Published in final edited form as:
Cancer Lett. 2015 February 28; 357(2): 549–556. doi:10.1016/j.canlet.2014.12.008.

Identification of epidermal growth factor receptor (*EGFR*) genetic variants that modify risk for head and neck squamous cell carcinoma

Christopher Fung[a], **Pei Zhou**[a], **Sonali Joyce**[a], **Kerry Trent**[b], **Jian-Min Yuan**[b], **Jennifer R. Grandis**[a,c], **Joel L. Weissfeld**[b], **Marjorie Romkes**[d], **Daniel E. Weeks**[e], and **Ann Marie Egloff**[a,f,*]

[a]Department of Otolaryngology, University of Pittsburgh, Pittsburgh, PA, USA

[b]Department of Epidemiology, University of Pittsburgh, Pittsburgh, PA, USA

[c]Department of Pharmacology & Chemical Biology, University of Pittsburgh, Pittsburgh, PA, USA

[d]Department of Medicine, University of Pittsburgh, Pittsburgh, PA, USA

[e]Department of Human Genetics and Biostatistics, University of Pittsburgh, Pittsburgh, PA, USA

[f]Department of Microbiology and Molecular Genetics, University of Pittsburgh, Pittsburgh, PA, USA

Abstract

EGFR polymorphisms have not been thoroughly evaluated for association with head and neck squamous cell carcinoma (HNSCC) risk. We genotyped 578 HNSCC patients and 588 cancer-free controls for 60 *EGFR* single nucleotide polymorphisms (SNPs) and tested associations with HNSCC risk.

EGFR intronic SNPs rs12535536, rs2075110, rs1253871, rs845561 and rs6970262 and synonymous SNP rs2072454 were associated with HNSCC risk among all subjects ($p < 0.05$). SNPs rs12538371, rs845561, and rs6970262 were significantly associated with HNSCC risk ($p < 0.05$) among never tobacco users. We identified EGFR variants that likely modify risk for HNSCC including three variants that contribute to tobacco-independent risk.

Keywords

Case–control study; Genetic polymorphism; Tobacco-independent; Nonsmoker; Head and neck cancer risk

*Corresponding author. Tel.: +1 412 648 9545. egloffam@upmc.edu (A.M. Egloff).

Conflict of interest
The authors declare no conflict of interest.

Introduction

Head and neck cancers, which occur primarily in the oral cavity, pharynx and larynx, are the sixth most common cause of cancer death worldwide with approximately 600,000 new cases and 300,000 associated deaths annually [1–3]. More than 90% of these are head and neck squamous cell carcinoma (HNSCC) [4]. HNSCC presents a challenging clinical problem due to the late stage of initial presentation, high rates of recurrence and metastasis and disease- and treatment-associated morbidities. Defining at-risk populations and gaining insights regarding specific molecular contributors to HNSCC development and progression are important steps toward developing more effective prevention, early diagnostic and treatment regimens.

Epidermal growth factor receptor (EGFR), which is overexpressed in the majority of HNSCCs [5–7], has remained relatively understudied regarding HNSCC risk. *EGFR*, which is located on the short arm of chromosome 7, encodes a receptor tyrosine kinase known to play a critical role in mediating the aberrant extracellular growth signals common to several cancers including HNSCC [8]. Polymorphisms in *EGFR* have previously been shown to be associated with differential risk of developing glioma [9–12], lung cancer [13–15], esophageal cancer [14], and renal cancer [16]. To date, studies investigating the relationship between HNSCC and *EGFR* polymorphisms have generally been restricted to putative promoter/enhancer elements and selected non-synonymous coding changes [17–23]. As yet, an extensive assessment of EGFR genetic variants and their association with HNSCC risk has not yet been reported, and extensive evaluations of EGFR genetic variants regarding cancer risk have only been reported for glioma [9,10]. Of the limited studies that have investigated the relationship between *EGFR* polymorphisms and HNSCC, only one has shown an association between the intron 1 CA-repeat polymorphism and HNSCC risk [20].

While *EGFR* activating mutations are rare in HNSCC cancers [24,25], *EGFR* gene amplification has been reported in a subset of HNSCC [26–28]. EGFR protein levels are elevated in many HNSCC tumors, and high EGFR tumor protein levels indicate poor prognosis [28,29]. In order to gain further insights regarding the role(s) of EGFR in HNSCC we sought to determine if *EGFR* polymorphisms were associated with HNSCC risk.

Tobacco and alcohol uses have been widely recognized to be principal HNSCC etiologic factors [30]. However, only a minority of smokers develop HNSCC. Furthermore, approximately 20% of HNSCC patients are lifelong never smokers [31]. More recently, infection with human papillomavirus (HPV) has been identified as a contributing risk factor to HNSCC. Although HPV-positive (HPV+) tumors are enriched among never smoking cases, HPV is an etiologic agent for HNSCC among tobacco users and never tobacco users [32,33]. HPV + HNSCC tumors are predominantly found in the oropharynx, with lower rates of positivity in the hypopharynx, oral cavity and larynx. Therefore, different environmental risk factors for HNSCC exist with likely different genetic susceptibility factors.

Positive family history of HNSCC has been associated with increased HNSCC risk even after adjusting for tobacco and alcohol use, suggesting a genetic component for HNSCC risk

may exist [34]. We hypothesized that *EGFR* genetic variants would be associated with HNSCC risk. EGFR variants related to tobacco use risk were hypothesized to be specifically associated with HNSCC risk among tobacco users. We further hypothesized that EGFR variants associated with HNSCC risk unrelated to tobacco use would be more readily discerned among never tobacco users where the obscuring effects of tobacco-related risk would be minimized. Here we report the testing of single nucleotide polymorphism (SNP) variants spanning *EGFR* for association with HNSCC using a case–control study design including approximately 600 HNSCC cases and 600 cancer-free control subjects.

Materials and methods

Study populations

From 2000 to 2009, patients aged 18–79 years with pathologically confirmed HNSCC were enrolled into this University of Pittsburgh Head and Neck Specialized Program of Research Excellence (SPORE)-sponsored study within 1 year of diagnosis. 610 HNSCC cases with blood-derived DNA available for analysis were successfully frequency-matched by age category (10-year strata), sex and race to 633 cancer-free control subjects enrolled during the same time period [35]. White subjects constituted approximately 97% of matched cases (n = 596) and controls (n = 612). White subjects with ≥95% genotyping success rates included 578 HNSCC cases and 588 controls that comprised the study subjects (Table 1). Upon enrollment, subjects donated peripheral blood and completed an administered questionnaire. All procedures were approved by the University of Pittsburgh Institutional Review Board.

Questionnaire data

Demographic and risk factor data for Pittsburgh subjects were obtained by using the interviewer-administered questionnaire previously described [35]. Age was defined as age at diagnosis for cases and age at interview for control subjects. Tobacco and alcohol use pack-year and drink-year metrics were as previously described [35].

SNP selection and genotyping assays

SNPs were selected based on previous reports of association with cancer, functional significance, or as HapMap project tagging SNPs for subjects with Northern and Western European ancestry (CEU). Tagging SNPs were selected using the publicly available Tagger program of the HapMap project [36]. SNPs representing linkage disequilibrium blocks with $r^2 \geq 0.8$ and a minor allele frequency ≥5% were selected for genotyping the longest EGFR gene transcript (variant 1) and 5000 base pairs upstream and downstream. In total, 68 SNPs were selected for genotyping (Appendix S1: Supplementary Table S1). DNA was extracted from peripheral whole blood using commercial kits and stored at −80 °C as previously described [35]. Genotyping was performed using the MassArray iPlex Gold system (Sequenom); PCR primers and extension primers for individual SNPs were designed using MassArray 3.0 software (Sequenom). Samples were arrayed in 384-well plates for analysis, and each plate contained a 5% replication of samples and one water negative control.

HPV and P16 tumor status

HNSCC tumor HPV status by *in situ* hybridization (ISH) and cyclin-dependent kinase inhibitor 2A (P16) tumor status by immunohistochemical (IHC) staining were determined as previously described [28]. HPV ISH and P16 IHC data were available through the Pittsburgh Head and Neck Organ-specific electronic database for 128 and 146 tumors, respectively [37].

Statistical analysis

STATA 12 (StataCorp) was used to perform all statistical analyses. Hardy–Weinberg equilibrium (HWE) for control subjects was assessed using Fisher's exact test. SNPs with a distribution of alleles out of HWE ($p < 0.05$) were excluded from further analysis as were SNPs with call rates less than 95% (Appendix S1: Supplementary Table S1). Association between SNPs and HNSCC was assessed using Fisher's exact test for genotype association (2 degrees of freedom) and the Cochran–Armitage test of trend using the *genass* command in STATA [38]. Adjusted odds ratios (aOR) and 95% confidence intervals (95CI) were estimated using multivariable logistic regression (MLR) models adjusted for age and sex and candidate variables if significant. Candidate variables including body mass index (BMI), tobacco use, cigarette pack-years category, alcohol drink-years category and cigarette by alcohol interaction (Table 1) were tested for significance in univariable logistic regression models, and significant variables (Wald $p < 0.05$) were retained in MLR models if associated Wald $p < 0.05$. SNP association with HNSCC risk was individually tested using MLR models stratified by tobacco use (ever versus never) and for HPV+ and HPV− tumors separately.

SNP function prediction

Prediction of variant impact on function was performed using the tool provided by the National Institutes of Environmental Health Sciences (http://snpinfo.niehs.nih.gov/snpinfo/snpfunc.htm) and the RegulomeDB database (http://regulome.stanford.edu) [39,40].

Results

HNSCC case–control subjects exhibited typical characteristics

There were no statistically significant differences in median age or sex distribution between successfully genotyped cases and cancer-free controls (Table 1). HNSCC occurs primarily in male patients, and the percentage of male cases among the cases (79.8% for the larynx and 70.1% for the oral cavity and pharynx) was similar to the percentage of male cases that would be expected based on prevalence data for whites in the U.S. during our recruitment period (79.9% for the larynx and 65.7% for the oral cavity and pharynx) [4].

Never tobacco use and never alcohol use rates reported in a large pooled analysis of multiple studies (59.6% and 27%, respectively) were lower than the never tobacco and never alcohol use rates among cases and controls (Table 1) [30]. Nonetheless, our HNSCC cases were more likely to have used tobacco and have consumed more cigarettes and alcohol than cancer-free control subjects.

Compared to ever tobacco-using HNSCC patients, never tobacco-using HNSCC patients were more often female and less likely to report positive alcohol use histories (Appendix S1: Supplementary Table S2). Anatomical tumor sites differed significantly for never versus ever tobacco users with more oral cavity and oropharynx cancers and fewer laryngeal cancers among never tobacco users.

Genotyping quality

Genotype calls were made in 95% of samples for 60 of the 68 SNPs assayed (Appendix S1: Supplementary Table S1). Successfully genotyped SNPs had similar minor allele frequencies among control subjects as the HapMap database (Appendix S1: Supplementary Table S1), supporting the validity of the assay results and the appropriateness of the control population. The 5′-untranslated region SNP rs712829, which was previously reported to be associated with lung cancer risk [41], was among 8 SNPs that failed in our study (Appendix S1: Supplementary Table S1). Of the 596 white cases and 612 white controls, 578 (97%) cases and 588 (96%) controls were successfully genotyped with call rates of at least 95%, and these constituted our study population. As a measure of assay reliability, 5% of samples were randomly selected for replicate plating, and these sample genotypes were 100% concordant between duplicate pairs for successfully called genotypes. The final case–control study population did not demonstrate statistically significant differences in age or sex between HNSCC cases and controls, an indication genotyping failure was sufficiently infrequent and/or random as to not negate frequency-matching (Table 1).

Ten EGFR SNPs were associated with HNSCC rsik by genotype or trend test

We sought to identify risk-associated polymorphisms using Fisher’s exact genotype test or the Cochran–Armitage test of trend to test significance as an initial threshold for further evaluation. This combination of two statistical tests was demonstrated to provide near optimal power under a number of inheritance modes [42]. We identified 8 *EGFR* SNPs associated with HNSCC risk in our unstratified analysis and 2 additional *EGFR* SNPS (1 among never tobacco users, 1 among ever tobacco users) (Table 2 and Appendix S1: Supplementary Fig. S1). One of these SNPs was a synonymous SNP in exon 4 (rs2072454). The other nine SNPs were intronic.

EGFR variants were associated with HNSCC risk after controlling for known risk factors

In order to estimate the contribution of EGFR SNPs after controlling for risk factors, we developed MLR models. In the unstratified analysis, age, sex, and alcohol and tobacco use as well as their combined interaction were significantly associated with HNSCC risk and retained in the final model. Six *EGFR* SNPs (rs12525536, rs2072454, rs2075110, rs12538371, rs845561 and rs6970262) were associated with HNSCC risk after adjusting for these risk factors in the unstratified analysis (Table 2). Of these SNPs, rs2072454 and rs2075110 were in linkage disequilibrium (LD) among control subjects ($r^2 = 0.98$).

Among tobacco users, only rs17586365 rare minor allele homozygotes, of which there were few (7 controls and 1 case), had significantly reduced HNSCC risk after adjusting for risk factors. SNPs rs12538371, rs845561 and rs6970262, residing in introns 15, 20 and 21,

respectively, were associated with HNSCC risk among never tobacco users after adjustment for age and sex (Table 2).

The increasing number of tobacco-independent risk alleles, defined as rs12538371 [C], rs845661 [C] and rs6970262 [G], was associated with significantly increasing HNSCC risk among never tobacco users but not ever tobacco users (Analysis 1, Table 3). This same trend was also noted in the unstratified analysis, though the magnitude of the risk was not as great as when the analysis was confined to never tobacco users, suggesting that the contribution to risk in the unstratified analysis was likely driven primarily by never tobacco users.

A summation of risk alleles for all SNPs that remained significantly associated with HNSCC in any of the MLR models in any stratum indicated that the increasing number of risk alleles was associated with increasing HNSCC risk in the unstratified analysis and in both tobacco strata (Analysis 2, Table 3). In this analysis, risk alleles for both tobacco-independent and tobacco-related HNSCC were included among the summed alleles. rs2075110 was excluded from this analysis because of its LD with rs2072454. It is of note that when tobacco-independent risk alleles were omitted, an increasing risk of HNSCC with increasing number of risk alleles was associated only among tobacco users and not among the never tobacco users (Analysis 3, Table 3).

Tobacco-independent SNPs rs6970202 and rs12538371 were associated with HPV+ HNSCC risk

In order to gain insights into genetic variants regarding their roles in HPV+ versus HPV− HNSCC risk, we evaluated *EGFR* SNPs that were associated with HNSCC risk in MLR models for significant differences in genotype frequencies between patients with HPV+ HNSCC and patients with HPV− HNSCC. Our ability to assess differences with HPV tumor status was limited because HPV tumor status was available for only 128 patient tumors (22%), and HPV evaluated cases were younger, more often had oropharyngeal tumors and less often had laryngeal tumors (Appendix S1: Supplementary Table S4). As would be expected of unselected HNSCC cases, 68% of oropharynx tumors tested were HPV positive, and HPV and P16 status were concordant for the majority of tumors analyzed for both (101/118 (86%)). In support of tobacco and HPV being independent risk factors, of the 104 HPV evaluated HNSCC cancers in ever tobacco users, 29 (28%) were HPV+ by ISH. Similarly, of the 24 HPV evaluated among never tobacco users, 8 (33%) were HPV+.

We observed that of the EGFR SNPs associated with HNSCC risk, only rs12538371 and rs6970262 had significantly different genotype frequencies between HPV+ and HPV− HNSCC cases ($p = 0.001$ and $p = 0.022$, respectively). When we restricted our analyses to those tumors that were HPV ISH positive and P16 IHC positive, thereby defining the tumors most likely to have functioning HPV components, we observed similar results in that only rs12538371 and rs6970262 had significantly different genotype frequencies for HPV+P16+ versus HPV− HNSCC. No other SNPs associated with HNSCC risk differed in genotype frequency between HPV+P16+ and HPV− HNSCC cases (all $p > 0.05$). In MLR models, rs12538371 and rs6970262 were independently associated with HPV+ HNSCC risk but not HPV− HNSCC risk (Table 4). Three or more risk alleles in these two SNPs increased risk for HPV+P16+ HNSCC in a univariate model compared to 0 or 1 risk alleles (aOR = 7.94

(95CI = 2.90–21.70)). Two risk alleles were associated with a trend toward increased HPV +P16+ HNSCC risk compared to 0 or 1 risk allele (aOR = 2.27 (95CI = 0.98–6.22)).

Potential modifiers of tobacco-independent HNSCC risk

We explored tobacco-independent HNSCC risk SNPs (rs12538371, rs845561 and rs6970262) using the UCSC Genome Brower (http://genome.ucsc.edu/) [43]. SNP rs1238371 resides near the last shared exon of EGFR splice variants 1, 2 and 4 (Appendix: Supplementary Fig. S2). Of these variants, only the variant 1 transcript contains sequences encoding the kinase domain of EGFR. Other than the noteworthy location of this SNP, no distinguishing characteristics were noted in proximity to this SNP. Two regions of histone H3 Lysine27 acetylation (H3K27Ac), a marker of enhancer elements [44], were observed near the intron 21 SNP rs6970262 in a subset of ENCODE cell lines (Appendix: Supplementary Fig. S2). Enhanced H3K27Ac was also observed in intron 20 near SNP rs845561 in ENCODE cell lines (Appendix: Supplementary Fig. S2). EGFR intron 1 contains several regions of enhanced H3K27Ac. However, apart from intron 1, there exists a paucity of these histone modifications except for small islands within introns 20, 21 and 22 (Appendix: Supplementary Fig. S3).

The ENCODE project validated the expression of a non-coding RNA, EGFR-AS1 transcribed anti-sense to *EGFR*. SNP rs6970262 resides approximately 3.8 kb upstream of the transcription start of *EGFR-AS1*; rs845561 resides within intron 1 of *EGFR-AS1*, which has two exons. We detected the expression of *EGFR-AS1* in three HNSCC cell lines (Appendix: Supplementary Fig. S4).

Web-based analytical tools predicted that of the three tobacco-independent SNPs, only rs6970262 was predicted to have functional consequence by the NIEHS SNP prediction tool as a transcription factor-binding (TFB) site. Several transcription factors had predicted altered binding; among the top scoring were GRE and CDPCR-3. The RegulomeDB analysis indicated that DNase1 hypersensitivity sites were present in several cell lines for rs6970262 but only one cell line each for rs1253871 and rs845561. RegulomeDB positional weight matrices identified a possible c-FOS binding site at rs845561; no other binding motifs were identified. Histone modifications including H3K27Ac, H4K20me1, associated with gene expression, and H3K36me3, an intragenic mark of active transcription, were among the most frequent histone modifications for all 3 SNPs, implicating chromatin modification in these regions as important for EGFR or EGFR-AS1 expression regulation [45,46].

Discussion

Our study is the first to perform a comprehensive analysis of common *EGFR* genetic variants associated with HNSCC risk. We have shown that in a large Pittsburgh-based case–control study, 7 *EGFR* SNPs were associated with differential risk of developing HNSCC after adjustment for known risk factors. Two of these SNPs, rs2072454 and rs2075110, were in LD $r^2 \geq 0.98$ with each other. We characterized 1 SNP as tobacco-dependent (rs17586365), 3 SNPs as tobacco-independent (rs12538371, rs845561 and rs6970262) and 2 as associated with HPV+ HNSCC risk (rs1253871 and s6970262).

rs2072454 has been most studied of the identified SNPs of interest. rs2072454 is located within 2 bases of an exon–intron junction and has been predicted to affect mRNA splicing [40]. Although rs2072454 was reported to be not associated with breast [47] or lung cancer risk [13], our study suggests that individuals with homozygous minor (T) allele may be at increased risk for HNSCC.

Intron 1 sequence elements of genes have been generally recognized as regulators of transcription. A frequently studied genetic variant of *EGFR* is the CA dinucleotide repeat polymorphism located in intron 1. This polymorphism has been associated with differential transcription of *EGFR* and risk of oral cavity and lung cancers [15,20,48,49]. Neither rs17586365 nor rs12535536 was in LD with SNPs flanking the CA dinucleotide repeat polymorphism in the CEU population. Also, rs17586365 and rs12535536 were not in LD with SNP rs712829, another well-studied polymorphism shown to be involved in the transcriptional regulation of *EGFR* but not successfully assessed in our study [41], suggesting rs17586365 or genomic elements in LD may have functions distinct from these variants. Interestingly, rs17586365 was in LD ($r^2 = 0.86$) with identified glioma risk SNP rs11979158 also located within *EGFR* intron 1 [50].

Other studies have demonstrated that SNP rs11543848, a non-synonymous coding SNP in the ligand binding domain of EGFR (R521K), was associated with a differential risk of colorectal cancer as well as a differential response to EGFR targeted therapy in several cancers including HNSCC [21,23,51,52]. Differential risk and therapy response were thought to be due to reduced ligand binding by the variant protein [53]. However, in our study, rs11543848 was not significantly associated with HNSCC risk. Several studies have also reported similar negative findings for this variant in other malignancies [13,49,54,55]. These discordant findings may reflect the diversity of pathology of the cancers being studied and/or the heterogeneity of the populations.

We postulated that genetic risk factors identified among never tobacco users would have implications for HNSCC risk among tobacco users; our identification of HPV+ HNSCC risk EGFR SNPs among never tobacco users supports this hypothesis. We also hypothesized that SNPs specific to tobacco-related HNSCC risk would be identified, and our identification of rs17586365, which is in strong LD with a known glioma risk-associated genetic variant, as a modifier of HNSCC risk among tobacco users supports this. We speculate that associations with HNSCC risk for these SNPs were not identified in the genome-wide association study (GWAS) of upper aerodigestive cancers by McKay et al. [56] because subset analyses included heavy smokers or drinkers but not nonsmokers, where we observed our strongest associations. Also, combinations of risk SNPs as presented in Table 3 were required to reach substantial risk, and the GWAS study sought to identify highly associated single genetic variants.

Differential splicing has been implicated by our results, and EGFR isoforms have been postulated to play a variety of roles in the pathogenesis and therapeutic response of several malignancies [57]. The expression of HPV 16 E6 and E7 genes in keratinocytes has been reported to augment the expression of specific vascular endothelial growth factor splice variants *in vitro* [58]. Therefore, it is of interest to note that rs1253871 was associated with

HPV+ HNSCC, suggesting a possible interplay between HPV infection and *EGFR* transcript splicing regulators at or near this SNP.

Variants rs845561 and rs6970262 were especially of interest following ENCODE reports describing histone H3 modifications consistent with enhancer elements in close proximity [40]. While neither of these two SNPs has been experimentally confirmed to affect transcription, the possibility exists that they serve a regulatory role in *EGFR* expression. Whether the putative intron 20, 21 and 22 enhancer elements and/or expression of the *EGFR-AS1* antisense non-coding RNA transcript in this same region contributes specifically to enhanced expression of *EGFR* during the development and/or progression of tobacco-independent HNSCC remains to be investigated.

Our study has several limitations including the inability to assess all *EGFR* SNPs of interest and the small proportion of tumors with known HPV status. Validation of associations in an independent case– control study population would further substantiate findings. However, given these limitations, our evaluation of *EGFR* genetic variants has revealed heretofore unappreciated potential genetic elements of *EGFR* regulation and contribution to HNSCC risk.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

Funding Sources: This work was funded by K07CA137140 (AME) and P50CA097190 (JRG) grants from the NCI/NIH, the Howard Hughes Medical Fellows Program (CF) and NIH grant UL1 RR024153.

Appendix: Supplementary material

Supplementary data to this article can be found online at doi:10.1016/j.canlet.2014.12.008.

References

1. Boyle, P.; Levin, B., et al. International Agency for Research on Cancer. International Agency for Research on Cancer. Lyon Geneva: Distributed by WHO Press; 2008. World cancer report 2008.
2. Kamangar F, Dores GM, Anderson WF. Patterns of cancer incidence, mortality, and prevalence across five continents: defining priorities to reduce cancer disparities in different geographic regions of the world. J. Clin. Oncol. 2006; 24:2137–2150. [PubMed: 16682732]
3. Parkin DM, Bray F, Ferlay J, et al. Global cancer statistics. CA Cancer J. Clin. 2002; 55(2005):74–108. [PubMed: 15761078]
4. Howlander, N.; Noone, AM.; Krapcho, M., et al. SEER Cancer Statistics Review, 1975–2009. National Cancer Institute; 2012. (Vintage 2009 Populations)
5. Bei R, Budillon A, Masuelli L, et al. Frequent overexpression of multiple ErbB receptors by head and neck squamous cell carcinoma contrasts with rare antibody immunity in patients. J. Pathol. 2004; 204:317–325. [PubMed: 15476268]
6. Grandis JR, Tweardy DJ. Elevated levels of transforming growth factor alpha and epidermal growth factor receptor messenger RNA are early markers of carcinogenesis in head and neck cancer. Cancer Res. 1993; 53:3579–3584. [PubMed: 8339264]

7. Ongkeko WM, Altuna X, Weisman RA, et al. Expression of protein tyrosine kinases in head and neck squamous cell carcinomas. Am. J. Clin. Pathol. 2005; 124:71–76. [PubMed: 15923166]
8. Fung C, Grandis JR. Emerging drugs to treat squamous cell carcinomas of the head and neck. Expert Opin. Emerg. Drugs. 2010; 15:355–373. [PubMed: 20557270]
9. Andersson U, Schwartzbaum J, Wiklund F, et al. A comprehensive study of the association between the EGFR and ERBB2 genes and glioma risk. Acta Oncol. 2010; 49:767–775. [PubMed: 20446891]
10. Hou WG, Ai WB, Bai XG, et al. Genetic variation in the EGFR gene and the risk of glioma in a Chinese Han population. PLoS ONE. 2012; 7:e37531. [PubMed: 22662167]
11. Sanson M, Hosking FJ, Shete S, et al. Chromosome 7p11.2 (EGFR) variation influences glioma risk. Hum. Mol. Genet. 2011; 20:2897–2904. [PubMed: 21531791]
12. Schwartzbaum JA, Xiao Y, Liu Y, et al. Inherited variation in immune genes and pathways and glioblastoma risk. Carcinogenesis. 2010; 31:1770–1777. [PubMed: 20668009]
13. Choi JE, Park SH, Kim KM, et al. Polymorphisms in the epidermal growth factor receptor gene and the risk of primary lung cancer: a case-control study. BMC Cancer. 2007; 7:199. [PubMed: 17956637]
14. Jou YS, Lo YL, Hsiao CF, et al. Association of an EGFR intron 1 SNP with never-smoking female lung adenocarcinoma patients. Lung Cancer. 2009; 64:251–256. [PubMed: 19026460]
15. Zhang W, Weissfeld JL, Romkes M, et al. Association of the EGFR intron 1 CA repeat length with lung cancer risk. Mol. Carcinog. 2007; 46(5):372–380. [PubMed: 17219440]
16. Dong LM, Brennan P, Karami S, et al. An analysis of growth, differentiation and apoptosis genes with risk of renal cancer. PLoS ONE. 2009; 4:e4895. [PubMed: 19603096]
17. Bandres E, Barricarte R, Cantero C, et al. Epidermal growth factor receptor (EGFR) polymorphisms and survival in head and neck cancer patients. Oral Oncol. 2007; 43:713–719. [PubMed: 17112774]
18. Etienne-Grimaldi MC, Pereira S, Magne N, et al. Analysis of the dinucleotide repeat polymorphism in the epidermal growth factor receptor (EGFR) gene in head and neck cancer patients. Ann. Oncol. 2005; 16:934–941. [PubMed: 15829495]
19. Jin T, Yu CX, Lei DP, et al. Identification of epidermal growth factor receptor (EGFR) exon 20 single nucleotide polymorphism in Chinese squamous cell carcinoma of head and neck (SCCHN). Acta Otolaryngol. 2009; 129:1306–1312. [PubMed: 19863329]
20. Kang D, Gridley G, Huang WY, et al. Microsatellite polymorphisms in the epidermal growth factor receptor (EGFR) gene and the transforming growth factor-alpha (TGFA) gene and risk of oral cancer in Puerto Rico. Pharmacogenet. Genomics. 2005; 15:343–347. [PubMed: 15864136]
21. Klinghammer K, Knodler M, Schmittel A, et al. Association of epidermal growth factor receptor polymorphism, skin toxicity, and outcome in patients with squamous cell carcinoma of the head and neck receiving cetuximab-docetaxel treatment. Clin. Cancer Res. 2010; 16:304–310. [PubMed: 20028750]
22. Krohn V, Wiegand S, Werner JA, et al. EGFR codon 497 polymorphism – implications for receptor sensitivity to inhibitors in HNSCC cell lines. Anticancer Res. 2011; 31:59–65. [PubMed: 21273581]
23. Stoehlmacher-Williams J, Obermann L, Ehninger G, et al. Polymorphisms of the epidermal growth factor receptor (EGFR) and survival in patients with advanced cancer of the head and neck (HNSCC). Anticancer Res. 2012; 32:421–425. [PubMed: 22287728]
24. Lemos-Gonzalez Y, Paez de la Cadena M, Rodriguez-Berrocal FJ, et al. Absence of activating mutations in the EGFR kinase domain in Spanish head and neck cancer patients. Tumour Biol. 2007; 28:273–279. [PubMed: 17962724]
25. Loeffler-Ragg J, Witsch-Baumgartner M, Tzankov A, et al. Low incidence of mutations in EGFR kinase domain in Caucasian patients with head and neck squamous cell carcinoma. Eur. J. Cancer. 2006; 42:109–111. [PubMed: 16324836]
26. Chung CH, Ely K, McGavran L, et al. Increased epidermal growth factor receptor gene copy number is associated with poor prognosis in head and neck squamous cell carcinomas. J. Clin. Oncol. 2006; 24:4170–4176. [PubMed: 16943533]

27. Temam S, Kawaguchi H, El-Naggar AK, et al. Epidermal growth factor receptor copy number alterations correlate with poor clinical outcome in patients with head and neck squamous cancer. J. Clin. Oncol. 2007; 25:2164–2170. [PubMed: 17538160]
28. Wheeler S, Siwak DR, Chai R, et al. Tumor epidermal growth factor receptor and EGFR PY1068 are independent prognostic indicators for head and neck squamous cell carcinoma. Clin. Cancer Res. 2012; 18:2278–2289. [PubMed: 22351687]
29. Grandis JR, Melhem MF, Gooding WE, et al. Levels of TGF-alpha and EGFR protein in head and neck squamous cell carcinoma and patient survival. J. Natl. Cancer Inst. 1998; 90:824–832. [PubMed: 9625170]
30. Lubin JH, Purdue M, Kelsey K, et al. Total exposure and exposure rate effects for alcohol and smoking and risk of head and neck cancer: a pooled analysis of case-control studies. Am. J. Epidemiol. 2009; 170:937–947. [PubMed: 19745021]
31. Hashibe M, Brennan P, Benhamou S, et al. Alcohol drinking in never users of tobacco, cigarette smoking in never drinkers, and the risk of head and neck cancer: pooled analysis in the International Head and Neck Cancer Epidemiology Consortium. J. Natl. Cancer Inst. 2007; 99:777–789. [PubMed: 17505073]
32. Gillison ML, Souza GD, Westra W, et al. Distinct risk factor profiles for human papillomavirus type 16-positive and human papillomavirus type 16-negative head and neck cancers. J. Natl. Cancer Inst. 2008; 100:407–420. [PubMed: 18334711]
33. Kreimer AR, Clifford GM, Boyle P, et al. Human papillomavirus types in head and neck squamous cell carcinomas worldwide: a systematic review. Cancer Epidemiol. Biomarkers Prev. 2005; 14:467–475. [PubMed: 15734974]
34. Negri E, Boffetta P, Berthiller J, et al. Family history of cancer: pooled analysis in the International Head and Neck Cancer Epidemiology Consortium. Int. J. Cancer. 2009; 124:394–401. [PubMed: 18814262]
35. Buch SC, Nazar-Stewart V, Weissfeld JL, et al. Case-control study of oral and oropharyngeal cancer in whites and genetic variation in eight metabolic enzymes. Head Neck. 2008; 30:1139–1147. [PubMed: 18642288]
36. Frazer KA, Ballinger DG, Cox DR, et al. A second generation human haplotype map of over 3.1 million SNPs. Nature. 2007; 449:851–861. [PubMed: 17943122]
37. Amin W, Kang HP, Egloff AM, et al. An informatics supported web-based data annotation and query tool to expedite translational research for head and neck malignancies. BMC Cancer. 2009; 9:396. [PubMed: 19912644]
38. Shephard, N. GENASS: Stata Module to Perform Genetic Case-Control Association Tests. Chestnut Hill: Boston College Department of Economics; 2005.
39. Boyle AP, Hong EL, Hariharan M, et al. Annotation of functional variation in personal genomes using RegulomeDB. Genome Res. 2012; 22:1790–1797. [PubMed: 22955989]
40. Xu Z, Taylor JA. SNPinfo: integrating GWAS and candidate gene information into functional SNP selection for genetic association studies. Nucleic Acids Res. 2009; 37:W600–W605. [PubMed: 19417063]
41. Liu W, Innocenti F, Wu MH, et al. A functional common polymorphism in a Sp1 recognition site of the epidermal growth factor receptor gene promoter. Cancer Res. 2005; 65:46–53. [PubMed: 15665278]
42. Kuo CL, Feingold E. What's the best statistic for a simple test of genetic association in a case-control study? Genet. Epidemiol. 2010; 34:246–253. [PubMed: 20025064]
43. Rosenbloom KR, Dreszer TR, Long JC, et al. ENCODE whole-genome data in the UCSC Genome Browser: update. Nucleic Acids Res. 2012; 40(2012):D912–D917. [PubMed: 22075998]
44. Creyghton MP, Cheng AW, Welstead GG, et al. Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proc. Natl. Acad. Sci. U.S.A. 2010; 107:21931–21936. [PubMed: 21106759]
45. Beck DB, Oda H, Shen SS, et al. PR-Set7 and H4K20me1: at the crossroads of genome integrity, cell cycle, chromosome condensation, and transcription. Genes Dev. 2012; 26:325–337. [PubMed: 22345514]

46. Chantalat S, Depaux A, Hery P, et al. Histone H3 trimethylation at lysine 36 is associated with constitutive and facultative heterochromatin. Genome Res. 2011; 21:1426–1437. [PubMed: 21803857]
47. Gorlov, I.; Gorlova, O.; Jupe, E. Genetic Polymorphisms in EGFR Gene: No Evidence for an Association with the Risk of Breast Cancer; Thirty-Second Annual CTRC-AACR San Antonio Breast Cancer Symposium; 2009.
48. Gebhardt F, Zanker KS, Brandt B. Modulation of epidermal growth factor receptor gene transcription by a polymorphic dinucleotide repeat in intron 1. J. Biol. Chem. 1999; 274:13176–13180. [PubMed: 10224073]
49. Hong YS, Deming SL, Gao YT, et al. A two-stage case-control study of EGFR polymorphisms and breast cancer risk. Cancer Epidemiol. Biomarkers Prev. 2009; 18:680–683. [PubMed: 19190167]
50. Rajaraman P, Melin BS, Wang Z, et al. Genome-wide association study of glioma and meta-analysis. Hum. Genet. 2012; 131:1877–1888. [PubMed: 22886559]
51. Goncalves A, Esteyries S, Taylor-Smedra B, et al. A polymorphism of EGFR extracellular domain is associated with progression free-survival in metastatic colorectal cancer patients receiving cetuximab-based treatment. BMC Cancer. 2008; 8:169. [PubMed: 18544172]
52. Martinelli M, Ugolini G, Scapoli L, et al. The EGFR R521K polymorphism influences the risk to develop colorectal cancer. Cancer Biomark. 2010; 8:61–65. [PubMed: 21896992]
53. Moriai T, Kobrin MS, Hope C, et al. A variant epidermal growth factor receptor exhibits altered type alpha transforming growth factor binding and transmembrane signaling. Proc. Natl. Acad. Sci. U.S.A. 1994; 91:10217–10221. [PubMed: 7937865]
54. Kallel I, Rebai M, Khabir A, et al. Genetic polymorphisms in the EGFR (R521K) and estrogen receptor (T594T) genes, EGFR and ErbB-2 protein expression, and breast cancer risk in Tunisia. J. Biomed. Biotechnol. 2009; 2009:753683. [PubMed: 19636371]
55. Ma F, Sun T, Shi Y, et al. Polymorphisms of EGFR predict clinical outcome in advanced non-small-cell lung cancer patients treated with Gefitinib. Lung Cancer. 2009; 66:114–119. [PubMed: 19201048]
56. McKay JD, Truong T, Gaborieau V, et al. A genome-wide association study of upper aerodigestive tract cancers conducted within the INHANCE consortium. PLoS Genet. 2011; 7:e1001333. [PubMed: 21437268]
57. Wilken JA, Baron AT, Maihle NJ. The epidermal growth factor receptor conundrum. Cancer. 2010; 117:2358–2360. [PubMed: 24048781]
58. Toussaint-Smith E, Donner DB, Roman A. Expression of human papillomavirus type 16 E6 and E7 oncoproteins in primary foreskin keratinocytes is sufficient to alter the expression of angiogenic factors. Oncogene. 2004; 23:2988–2995. [PubMed: 14968115]

Table 1

Study population characteristics.

Characteristic	HNSCC cases (N = 578)	Controls (N = 588)	p*
Age, years			
Median (range)	59.9 (20.4–79.9)	60.2 (20.9–83.0)	0.408†
BMI			
Median (range)	27.1(15.3–61.8)	27.4 (16.4–59.4)	0.011†
Sex, N (%)			
Men	417 (72.1)	441 (75.0)	0.269‡
Women	161 (27.9)	147 (25.0)	
Tobacco//, N (%)			
Never	103 (17.8)	230 (39.1)	<0.001¶
Ever	475 (82.2)	358 (60.9)	
Cigarette only	326 (56.4)	229 (38.9)	
Pipe only	5 (0.9)	3 (0.5)	
Cigar only	5 (0.9)	6 (1.0)	
Chew only	13 (2.2)	8 (1.4)	
Cigarette and other tobacco	124 (21.5)	104 (17.7)	
No cigarette, combination other	2 (0.3)	8 (1.4)	
Cigarette, N (%)			
Never 0 pack-year	128 (22.1)	255 (43.4)	<0.001‡
>0 and <25 pack-year	121 (20.9)	173 (29.4)	
25–50 pack-year	172 (29.8)	102 (17.3)	
>50 pack-year	157 (27.2)	56 (9.5)	
Unknown quantity	0 (0)	2 (0.3)	
Alcohol, N (%)			
0 drink-year	105 (18.2)	131 (22.3)	<0.001‡
>0 and <30 drink-year	185 (32.0)	295 (50.2)	
30 + drink-year	282 (48.8)	159 (27.0)	
Unknown quantity	5 (0.9)	3 (0.5)	
Unknown	1 (0.2)	0 (0)	
Cigarette and alcohol use, N (%)			
Never smoker, < 30 drink-year	107 (18.5)	209 (35.5)	<0.001‡
Never smoker, 30 + drink-year	19 (3.3)	46 (7.8)	
<25 pack-year, <30 drink-year	66 (11.4)	121 (20.6)	
<25 pack-year, 30+ drink-year	55 (9.5)	51 (8.7)	
25–50 pack-year, < 30 drink-year	68 (11.8)	63 (10.7)	
25–50 pack-year, 30 + drink-year	103 (17.8)	37 (6.3)	
>50 pack-year, <30 drink-year	48 (8.3)	32 (5.4)	
>50 pack-year, 30+ drink-year	106 (18.3)	24 (4.1)	
Never smoker, unknown pack-year/drink-year	2 (0.3)	0 (0)	

Characteristic	HNSCC cases (N = 578)	Controls (N = 588)	p*
Ever smoker, unknown pack-year/drink-year	4 (0.7)	5 (0.9)	
Cancer site, N (%)			
Oral cavity	278 (48.1)	–	–
Oropharynx	134 (23.2)	–	
Hypopharynx	27 (4.7)	–	
Larynx	119 (20.6)	–	
Nasopharynx	6 (1.0)	–	
Other head/neck	14 (2.4)	–	
HPV			
Positive	37 (6.4)	–	
Negative	91 (15.7)	–	
Not evaluated	450 (77.9)	–	

*HNSCC cases versus controls.

†Rank-sum test.

‡Chi-square test.

¶Chi-square test Ever versus Never tobacco users.

//Cigarette, pipe, cigar, or chew use.

Table 2

Tests of association for all subjects and stratified by tobacco use for SNPs associated with HNSCC by genotype or trend test.

SNP Number/ SNP ID[¶]/Gene Location	Genotype	All Subjects					All Subjects					Never Tobacco Users				
		HNSCC Cases (N = 578) N (%)	Controls (N = 588) N (%)	p-geno	p-trend	OR (95% CI)[†]	HNSCC Cases (N = 475) N (%)	Controls (N = 358) N (%)	p-geno	p-trend	OR (95% CI)[†]	HNSCC Cases (N = 103) N (%)	Controls (N = 230) N (%)	p-geno	p-trend	OR (95% CI)[‡]
8	GG	430 (77.2)	414 (74.6)			Ref	351 (76.8)	244 (73.5)			Ref	79 (79.0)	170 (76.2)			Ref
rs17586365	GA	123 (22.1)	130 (23.4)			0.91 (0.67–1.24)	105 (23.0)	81 (24.4)			0.95 (0.66–1.35)	18 (18.0)	49 (22.0)			0.79(0.43–1.45)
Intron 1	AA	4 (0.7)	11 (2.0)	0.151	0.161	0.45 (0.14–1.49)	1 (0.2)	7 (2.1)	**0.028**	0.116	**0.12 (0.01–0.97)**	3 (3.0)	4 (1.8)	0.543	0.788	1.75 (0.37–8.19)
	GA/AA	127 (23.2)	141 (25.4)			0.88 (0.65–1.18)	106 (23.2)	88 (26.5)			0.88 (0.62–1.24)	21 (21.0)	53 (23.8)			0.86 (0.48–1.53)
13	AA	255 (44.1)	292 (49.7)			Ref	213 (44.8)	180 (50.3)			Ref	42 (40.8)	112 (48.7)			Ref
rs12535536	AG	265 (45.8)	249 (42.3)			1.28 (0.99–1.66)	211 (44.4)	150 (41.9)			1.23 (0.91–1.68)	54 (52.4)	99 (43.0)			1.48 (0.90–2.42)
Intron 1	GG	58 (10.0)	47 (8.0)	0.133	**0.045**	1.39 (0.88–2.18)	51 (10.7)	28 (7.8)	0.186	0.067	1.59 (0.93–2.73)	7 (6.8)	19 (8.3)	0.293	0.386	0.95 (0.37–2.44)
	AG/GG	323 (55.9)	296 (50.3)			**1.29 (1.01–1.66)**	262 (55.2)	178 (49.7)			1.29 (0.96–1.73)	61 (59.2)	118 (51.3)			1.39 (0.86–2.24)
27	AA	268 (46.5)	229 (38.9)			Ref	225 (47.6)	149 (41.6)			Ref	43 (41.7)	80 (34.8)			Ref
rs6593206	AC	235 (40.8)	279 (47.4)			0.78 (0.59–1.02)	188 (39.7)	165 (46.1)			0.79 (0.58–1.09)	47 (45.6)	114 (49.6)			0.77 (0.46–1.29)
Intron 1	CC	73 (12.7)	80 (13.6)	**0.030**	**0.034**	0.82 (0.56–1.22)	60 (12.7)	44 (12.3)	0.169	0.248	0.92 (0.58–1.48)	13 (12.6)	36 (15.7)	0.464	0.218	0.64 (0.30–1.35)
	AC/CC	308 (53.5)	359 (61.1)			0.79 (0.62–1.02)	248 (52.4)	209 (58.4)			0.82 (0.61–1.11)	60 (58.3)	150 (65.2)			0.74 (0.46–1.20)
31	AA	378 (65.6)	345 (58.9)			Ref	311 (65.6)	207 (57.8)			Ref	67 (65.7)	138 (60.5)			Ref
rs4947979[§]	AG	174 (30.2)	219 (37.4)			0.78 (0.59–1.02)	141 (29.7)	141 (39.4)			0.74 (0.54–1.00)	33 (32.4)	78 (34.2)			0.84 (0.50–1.40)
Intron 1	GG	24 (4.2)	22 (3.8)	**0.036**	0.057	0.94 (0.49–1.79)	22 (4.6)	10 (2.8)	**0.009**	0.134	1.37 (0.61–3.09)	2 (2.0)	12 (5.3)	0.352	0.216	0.35 (0.07–1.61)
	AG/GG	198 (34.4)	241 (41.1)			0.79 (0.61–1.03)	163 (34.4)	151 (42.2)			0.78 (0.58–1.06)	35 (34.3)	90 (39.5)			0.78 (0.47–1.27)
35	AA	455 (78.7)	437 (74.4)			Ref	372 (78.3)	259 (72.5)			Ref	83 (80.6)	178 (77.4)			Ref
rs6969570[§]	AG	112 (19.4)	145 (24.7)			0.76 (0.56–1.03)	94 (19.8)	95 (26.6)			0.76 (0.53–1.07)	18 (17.5)	50 (21.7)			0.74 (0.40–1.36)
Intron 1	GG	11 (1.9)	5 (0.9)	**0.034**	0.235	2.11 (0.69–6.42)	9 (1.9)	3 (0.8)	**0.036**	0.151	2.18 (0.56–8.52)	2 (1.9)	2 (0.9)	0.441	0.689	1.96 (0.27–14.41)
	AG/GG	123 (21.3)	150 (25.6)			0.81 (0.60–1.09)	103 (21.7)	98 (27.5)			0.80 (0.57–1.12)	20 (19.4)	52 (22.6)			0.79 (0.44–1.42)
36	TT	155 (26.8)	138 (23.5)			Ref	126 (26.5)	89 (24.9)			Ref	29 (28.2)	49 (21.3)			Ref
rs2072454[//]	TC	310 (53.6)	305 (51.9)			0.87 (0.65–1.18)	259 (54.5)	187 (52.2)			0.93 (0.66–1.33)	51 (49.5)	118 (51.3)			0.72 (0.41–1.28)
Exon 4	CC	113 (19.6)	145 (24.6)	0.086	**0.036**	**0.67 (0.47–0.96)**	90 (18.9)	82 (22.9)	0.372	0.237	0.68 (0.44–1.04)	23 (22.3)	63 (27.4)	0.344	0.153	0.66 (0.34–1.29)
Asn158Asn	TC/CC	423 (73.2)	450 (76.5)			0.81 (0.61–1.08)	349 (73.5)	269 (75.1)			0.85 (0.61–1.19)	74 (71.8)	181 (78.7)			0.70 (0.41–1.20)
37	TT	156 (27.0)	140 (23.8)			Ref	127 (26.8)	91 (25.4)			Ref	29 (28.2)	49 (21.3)			Ref

SNP Number/ SNP ID[¶]/Gene Location	Genotype	All Subjects					All Subjects					Never Tobacco Users				
		HNSCC Cases (N = 578) N (%)	Controls (N = 588) N (%)	p-geno	p-trend	OR (95% CI)[†]	HNSCC Cases (N = 475) N (%)	Controls (N = 358) N (%)	p-geno	p-trend	OR (95% CI)[†]	HNSCC Cases (N = 103) N (%)	Controls (N = 230) N (%)	p-geno	p-trend	OR (95% CI)[‡]
rs2075110[//]	TC	310 (53.7)	306 (52.0)			0.88 (0.66–1.19)	258 (54.4)	187 (52.2)			0.95 (0.67–1.34)	52 (50.5)	119 (51.7)			0.73 (0.41–1.29)
Intron 5	CC	111 (19.2)	142 (24.1)	0.099	**0.043**	**0.65 (0.47–0.98)**	89 (18.8)	80 (22.3)	0.448	0.299	0.70 (0.46–1.08)	22 (21.4)	62 (27.0)	0.312	0.133	0.64 (0.32–1.26)
	TC/CC	421 (72.9)	448 (76.1)			0.82 (0.62–1.09)	347 (73.2)	267 (74.5)			0.87 (0.62–1.22)	74 (71.9)	181 (78.7)			0.70 (0.41–1.20)
44	TT	454 (78.5)	495 (84.2)			Ref	378 (79.6)	301 (84.1)			Ref	76 (73.8)	194 (84.3)			Ref
rs12538371	TC	120 (20.8)	88 (15.0)			**1.47 (1.06–2.04)**	93 (19.6)	52 (14.5)			1.38 (0.93–2.04)	27 (26.2)	36 (15.7)			**1.82 (1.02–3.22)**
Intron 15	CC	4 (0.7)	5 (0.9)	**0.029**	**0.024**	0.64 (0.16–2.61)	4 (0.8)	5 (1.4)	0.128	0.183	0.64 (0.16–2.56)	0 (0)	0 (0)	**0.033**	**0.023**	-
	TC/CC	124 (21.5)	93 (15.8)			**1.42 (1.03–1.95)**	97 (20.4)	57 (15.9)			1.31 (0.89–1.92)	27 (26.2)	36 (15.7)			**1.82 (1.02–3.22)**
51	TT	327 (56.8)	362 (61.9)			Ref	271 (57.3)	215 (60.4)			Ref	56 (54.4)	147 (64.2)			Ref
rs845561	TC	220 (38.2)	201 (34.4)			1.25 (0.96–1.63)	181 (38.3)	127 (35.7)			1.18 (0.87–1.61)	39 (37.9)	74 (32.3)			1.45 (0.88–2.40)
Intron 20	CC	29 (5.0)	22 (3.8)			1.69 (0.90–3.17)	21 (4.4)	14 (3.9)	0.671	0.373	1.27 (0.59–2.72)	8 (7.8)	8 (3.5)	0.108	**0.043**	**3.09 (1.08–8.86)**
	TC/CC	249 (43.2)	223 (38.1)	0.174	0.060	**1.29 (1.00–1.66)**	202 (42.7)	141 (39.6)			1.19 (0.88–1.60)	47 (45.6)	82 (35.8)			1.59 (0.98–2.58)
55	GG	239 (41.3)	205 (34.9)			Ref	192 (40.4)	130 (36.3)			Ref	47 (45.6)	75 (32.6)			Ref
rs6970262	GA	264 (45.7)	294 (50.0)			0.77 (0.58–1.00)	217 (45.7)	172 (48.0)			0.86 (0.62–1.18)	47 (45.6)	122 (53.0)			**0.59 (0.36–0.98)**
Intron 21	AA	75 (13.0)	89 (15.1)	0.069	**0.030**	0.69 (0.47–1.02)	66 (13.9)	56 (15.6)	0.456	0.225	0.80 (0.51–1.26)	9 (8.7)	33 (14.3)	0.056	**0.017**	**0.40 (0.18–0.94)**
	GA/AA	339 (58.7)	383 (65.1)			**0.75 (0.58–0.97)**	283 (59.6)	228 (63.7)			0.85 (0.63–1.14)	56 (54.4)	155 (67.4)			**0.55 (0.34–0.89)**

[†]Adjusted for sex, age, packyear category, drinkyear category and packyear–drinkyear category interaction.

[‡]Adjusted for sex and age.

[§]rs4947979 and rs6969570 LD CEU HapMap r^2 = 0.561; LD among controls r^2 = 0.52.

[//]rs2072454 and rs28110 LD CEU HapMap r^2 = 1.00; LD among controls r^2 = 0.98.

[¶]LD r^2 provided if $r^2 \geq 0.5$.

Cancer Lett. Author manuscript; available in PMC 2016 February 28.

Table 3

Tests of association by number of SNP risk-associated alleles by tobacco use status.

Analysis	SNP and risk allele	N risk alleles	All subjects			Ever tobacco users			Never tobacco users		
			N	Adjusted OR† (95% CI)	P trend	N	Adjusted OR† (95% CI)	P trend	N	Adjusted OR‡ (95% CI)	P trend
1. Sig. tobacco independent SNPs	rs12538371[C] +	0	154	Ref.		118	Ref.		36	Ref.	
	rs845561[C] +	1	291	1.06 (0.69–1.62)		197	0.98 (0.60–1.61)		94	1.65 (0.64–4.26)	
	rs6970262[G]	2	363	1.15 (0.76–1.73)		254	1.04 (0.65–1.66)		109	1.79 (0.70–4.54)	
		3	244	1.44 (0.93–2.22)		180	1.42 (0.86–2.34)		64	1.90 (0.70–5.18)	
		4	109	2.11 (1.23–3.62)	**0.002**	80	1.29 (0.69–2.39)	0.114	29	8.62 (2.75–27.08)	**0.001**
2. All Sig. SNPs	rs17586365[G] +	0–3	161	Ref.		122	Ref.		39	Ref.	
	rs12535536[G] +	4	219	1.62 (1.03–2.55)		148	1.75 (1.04–2.93)		71	1.66 (0.59–4.71)	
	rs2072454[T] +	5	324	2.12 (1.39–3.24)		233	2.13 (1.33–3.43)		91	2.45 (0.91–6.59)	
	rs12538371[C] +	6	227	2.47 (1.57–3.87)		155	2.39 (1.43–3.99)		72	3.04 (1.11–8.33)	
	rs845561[C] +	7	140	2.68 (1.61–4.42)		100	2.18 (1.22–3.87)		40	5.08 (1.72–15.03)	
	rs6970262[G]	8–10	90	3.09 (1.73–5.50)	**2.37×10^{-6}**	71	3.05 (1.59–5.85)	**3.57×10^{-4}**	19	3.96 (1.10–14.21)	**0.001**
3. Sig. non-tobacco independent SNPs	rs17586365[G] +	0–2	287	Ref.		199	Ref.		88	Ref.	
	rs12535536[G] +	3	372	1.77 (1.26–2.47)		278	2.06 (1.39–3.06)		94	1.15 (0.59–2.24)	
	rs2072454[T]	4	323	1.85 (1.30–2.62)		221	2.06 (1.36–3.13)		102	1.46 (0.77–2.77)	
		5–6	184	2.04 (1.36–3.06)	**3.51×10^{-4}**	135	2.17 (1.35–3.48)	**0.001**	49	1.82 (0.85–3.88)	0.088

†Adjusted for sex, age, packyear category, drinkyear category and packyear–drinkyear category interaction.

‡Adjusted for sex and age.

Cancer Lett. Author manuscript; available in PMC 2016 February 28.

Table 4

Tests of association by tumor HPV status.

SNP number/SNP ID¶/gene location	Genotype	Controls (N = 588) N (%)	HPV+ HNSCC cases (N = 37) N (%)	Adjusted OR (95% CI)†	HPV– HNSCC cases (N = 91) N (%)	Adjusted OR (95% CI)†
8	GG	414 (74.6)	28 (77.8)	Ref	72 (80.9)	Ref
rs17586365	GA	130 (23.4)	7 (19.4)	0.82 (0.34–1.99)	16 (18.0)	0.72 (0.39–1.32)
Intron 1	AA	11 (2.0)	1 (2.8)	1.46 (0.17–12.65)	1 (1.1)	0.61 (0.07–5.03)
	GA/AA	141 (25.4)	8 (22.2)	0.87 (0.38–2.02)	17 (19.1)	0.71 (0.39–1.28)
13	AA	292 (49.7)	17 (45.9)	Ref	40 (44.0)	Ref
rs12535536	AG	249 (42.3)	16 (43.2)	1.21 (0.58–2.49)	40(44.0)	1.20 (0.73–1.98)
Intron 1	GG	47 (8.0)	4 (10.8)	0.89 (0.58–6.14)	11 (12.1)	1.71 (0.78–3.75)
	AG/GG	296 (50.3)	20 (54.1)	1.30 (0.65–2.59)	51 (56.0)	1.29 (0.80–2.06)
36	TT	138 (23.5)	6 (16.2)	Ref	29 (31.9)	Ref
rs2072454//	TC	305 (51.9)	21 (56.8)	1.71 (0.66–4.43)	45 (49.5)	**0.55 (0.32–0.96)**
Exon 4	CC	145 (24.7)	10 (27.0)	1.59 (0.55–4.61)	17 (18.7)	**0.46 (0.23–0.91)**
Asn158Asn	TC/CC	450 (76.5)	31 (83.8)	1.67 (0.67–4.17)	62 (68.1)	**0.52 (0.31–0.86)**
37	TT	140 (23.8)	6 (16.2)	Ref	30 (33.0)	Ref
rs2075110//	TC	306 (52.0)	21 (56.8)	1.72 (0.67–4.47)	44 (48.4)	**0.54 (0.31–0.92)**
Intron 5	CC	142 (24.1)	10 (27.0)	1.61 (0.56–4.67)	17 (18.7)	**0.47 (0.24–0.92)**
	TC/CC	448 (76.2)	31 (83.8)	1.69 (0.68–4.22)	61 (67.0)	**0.51 (0.31–0.86)**
44	TT	495 (84.2)	22 (59.5)	Ref	79 (86.8)	Ref
rs12538371	TC	88 (15.0)	15 (40.5)	**4.19 (2.02–8.73)**	11 (12.1)	0.71 (0.35–1.45)
Intron 15	CC	5 (0.9)	0 (0.0)	-	1 (1.1)	1.04 (0.11–9.77)
	TC/CC	93 (15.8)	15 (40.5)	**4.01 (1.93–8.32)**	12 (13.2)	0.73 (0.37–1.45)
51	TT	362 (61.9)	21 (56.8)	Ref	47 (51.6)	Ref
rs845561	TC	201 (34.4)	13 (35.1)	0.99 (0.48–2.09)	40 (44.0)	1.48 (0.90–2.40)
Intron 20	CC	22 (3.8)	3 (8.1)	3.09 (0.81–11.81)	4 (4.4)	1.68 (0.51–5.55)
	TC/CC	223 (38.1)	16 (43.2)	1.15 (0.58–2.31)	44 (48.4)	1.49 (0.93–2.40)
55	GG	205 (34.9)	22 (59.5)	Ref	32 (35.2)	Ref
rs6970262	GA	294 (50.0)	13 (35.1)	**0.48 (0.23–0.99)**	41 (45.1)	1.02 (0.60–1.79)
Intron 21	AA	89 (15.1)	2 (5.4)	**0.21 (0.05–0.93)**	18 (19.8)	1.63 (0.83–3.19)

SNP number/SNP ID¶/gene location	Genotype	Controls (N = 588) N (%)	HPV+ HNSCC cases (N = 37) N (%)	Adjusted OR (95% CI)†	HPV– HNSCC cases (N = 91) N (%)	Adjusted OR (95% CI)†
	GA/AA	383 (65.1)	15 (40.5)	**0.41 (0.20–0.82)**	59 (64.8)	1.16 (0.70–1.90)

†Adjusted for sex, age, packyear category, drinkyear category and packyear–drinkyear category interaction.

//rs2072454 and rs2075110 LD CEU HapMap r^2 = 1.00; LD among controls r^2 = 0.98.

¶LD r^2 provided if $r^2 \geq 0.5$.