Gene 526 (2013) 223–227

Contents lists available at SciVerse ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

# Single nucleotide polymorphism in *hMLH1* promoter and risk of tobacco-related oral carcinoma in high-risk Asian Indians

Ritu Jha [a], Poonam Gaur [a], Suresh Chandra Sharma [b], Satya Narayan Das [a,*]

[a] Department of Biotechnology, All India Institute of Medical Sciences, Ansari Nagar, New Delhi 110029, India
[b] Department of Otorhinolaryngology, All India Institute of Medical Sciences, Ansari Nagar, New Delhi 110029, India

ARTICLE INFO

*Article history:*
Accepted 6 May 2013
Available online 31 May 2013

*Keywords:*
*hMLH1*
Oral cancer
SNP
Asian Indians

ABSTRACT

*hMLH1* is a member of mismatch repair genes (MMR) that plays a crucial role in correcting replication errors, cell cycle arrest, apoptosis and oxidative stress. We explored the risk associated with *hMLH1* −93 A>G (rs 1800734) single nucleotide polymorphism (SNP) with the oral squamous cell carcinoma (OSCC) in Asian Indians. We genotyped 242 patients with tobacco-related OSCC and 205 healthy controls by polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP) technique. The frequency of AA genotype was found to be significantly ($P_c < 0.0006$) lower in patients as compared to the controls (21.49% vs. 47.8%) while GG genotype showed significantly higher ($P_c < 0.0006$) prevalence in patients as compared to the healthy controls (41.32% vs. 13.66%). In logistic regression analysis AG (adjusted OR = 1.95, 95% CI = 0.72–5.26) and GG genotype (adjusted OR = 4.5, 95% CI = 1.54–13.16, $P = 0.006$) appeared susceptible when compared with the wild-type AA genotype. The allelic distribution showed that variant G allele is significantly higher ($P_c < 0.0004$) in patients and associated with increased risk (adjusted OR = 2.36, 95% CI = 1.33–4.19, $P = 0.003$) as compared to the wild-type A allele. Altogether, our results suggest that the *hMLH1* −93 A>G polymorphism is associated with the higher risk of tobacco-related OSCC in Asian Indians and could be useful in screening population at a higher risk.



## 1. Introduction

Oral cancer is the sixth most common cancer worldwide with a global annual incidence rate of over 640 000, of which 62% arise in developing countries including India and Southeast Asia (Ferlay et al., 2010). The primary etiological factors include usage of tobacco/areca nut/betel leaf, marijuana, human papilloma virus, Epstein–Barr virus infections, alcohol, poor diet, exposure to radiation and genetic predisposition (Sturgis et al., 2004).

Oral squamous cell carcinoma (OSCC) is preceded by benign lesions such as leukoplakia and submucous fibrosis which are specifically caused by N-nitrosamines (TSNAs)/benzo-pyrene like classical carcinogens found in tobacco (Hecht, 2003; Tang et al., 1992). These carcinogens produce pyridyl hydroxybutyl/benzo(a)pyrene diol epoxide (BPDE) adducts in DNA and induce frequent mutations, DNA damage and genomic-wide instability. These events subsequently lead to activation of oncogenes, inactivation of tumor suppressor genes and ultimately cancer development (Barnes, 2002). The pivotal set of genes in this regard is the DNA repair system that plays crucial role in maintaining genomic integrity during DNA replication, correcting post-replicative errors, random mutations, oxidative stress and the aging process (Dixon and Kopras, 2004). There are 5 major DNA repair pathways in humans namely nucleotide excision repair, base excision repair, mismatch repair, and homologous recombinational and non homologous end joining which together involve approximately 70 genes (Bernstein et al., 2002). Any genetic or epigenetic alteration in these genes may have a serious implication in one's DNA repair capacity (DRC) (Qiao et al., 2002), which in turn changes the susceptibility to cancer development (Goode et al., 2002).

hMLH1 is a key component of the mismatch repair system that plays crucial role in recognition of nucleotide mismatch and together with MSH2 recruits whole repair machinery to the error site (Ilyas et al., 1999). Besides, it is important for other cellular processes such as cell cycle arrest, oxidative stress and apoptosis (Jiricny, 2006). A number of studies have explored the association of its profound SNPs (Wehner et al., 1997) with the susceptibility of developing various human malignancies including lung (Lo et al., 2011; Park et al., 2004), breast (Lee et al., 2005), ovarian (Harley et al., 2008), endometrial (Beiner et al., 2006), colorectal (Raptis et al., 2007) and hereditary non-polyposis colorectal cancer syndrome (HNPCC) (Mitchell et al., 2002). Out of all, *hMLH1* −93 A>G has been studied extensively because of its prevalence and functional significance. *hMLH1* gene is composed of 19 exons spanning

*Abbreviations:* hMLH1, human MutL homolog 1; MMR, mismatch repair; SNP, single nucleotide polymorphism; OSCC, oral squamous cell carcinoma; PCR-RFLP, polymerase chain reaction-restriction fragment length polymorphism; BPDE, benzo(a)pyrene diol epoxide; DRC, DNA repair capacity; MSH2, MutS homolog 2; HNPCC, hereditary non-polyposis colorectal cancer syndrome; GT-IIB, GT-motif 2B; NF-IL6, interleukin 6-regulatory nuclear factor; UICC, Union for International Cancer Control.

* Corresponding author. Tel.: +91 11 26593548; fax: +91 11 26589286.
*E-mail addresses:* ritujha.aiims@gmail.com (R. Jha), gaur.poonam5@gmail.com (P. Gaur), suresh6sharma@yahoo.com (S.C. Sharma), satyandas@gmail.com, satyandas@hotmail.com (S.N. Das).

0378-1119/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.gene.2013.05.014

a region of 57360 bp on chromosome 3p21.3 (GenBank accession no. AC011816). Its core promoter is expanded from transcription start site to −184 nucleotides and constitutes consensus binding sites for two transcription factors *GT-IIB* (GT-motif 2B) and *NF-IL6* (interleukin 6-regulatory nuclear factor) (Kinoshita et al., 1992; Zenke et al., 1986). *hMLH1* −93 A>G is a single nucleotide polymorphism located in its promoter region and constitutes the putative consensus sequence for the binding of a crucial transcription factor *AP4* (nCnnCAGCTG from −102 to −93) that potentially regulates its promoter activity (Ito et al., 1999). Functional studies have revealed the potential role of this polymorphism in modulating promoter activity and transcription rate of the gene. It has been demonstrated that the binding efficiency of transcription factor and the promoter activity is increased with −93 G allele in colorectal and endometrial cancer cell lines (Perera et al., 2011). Thereby it is suggested to have a functional impact on DRC and thus may have an association with the cancer susceptibility.

Several studies have documented the association of *hMLH1* −93 A>G SNP with different cancers but its role in OSCC remains to be elucidated. Therefore the present study has been designed to assess the risk of *hMLH1* −93 A>G polymorphism in developing oral cancer. Indeed we found a strong risk associated with the individuals carrying the variant G allele as compared to the wild type A allele.

## 2. Materials and methods

### 2.1. Study population

This case control study included 242 patients with OSCC and 205 age, gender and ethnicity matched healthy volunteers. The cases were patients with histopathologically proven OSCC enrolled in Dr. B.R. Ambedkar Institute-Rotary Cancer Hospital of the All India Institute of Medical Sciences, New Delhi, India. Clinical staging of the tumor was performed as per UICC criteria (Sobin, 2002). Patients with prior history of any other form of malignancy, chronic disease and occupational radiation exposures were excluded, however, the selection was unbiased with age, gender and tumor stage. Controls were healthy volunteers and mostly the patient's attendants. The study protocol was approved by the ethics committee of AIIMS and a written informed consent was obtained from each study subject.

### 2.2. hMLH1 −93 A>G genotyping by PCR-RFLP

Ten milliliters venous blood was drawn aseptically by venipuncture. Genomic DNA was extracted from it using sodium perchlorate method as described earlier (Baniasadi et al., 2006) and stored at −20 °C until further use. The genotyping for *hMLH1* −93 A>G SNP was done using PCR-RFLP assay using *PvuII* endonuclease. Briefly, the genomic DNA was amplified with specific primers (F, 5′-ccgagctcctaaaaacgaac-3′) and (R, 5′-ctggccgctggataacttc-3′) to get a 387 bp amplicon containing the −93 *hMLH1* promoter site. The 20 μl PCR reaction contained 50 ng DNA, 25 pM of each primers, 0.2 mM of each dNTPs, 1× PCR buffer, 1.5 mM $MgCl_2$ and 1 U *Taq* Polymerase. The PCR cycling conditions consisted of an initial denaturation step of 95 °C for 10 min, followed by 39 cycles of 95 °C for 30 s, 54 °C for 45 s, 72 °C for 45 s and a final extension of 72 °C for 10 min. The PCR products were checked for specificity on 1% agarose gel. The 5 μl of specifically positive PCR products was digested overnight with 10 units of *PvuII* (New England BioLabs, Beverly, MA) at 37 °C and digested DNA fragments were resolved on 3% agarose gel. The homozygous AA genotype gave the undigested fragment of 387 bp while the homozygous GG genotype yielded 2 digested fragments of sizes 207 and 180 bp. Thus AG heterozygous condition revealed 3 bands of sizes 387, 207 and 180 bp. The randomly selected PCR amplified DNA samples for each genotype were cross checked by DNA sequencing and the results were found to be 100% concordant.

### 2.3. Statistical analysis

The frequencies of various genotypes and alleles between the patients and healthy controls were compared by chi-square ($\chi^2$) test using graph pad software (http://www.graphpad.com). Bonferroni's inequality method was applied to obtain corrected *P* values ($P_c$) for genotypes and alleles. $P_c < 0.05$ at 95% confidence interval (CI) was considered significant. Hardy–Weinberg equilibrium was tested by a goodness-of-fit chi-square test with one degree of freedom (http://www.oege.org/software/hwe-mr-calc.shtml) and unconditional logistic regression was used to estimate the risk of mutant genotype or allele with respect to wild type. The tobacco load adjusted and unadjusted odd ratios (OR) and 95% CI were determined using STATA/IC-11.2 software.

## 3. Results

### 3.1. Clinico-epidemiological features of study subjects

The clinico-epidemiological features of the study subjects have been mentioned in Table 1. The age of the patients ranged from 20 to 90 yrs with the mean (±SD) age of 49.9 (±12.7) yrs. Whereas the mean (±SD) age of control group was 45.2 ± 10.3 yrs and out of total 205, 170 (82.9%) were males and only 35 (17.1%) were females. In concordance with our previous studies, majority of the patients were males (84.3%) while only 15.7% were females. Tongue (48.8%) was found to be the most affected site of lesion followed by the buccal mucosa (24.4%), and retro molar trigone (11.1%) whereas the alveolus, gingivo-buccal sulcus, lips and maxilla were the less affected sites. When patients were classified according to the TNM staging, majority of them had larger tumor size ($T_3$ and $T_4$; 73.6%) and lymph node involvement ($N_+$; 55.0%). Thus only 17.3% of the

**Table 1**
Clinico-epidemiological features of study subjects.

| Characteristics | No. (%) | |
|---|---|---|
| | Patients | Controls |
| Age (yr) | | |
| Range | 20–90 | 20–55 |
| Mean ± SD | 49.9 ± 12.7 | 45.2 ± 10.3 |
| Gender | | |
| Male | 204 (84.3) | 170 (82.9) |
| Female | 38 (15.7) | 35 (17.1) |
| Site | | |
| Tongue | 118 (48.8) | |
| Buccal mucosa | 59 (24.4) | |
| Retro molar trigone | 27 (11.1) | |
| Alveolus | 13 (5.4) | |
| Gingivo-buccal sulcus | 13 (5.4) | |
| Lips | 11 (4.5) | |
| Maxilla | 1 (0.4) | |
| Tumor size | | |
| $T_1 + T_2$ | 64 (26.4) | |
| $T_3 + T_4$ | 178 (73.6) | |
| Lymph node involvement | | |
| $N_0$ | 109 (45.0) | |
| $N_+$ | 133 (55.0) | |
| Pathological classification | | |
| Squamous cell carcinoma | 242 (100.0) | |
| Histological classification | | |
| Well differentiated | 229 (94.6) | |
| Moderately differentiated | 8 (3.3) | |
| Poorly differentiated | 5 (2.1) | |
| Clinical stage | | |
| Early stage (I + II) | 42 (17.3) | |
| Late stage (III + IV) | 200 (82.7) | |
| Tobacco usage | | |
| Nil | 27 (11.2) | 205 (100.0) |
| <5 yrs | 65 (26.9) | 0 |
| ≥5 yrs | 150 (61.9) | 0 |