con una mayor expresión génica, previamente asociadas con un aumento en la inflamación. Nuestro estudio mostró como la variante en *IL6* produjo un riesgo incrementado de desarrollar carcinoma de faringe y cavidad oral, mientras que *IL2* rs2069762 disminuyó la susceptibilidad a cáncer oral, estando esta citocina más implicada en la autoinmunidad que en la inflamación protumoral.

Mdm2 es un regulador negativo de la proteína supresora de tumores p53, y su SNP rs2279744 en la zona promotora atenúa esta inhibición. Nuestros datos confirman el efecto visto en otros tumores donde el alelo variante se asocia con un riesgo incrementado a desarrollar cáncer laríngeo, debido a una menor actividad supresora mediada por p53.

El polimorfismo Pro72Arg rs1042522 en *TP53* es uno de los más estudiados, relacionándose con un incremento al desarrollo de tumores, aunque sus resultados son debatidos. Este SNP es esencial en el aumento de la respuesta apoptótica. Nuestros resultados mostraron un riesgo menor a desarrollar tumores de faringe y laringe en portadores del alelo antiapoptótico Pro72 y, aunque estos resultados son contrarios al aumento de riesgo establecido en otros trabajos, podría deberse tanto al desequilibrio en su distribución (HWE $p<0.05$), como en la mayor supervivencia del alelo Pro previamente descrita, aumentando el número de individuos en nuestra muestra retrospectiva.

En nuestra serie, el polimorfismo rs2279115, localizado en la zona promotora del gen anti-apoptótico *BCL2* aparece asociado con un mayor riesgo de desarrollar cáncer oral, confirmando resultados previos en CECC. Este SNP en el promotor P2 produce un aumento en la unión a proteínas nucleares disminuyendo la actividad del promotor activador P1 y, como consecuencia, una disminución de la apoptosis.

Por último, dos polimorfismos en el gen *NFE2L2* se asociaron con menor riesgo a desarrollar cáncer laríngeo y faríngeo. *NFE2L2* rs1303586 se asoció con menor riesgo al desarrollo de ambos tumores, mientras que rs2706110 se relacionó con menor riesgo únicamente en carcinoma de laringe. Aunque hasta el momento no se han presentado análisis funcionales de estos SNPs, nuestra hipótesis es que incrementarían la inducción de los genes antioxidativos ante estrés, disminuyendo así el riesgo en CECC.

Por lo tanto, este estudio muestra la asociación entre diferentes polimorfismos en genes implicados en reparación de DNA, inflamación, oxidación y apoptosis con un riesgo

diferencial a desarrollar CECC. Así mismo, las diferencias encontradas en este estudio según la localización tumoral ponen de manifiesto la heterogeneidad subyacente en estos cánceres.

Identification of polymorphisms associated with head and neck squamous cell carcinoma susceptibility in the Spanish population

Javier Fernández-Mateos[1,2,3,4], Raquel Seijas-Tamayo[1,2], Juan Carlos Adansa Klain[1,2], Miguel Pastor Borgoñón[5] , Elisabeth Pérez-Ruiz[6], Ricard Mesía[7], Elvira del Barco[1,2], Carmen Salvador Coloma[5], Antonio Rueda Dominguez[6], Javier Caballero Daroqui[5], Encarnación Fernández Ruiz[8], Rogelio González-Sarmiento[2,3,4,*],Juan Jesús Cruz-Hernández[1,2,3,4,*]

[1]Medical Oncology Service, University Hospital of Salamanca-IBSAL, Salamanca, 37007 Spain

[2]Biomedical Research Institute of Salamanca (IBSAL), SACYL-University of Salamanca-CSIC, Salamanca, 37007, Spain.

[3]Molecular Medicine Unit- IBSAL, Department of Medicine, University of Salamanca, 37007, Spain

[4]Institute of Molecular and Cellular Biology of Cancer (IBMCC), University of Salamanca-CSIC, Salamanca, 37007, Spain

[5]Medical Oncology Service, Hospital Universitario Politécnico La Fe, Valencia, 46026, Spain

[6]Division of Medical Oncology, Oncology department, Agencia Sanitaria Hospital Costa del Sol de Marbella, 29603, Spain

[7]Medical Oncology Department, Universitat de Barcelona, IDIBELL, Institut Català d'Oncologia, L'Hospitalet de Llobregat, Barcelona, 08908, Spain

[8]Otolaryngology Department, Agencia Sanitaria Hospital Costa del Sol de Marbella, 29603, Spain

Abstract

Head and neck squamous cell carcinoma (HNSCC) is the sixth most common cancer worldwide. Despite tobacco, alcohol and viral infection are the most common etiopathogenic factors, genetic predisposition is also an important cause. Not all individuals expose to the environmental carcinogens develop the illness so, the identification of polymorphism could define those variant associated with HNSCC susceptibility. In a multicentre case-control study 126 laryngeal, 100 pharyngeal and 70 oral cavity HPV- tumours with paired controls in sex, age and tobacco and alcohol consumption were included. A gene candidate association study of 29 selected SNPs in genes implied in important carcinogenic pathways was perfomed by TaqMan probes or RFLP assay in germinal DNA. The results showed that nine SNPs in genes implicated in DNA repair (*ERCC1* rs11615 and *ERCC2* rs13181), inflammation (*IL2* rs2069762 and *IL6* rs1800795), oxidation (*NFE2L2* rs13035806 and rs2706110) and apoptosis (*TP53* rs1042522, *MDM2* rs2279744 and *BCL2* rs2279115) were associated with HNSCC susceptibility in a Spanish population. The fact that these results change according to tumour location, confirm the diversity of the tumours grouped under the term of HNSCC.

Keywords: Head and neck squamous cell carcinoma, polymorphism, SNP, association study, susceptibility

Introduction

Head and neck cancer (HNC) includes a set of diverse neoplasms located in the lips, oral cavity, pharynx, larynx, salivary glands and thyroid glands among others[1]. Approximately 600.000 new cases are diagnosed per year, being the sixth cancer type on incidence worldwide. Treatment of early stages includes surgery and/or radiotherapy, while locally advanced tumours are also treated with chemotherapy and biological therapies[2]. Only the 40-50% reach five-years survival rate[3] causing and annual death of 271.000 patients[4,5]. Regarding its histological type, not all the HNC are the same, prevailing the squamous cell carcinomas[1].

Head and neck squamous cell carcinoma (HNSCC) development is a multifactorial process associated with a variety of risk factors. At least 75% HNSCC are attributable to the combination of cigarette smoking and alcohol drinking, the most classic carcinogens[6]. Diverse epidemiological studies have also revealed the existence of other environmental and genetic related factors. Similar to other tumours, viral aetiology has shown an implication in the development of HNSCC predominating Epstein-Barr virus (EBV) infection in nasopharynx and human papillomavirus (HPV), mainly subtype 16, in oral cavity and oropharynx tumours[6]. The carcinogenesis procedure triggered by viral infection define a different entity to that caused by tobacco and alcohol[8] allowing the classification of HNSCC in two main prognostic and therapeutic groups, in which HPV negative tumours are associated with an aggressive course and worst prognosis than HPV positive ones[9]. Despite the defined role of environmental factors there is also an evidence of familial aggregation and increased cancer risks amongst HNSCC relatives[9], suggesting the existence of genetic predisposition factors[10].

Not all individuals exposed to these carcinogens will suffer the disease so the identification of genetic variants in important signalling pathways could help to define tumour susceptibility as well as differences in response and toxicity to treatment. HNSCC carcinogenesis involves different pathways: carcinogen metabolism, DNA repair, cell cycle, immunity and inflammation[11–13]. Single Nucleotide Polymorphism (SNP) is the most abundant form of genetic variation, becoming an ideal genetic susceptibility marker[1].

Previously, diverse studies have evaluated the relationship between genetic variants and susceptibility to HNSCC[14–17]. However, the results were inconsistent because they were conducted in different populations with heterogeneity in the study design regarding risk factors, race or ethnicity.

In this context, we present the first association study of polymorphisms in genes involve in the main cancer pathways and susceptibility to develop HNSCC within a Spanish population whose control group was totally paired in most important risk factors (tobacco and alcohol consumption). Due to the equality in the risk factors between both groups, the importance lies on the genotype, assuming the result into a greater extent to the SNPs.

Material and methods

Study population

TTCC-2010-05 was an observational multicentre study conducted in 19 Spanish centres, all of them belonging to the Spanish Group of Head and Neck Cancer Treatment (TTCC) coordinated by the Medical Oncology Department of the University Hospital of Salamanca between January 2012 and December 2014. Epidemiological details have been previously described[18].

The inclusion criteria were histologically confirmed HPV negative HNSCC patients from larynx, oro/hypopharynx and oral cavity carcinomas. They were recruited in Oncology, Radiotherapy and Otorhinolaryngology departments. Controls were hospitalized patients without tumour history and paired by age, sex, smoking and alcoholism habit with the HNSCC cases. They were captured in Pneumology, Radiotherapy, Otorhinolaryngology and Internal Medicine departments. Only Spanish population were permitted, avoiding ethnicity bias.

Taking into account incidence in Spain, 10% of possible lost and duration of the study, initial calculations of recruitment were of 440 patients in each group. A total of 459 patients and 259 controls were included.

In this study the variables were polymorphisms in oncogenes, tumour suppressor genes, genes implicated in DNA reparation, inflammation, carcinogen metabolism and apoptosis, together with some risk factors collected in the socio-demographic (6 questions) and the data informed by patients (19 questions) questionnaires. The information of both questionnaires were collected by auto-application, being supervised by the members of the research team with the objective of correct filled. Clinicopathologic data as well as response and specific toxicity to treatment were collected by oncologists in the case report form questionnaire (CRF).

The study was approved by the University Hospital of Salamanca and the local ethics committees. All participants were previously informed and signed the provided written informed consent designed for this project. All data were treated with the security measures established in compliance with the Protection of Personal Data Organic Law 15/1999, 13th December, and safe-keeping by the University Hospital of Salamanca in its specific hospital server.

Selection of polymorphism

Candidate SNPs selection was done according to at least two of the following criteria: >5% allele frequency in Caucasian/European population, previously defined association with HNSCC susceptibility, earlier related different response or toxicity to chemotherapy or radiotherapy. At the initial stages of the project design, a huge search was performed in available databases using keywords as SNPs, susceptibility, HNSCC, response, toxicity, selecting only those with statistically significant results in other populations[12,14–17,19]. SNPs with some published evidence of functionality were preferred selected (Table 1).

DNA isolation and genotyping

DNA was extracted from leukocytes of peripheral blood by phenol-chloroform method. Genotyping was performed using the TaqMan® Allelic Discrimination Assay (Applied Biosystems, Foster, CA) in those SNPs where the probes are available. 40 ng/µl of DNA sample were added to 6.25 µl of Taqman® Universal PCR Master Mix and it was combined with specific forward and reverse primers, and allele-specific VIC (allele 1) and FAM (allele 2) labelled probes. The assay was performed in a 96 well plate and the detection was measured in the Step One Plus Real-Time PCR System Thermal Cycling Block (Applied Biosystems, Foster, CA). Negative and positive controls were always added. 5% of random samples were re-genotyping to ensure the reproducibility.

In those candidate SNPs where TaqMan® probes were not available, genotyping was analyzed by polymerase chain reaction- restriction fragment length polymorphism (PCR-RFLP). Specific oligonucleotides were designed to amplify the polymorphic sequences and digestion was made by the specific restriction enzymes. The PCR products were run on 3% Syber-safe stained agarose gel and visualized under UV light.

Statistical analysis

The statistical analysis to associate the relation between the different clinical and molecular variables was analyzed by cross tabs and the X^2 test of Pearson. The Odds ratio (OR) and 95% confidence intervals were calculated by logistic regression analysis. Quantitative variable distribution was analyzed by ANOVA test in those examples where the sample followed a parametric distribution ($p>0.05$ in Levene´s test); while in those with a non-parametric distribution; a Mann Whitney U test was applied. Hardy-Weinberg equilibrium (HWE) was tested in control population by X^2 test. It was considered the existence of statistically significant differences where the two sided P-value was <0.05. Only TP53 rs1042522 and APEX rs1130409 were in disequilibrium ($p_{HWE}<0.05$).

Because of the lower inclusion of matched controls, the statistical analysis was realized matching the group by the Propensity Score method (PS). This potent matching technique allowed us to equate groups in a cohort study through a logistic regression introducing the confounders as predictive variables[20]. Groups were matched according to: packs per year consumed (PPY): no smokers, <20PPY and >20PPY, standard unit of alcohol per week (SDU/week): <14SDU/week and >14SDU/week and sex. Quantitative age was not included in the PS and it was introduced in the logistic regression as adjustment variable only in laryngeal cancer where the age between both groups was statistically significant different.

These analyses were performed with the statistical software SPSS v.21.0 (IBM-SPSS Inc., Chicago).

Results

After the application of the Propensity Score method 126 larynx, 100 pharynx and 70 oral cavity squamous cell carcinomas were totally paired with the control group. Descriptive study of the analysis by locations did not show any statistically significant differences between sex, age, tobacco and alcohol intake respect the control group (Table 2). Only age between laryngeal tumour and control was significant ($p<0.05$) so this variable was included in the logistic regression as an adjustment variable.

The laryngeal cancer susceptibility study showed a different distribution in *TP53*, *MDM2*, *ERCC1*, *ERCC2*, *IL6* and *NRF2* selected SNPs (Supplementary table 1). Less common genotypes in *ERCC1* rs11615 (p=0.011, OR=0.288 (0.110-0.751) in recessive model) and *ERCC2* rs13181 (p=0.046, OR=0.375 (0.143-0.982) in codominant model) were associated with lower risk to suffer laryngeal cancer. This also happened with *TP53* rs1042522 in which its mutant allele C was associated with a decrease risk to develop laryngeal cancer (p=0.002, OR=0.286 (0.119-0.607) in recessive model). Variant allele in *MDM2* rs2279744, p=0.029 OR=2.413 (1.094-5.323) in recessive model, and *IL6* rs1800795 p=0.002 OR=2.394 (1.376-4.163) in dominant model, were related with a higher risk to suffer this tumour. Lastly, rs1303586 variant genotypes GA+AA and CT+TT in rs2706110, both in the *NRF2* gene, were associated with a lower risk to suffer this carcinoma, p=0.035 (OR=0.478 (0.240-0.949)) and p=0.518 (OR=0.518 (0.299-0.900)) respectively (Table 3).

Pharyngeal squamous carcinoma showed an association between CC genotype in *TP53* rs1042522 and a lower probability to suffer this illness, p=0.001, OR=0.124 (0.035-0.476) in recessive model. Also, an association between *NRF2* rs2706110 less common allele gentoypes CC+CT and lower risk to develop pharyngeal carcinoma was also found, p=0.043, OR=0.552 (0.311-0.982) (Supplementary table 2 and table 4).

Finally, the analysis in oral cavity squamous carcinoma showed significant differences in the distribution of the polymorphism in inflammatory and apoptotic genes (Supplementary table 3). Inflammatory gene *IL2* rs2069762 variant was associated with a lower risk of oral cavity cancer (GG p=0.039, OR=0.300 (0.096-0.940)) while the variants CG+GG in the pro-inflammatory *IL6* rs1800795 were related with increased oral cavity susceptibility, p=0.018, OR=2.265 (1.148-4.467). CA+AA genotypes in *BCL2* rs2279115 were related with a higher risk to develop from oral carcinoma p=0.010, OR=2.753 (1.273-5.952) in dominant model (Table 5).

Discussion

Not all individuals exposed to the most classic carcinogens (tobacco and alcohol) develop HNSCC. Several susceptibility studies have identified SNPs in carcinogenesis-related pathways but their results are controversial due to the design and the risk factors misunderstanding. In this multicentre case-control study we examined the association between some polymorphisms and HNSCC susceptibility in a Spanish cohort with a control group totally paired by its risk factors considering their genetic background on its own.

Analysis of laryngeal squamous cell carcinoma showed an association with lower susceptibility risk in the *ERCC1* rs11615 and *ERCC2* rs13181 SNPs. These genes are part of the nucleotide excision repair (NER) pathway, implicated in the repair of adducts produced by the tobacco, DNA oxidative damage, alkylating agents damage and thymine dimers[21]. *ERCC1* Asn118Asn (rs11615; c.354G>A) and *ERCC2* Lys751Gln (rs13181; c.2251A>G) variants have been related with an increased risk to develop cancer in a huge variety of populations[22], among them in HNSCC cancer[23,24]. These polymorphisms were linked to a minor mRNA level or with a deficient DNA repair capacity respectively[25,26]. rs11615 and rs13181 variant genotypes have been also associated to a better response and longer survival for higher damage accumulation in patients treated with platinum[27], triggering cell death. A possible explanation of these results in our series could be the retrospective nature of our study, where those mutated allele carriers with less reparation and more aggressive illness would have been died along the follow-up and most of longest survivors with common allele would have been included in our series. This distribution is confirmed in our data.

Inflammation has been considered an important factor in the pathogenesis of human cancer[28,29], with a special interest in the context of oral cancer[30]. Inflammation is mediated by cytokines, being IL-2 and pro-inflammatory IL-6 two of the most important ones. SNPs in interleukin genes can influence in the magnitude of the inflammatory response, contributing into a different reaction[31]. rs1800795 -174C variant in the promoter of the *IL6* gene is related with lower level of serum proteins while -174G corresponds with a higher expression, increasing the inflammatory response[32]. Cytokine IL-2 is important in the activated T-lymphocytes proliferation and the phagocytes activation to destroy the uptaken material (Th1). The SNP -330G>T (rs2069762) in the gene promoter has been identified with a differentiating protein production in healthy subject, in which G allele produces a rise in the *IL2* gene expression, whereas GT/TT carriers have a decrease *IL2* expression skewing the balance Th1/Th2 towards Th2[33]. GT/TT genotypes were related with a certain susceptibility to suffer from diverse illnesses such as multiple sclerosis and minor risk to develop gastric cancer[33]. Our result with *IL6* variant can be related with the carcinogenesis induced by inflammatory process[29] in which G allele in *IL6* rs1800975 was associated with a higher gene expression and largest inflammatory response, increasing laryngeal and oral tumour susceptibility. However, GG genotype in *IL2* rs2069762 was associated with lower oral cavity carcinoma risk, in contrast with reported associations[34]. This result could be explained by the main role of IL-2 in self-reactive cells elimination[35], decreasing the antitumour efficacy by immune system.

Mdm2 is an important negative regulator of the tumour suppressor protein p53. Both form a complex in which Mdm2 attenuates p53 through proteasomal degradation by ubiquitinylation, while p53 induces *MDM2* transcription in response to genotoxic stress[36]. SNP rs2279744 309T>G, is located in the promoter P2 and increase *MDM2* expression by improving the binding affinity with Sp1 transcription factor, attenuating TP53 suppressor pathway[37]. Several correlations between *MDM2* rs2279744 and higher risk of HNSCC have been reported[38]. Our data confirm these results, showing a higher risk to suffer from laryngeal cancer in those patients with the GG genotype, related with an increase *MDM2* expression and lower p53 tumor suppressing activity.

It was found an association between the SNP in *TP53* rs1042522 and a higher susceptibility to develop pharyngeal and laryngeal cancer. *TP53* is a tumour suppressor gene and it is one of the main regulators of the genome integrity, cell cycle and apoptosis mechanism[39]. The polymorphism c.215C>G in the exon 4 of *TP53* produces a change of proline by arginine in the codon 72 (Pro72Arg) of the protein, being this domain essential in the apoptotic response and the carcinogenesis inhibition. In this sense, the arginine allele in the codon 72 is an apoptotic inductor more powerful than the proline one[40,41]. Lot of studies have related the Pro72Arg polymorphism with an increased risk to develop gastric, oesophageal, bladder cancer[42,43], although its role is controversial. In HNSCC, it has been only associated in isolated studies[17,44]. Though our results showed the contrary relationship, the explanation of this fact is difficult to define because SNP distribution was in disequilibrium (p_{HWE}>0.05). One plausible explanation could be for the influence of this SNP in survival, in which Pro72 was associated with longer survival after life-threatening diseases[40], tendency also shown in our series in laryngeal cancer (data not shown). So, our retrospective protocol would have selected those longest survivors, increasing the number of CC (Pro72) variant.

In addition to previous results, we found a statistically significant result in the polymorphism of the anti-apoptotic gene *BCL2* and oral cavity cancer. *BCL2*-938C>A (rs2279115) polymorphism is found in the gene promoter P2, which acts as a negative regulator element, decreasing the activity of the promoter P1[45]. C allele is related with a rise in the P2 promoter activity joining to nuclear proteins. The presence of C allele reduces the activity of P1 and the Bcl-2 protein expression increasing apoptosis. Our results showed similar results to that reported on breast cancer and acute myeloid leukaemia[46] where the presence of the A allele (CA+AA) increased susceptibility to these tumours due to an anti-apoptotic effect, being related with a higher Bcl-2 protein expression. This effect was also reported by Chen et al where the rs2279115 polymorphism was associated with an increase risk to HNSCC development[47].

NFE2L2 codifies for a transcription factor family (Nrf2) that induce many antioxidative genes under oxidative stress. SNPs in this gene have been associate to cancer risk[48]. In our sample, *NFE2L2* rs2706110 and rs1303586 less common allele genotypes were associated with lower risk to develop laryngeal cancer while in pharyngeal cancer only rs1303586 was found. This effect was different from that reported on a higher risk to suffer from breast cancer, related with AA genotype[49]. Functional analyses of these SNPs have not yet been described but our hypothesis is that these changes could increase antioxidative genes induction under stress, produced at high level in HNSCC by tobacco and alcohol consumption. Besides, the fact that Nrf2 and inflammation pathways are related[50], could enhance our results.

Overall, this study has shown the association between some polymorphisms in DNA repair, inflammation, antioxidative and apoptosis genes with different susceptibility to develop HNSCC. The characteristics of the control group favoured that this results were caused by the genetic background, avoiding confounder variables. In the same way, the differences found in this association study according to the location corroborate the heterogeneity in these tumours included under the same term of head and neck squamous cell carcinoma. The statistical power of this study is limited due to the moderate number of analyzed patients so; studies in larger groups should be done and would be necessary to confirm these results.

Interest conflicts

None of the authors who appear in this study have a potential interest conflict related with the manuscript.

Acknowledgement

This study was supported by the health research program of the “Instituto de Salud Carlos III” (PI11/00519, PI13/01741 and PIE14/00066) co funded with FEDER founds and for the Health Regional Management of the Junta de Castilla y León (GRS630/A11). J. Fernández-Mateos was partially supported by a predoctoral research grant from the Consejería de Educación—Junta de Castilla y León and the European Social Fund to CC-B (EDU/1084/2012). Moreover the authors thank the participation to all the collaborators members of the Spanish Head and Neck Cancer Cooperative Group (TTCC) group and the Institute of Biomedical Research of Salamanca (IBSAL).

References

1. Ganci, F. *et al.* Molecular Genetics and Biology of Head and Neck Squamous Cell Carcinoma: Implications for Diagnosis, Prognosis and Treatment. *Dr. Mark Agulnik (Ed.), In Tech*, doi: 10.50772/31956 (2012).
2. Argiris, A., Karamouzis, M. V, Raben, D. & Ferris, R. L. Head and neck cancer. *Lancet (London, England)* **371,** 1695–709 (2008).
3. Leemans, C. R., Braakhuis, B. J. M. & Brakenhoff, R. H. The molecular biology of head and neck cancer. *Nat. Rev. Cancer* **11,** 9–22 (2011).
4. Perez-Ordoñez, B., Beauchemin, M. & Jordan, R. C. K. Molecular biology of squamous cell carcinoma of the head and neck. *J. Clin. Pathol.* **59,** 445–53 (2006).
5. Torre, L. a *et al.* Global cancer statistics, 2012. *CA. Cancer J. Clin.* **65,** 87–108 (2015).
6. Hashibe, M. *et al.* Alcohol drinking in never users of tobacco, cigarette smoking in never drinkers, and the risk of head and neck cancer: pooled analysis in the International Head and Neck Cancer Epidemiology Consortium. *J. Natl. Cancer Inst.* **99,** 777–89 (2007).
7. Ragin, C. C. R., Modugno, F. & Gollin, S. M. The Epidemiology and Risk Factors of Head and Neck Cancer : a Focus on Human Papillomavirus. *J. Dent. Res.* **86**, 104–14 (2006).
8. Cardesa, A. & Nadal, A. Carcinoma of the head and neck in the HPV era. Carcinoma of the head and neck in the HPV era. *Acta Dermatovenerol Alp Pannonica Adriat*. **20**, 161–173 (2011).
9. Jefferies, S. *et al.* The role of genetic factors in predisposition to squamous cell cancer of the head and neck. *Br. J. Cancer.* **79,** 865–867 (1999).
10. Lacko, M. *et al.* Genetic Susceptibility to Head and Neck Squamous Cell Carcinoma. *Int. J. Radiat. Oncol.* **89,** 38–48 (2014).
11. Hanahan, D. & Weinberg, R. A. Hallmarks of Cancer: The Next Generation. *Cell* **144,** 646–674 (2011).
12. Azad, A. K. *et al.* Genetic sequence variants and the development of secondary primary cancers in patients with head and neck cancers. *Cancer* **118,** 1554–1565 (2012).
13. Hiyama, T., Yoshihara, M., Tanaka, S. & Chayama, K. Genetic polymorphisms and head and neck cancer risk (Review). *Int. J. Oncol.* **32,** 945–73 (2008).
14. Canova, C. *et al.* Genetic associations of 115 polymorphisms with cancers of the upper aerodigestive tract across 10 european countries: The ARCAGE project. *Cancer Res.* **69,** 2956–2965 (2009).
15. Chuang, S.-C. *et al.* Sequence Variants and the Risk of Head and Neck Cancer: Pooled Analysis in the INHANCE Consortium. *Front. Oncol.* **1,** 1–15 (2011).
16. Hiyama, T., Yoshihara, M., Tanaka, S. & Chayama, K. Genetic polymorphisms and head and neck cancer risk (Review). *Int. J. Oncol.* **32,** 945–73 (2008).
17. Brunotto, M., Zarate, a M., Bono, a, Barra, J. L. & Berra, S. Risk genes in head and neck cancer: a systematic review and meta-analysis of last 5 years. *Oral Oncol.* **50,** 178–88 (2014).
18. Seijas-Tamayo, R. *et al.* Epidemiological characteristics of a Spanish cohort of patients diagnosed with squamous cell carcinoma of head and neck: distribution of risk factors by tumor location. *Clin. Transl. Oncol.* **18,** 1114–1122 (2016).
19. Canova, C. *et al.* Genetic Associations of 115 Polymorphisms with Cancers of the Upper Aerodigestive Tract across 10 European Countries: The ARCAGE Project. *Cancer Res.* **69,** 2956–2965 (2009).
20. Streiner, D. L. & Norman, G. R. The pros and cons of propensity scores. *Chest* **142,** 1380–2 (2012).
21. Gugatschka, M., Dehchamani, D., Wascher, T. C., Friedrich, G. & Renner, W. DNA repair gene ERCC2 polymorphisms and risk of squamous cell carcinoma of the head and neck. *Exp. Mol. Pathol.* **91,** 331–4 (2011).

22. Gómez-Díaz, B. *et al.* Analysis of ERCC1 and ERCC2 gene variants in osteosarcoma, colorectal and breast cancer. *Oncol. Lett.* **9,** 1657–1661 (2015).

23. Li, X. *et al.* Association of single nucleotide polymorphisms of nucleotide excision repair genes with laryngeal cancer risk and interaction with cigarette smoking and alcohol drinking. *Tumour Biol.* **35,** 4659–65 (2014).

24. Gugatschka, M., Dehchamani, D., Wascher, T. C., Friedrich, G. & Renner, W. DNA repair gene ERCC2 polymorphisms and risk of squamous cell carcinoma of the head and neck. *Exp. Mol. Pathol.* **91,** 331–334 (2011).

25. Lunn, R. M. *et al.* XPD polymorphisms: effects on DNA repair proficiency. *Carcinogenesis* **21,** 551–555 (2000).

26. Yu, J. J. *et al.* Comparison of two human ovarian carcinoma cell lines (A2780/CP70 and MCAS) that are equally resistant to platinum, but differ at codon 118 of the ERCC1 gene. *Int. J. Oncol.* **16,** 555–560 (2000).

27. Dong, J. *et al.* Potentially functional polymorphisms in DNA repair genes and non-small-cell lung cancer survival: a pathway-based analysis. *Mol. Carcinog.* **51,** 546–52 (2012).

28. Grivennikov, S. I., Greten, F. R. & Karin, M. Immunity, Inflammation, and Cancer. *Cell* **140,** 883–899 (2011).

29. Kundu, J. K. & Surh, Y.-J. Inflammation: gearing the journey to cancer. *Mutat. Res.* **659,** 15–30 (2008).

30. Degenhardt, K. *et al.* Autophagy promotes tumor cell survival and restricts necrosis, inflammation, and tumorigenesis. *Cancer Cell* **10,** 51–64 (2006).

31. Slattery, M. L. *et al.* Genetic variants in interleukin genes are associated with breast cancer risk and survival in a genetically admixed population: The breast cancer health disparities study. *Carcinogenesis* **35,** 1750–1759 (2014).

32. Serefoglou, Z., Yapijakis, C., Nkenke, E. & Vairaktaris, E. Genetic association of cytokine DNA polymorphisms with head and neck cancer. *Oral Oncol.* **44,** 1093–9 (2008).

33. Wu, J. *et al.* Promoter polymorphisms of IL2, IL4, and risk of gastric cancer in a high-risk Chinese population. *Mol. Carcinog.* **48,** 626–32 (2009).

34. Feller, L., Altini, M. & Lemmer, J. Inflammation in the context of oral cancer. *Oral Oncol.* **49,** 887–892 (2013).

35. Hoyer, K. K., Dooms, H., Barron, L. & Abbas, A. K. Interleukin-2 in the development and control of inflammatory disease. *Immunol. Rev.* **226,** 19–28 (2008).

36. Gansmo, L. B. *et al.* Associations between the MDM2 promoter P1 polymorphism del1518 (rs3730485) and incidence of cancer of the breast, lung, colon and prostate. *Oncotarget* **7,** 28637–46 (2016).

37. Yang, X. *et al.* Association of MDM2 promoter T309G polymorphism with oral cancer risk: A meta-analysis of 3,536 subjects. *Mol. Clin. Oncol.* **5,** 175–180 (2016).

38. Yu, H., Li, H., Zhang, J. & Liu, G. Influence of MDM2 polymorphisms on squamous cell carcinoma susceptibility: a meta-analysis. *Onco. Targets. Ther.* **9,** 6211–6224 (2016).

39. Levine, A. J. & Oren, M. The first 30 years of p53: growing ever more complex. *Nat. Rev. Cancer* **9,** 749–58 (2009).

40. Bojesen, S. E. & Nordestgaard, B. G. The common germline Arg72Pro polymorphism of p53 and increased longevity in humans. *Cell Cycle* **7,** 158–163 (2014).

41. Dumont, P., Leu, J. I.-J., Della Pietra, A. C., George, D. L. & Murphy, M. The codon 72 polymorphic variants of p53 have markedly different apoptotic potential. *Nat. Genet.* **33,** 357–65 (2003).

42. Naccarati, A. *et al.* Genotype and haplotype analysis of TP53 gene and the risk of pancreatic cancer: an association study in the Czech Republic. *Carcinogenesis* **31,** 666–70 (2010).

43. Khan, M. H., Khalil, A. & Rashid, H. Evaluation of the p53 Arg72Pro polymorphism and its association with cancer risk: a HuGE review and meta-analysis. *Genet. Res. (Camb).* **97,** e7 (2015).

44. Sourvinos, G., Rizos, E. & Spandidos, D. A. p53 codon 72 polymorphism is linked to the development and not the progression of benign and malignant laryngeal tumours. *Oral Oncol*. **37,** 572–578 (2001).

45. Lehnerdt, G. F. *et al.* The regulatory BCL2 promoter polymorphism (-938C>A) is associated with relapse and survival of patients with oropharyngeal squamous cell carcinoma. *Ann. Oncol.* **20,** 1094–9 (2009).

46. Cingeetham, A. *et al.* Influence of BCL2-938C>A and BAX-248G>A promoter polymorphisms in the development of AML: case-control study from South India. *Tumour Biol.* (2015). doi:10.1007/s13277-015-3457-4

47. Chen, K. *et al.* Single-nucleotide polymorphisms at the TP53-binding or responsive promoter regions of BAX and BCL2 genes and risk of squamous cell carcinoma of the head and neck. *Carcinogenesis* **28,** 2008–12 (2007).

48. Jaramillo, M. C. & Zhang, D. D. The emerging role of the Nrf2-Keap1 signaling pathway in cancer. *Genes and Development* **27,** 2179–2191 (2013).

49. Hartikainen, J. M. *et al.* Genetic polymorphisms and protein expression of NRF2 and sulfiredoxin predict survival outcomes in breast cancer. *Cancer Res.* **72,** 5537–5546 (2012).

50. Nair, S., Doh, S. T., Chan, J. Y., Kong, A.-N. & Cai, L. Regulatory potential for concerted modulation of Nrf2- and Nfkb1-mediated gene expression in inflammation and carcinogenesis. *Br. J. Cancer* **99,** 2070–2082 (2008).

Table 1. SNPs selected in the study.

FUNCTION	GENE	RS	ID	Change
Oncogenes and tumor supressor genes	*TP53*	1042522	C_2403545_10	Pro72Arg
	MDM2	2279744	PCR-RFLP	Thr309Gly
	KRAS-LC6	61764370	PCR-RFLP	Intronic
	EGFR	2227983	C_16170352_20	Lys521Arg
Base excision repair (BER)	*XRCC1*	25487	C_622564_10	Gln399Arg
		1799782	C_11463404_10	Arg194Trp
	APEX	1130409	C_8921503_10	Asp148Glu
Nucleotide excision repair (NER)	*ERCC2(XPD)*	13181	C_3145033_10	Lys751Gln
	ERCC1	11615	C_2532959_10	Asn118Asn
	XPC	2228000	C_16018061_10	Ala499Val
Double-strand break repair genes	*XRCC3*	861539	C_8901525_10	Thr241Met
		1799794	C_2983904_10	c.-316A>G
	KU70	2267437	C_15872242_20	c.-731C>G
Inflammatory genes	*IL1*	16944	C_1839943_10	c.-598T>C
	IL2	2069762	C_15859930_10	c.-385T>G
	IL6	1800795	C_1839697_20	c.-237C>G
	IL10	1800872	C_1747363_10	c.-627A>C
	TNFA	361525	C_2215707_10	c.-418G>A
Apoptotic genes	*NOD2*	2066844	C_11717468_20	Arg702Trp
		2066845	C_11717466_20	Arg908Gly
	BAX	4645878	C_27848291_10	Intronic
	BCL2	2279115	C_3044428_30	Intronic
Carcinogen Metabolism genes	*CYP3A5*	776746	C_26201809_30	Intronic
	GSTP1	1695	C_3237198_20	Ile105Val
	GSTT1	N/A	PCR	Null/present
	GSTM1	N/A	PCR	Null/present
	NFE2L2 (NRF2)	13035806	C_11745134_10	3'-region
		2706110	C_11745133_10	3'-region
	KEAP1	1048290	C_9323035_10	Leu471Leu

Table 2. Descriptive case-control study.

	LARYNX N=126		CONTROL N=126		P-value	PHARYNX N=100		CONTROL N=100		P-value	ORAL CAVITY N=70		CONTROL N=70		P-value
Characteristics	N	%	N	%		N	%	N	%		N	%	N	%	
Age (years)	63.02±8.566		56.30±12.803		**0.000**	59.96±8.41		59.52±10.044		0.742	60.92±10.008		62.24±8.88		0.412
Sex															
Female	13	10.3	13	10.3	1.000	20	20.0	22	22.0	0.728	16	22.9	17	22.9	1.000
Male	113	89.7	113	89.7		80	80.0	78	78.0		54	77.1	54	77.1	
Tobacco smoking															
Never	7	5.5	7	5.5		7	7.0	8	8.0		7	10.0	7	10.0	
<20 PPY	20	15.9	22	17.5	0.944	22	22.0	23	23.0	0.943	12	17.1	12	17.1	1.000
>20 PPY	99	78.6	97	77.0		71	71.0	69	69.0		51	72.9	51	72.9	
Missing	0	0	0	0		0	0	0	0		0	0	0	0	
Alcohol drinking															
Never	53	42.1	51	40.5		26	26.0	27	27.0		23	32.9	23	32.9	
<14 SDU/week	28	22.2	31	24.6	0.904	30	30.0	30	30.0	0.985	19	27.1	19	27.1	1.000
>14 SDU/week	45	35.7	44	34.9		44	44.0	43	43.0		28	40.0	28	40.0	
Missing	0	0	0	0		0	0	0	0		0	0	0	0	

P-values related to controls. Statistically significant results in bold.

Table 3. Statistically significant SNPs in laryngeal cancer.

	Genotype	Larynx		Control		P-value*	OR (CI 95%)
SNPs		**N**	**%**	**N**	**%**		
***TP53* rs1042522**	GG	61	48.4	62	49.2	Ref.	1.00
	GC	54	42.9	37	29.4	0.165	1.505 (0.846-2.677)
	CC	11	8.7	27	21.4	**0.008**	**0.319 (0.136-0.745)**
Recessive	GG+GC	115	91.3	99	78.6	Ref.	1.00
	CC	11	8.7	27	21.4	**0.002**	**0.268 (0.119-0.607)**
Dominant	GG	61	48.4	62	49.2	Ref.	1.00
	GC+CC	65	51.6	64	50.8	0.596	0.986 (0.587-1.654)
***MDM2* rs2279744**	TT	44	34.9	62	49.2	Ref.	1.00
	TG	57	45.2	53	42.1	0.279	1.364 (0.778-2.392)
	GG	25	19.8	11	8.7	**0.015**	**2.826 (1.219-6.552)**
Recessive	TT+TG	101	80.2	115	91.3	Ref.	1.00
	GG	25	19.8	11	8.7	**0.029**	2.413 (1.094-5.323)
Dominant	TT	44	34.9	62	49.2	Ref.	1.00
	TG+GG	82	65.1	64	50.8	0.075	1.616 (0.953-2.742)
***ERCC1* rs11615**	TT	53	42.1	45	35.7	Ref.	1.00
	TC	67	53.2	58	46.0	0.872	0.956 (0.550-1.661)
	CC	6	4.8	23	18.3	**0.013**	**0.281 (0.103-0.768)**
Recessive	TT+TC	120	95.2	103	81.7	Ref.	1.00
	CC	6	4.8	23	18.3	**0.011**	**0.288 (0.110-0.751)**
Dominant	TT	53	42.1	45	35.7	Ref.	1.00
	TC+CC	73	57.9	81	64.3	0.354	0.778 (0.457-1.324)
***ERCC2* rs13181**	TT	72	57.1	52	41.3	Ref.	1.00
	TG	46	36.5	58	46.0	0.247	0.720 (0.413-1.255)
	GG	8	6.3	16	12.7	**0.046**	**0.375 (0.143-0.982)**
Recessive	TT+TG	118	93.7	110	87.3	Ref.	1.00
	GG	8	6.3	16	12.7	0.079	0.433 (0.170-1.102)
Dominant	TT	72	57.1	52	41.3	Ref.	1.00
	TG+GG	54	42.9	74	58.7	0.093	0.638 (0.377-1.078)
***IL6* rs1800795**	CC	43	34.1	62	50.8	Ref.	1.00
	CG	64	50.8	46	37.7	**0.003**	**2.471 (1.372-4.452)**
	GG	19	15.1	14	11.5	0.070	2.164 (0.938-4.991)
Recessive	CC+CG	107	84.9	108	88.5	Ref.	1.00
	GG	19	15.1	14	11.5	0.444	1.351 (0.625-2.921)
Dominant	CC	43	34.1	62	50.8	Ref.	1.00
	CG+GG	83	65.9	60	49.2	**0.002**	**2.394 (1.376-4.163)**
***NRF2* rs13035806**	GG	109	87.2	95	76.0	Ref.	1.00
	GA	14	11.2	29	23.2	**0.019**	**0.424 (0.207-0.869)**
	AA	2	1.6	1	0.8	0.520	2.235 (0.193-25.903)
Recessive	GG+GA	123	98.4	124	99.2	Ref.	1.00
	AA	2	1.6	1	0.8	0.444	2.600 (0.225-30.064)
Dominant	GG	109	87.2	95	76.0	Ref.	1.00
	GA+AA	16	12.8	30	24.0	**0.035**	**0.478 (0.240-0.949)**
***NRF2* rs2706110**	CC	92	73.6	72	57.1	Ref.	1.00
	CT	24	19.2	47	37.3	**0.005**	**0.425 (0.233-0.775)**
	TT	9	7.2	7	5.6	0.732	1.207 (0.411-3.541)
Recessive	CC+CT	116	92.8	119	94.4	Ref.	1.00
	TT	9	7.2	7	5.6	0.403	1.574 (0.544-4.560)
Dominant	CC	92	73.6	72	57.1	Ref.	1.00
	CT+TT	33	26.4	54	42.9	**0.020**	**0.518 (0.299-0.900)**

**P-values adjusted by age. Statistically significant results in bold.*

Table 4. Statistically significant SNPs in pharyngeal cancer.

	Genotype	Pharynx		Control		P-value	OR (CI 95%)
SNPs		**N**	**%**	**N**	**%**		
TP53 **rs1042522**	GG	53	53.0	47	47.0	Ref.	1.00
	GC	44	44.0	33	33.0	0.583	1.182 (0.650-2.151)
	CC	3	3.0	20	20.0	**0.002**	**0.133 (0.037-0.476)**
Recessive	GG+GC	97	97.0	80	80.0	Ref.	1.00
	CC	3	3.0	20	20.0	**0.001**	**0.124 (0.035-0.431)**
Dominant	GG	53	53.0	47	47.0	Ref.	1.00
	GC+CC	47	47.0	53	53.0	0.396	0.786 (0.451-1.370)
NRF2 **rs2706110**	CC	68	68.0	54	54.0	Ref.	1.00
	CT	25	25.0	41	41.0	**0.020**	**0.484 (0.262-0.893)**
	TT	7	7.0	5	5.0	0.863	1.112 (0.334-3.698)
Recessive	CC+CT	93	93.0	95	95.0	Ref.	1.00
	TT	7	7.0	5	5.0	0.553	1.430 (0.438-4.667)
Dominant	CC	68	68.0	54	54.0	Ref.	1.00
	CT+TT	32	32.0	46	46.0	**0.043**	**0.552 (0.311-0.982)**

Statistically significant results in bold.

Table 5. Statistically significant SNPs in oral cavity carcinoma.

SNPs	Genotype	Oral cavity		Control		P-value	OR (CI 95%)
		N	%	N	%		
***IL2* rs2069762**	TT	43	61.4	31	44.3	/	1.00
	TG	22	31.4	27	38.6	0.152	0.587 (0.284-1.217)
	GG	5	7.1	12	17.1	**0.039**	**0.300 (0.096-0.940)**
Recessive	TT+TG	65	92.9	58	82.9	/	1.00
	GG	5	7.1	12	17.1	0.078	0.372 (0.124-1.119)
Dominant	TT	43	61.4	31	44.3	/	1.00
	TG+GG	27	38.6	39	55.7	**0.043**	**0.499 (0.254-0.979)**
***IL6* rs1800795**	CC	25	35.7	39	55.7	/	1.00
	CG	33	47.1	23	32.9	**0.031**	**2.238 (1.077-4.653)**
	GG	12	17.1	8	11.4	0.104	2.340 (0.839-6.528)
Recessive	CC+CG	58	82.9	62	88.6	/	1.00
	GG	12	17.1	8	11.4	0.337	1.603 (0.612-4.203)
Dominant	CC	25	35.7	39	55.7	/	1.00
	CG+GG	45	64.3	31	44.3	**0.018**	**2.265 (1.148-4.467)**
***BCL2* rs2279115**	CC	13	18.6	27	38.6	/	1.00
	CA	43	61.4	30	42.9	**0.008**	**2.977 (1.325-6.688)**
	AA	14	20.0	13	18.6	0.116	2.237 (0.820-6.103)
Recessive	CC+CA	56	80.0	57	81.4	/	1.00
	AA	14	20.0	13	18.6	0.830	1.096 (0.473-2.540)
Dominant	CC	13	18.6	27	38.6	/	1.00
	CA+AA	57	81.4	43	61.4	**0.010**	**2.753 (1.273-5.952)**

Statistically significant results in bold.

Supplementary table 1. Distribution between selected SNPs in different candidate genes in laryngeal tumours (Ca) and controls (Co)*.

			Homozygotes common alleles		Heterozygotes				Homozygotes rarer allele			
FUNCTION	**GENE**	**RS**	**Ca**	**Co**	**Ca**	**Co**	**P-value**	**OR (95% CI)**	**Ca**	**Co**	**P-Value**	**OR (95% CI)**
Oncogenes y tumor suppressor genes	*p53*	1042522	61	62	54	37	0.165	1.505 (0.846-2.677)	11	27	**0.008**	0.319 (0.136-0.745)
	MDM2	2279744	44	65	57	53	0.279	1.364 (0.778-2.392)	25	11	**0.015**	2.826 (1.219-6.552)
	KRAS-LC6	rs61764370	97	87	27	33	0.382	0.762 (0.414-1.401)	2	6	0.247	0.367 (0.068-1.998)
	EGFR	2227983	78	72	41	44	0.514	0.830 (0.474-1.453)	7	10	0.283	0.536 (0.197-1.608)
Base excision repair (BER)	*XRCC1*	25487	58	57	50	55	0.794	0.929 (0.533-1.617)	18	14	0.724	1.160 (0.509-2.646)
		1799782	106	109	20	14	0.371	1.416 (0.663-3.012)	0	2	0.999	0.000 (0.000-)
	APEX	1130409	38	42	65	49	0.308	1.366 (0.750-2.490)	23	35	0.546	0.803 (0.394-1.637)
Nucleotide excision repair (NER)	*ERCC2=XPD*	13181	72	52	46	58	0.247	0.720 (0.413-1.255)	8	16	**0.046**	0.375 (0.143-0.982)
	ERCC1	11615	53	45	67	58	0.872	0.956 (0.550-1.661)	6	23	**0.013**	0.281 (0.103-0.768)
	XPC	2228000	62	61	57	52	0.530	1.190 (0.692-2.046)	7	13	0.384	0.636 (0.230-1.763)
Double-strand break repair genes	*XRCC3*	861539	49	47	59	58	0.721	0.902 (0.512-1.590)	18	21	0.504	0.766 (0.352-1.671)
		1799794	78	78	43	40	0.593	1.166 (0.664-2.047)	5	8	0.789	0.849 (0.255-2.823)
	KU70	2267437	39	44	65	58	0.995	0.998 (0.550-1.811)	22	24	0.780	0.898 (0.422-1.909)
Inflammatory genes	*IL1*	16944	56	58	58	56	0.849	1.054 (0.612-1.817)	12	12	0.930	0.959 (0.382-2.408)
	IL2	2069762	62	67	56	45	0.307	1.333 (0.768-2.311)	8	14	0.132	0.472 (0.178-1.253)
	IL6	1800795	43	62	64	46	**0.003**	2.471 (1.372-4.452)	19	14	0.070	2.164 (0.938-4.991)
	IL10	1800872	75	71	45	47	0.557	0.849 (0.491-1.467)	6	8	0.540	0.694 (0.216-2.229)
	TNFA	361525	104	103	19	21	0.694	0.866 (0.422-1.777)	3	2	0.632	1.565 (0.250-9.780)
Apoptotic genes	*NOD2*	2066844	110	109	16	17	0.876	0.940 (0.433-2.043)	0	0	-	-
		2066845	112	119	4	7	0.277	0.485 (0.131-1.789)	0	0	-	-
	BAX	4645878	92	93	31	31	0.806	1.078 (0.592-1.964)	2	2	0.986	1.017 (0.146-7.065)
	BCL2	2279115	37	40	66	55	0.124	1.614 (0.878-2.969)	23	31	0.896	1.050 (0.501-2.203)
Carcinogen Metabolism genes	*CYP3A5*	rs776746	116	110	8	14	0.246	0.573 (0.223-1.468)	1	1	0.914	1.167 (0.072-19.029)
	GSTP1	1695	65	62	54	51	0.805	1.071 (0.622-1.842)	12	6	0.254	0.535 (0.183-1.568)
	GSTT1	N/A	110	109	15	17	0.483	1.322 (0.606-2.882)	-	-	-	-
	GSTM1	N/A	66	59	59	67	0.337	1.289 (0.767-2.168)	-	-	-	-
	NFE2L2 (NRF2)	13035806	109	95	14	29	**0.019**	0.424 (0.207-0.869)	2	1	0.520	2.235 (0.193-25.903
		2706110	92	72	24	47	**0.005**	0.425 (0.233-0.775)	9	7	0.732	1.207(0.411-3.541)
	KEAP1	1048290	50	53	58	57	0.768	1.008 (0.622-1.900)	17	16	0.541	1.290 (0.569-2.925)

**P-values adjusted by age and related to controls. Statistically significant results in bold.*

Supplementary table 2. Distribution between selected SNPs in different candidate genes in pharyngeal tumours (Ca) and controls (Co).

			Homozygotes common alleles		Heterozygotes				Homozygotes rarer allele			
FUNCTION	**GENE**	**RS**	**Ca**	**Co**	**Ca**	**Co**	**P-value**	**OR (95% CI)**	**Ca**	**Co**	**P-Value**	**OR (95% CI)**
Oncogenes y tumor suppressor genes	*p53*	1042522	53	47	44	33	0.583	1.182 (0.650-2.151)	3	20	**0.002**	0.133 (0.037-0.476)
	MDM2	2279744	49	49	37	42	0.675	0.881 (0.487-1.595)	14	9	0.350	1.556 (0.616-3.928)
	KRAS-LC6	rs61764370	68	69	30	27	0.704	1.127 (0.607-2.093)	2	4	0.442	0.507 (0.090-2.862)
	EGFR	2227983	58	54	36	37	0.743	0.906 (0.502-1.634)	6	9	0.394	0.621 (0.207-1.860)
Base excision repair (BER)	*XRCC1*	25487	40	44	47	44	0.594	1.175 (0.649-2.127)	13	12	0.701	1.192 (0.487-2.913)
		1799782	90	87	10	12	0.634	0.806 (0.3311.961)	0	1	1.000	0.000 (0.000-)
	APEX	1130409	29	35	45	39	0.320	1.393 (0.725-2.675)	26	26	0.615	1.207 (0.580-2.513)
Nucleotide excision repair (NER)	*ERCC2=XPD*	13181	48	49	36	40	0.782	0.919 (0.504-1.676)	16	11	0.370	1.485 (0.625-3.526)
	ERCC1	11615	39	36	48	51	0.646	0.869 (0.477-1.584)	13	13	0.860	0.923 (0.378-2.253)
	XPC	2228000	48	54	46	38	0.296	1.362 (0.763-2.431)	6	8	0.768	0.844 (0.273-2.606)
Double-strand break repair genes	*XRCC3*	861539	36	35	43	49	0.616	0.853 (0.459-1.586)	21	16	0.550	1.276 (0.574-2.839)
		1799794	56	59	36	38	0.722	1.112 (0.620-1.995)	5	6	0.711	1.264 (0.365-4.377)
	KU70	2267437	39	33	47	49	0.504	0.812 (0.440-1.497)	14	18	0.328	0.658 (0.285-1.522)
Inflammatory genes	*IL1*	16944	47	52	43	39	0.506	1.220 (0.679-2.192)	10	9	0.681	1.229 (0.460-3.286)
	IL2	2069762	46	52	41	37	0.459	1.253 (0.690-2.273)	13	11	0.526	1.336 (0.546-3.272)
	IL6	1800795	46	48	45	35	0.336	1.342 (0.737-2.442)	9	13	0.498	0.722 (0.282-1.852)
	IL10	1800872	60	57	37	38	0.792	0.925 (0.518-1.652)	3	5	0.456	0.570 (0.130-2.495)
	TNFA	361525	83	83	17	15	0.746	1.133 (0.531-2.419)	0	2	0.999	0.000 (0.000-)
Apoptotic genes	*NOD2*	2066844	82	86	17	14	0.538	1.274 (0.590-2.749)	1	0	1.000	1694278518 (000-)
		2066845	96	95	4	5	0.734	0.792 (0.206-3.039)	0	0	-	-
	BAX	4645878	83	74	15	23	0.141	0.581 (0.282-1.197)	2	3	0.575	0.594 (0.097-3.655)
	BCL2	2279115	30	33	46	42	0.573	1.205 (0.630-2.302)	24	25	0.886	1.056 (0.500-2.229)
Carcinogen Metabolism genes	*CYP3A5*	rs776746	89	89	10	10	1.000	1.000 (0.397-2.520)	0	0	-	-
	GSTP1	1695	49	40	41	46	0.293	0.728 (0.402-1.317)	7	12	0.155	0.476 (0.171-1.322)
	GSTT1	N/A	84	86	16	14	0.692	0.855 (0.393-1.860)	-	-	-	-
	GSTM1	N/A	46	49	54	51	0.671	0.887 (0.509-1.545)	-	-	-	-
	NFE2L2 (NRF2)	13035806	78	75	18	25	0.292	0.692 (0.349-1.372)	3	0	0.999	-
		2706110	68	54	25	41	**0.020**	0.484 (2.262-0.893)	7	5	0.863	1.112 (0.334-3.698)
	KEAP1	1048290	42	44	49	46	0.713	1.116 (0.623-2.000)	9	10	0.908	0.943 (0.349-2.550)

P-values related to controls. Statistically significant results in bold.

Supplementary table 3. Distribution between selected SNPs in different candidate genes in tumours from oral cavity (Ca) and controls (Co).

			Homozygotes common alleles		Heterozygotes				Homozygotes rarer allele			
FUNCTION	**GENE**	**RS**	**Ca**	**Co**	**Ca**	**Co**	**P-value**	**OR (95% CI)**	**Ca**	**Co**	**P-Value**	**OR (95% CI)**
Oncogenes y tumor suppressor genes	*p53*	1042522	32	38	30	21	0.156	1.696 (0.818-3.518)	8	11	0.779	0.864 (0.310-2.407)
	MDM2	2279744	30	30	28	34	0.593	0.824 (0.404-1.678)	12	6	0.218	2.000 (0.664-6.026)
	KRAS-LC6	rs61764370	40	49	26	18	0.126	1.769 (0.851-3.678)	4	3	0.536	1.633 (0.345-7.727)
	EGFR	2227983	41	35	25	30	0.338	0.711 (0.354-1.428)	4	5	0.591	0.683 (0.170-2.742)
Base excision repair (BER)	*XRCC1*	25487	31	31	30	30	1.000	1.000 (0.492-2.034)	9	9	1.000	1.000 (0.350-2.856)
		1799782	62	62	8	8	1.000	1.000 (0.353-2.833)	0	0	-	-
	APEX	1130409	18	25	39	26	0.066	2.083 (0.952-4.559)	13	19	0.914	0.950 (0.375-2.408)
Nucleotide excision repair (NER)	*ERCC2=XPD*	13181	24	33	36	27	0.101	1.833 (0.888-3.785)	10	10	0.541	1.371 (0.495-3.821)
	ERCC1	11615	19	22	43	36	0.401	1.383 (0.649-2.948)	8	12	0.640	0.772 (0.261-2.284)
	XPC	2228000	36	37	27	26	0.857	1.067 (0.526-2.165)	7	7	0.963	1.028 (0.327-3.226)
Double-strand break repair genes	*XRCC3*	861539	27	32	29	28	0.582	1.228 (0.592-2.546)	14	10	0.301	1.659 (0.636-4.332)
		1799794	39	42	26	23	0.587	1.217 (0.598-2.477)	5	5	0.912	1.077 (0.289-4.007)
	KU70	2267437	19	22	35	34	0.657	1.192 (0.550-2.585)	16	14	0.561	1.323 (0.515-3.401)
Inflammatory genes	*IL1*	16944	27	37	36	27	0.093	1.827 (0.904-3.693)	7	6	0.443	1.599 (0.483-5.297)
	IL2	2069762	43	31	22	27	0.152	0.587 (0.284-1.217)	5	12	**0.039**	0.300 (0.096-0.940)
	IL6	1800795	25	39	33	23	**0.031**	2.238 (1.077-4.653)	12	8	0.104	2.340 (0.839-6.528)
	IL10	1800872	29	37	37	29	0.165	1.628 (0.819-3.237)	4	4	0.745	1.276 (0.294-5.542)
	TNFA	361525	61	60	8	9	0.796	0.874 (0.316-2.417)	1	1	0.991	0.984 (0.060-16.088)
Apoptotic genes	*NOD2*	2066844	62	63	8	7	0.785	1.161 (0.397-3.397)	0	0	-	-
		2066845	68	67	2	3	0.651	0.657 (0.106-4.057)	0	0	-	-
	BAX	4645878	51	51	19	18	0.888	1.056 (0.497-2.240)	0	1	1.000	0.000 (0.000-)
	BCL2	2279115	13	27	43	30	**0.008**	2.977 (1.325-6.688)	14	13	0.116	2.237 (0.820-6.103)
Carcinogen Metabolism genes	*CYP3A5*	rs776746	61	63	9	6	0.432	1.549 (0.520-4.614)	0	0	-	-
	GSTP1	1695	38	33	29	28	0.766	0.899 (0.448-1.808)	3	8	0.118	0.326 (0.080-1.329)
	GSTT1	N/A	8	11	62	59	0.461	1.445 (0.543-3.842)	-	-	-	-
	GSTM1	N/A	42	34	28	36	0.176	0.630 (0.322-1.230)	-	-	-	-
	NFE2L2 (NRF2)	13035806	56	54	12	16	0.448	0.723 (0.313-1.670)	2	0	0.999	-
		2706110	45	39	22	29	0.241	0.657 (0.326-1.325)	3	2	0.780	1.300 (0.206-8.184)
	KEAP1	1048290	24	27	39	36	0.586	1.219 (0.598-2.485)	7	7	0.845	1.125 (0.345-3.673)

P-values related to controls. Statistically significant results in bold.