

This Provisional PDF corresponds to the article as it appeared upon acceptance. Fully formatted PDF and full text (HTML) versions will be made available soon.

Association between novel PLCE1 variants identified in published esophageal cancer genome-wide association studies and risk of squamous cell carcinoma of the head and neck

BMC Cancer 2011, **11**:258 doi:10.1186/1471-2407-11-258

Hongxia Ma (hma@mdanderson.org)
Li-E Wang (lwang@mdanderson.org)
Zhensheng Liu (zhliu@mdanderson.org)
Erich M Sturgis (esturgis@mdanderson.org)
Qingyi Wei (qwei@mdanderson.org)

ISSN 1471-2407

Article type Research article

Submission date 23 December 2010

Acceptance date 20 June 2011

Publication date 20 June 2011

Article URL <http://www.biomedcentral.com/1471-2407/11/258>

Like all articles in BMC journals, this peer-reviewed article was published immediately upon acceptance. It can be downloaded, printed and distributed freely for any purposes (see copyright notice below).

Articles in BMC journals are listed in PubMed and archived at PubMed Central.

For information about publishing your research in BMC journals or any BioMed Central journal, go to

<http://www.biomedcentral.com/info/authors/>

Association between novel *PLCE1* variants identified in published esophageal cancer genome-wide association studies and risk of squamous cell carcinoma of the head and neck

Hongxia Ma^{1,2}, Li-E Wang¹, Zhensheng Liu¹, Erich M. Sturgis^{1,3}, Qingyi Wei^{1,*}

Author Affiliations:

¹Departments of Epidemiology, The University of Texas M.D. Anderson Cancer Center, Houston, Texas 77030, USA

²Department of Epidemiology and Biostatistics, School of Public Health, Nanjing Medical University, Nanjing 210029, China

³Department of Head and Neck Surgery, The University of Texas M.D. Anderson Cancer Center, Houston, Texas 77030, USA

***Correspondence to:** Qingyi Wei, Department of Epidemiology, Unit 1365, The University of Texas MD Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, TX 77030. Phone: 713-792-3020; Fax: 713-563-0999; E-mail: qwei@mdanderson.org.

Running title: *PLCE1* variants and risk of head and neck cancer

Key words: *PLCE1*; polymorphism; SCCHN; risk; susceptibility

Abstract

Background: Phospholipase C epsilon 1 (PLCE1) (an effector of Ras) belonging to the phospholipase family plays crucial roles in carcinogenesis and progression of several cancers, including squamous cell carcinoma of the head and neck (SCCHN). A single nucleotide polymorphism (SNP, rs2274223) in *PLCE1* has been identified as a novel susceptibility locus in genome-wide association studies (GWAS) of esophageal squamous cell carcinoma (ESCC) and gastric cardia adenocarcinoma (GCA) that share similar risk factors with SCCHN. Therefore, we investigated the association between potentially functional SNPs in *PLCE1* and susceptibility to SCCHN.

Methods: We genotyped three potentially functional SNPs (rs2274223A/G, rs3203713A/G and rs11599672T/G) of *PLCE1* in 1,098 SCCHN patients and 1,090 controls matched by age and sex in a non-Hispanic white population.

Results: Although none of three SNPs was alone significantly associated with overall risk of SCCHN, their combined effects of risk alleles (rs2274223G, rs3203713G and rs11599672G) were found to be associated with risk of SCCHN in a locus-dose effect manner ($P_{\text{trend}}=0.046$), particularly for non-oropharyngeal tumors ($P_{\text{trend}}=0.017$); specifically, rs2274223 was associated with a significantly increased risk (AG vs. AA: adjusted OR=1.29, 95% CI=1.01-1.64; AG/GG vs. AA: adjusted OR=1.30, 95% CI=1.03-1.64), while rs11599672 was associated with a significantly decreased risk (GG vs. TT: adjusted OR=0.54, 95% CI=0.34-0.86; TG/GG vs. TT: adjusted OR=0.76, 95% CI=0.61-0.95).

Conclusions: Our findings suggest that *PLCE1* variants may have an effect on risk of SCCHN associated with tobacco and alcohol exposure, particularly for those tumors arising at non-oropharyngeal sites. These findings, although need to be validated by larger studies, are consistent with those in esophageal and gastric cancers.

Background

Squamous cell carcinoma of the head and neck (SCCHN) is the fifth most common cancer worldwide [1], which includes malignancies at the sites of the oral cavity, larynx and pharynx. It was estimated that there were 48,010 new cases and 11,260 deaths of SCCHN in the United States in 2010 [2]. Tobacco smoke and alcohol consumption are the well-recognized risk factors for SCCHN, and more recently human papillomavirus (HPV) is recognized as a major risk factor for the oropharyngeal cancer [3]. However, accumulative evidence also shows that genetic factors, including family history and polymorphisms in genes involved in multiple biological pathways, such as carcinogen metabolism, DNA repair, cell cycle regulation, apoptosis and other cellular processes, play important roles in the etiology of SCCHN [4, 5]. Although an increasing number of studies on genetic risk factors for SCCHN have been published, the exact genetic basis of susceptibility to SCCHN is still not well defined.

In the past few years, the wave of genome-wide association studies (GWASs) provided a more robust tool to find novel susceptibility loci or genes for cancer susceptibility by using high-throughput genotyping technology to interrogate a large number of tagging polymorphisms in a high density across the whole genome [6]. To date, more than 50 cancer GWASs have been published, including at least 15 different malignant tumors [7], which has greatly improved our understanding of genetic basis of human cancers. However, only one recent GWAS focused on SCCHN risk and identified five variants at 4q21, 12q24 and ADH gene cluster, significantly associated with risk of upper aerodigestive tract cancers (UATC) including SCCHN [8]. It may be a very small proportion of SNPs associated with SCCHN risk because of the strict criteria for the GWAS significance level ($P=10^{-7}$ or $P=10^{-8}$). Thus, further exploration for the genetic variants that did not reach the GWAS significance level in the development of SCCHN is needed.

In 2010, two large-scale genome-wide association studies [9, 10] simultaneously reported that a new and notable low-penetrance susceptibility locus (rs2274223) located in the phospholipase C epsilon 1 gene (*PLCE1*) was strongly associated with risk of esophageal squamous cell carcinoma (ESCC) and gastric cardia adenocarcinoma (GCA) in Chinese population. Rs2274223 is a non-synonymous SNP located in exon 26 of the *PLCE1* gene, causing the amino acid change from histidine to arginine. Additionally, one study also showed that the positive rates of the *PLCE1* protein in ESCC and GCA tissues were significantly higher than that in normal ones, suggesting a biologically plausible role of *PLCE1* in carcinogenesis of human cancers [9].

The *PLCE1* gene resides on chromosome 10q23 and encodes phospholipase C epsilon 1 (*PLCE1*), which belongs to the phospholipase family that catalyzes the hydrolysis of polyphosphoinositides to generate secondary messengers, such as inositol-1,4,5 trisphosphate and diacylglycerol. Therefore, *PLCE1* is involved in cell growth and differentiation [11]. Studies have reported that *PLCE1* functions as an effector of Ras and plays crucial roles in carcinogenesis and progression of several cancers, including cancers of the intestine [12], skin [13], bladder [14], colorectal [15] and head and neck [16]. All these findings further support the biological plausibility that genetic variations, such as single nucleotide polymorphisms (SNPs) in *PLCE1* that affect the gene expression or protein functions, may affect the risk of some cancers.

Studies have reported that SCCHN shares some similar risk factors with ESCC and GCA, such as tobacco smoke and alcohol consumption [17-19]. Thus, the exciting results in the GWAS of ESCC and GCA, and the strong biological evidence encouraged us to investigate the association between functional SNPs in *PLCE1* and susceptibility to SCCHN. In the present study, we performed genotyping analyses of three potentially functional SNPs in *PLCE1*

(rs2274223A/G, rs3203713A/G and rs11599672T/G) in non-Hispanic whites and assessed their associations with risk of SCCHN in our ongoing hospital-based case-control study of 1,098 SCCHN cases and 1,090 cancer-free controls matched to cases on age (± 5 years) and sex.

Methods

Study populations

Participant recruitment has been described elsewhere [20, 21]. In brief, newly-diagnosed SCCHN cancer patients were consecutively recruited from The University of Texas M.D. Anderson Cancer Center between October 1999 and October 2007. All cases were diagnosed with histologically confirmed SCCHN, and there were no age, sex or stage restrictions. Exclusion criteria included second SCCHN primary tumors, primary tumors of the nasopharynx or sinonasal tract, metastasized cancer from other organs, or any histopathologic diagnosis other than SCCHN. Additionally, patients who had received prior surgery (other than diagnostic biopsies), chemotherapy or radiation therapy before recruitment, and any blood transfusion during the preceding 6 months were also excluded.

Cancer-free controls were recruited from those visitors in outpatient clinics at M.D. Anderson Cancer Center, who were not genetically related to the patients or to each other, and were frequency-matched to the cases on age (± 5 years), sex and ethnicity. The inclusion criterion for controls was defined as the self-reported absence of prior history of cancer. After the first survey using a short questionnaire to find out whether potential subjects were willing to participate and qualified in this study, a written informed consent was obtained, and a structured questionnaire was administered by trained interviewers to collect demographic data and environmental exposure history, such as age, sex, ethnicity, smoking and alcohol consumption. After interview, approximately 30-ml venous blood sample was collected from each participant.

The response rates for cases and controls were approximately 90% and 85%, respectively. In the analysis, a total of 1,098 cases and 1,090 controls that completed the interview and donated a one-time blood sample were included. All subjects were non-Hispanic whites. This study was approved by the Institutional Review Board of M. D. Anderson.

SNP Selection

Besides the non-synonymous SNP (rs2274223) identified in the published GWAS, we also searched the NCBI dbSNP database (<http://www.ncbi.nlm.nih.gov/>, build 131) for other common [minor allele frequency (MAF) ≥ 0.05] and potentially functional SNPs located in the 5' near gene, 5'- and 3'-untranslated regions of *PLCE1* by using a set of tools at the website SNPinfo (<http://snpinfo.niehs.nih.gov/>) [22]. In addition, the LD analyses were performed to optimize SNP selection. We selected additional two potentially functional SNPs in *PLCE1*: rs3203713, located in the miRNA binding site (3'UTR) and rs11599672, located in the transcription factor binding site (TFBS in 5' near gene). Taken together, we included three SNPs (rs2274223A/G, rs3203713A/G and rs11599672T/G) in *PLCE1* for the final genotyping.

Genotyping

We extracted genomic DNA from the buffy-coat fraction of the blood samples using the Qiagen Blood DNA Mini Kit (Qiagen Inc., Valencia, Calif) according to the manufacturer's instruction and genotyped SNPs rs2274223 and rs3203713 using the TaqMan allelic discrimination assay on an ABI 7900 system (Applied Biosystems). Genotyping was performed without knowing the subjects' case or control status, and four negative controls (no DNA) and duplicated commercial positive controls (TaqMan Control Genomic DNA; Applied Biosystems) included in each 384-well plate were used for quality control. The concordance achieved 100% for the duplicates of 5%

of samples. Polymerase chain reaction-restriction fragment length polymorphisms (PCR-RFLP) method was also used to identify genotypes of rs11599672, because the Taqman assay was not applicable to this SNP. The antisense primer was introduced a mismatched A to replace T at +2bp from the polymorphic site to create an *SspI* restriction site (sense: 5'-GGAGAGACATTCTGTTGGGTGA-3'; antisense: 5'- CCTTCAAACCACCGCTGTAAT- 3'). A 155-bp fragment containing this T/G site was amplified in the PCR mixture consisted with approximately 20 ng of genomic DNA, 2 pmol of each primer, 0.1 mM each dNTP, 1× PCR buffer (50 mM KCl, 10 mM Tris HCl, and 0.1 % Triton X-100), 1.5 mM MgCl₂, and 0.5 unit of *Taq* polymerase. The PCR profile consisted of an initial melting step of 95°C for 5 min; 35 cycles of 95°C for 30 s, 59°C for 45 s, and 72°C for 1 min; and final extension step of 72°C for 10 min. The PCR product was then digested with the restriction enzyme *SspI* (New England BioLabs, Beverly, MA) overnight at 37°C and separated on 3% agarose gel. The T allele has the restriction site and produces two bands of 134- and 21-bp, whereas the G allele lacks the *SspI* restriction site, resulting in one band of 155-bp (**Figure 1**). Finally, 10% of the samples were randomly selected to perform the repeated assays, and the results were 100% concordant. Direct sequencing was also conducted to confirm the genotypes of rs11599672 (**Figure 2**).

Statistical analysis

We evaluated differences in selected demographic variables, risk factors, and frequencies of the *PLCE1* genotypes between cases and controls by using the χ^2 test and examined Hardy-Weinberg equilibrium by a goodness-of-fit χ^2 test to compare the observed genotype frequencies with the expected ones among the controls. We estimated associations between *PLCE1* variants and SCCHN risk by computing odds ratios (ORs) and 95% confidence intervals (CIs) from both

univariate and multivariate logistic regression analyses with adjustment for the known risk factors for SCCHN, such as age, sex, smoking status and alcohol use. Haplotype frequencies and individual haplotypes were generated by using SAS PROC HAPLOTYPE based on the observed genotypes. All statistical analyses were two sided, and $P < 0.05$ was considered statistically significant. All the statistical analyses were performed with Statistical Analysis System software (v.9.1 SAS Institute, Cary, NC).

Results

Characteristics of study subjects

The characteristics of cases and controls included in this study are summarized in **Table 1**. The mean age was 57.1 years (± 11.1) for cases and 56.6 years (± 11.0) for controls. There were no significant differences in the distributions of age and sex between cases and controls ($P = 0.676$ and 0.581 , respectively), suggesting the frequency-matching on age and sex was adequate. However, 72.2% and 72.9% of the SCCHN cases were smokers and drinkers, respectively, which were significantly higher than that of the controls (50.9% and 56.6%, respectively) ($P < 0.001$ for the both). Of the 1,098 cases, 559 (50.9%) had tumors of the oropharynx and 539 (49.1%) had tumors arising at non-oropharynx sites including oral cavity, hypopharynx and larynx. Furthermore, 272 cases (24.8%) presented with stage I-II and 826 cases (75.2%) presented with III-IV stage.

Overall associations between PLCE1 variants and risk of SCCHN

The genotype and allele distributions of three SNPs (rs2274223, rs3203713 and rs11599672) in cases and controls are shown in **Table 2**. The observed genotype frequencies for these three

polymorphisms were all in Hardy-Weinberg equilibrium in the controls ($P = 0.07, 0.98$ and 0.72 , respectively). The single locus analyses revealed that genotype distributions of these three polymorphisms were not significantly different between overall cases and controls ($P = 0.554$ for rs2274223, $P=0.860$ for rs3203713 and $P = 0.265$ for rs11599672, respectively). However, we found that the frequencies of variant rs2274223G and rs3203713G alleles (0.335 and 0.170, respectively) among the cases were slightly higher than those in the controls (0.320 and 0.162, respectively), while the frequency of the variant rs11599672G allele among the cases was slightly lower than that among the controls (0.277 vs. 0.297), suggesting rs2274223G, rs3203713G and rs11599672T alleles may be risk alleles to be considered in further analyses.

The LD analysis showed that two *PLCE1* polymorphisms rs2274223 and rs3203713 were in incomplete LD in our study population ($D'=0.99, r^2=0.40$), but these two SNPs were not in LD with rs11599672. To estimate possible joint effects and potential locus-locus interactions of *PLCE1* polymorphisms on risk of SCCHN, we then examined the combined effects of these three variants by the number of putative risk alleles (i.e. rs2274223G, rs3203713G and rs11599672T). As shown in **Table 2**, we trichotomized the subjects into three groups with “0-1”, “2-3” and “4-6” risk alleles. Compared with the group of “0-1” risk allele, there was a locus-dose effect as the risk allele number increased ($P_{\text{trend}}=0.046$). After adjustment for age, sex, smoking and alcohol status, the group of “2-3” risk alleles was borderline associated with risk of SCCHN (adjusted OR=1.20, 95% CI= 0.97-1.50, adjusted $P=0.096$) and the “4-6” risk allele group was significantly associated with risk of SCCHN (adjusted OR=1.31, 95% CI= 1.00-1.73, adjusted $P=0.049$). When “2-3” and “4-6” groups were combined for a larger number in the same stratum, the association between a larger number of risk alleles and risk of SCCHN remained statistically significant (adjusted OR=1.23, 95% CI=1.00-1.52, adjusted $P=0.048$).

Associations between *PLCE1* variants and risk of SCCHN by tumor sites

To investigate the modifying effects of *PLCE1* variants on risk of SCCHN with different tumor sites, we conducted the stratification analysis by oropharyngeal and non-oropharyngeal cancers. As shown in Table 2, rs2274223 variant genotypes were associated with a significantly increased risk of SCCHN only for non-oropharyngeal sites (AG vs. AA: adjusted OR=1.29, 95% CI=1.01-1.64, adjusted $P=0.042$; AG/GG vs. AA: adjusted OR=1.30, 95% CI=1.03-1.64; adjusted $P=0.025$), while rs11599672 variant genotypes were associated with a significantly decreased risk of SCCHN for this group of patients (GG vs. TT: adjusted OR=0.54, 95% CI=0.34-0.86, adjusted $P=0.009$; TG/GG vs. TT: adjusted OR=0.76, 95% CI=0.61-0.95, adjusted $P=0.018$). No association was observed for rs3203713 variant genotypes and risk of SCCHN in both subgroups. In addition, the locus-dose effect of combined risk alleles was also seen in SCCHN arising at non-oropharyngeal sites ($P_{\text{trend}}=0.017$). When the group of “0-1” risk allele was used as the reference, the group with “2-6” risk alleles had a significantly higher risk of SCCHN arising at non-oropharyngeal sites (adjusted OR=1.35, 95% CI=1.03-1.78; adjusted $P=0.033$). However, these associations were not found for oropharyngeal cancer.

Haplotype and stratification analyses

We further evaluated the combined effect of the three polymorphisms on risk of SCCHN arising at non-oropharyngeal sites by using the haplotype analysis (**Table 3**). A total of six haplotypes were derived from the observed genotypes, of which T_{rs11599672}A_{rs2274223}A_{rs3203713} was the most common haplotype in cases and controls with the frequencies of 53.6% and 53.0%, respectively. Compared with the common haplotype, the G_{rs11599672}A_{rs2274223}A_{rs3203713} haplotype was

associated with a 28% decreased risk of SCCHN arising at non- oropharyngeal sites (adjusted OR = 0.72, 95% CI = 0.56-0.92); in contrast, the T_{rs11599672}G_{rs2274223}A_{rs3203713} haplotype was associated with a 31% increased risk of SCCHN arising at non-oropharyngeal sites (adjusted OR = 1.31, 95% CI = 1.00-1.72). These associations were not observed for other haplotypes.

We then performed stratification analyses to evaluate the effects of variant genotypes on risk of SCCHN arising at non-oropharyngeal sites by age, sex, smoking status, alcohol use and disease stage (**Table 4**). The results showed that the risk effect of rs2274223 was more evident in the younger (adjusted OR = 1.44, 95% CI = 1.02-2.02), male (adjusted OR = 1.51, 95% CI = 1.14-2.00), smokers (adjusted OR = 1.33, 95% CI = 1.02-1.72), drinkers (adjusted OR = 1.69, 95% CI = 1.25-2.28) and subjects with early stage (adjusted OR = 1.40, 95% CI = 1.03-1.91), whereas the protective effect of rs11599672 was more prominent in the younger (adjusted OR = 0.71, 95% CI = 0.50-1.00), smokers (adjusted OR = 0.73, 95% CI = 0.56-0.94) and non-drinkers (adjusted OR = 0.67, 95% CI = 0.46-0.99). Additionally, the increased risk associated with the combined effect of risk alleles was also more pronounced among the younger (adjusted OR=1.68, 95% CI=1.09-2.59), male (adjusted OR=1.41, 95% CI=1.00-1.98), smokers (adjusted OR=1.65, 95% CI=1.21-2.26), drinkers (adjusted OR=1.42, 95% CI=1.00-2.03) and subjects with early-stage tumors (adjusted OR=1.46, 95% CI=1.01-2.13). However, no significant associations were found for rs3203713 variant genotypes and risk of SCCHN arising at non-oropharyngeal sites in every stratum (data not shown). Heterogeneity test showed that there was no significant heterogeneity ($P>0.05$) in every two stratum except for that of rs2274223 stratified by alcohol status ($P=0.005$). Interestingly, we detected a significant interaction between alcohol status (never and ever) and rs2274223 variant genotypes (AA and AG/GG) for risk of SCCHN arising at non-oropharyngeal sites, even after the Bonferroni correction ($P_{int}=0.004$).

Discussion

In this case-control study, we investigated the associations between three novel, potentially functional SNPs of *PLCE1* (rs2274223, rs3203713 and rs11599672) and risk of SCCHN in a non-Hispanic white population. Although we did not find evidence of a main effect of each *PLCE1* SNP on overall SCCHN risk, the joint effect of these variants appeared to contribute to risk of SCCHN in a dose-response manner, especially for cancers arising at non-oropharyngeal sites. Further, the subgroup analysis of SCCHN arising at non-oropharyngeal sites showed that variant genotypes of both rs2274223 and rs11599672 were independently associated with the risk. These findings suggested, for the first time, that potentially functional polymorphisms of *PLCE1* may play a role in the development of SCCHN, particularly of those tumors at non-oropharyngeal sites.

PLCE1 is known to be a direct downstream effector of small GTPases (Ras, Rap1 and Rap2), with the presence of two Ras-associating domains at its C terminus and a CDC25 guanine exchange factor (CDC25 gef) domain at its N terminus [11, 23-25]. Since mutations in the *RAS* gene family are associated with about 30% of all human cancers, several studies have investigated the possible role of *PLCE1* in cancer development and progression [13-16, 26]. It has been reported that *PLCE1* has an oncogenic role in carcinogenesis of several human cancers, including SCCHN, through distinct mechanisms, such as inflammation, binding to the Ras family small GTPase, and augmentation of angiogenesis [13-16, 26]. Recent evidence has demonstrated that *PLCE1* mutations might cause the nephritic syndrome [23] and diffuse mesangial sclerosis (DMS) [27]; however, few studies have investigated the association between genetic variants of *PLCE1* and risk of human cancers.

Most recently, for the first time, Wang *et al.* found that rs2274223, a nonsynonymous SNP of the *PLCE1* gene, was associated with an increased risk of ESCC and GCA in a Chinese population by a two-stage GWAS in 9,053 ESCC cases, 2,766 GCA cases and 11,013 controls [8]. Likewise, in another GWAS including 2,115 ESCC cases, 2,240 gastric cancer cases and 3,302 controls in another Chinese population, Abnet *et al.* also reported that rs2274223 was a notable signal for susceptibility to ESCC and GCA [10]. Despite the present study with 1,098 SCCHN cases and 1,090 controls may have a limited power to detect weak associations between polymorphisms of *PLCE1* including rs2274223 and overall risk of SCCHN compared with the published GWA studies [9, 10], our results did show that subjects carrying more risk alleles in *PLCE1* (rs2274223G, rs3203713G and rs11599672T) had a higher risk of SCCHN than those with zero to one risk allele, especially for SCCHN arising at non-oropharyngeal sites, suggesting a joint effect of these SNPs on susceptibility to SCCHN. Given only a modest effect of each SNP individually, evaluating their combined effects may help us better understand any role of *PLCE1* SNPs in cancer etiology. Furthermore, different carcinogenic mechanisms between esophageal and gastric cardia cancers and SCCHN or genetic difference in different populations may result in the discrepancy for the main effect of rs2274223 between our study on SCCHN and two GWASs of esophageal and gastric cancers.

Studies have shown that the oropharynx is the most common site for HPV-associated SCCHN [28, 29]; in contrast, SCCHN arising at non-oropharyngeal sites have much lower HPV seroprevalence [28, 30] and thus are more likely to be caused by smoking and alcohol drinking, similar to ESCC and GCA. In fact, 36% of our oropharyngeal cancer patients were never smokers compared to only 19% of patients with SCCHN arising at non-oropharyngeal sites. In our study, we found that rs2274223 and rs11599672 had an independent effect on risk of smoking and

drinking related SCCHN arising at non-oro-pharyngeal sites but not HPV-related oro-pharyngeal cancer. Even after Bonferroni corrections, the association of rs11599672 with non-oro-pharyngeal SCCHN and association of rs2274223 with drinking at non-oro-pharyngeal sites remained significant, suggesting different roles of these polymorphisms in the etiology of two different tumor subsites. Additionally, subgroup analyses restricted to oro-pharyngeal cancer did find a similar pattern of risk (though non-significant) associated with variant genotypes of rs2274223, rs11599672 and combined risk alleles among smokers but not among never smokers (data not shown). The findings from GWASs and our study implied that polymorphisms of *PLCE1* are likely to be associated with the development of human cancers related to tobacco and alcohol exposure, which will need the validation from large studies on different cancers.

Further, compared to the most common T_{rs11599672}A_{rs2274223}A_{rs3203713} haplotype, both T_{rs11599672}G_{rs2274223}A_{rs3203713} and G_{rs11599672}A_{rs2274223}A_{rs3203713} haplotypes had a significant association with risk of SCCHN arising at non-oro-pharyngeal sites, which further reflected the main effect of rs2274223 G and rs11599672 G alleles on risk of SCCHN arising at non-oro-pharyngeal sites. We also found that the effects of rs2274223 and rs11599672 on risk of SCCHN arising at non-oro-pharyngeal sites were segregated between the subgroups by smoking or drinking status, indicating the possibility of gene-environment interactions. Indeed, we detected a significant interaction between variant genotypes of rs2274223 and alcohol status with a relatively small sample size, supporting the role of a gene-environment interaction in the development of SCCHN.

It has been identified that rs2274223 is a non-synonymous SNP of *PLCE1*, and bioinformatics tools show that rs3203713 in *PLCE1* is located in the 3'UTR (<http://snpinfo.niehs.nih.gov/snfunc.htm>), which may affect the binding of miRNA and target

gene *PLCE1*. Further, rs11599672 is located in the transcription factor binding site (TFBS) of *PLCE1* (<http://snpinfo.niehs.nih.gov/snpfunc.htm>) and may result in the variation of transcription activity and expression of *PLCE1*. Considering the potentially functional significance of these SNPs, it is biologically plausible that *PLCE1* polymorphisms may contribute to the development of SCCHN. However, we cannot exclude the possibility that the findings from a subgroup analysis could be false positive because of a limited sample size. More rigorous studies with larger sample sizes, detailed HPV data and SNP functional relevance are warranted to replicate our findings and identify the underlying mechanism(s) of these SNPs in the etiology of SCCHN.

Some limitations of this study need to be addressed. Firstly, it is a hospital-based case-control study, and inherent selection bias cannot be completely excluded. However, the agreement of observed genotype distributions with Hardy-Weinberg equilibrium and similar allele frequencies of our controls with those reported in CEU populations from the dbSNP database (0.320 vs. 0.314 for rs2274223 G allele, 0.162 vs. 0.120 for rs3203713 G allele, and 0.297 vs. 0.347 for rs11599672, respectively) suggested that selection bias in terms of genotype distribution would not be substantial, if any. Secondly, it is uncertain whether the results from other populations such as Chinese populations in the reported GWASs [9, 10] are generalizable to our study population; yet similar findings from our non-Hispanic white subjects further support a similar biological plausibility of these SNPs. Thirdly, while we explored associations with smoking and drinking exposure, we were not able to stratify our data by HPV exposure. Finally, 3 SNPs were included in this study and we cannot rule out the possibility of false-positive associations because of multiple tests. Actually, only *P* values for the effect of rs11599672 variant genotypes on tumors arising at non-oropharyngeal sites and the association

of rs2274223 variant genotypes with drinking for non-oro-pharyngeal SCCHN risk remained significant after Bonferroni corrections.

Conclusions

In summary, in this hospital-based case-control study, we found that three *PLCE1* SNPs may have a joint effect on the risk of SCCHN, especially those arising at non-oro-pharyngeal sites, including oral cavity, hypopharynx or larynx, associated with smoking and alcohol exposure. Furthermore, those results from subgroup analysis suggest that variant genotypes of rs2274223 and rs11599672 may independently affect the risk of SCCHN arising at non-oro-pharyngeal sites. Additional larger studies in different populations are needed to validate our findings.

Abbreviations: OR, odds ratios; CI, confidence interval; *PLCE1*, phospholipase C epsilon 1 gene; GWAS, genome-wide association study; LD, linkage disequilibrium; MAF, minor allele frequency; PCR, polymerase chain reaction; SCCHN, squamous cell carcinoma of the head and neck; TFBS, transcription factor binding site

Competing interests:

The authors declare that they have no competing interests.

Authors' contributions

HXM, LEW and QYW designed the study and wrote the manuscript. HXM performed the experiments with the assistance from ZSL. EMS helped to revise the manuscript. All authors read and approved the final manuscript.

Acknowledgements and Funding

We thank Margaret Lung, Jessica Fiske, Shara Challa and Ana Neumann for their assistance in recruiting the subjects and gathering the questionnaire information, Yawei Qiao, Jianzhong He, Kejing Xu and Min Zhao for laboratory assistance, and Dakai Zhu for his technical support. This work was supported by National Institute of Health grants R01 CA131274 and R01 ES011740 (Q. Wei), P50 CA097007 (Scott Lippman), and P30 CA016672 (The University of Texas M. D. Anderson Cancer Center). Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the National Institutes of Health.

References

1. Marcu LG, Yeoh E: **A review of risk factors and genetic alterations in head and neck carcinogenesis and implications for current and future approaches to treatment.** *J Cancer Res Clin Oncol* 2009, **135**(10):1303-1314.
2. Jemal A, Siegel R, Xu J, Ward E: **Cancer statistics, 2010.** *CA Cancer J Clin* 2010, **60**(5):277-300.
3. Sturgis EM, Cinciripini PM: **Trends in head and neck cancer incidence in relation to smoking prevalence: an emerging epidemic of human papillomavirus-associated cancers?** *Cancer* 2007, **110**(7):1429-1435.
4. Neumann AS, Sturgis EM, Wei Q: **Nucleotide excision repair as a marker for susceptibility to tobacco-related cancers: a review of molecular epidemiological studies.** *Mol Carcinog* 2005, **42**(2):65-92.
5. Ho T, Wei Q, Sturgis EM: **Epidemiology of carcinogen metabolism genes and risk of squamous cell carcinoma of the head and neck.** *Head Neck* 2007, **29**(7):682-699.
6. Carlson CS, Eberle MA, Kruglyak L, Nickerson DA: **Mapping complex disease loci in whole-genome association studies.** *Nature* 2004, **429**(6990):446-452.
7. Stadler ZK, Thom P, Robson ME, Weitzel JN, Kauff ND, Hurley KE, Devlin V, Gold B, Klein RJ, Offit K: **Genome-Wide Association Studies of Cancer.** *J Clin Oncol* 2010, **28**(27):4255-4267.
8. McKay JD, Truong T, Gaborieau V, Chabrier A, Chuang SC, Byrnes G, Zaridze D, Shangina O, Szeszenia-Dabrowska N, Lissowska J, Rudnai P, Fabianova E, Bucur A, Bencko V, Holcatova I, Janout V, Foretova L, Laggiou P, Trichopoulos D, Benhamou S, Bouchardy C, Ahrens W, Merletti F, Richiardi L, Talamini R, Barzan L, Kjaerheim K,

- Macfarlane GJ, Macfarlane TV, Simonato L, *et al*: **A Genome-Wide Association Study of Upper Aerodigestive Tract Cancers Conducted within the INHANCE Consortium.** *PLoS Genet* 2011, **7**(3):e1001333.
9. Wang LD, Zhou FY, Li XM, Sun LD, Song X, Jin Y, Li JM, Kong GQ, Qi H, Cui J, Zhang LQ, Yang JZ, Li JL, Li XC, Ren JL, Liu ZC, Gao WJ, Yuan L, Wei W, Zhang YR, Wang WP, Sheyhidin I, Li F, Chen BP, Ren SW, Liu B, Li D, Ku JW, Fan ZM, Zhou SL, *et al*: **Genome-wide association study of esophageal squamous cell carcinoma in Chinese subjects identifies susceptibility loci at PLCE1 and C20orf54.** *Nat Genet* 2010, **42**(9):759-763.
10. Abnet CC, Freedman ND, Hu N, Wang Z, Yu K, Shu XO, Yuan JM, Zheng W, Dawsey SM, Dong LM, Lee MP, Ding T, Qiao YL, Gao YT, Koh WP, Xiang YB, Tang ZZ, Fan JH, Wang C, Wheeler W, Gail MH, Yeager M, Yuenger J, Hutchinson A, Jacobs KB, Giffen CA, Burdett L, Fraumeni JF Jr, Tucker MA, Chow WH *et al*: **A shared susceptibility locus in PLCE1 at 10q23 for gastric adenocarcinoma and esophageal squamous cell carcinoma.** *Nat Genet*, **42**(9):764-767.
11. Wing MR, Bourdon DM, Harden TK: **PLC-epsilon: a shared effector protein in Ras-, Rho-, and G alpha beta gamma-mediated signaling.** *Mol Interv* 2003, **3**(5):273-280.
12. Li M, Edamatsu H, Kitazawa R, Kitazawa S, Kataoka T: **Phospholipase Cepsilon promotes intestinal tumorigenesis of Apc(Min/+) mice through augmentation of inflammation and angiogenesis.** *Carcinogenesis* 2009, **30**(8):1424-1432.
13. Bai Y, Edamatsu H, Maeda S, Saito H, Suzuki N, Satoh T, Kataoka T: **Crucial role of phospholipase Cepsilon in chemical carcinogen-induced skin tumor development.** *Cancer Res* 2004, **64**(24):8808-8810.

14. Ou L, Guo Y, Luo C, Wu X, Zhao Y, Cai X: **RNA interference suppressing PLCE1 gene expression decreases invasive power of human bladder cancer T24 cell line.** *Cancer Genet Cytogenet*, **200**(2):110-119.
15. Wang X, Zbou C, Qiu G, Fan J, Tang H, Peng Z: **Screening of new tumor suppressor genes in sporadic colorectal cancer patients.** *Hepatogastroenterology* 2008, **55**(88):2039-2044.
16. Bourguignon LY, Gilad E, Brightman A, Diedrich F, Singleton P: **Hyaluronan-CD44 interaction with leukemia-associated RhoGEF and epidermal growth factor receptor promotes Rho/Ras co-activation, phospholipase C epsilon-Ca²⁺ signaling, and cytoskeleton modification in head and neck squamous cell carcinoma cells.** *J Biol Chem* 2006, **281**(20):14026-14040.
17. Mehanna H, Paleri V, West CM, Nutting C: **Head and neck cancer--Part 1: Epidemiology, presentation, and prevention.** *BMJ* 2010, **341**:c4684.
18. Morita M, Kumashiro R, Kubo N, Nakashima Y, Yoshida R, Yoshinaga K, Saeki H, Emi Y, Kakeji Y, Sakaguchi Y, Toh Y, Maehara Y: **Alcohol drinking, cigarette smoking, and the development of squamous cell carcinoma of the esophagus: epidemiology, clinical findings, and prevention.** *Int J Clin Oncol* 2010, **15**(2):126-134.
19. Steevens J, Schouten LJ, Goldbohm RA, van den Brandt PA: **Alcohol consumption, cigarette smoking and risk of subtypes of oesophageal and gastric cancer: a prospective cohort study.** *Gut* 2010, **59**(1):39-48.
20. Liu Z, Calderon JI, Zhang Z, Sturgis EM, Spitz MR, Wei Q: **Polymorphisms of vitamin D receptor gene protect against the risk of head and neck cancer.** *Pharmacogenet Genomics* 2005, **15**(3):159-165.

21. Huang YJ, Niu J, Liu Z, Wang LE, Sturgis EM, Wei Q: **The functional IGFBP7 promoter -418G>A polymorphism and risk of head and neck cancer.** *Mutat Res* 2010,**702**(1):32-39.
22. Xu Z, Taylor JA: **SNPinfo: integrating GWAS and candidate gene information into functional SNP selection for genetic association studies.** *Nucleic Acids Res* 2009,**37**(Web Server issue):W600-605.
23. Bunney TD, Baxendale RW, Katan M: **Regulatory links between PLC enzymes and Ras superfamily GTPases: signalling via PLCepsilon.** *Adv Enzyme Regul* 2009,**49**(1):54-58.
24. Kelley GG, Reks SE, Ondrako JM, Smrcka AV: **Phospholipase C(epsilon): a novel Ras effector.** *EMBO J* 2001, **20**(4):743-754.
25. Lopez I, Mak EC, Ding J, Hamm HE, Lomasney JW: **A novel bifunctional phospholipase c that is regulated by Galpha 12 and stimulates the Ras/mitogen-activated protein kinase pathway.** *J Biol Chem* 2001, **276**(4):2758-2765.
26. Sorli SC, Bunney TD, Sugden PH, Paterson HF, Katan M: **Signaling properties and expression in normal and tumor tissues of two phospholipase C epsilon splice variants.** *Oncogene* 2005, **24**(1):90-100.
27. Gbadegesin R, Hinkes BG, Hoskins BE, Vlangos CN, Heeringa SF, Liu J, Loirat C, Ozaltin F, Hashmi S, Ulmer F, Cleper R, Ettenger R, Antignac C, Wiggins RC, Zenker M, Hildebrandt F: **Mutations in PLCE1 are a major cause of isolated diffuse mesangial sclerosis (IDMS).** *Nephrol Dial Transplant* 2008, **23**(4):1291-1297.
28. Gillison ML, Koch WM, Capone RB, Spafford M, Westra WH, Wu L, Zahurak ML, Daniel RW, Viglione M, Symer DE, Shah KV, Sidransky D: **Evidence for a causal**

association between human papillomavirus and a subset of head and neck cancers. *J Natl Cancer Inst* 2000, **92**(9):709-720.

29. D'Souza G, Kreimer AR, Viscidi R, Pawlita M, Fakhry C, Koch WM, Westra WH, Gillison ML: **Case-control study of human papillomavirus and oropharyngeal cancer.** *N Engl J Med* 2007, **356**(19):1944-1956.
30. Kreimer AR, Clifford GM, Boyle P, Franceschi S: **Human papillomavirus types in head and neck squamous cell carcinomas worldwide: a systematic review.** *Cancer Epidemiol Biomarkers Prev* 2005, **14**(2):467-475.

Table 1. Distribution of selected variables in SCCHN cases and cancer-free controls

Variables	Cases No. (%)	Controls No. (%)	<i>P</i> ^a
All subjects	1,098 (100%)	1,090 (100%)	
Age, yr			0.676
≤57 (median)	586 (53.4)	572 (52.5)	
>57 (median)	512 (46.6)	518 (47.5)	
Sex			0.581
Females	271 (24.7)	258 (23.7)	
Males	827 (75.3)	832 (76.3)	
Smoking status			<0.001
Never	305 (27.8)	535 (49.1)	
Ever	793 (72.2)	555 (50.9)	
Alcohol status			<0.001
Never	298 (27.1)	473 (43.4)	
Ever	800 (72.9)	617 (56.6)	
Tumor site			
Oropharynx	559 (50.9)		
Non-oropharynx	539 (49.1)		
Stage			
I-II	272 (24.8)		
III-IV	826 (75.2)		

^a Two-sided χ^2 test.

Table 2. Logistic regression analysis for associations between *PLCE1* variant genotypes and SCCHN risk

Locus	Genotype	Controls (%) (N= 1,090)	Overall (N= 1,098)		Oropharynx (N= 559)		Non-oropharynx (N=539)	
			Cases (%)	OR (95%CI) ^a	Cases (%)	OR (95%CI) ^a	Cases (%)	OR (95%CI) ^a
<i>PLCE1</i> rs2274223								
	AA	504 (46.3)	477 (43.5)	1.00	253 (45.3)	1.00	224 (41.6)	1.00
	AG	474 (43.5)	506 (46.1)	1.14 (0.95-1.38)	248 (44.4)	1.09 (0.87-1.36)	258 (48.0)	1.29 (1.01-1.64)
	GG	111 (10.2)	114 (10.4)	1.20 (0.88-1.62)	58 (10.3)	1.15 (0.80-1.64)	56 (10.4)	1.38 (0.93-2.06)
	AG/GG	585 (53.7)	620 (56.5)	1.15 (0.97-1.38)	306 (54.7)	1.10 (0.89-1.36)	314 (58.4)	1.30 (1.03-1.64)
	G allele	0.320	0.335					
<i>PLCE1</i> rs3203713								
	AA	759 (70.0)	753 (68.8)	1.00	391 (70.1)	1.00	362 (67.4)	1.00
	AG	298 (27.5)	311 (28.4)	1.07 (0.88-1.31)	146 (26.2)	0.98 (0.77-1.25)	165 (30.7)	1.27 (0.99-1.64)
	GG	27 (2.5)	31 (2.8)	1.28 (0.74-2.21)	21 (3.7)	1.74 (0.96-3.16)	10 (1.9)	0.80 (0.36-1.79)
	AG/GG	325 (30.0)	342 (31.2)	1.09 (0.90-1.32)	167 (29.9)	1.04 (0.83-1.31)	175 (32.6)	1.23 (0.96-1.58)
	G allele	0.162	0.170					
<i>PLCE1</i> rs11599672								
	TT	519 (48.3)	554 (50.8)	1.00	270 (48.8)	1.00	284 (52.9)	1.00
	TG	473 (44.0)	469 (43.0)	0.94 (0.78-1.13)	246 (44.5)	1.03 (0.83-1.28)	223 (41.5)	0.81 (0.64-1.02)
	GG	82 (7.7)	67 (6.2)	0.72 (0.50-1.03)	37 (6.7)	0.89 (0.58-1.37)	30 (5.6)	0.54 (0.34-0.86)
	TG/GG	555 (51.7)	536 (49.2)	0.91 (0.76-1.08)	283 (51.2)	1.01 (0.82-1.25)	253 (47.1)	0.76 (0.61-0.95)
	G allele	0.297	0.277					
Combined effect of risk alleles ^b								
Trichotomy								
	0-1	263 (24.6)	235 (21.6)	1.00	124 (22.5)	1.00	111 (20.8)	1.00
	2-3	607 (56.9)	630 (58.0)	1.20 (0.97-1.50)	320 (58.0)	1.14 (0.88-1.48)	310 (58.1)	1.29 (0.97-1.73)
	4-6	197 (18.5)	221 (20.4)	1.31 (1.00-1.73)	108 (19.6)	1.23 (0.89-1.70)	113 (21.2)	1.54 (1.08-2.20)
				$P_{\text{trend}}=0.046$		$P_{\text{trend}}=0.210$		$P_{\text{trend}}=0.017$
Dichotomy								
	0-1	263 (24.6)	235 (21.6)	1.00	124 (22.5)	1.00	111 (20.8)	1.00
	2-6	804 (75.4)	851 (78.4)	1.23 (1.00-1.52)	428 (77.5)	1.16 (0.90-1.49)	423 (79.2)	1.35 (1.03-1.78)

^a Adjusted for age, sex, smoking and alcohol status. The SNP calling rates were all >98% with 2 samples failed in rs2274223, 9 samples in rs3203713 and 24 samples in rs11599672.

^b The risk alleles: rs2274223G, rs3203713G and rs11599672T.

Table 3. *PLCE1* haplotype and risk of SCCHN arising at non-orpharyngeal sites

<i>PLCE1</i> haplotypes ^a	Haplotype frequencies				Crude OR (95% CI)	Adjusted OR ^b (95% CI)	<i>P</i> ^b
	Cases (N =1,068)		Controls (N =2,134)				
	N	%	N	%			
TAA	572	53.6	1131	53.0	1.00	1.00	
GAA	128	12.0	318	14.9	0.80 (0.63-1.00)	0.72 (0.56-0.92)	0.008
TGA	119	11.1	198	9.3	1.19 (0.93-1.52)	1.31 (1.00-1.72)	0.049
GGG	87	8.1	179	8.4	0.96 (0.73-1.27)	0.96 (0.71-1.29)	0.771
TGG	96	9.0	170	8.0	1.12 (0.85-1.46)	1.10 (0.82-1.46)	0.528
GGA	66	6.2	138	6.5	0.95 (0.69-1.29)	0.85 (0.61-1.18)	0.334

^a The alleles of haplotypes were arrayed as the location of the SNPs in *PLCE1* stand from 5' to 3' (e.g. TAA denotes T_{rs11599672}A_{rs2274223}A_{rs3203713}).

^b Adjusted for age, gender, smoking and alcohol status in logistic models.

Figure Legends

Figure 1. PCR-based genotyping for rs11599672

Figure 2. Direct sequencing results for the *PLCE1* rs11599672 T>G (TT, TG and GG)

Table 4. Stratification analysis for associations between *PLCE1* variant genotypes and risk of SCCHN arising at non-oro-pharyngeal sites

Variables	rs2274223 (cases/controls)		Adjusted OR ^a (95%CI)		rs11599672 (cases/controls)		Adjusted OR ^a (95%CI)		Combined effect of risk alleles ^b (cases/controls)		Adjusted OR ^a (95%CI)	
	AA	AG/GG	TT	TG/GG	0-1	2-6	0-1	2-6				
Age, yr												
≤57(median)	96/286	133/285	133/284	97/279	0.71 (0.50-1.00)	39/140	189/417	1.68 (1.09-2.59)				
>57(median)	128/218	181/300	151/235	156/276	0.88 (0.64-1.20)	72/123	234/387	1.18 (0.82-1.70)				
Gender												
Females	84/119	100/139	96/110	89/145	0.73 (0.48-1.10)	44/74	140/177	1.30 (0.81-2.08)				
Males	140/385	214/446	188/409	164/410	0.84 (0.64-1.10)	67/189	283/627	1.41 (1.00-1.98)				
Smoking status												
Never	46/236	57/298	54/266	48/261	0.85 (0.55-1.31)	27/116	74/407	0.81 (0.50-1.33)				
Ever	178/268	257/287	230/253	205/294	0.73 (0.56-0.94)	84/147	349/397	1.65 (1.21-2.26)				
Alcohol status												
Never	79/210	76/262	86/223	69/244	0.67 (0.46-0.99)	34/112	120/352	1.33 (0.83-2.11)				
Ever	145/294	238/323	198/296	184/311	0.86 (0.64-1.15)	77/151	303/452	1.42 (1.00-2.03)				
Stage												
I-II	87/504	137/585	114/519	111/555	0.84(0.62-1.14)	44/263	178/804	1.46 (1.01-2.13)				
III-IV	137/504	177/585	170/519	142/555	0.78(0.59-1.03)	67/263	245/804	1.32 (0.94-1.85)				

^a Adjusted for age, sex, smoking and alcohol status (the stratified factor in each stratum excluded).

^b The risk alleles: rs2274223G, rs3203713G and rs11599672T.

TG

TG

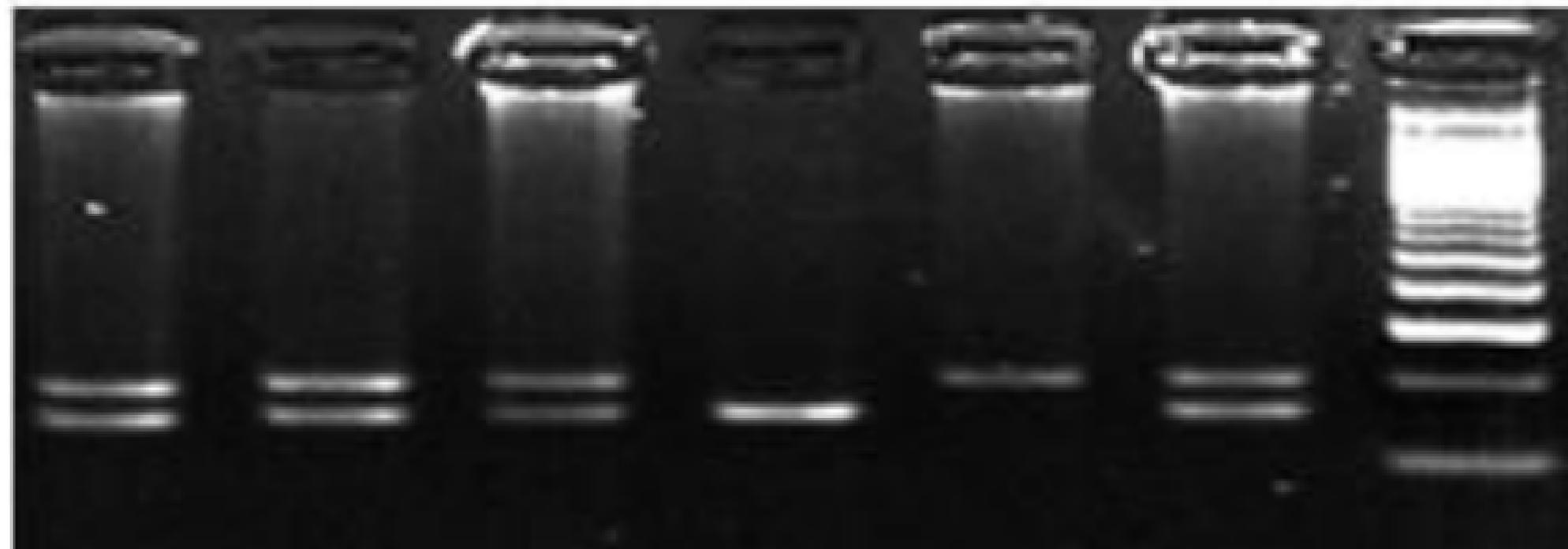
TG

TT

GG

TG

Marker

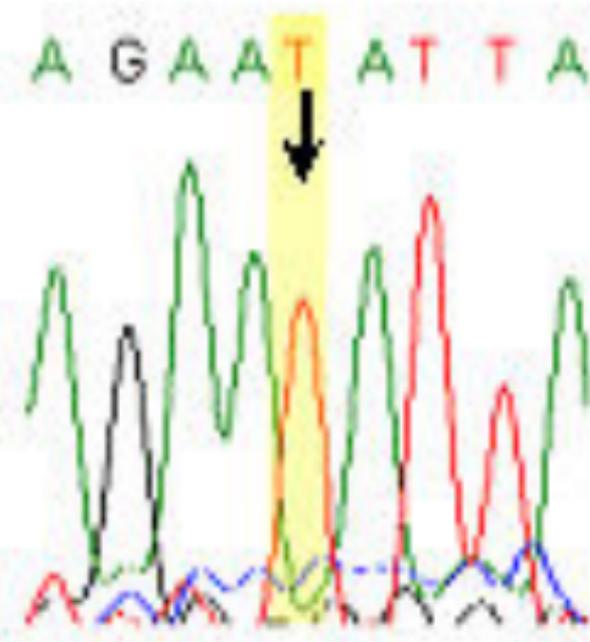


← 150bp

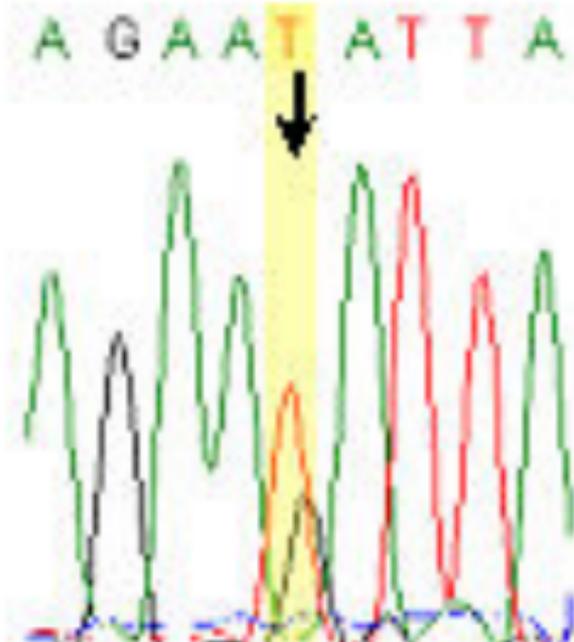
← 100bp

Figure 1

TT



TG



GG

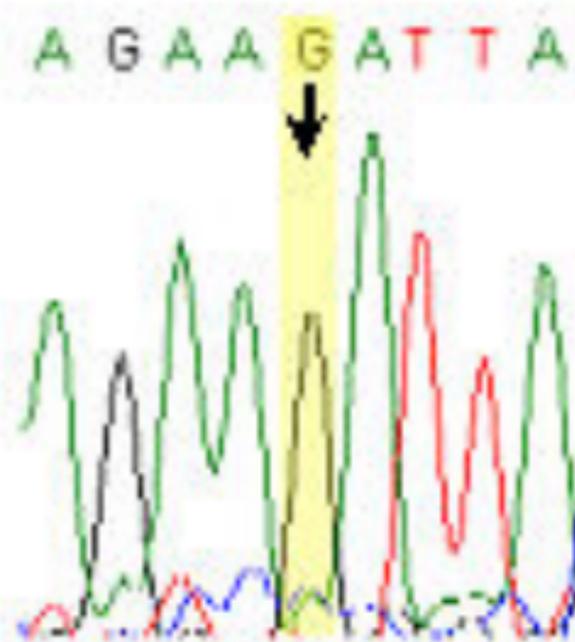


Figure 2