

RESEARCH ARTICLE

Evaluation of Risk Factors for Nasopharyngeal Carcinoma in a High-risk Area of India, the Northeastern Region

Deepak Singh Lourembam¹, Asem Robinson Singh¹, T. Dhaneshor Sharma², Th Sudheeranjan Singh³, Thiyam Ramsing Singh^{1*}, Lisam Shanjukumar Singh^{1*}

Abstract

Northeastern India is a major nasopharyngeal carcinoma (NPC) high risk-area although the rest of the country has very low incidence. A case-control study of 105 NPC cases and 115 controls was conducted to identify the potential risk factors for NPC development in this region. Information was collected by interviewer about socio-demographic characteristics, cigarette smoking, alcohol consumption, dietary history, occupational history, and a family history of cancer. Epstein-Barr viral load was assayed from the blood DNA by real time PCR. Associations between GSTs genotypes, cytochrome P450 family including CYP1A1, CYP2E1 and CYP2A6 polymorphisms and susceptibility to relationship between the diseases were studied using PCR-RFLP assay. Results indicate that Epstein-Barr virus load was significantly higher in patients compared to controls ($p < 0.0001$). Furthermore, concentration of blood EBV-DNA was significantly higher in advanced stage disease (Stage III and IV) than in early stage disease (Stage I and II) ($p < 0.05$). Presence of CYP2A6 variants that reduced the enzyme activity was significantly less frequent in cases than controls. Smoked meat consumption, exposure to smoke, living in poorly ventilated house and alcohol consumption were associated with NPC development among the population of Northeastern India. Thus, overall our study revealed that EBV viral load and genetic polymorphism of CYP2A6 along with living practices which include smoked meat consumption, exposure to smoke, living in poorly ventilated houses and alcohol consumption are the potential risk factors of NPC in north eastern region of India. Understanding of the risk factors and their role in the etiology of NPC are helpful for preventive measures and screening.

Keywords: Nasopharyngeal carcinoma - epstein-barr virus - viral load - genetic polymorphism - risk factor

Asian Pac J Cancer Prev, 16 (12), 4927-4935

Introduction

The distinct geographic variation in the global incidence of nasopharyngeal carcinoma (NPC) reflects a complex etiology involving viral, environmental, and genetic components. Although geographic regions have generally been classified as high- or low-incidence areas, the racial/ethnic distribution of NPC within regions is far from uniform (Chang and Adami, 2006). In the southeastern Chinese province of Guangdong, where the overall NPC incidence rate is >20 per 100,000 person-years among males, rates in Cantonese speakers are double those in other dialect groups such as the Hakka, Hokkien, and Chiu Chau (Parkin et al., 2002; Chang and Adami, 2006). Likewise, in the Malaysian state of Selangor, rates in Chinese residents have historically been highest among Cantonese, intermediate among Khek, and lowest among Hokkien and Teochiu (Parkin et al., 2002). In the United States, rates are highest among Chinese Americans, followed distantly by Filipino Americans then Japanese Americans, Blacks, Hispanics, and finally Whites (Burt

et al., 1992). In Southeast Asia, NPC risk seems to vary with degree of racial and social admixture with southern Chinese. Incidence is low among Singapore Indians who have had practically no intermingling with Chinese, but much higher in the Thai, Macaonese, and Malay indigenous populations, which have a history of intermarriage with Chinese ancestors (Ho, 1976). Similarly, rates in Ho Chi Minh City are roughly half those in Hanoi, where a higher proportion of the population is of ethnic Chinese descent (Nguyen et al., 1998). Close ties have existed between Japan and China for thousands of years, but mainly with northern China, and the incidence of NPC in Japan is low (Chang and Adami, 2006). In India, NPC is generally rare in most regions. For instance, in Mumbai (West India), the incidence rate is cited as 0.3 per 100,000 person-year (Bhatia and Singh, 1981). However, it is high in most part of North-eastern India namely Nagaland, Mizoram and Manipur, where the incidence rates are 19.4, 6 and 5 per 100,000 person/year respectively (Sharma et al., 2011). The North-eastern region of India is bordering West China. The populations of Northeastern region of India

¹Department of Biotechnology, Manipur University, Canchipur, ²Department of Radiotherapy, ³Department of Otorhinolaryngology, Regional Institute of Medical Sciences, Imphal, India *For correspondence: shanju.lisam@gmail.com, ramsing1thiyam@gmail.com

are believed to be migrated from East and Southeast Asia and they have the habit of eating smoked fish and meat (Kataki et al., 2011; Sharma et al., 2011).

In almost all populations surveyed, the incidence of NPC is 2 to 3 fold higher in males than in females. In most low-risk populations, NPC incidence increases monotonically with increasing age. In contrast, in high-risk groups, the incidence peaks around ages 50 to 59 years and declines thereafter, suggesting the involvement of exposure to carcinogenic agents early in life. Likewise, the minor incidence peak observed among adolescents and young adults in Southeast Asia, the Middle East/North Africa, and the United States is consistent with exposure to a common agent in early life. (Parkin et al., 2002; Chang and Adami, 2006)

The aetiology of NPC is complex, and includes a host of viral, genetic and environmental factors (Lakhanpal et al., 2014; Lung et al., 2014; Tsao et al., 2014; Chang and Adami, 2006). The collective evidence strongly indicates a causal role of Epstein-Barr virus (EBV) in the development of NPC; early-life infection, which is typical of high-incidence areas, may be critical (Tsang et al., 2014; Tsao et al., 2015). EBV DNA has been detected in the plasma, serum (Shotelersuk et al., 2000; Ji et al., 2014; Yip et al., 2014) or unfractionated whole blood (Stevens et al., 1999; Stevens et al., 2001b; Adham et al., 2013) of NPC patients and suggested to be a sensitive and specific molecular marker for both diagnosis and prognosis. Unfractionated whole blood is strongly preferred since it combines all blood compartments that may harbor EBV and it best reflects the absolute viral burden in the patient's circulation (Stevens et al., 2001a). EBV load correlated with NPC tumor stage, recurrence and survival (Lo et al., 2000; Lin et al., 2001; Lin et al., 2004) showed that EBV load correlates with response to treatment, likelihood of relapse and survival among patients with advanced NPC. However, EBV alone is not a sufficient cause of NPC because virtually all adults worldwide are infected with the virus, yet only a small proportion of individuals develop NPC. Therefore, it is apparent that environmental and/or genetic cofactors also contribute to NPC risk. It is believed that family history of NPC case, consumption of salted fish, smoked fish/meat, cigarette smoking, alcohol, use of traditional herbal medicines, expose to smoke from wood fires in chimneyless homes and occupational exposure to fumes, smokes, dusts, or chemicals are the environmental cofactors (Nor Hashim et al., 2012; Ghosh et al., 2014). However, involvements of these environmental factors in the cause of NPC are inconsistent and/or weakly associated in other studies taken placed in several NPC endemic and non-endemic arrears. Again, inconsistent findings may be due to differences in study characteristics, as well as chance or confounding factors (Chang and Adami, 2006). Several genetic polymorphisms and chromosomal abnormalities have been identified for NPC susceptibility loci (Chang and Adami, 2006). Cytochrome P450 (CYP) enzymes, phase I metabolic isozymes, play a key role in the metabolism of drugs and environmental chemicals. Several CYP enzymes metabolically activate procarcinogens to genotoxic intermediates (Go et al., 2015). Phenotyping analyses revealed an association

between CYP enzyme activity and the risk to develop several forms of cancer (Bozina et al., 2009). Research carried out in the last decade demonstrated that several CYP enzymes are polymorphic due to single nucleotide polymorphisms, gene duplications and deletions. Consistent evidences for association between CYP polymorphisms and lung, head and neck, and other types of cancer were reported (Gattas et al., 2006; Huang et al., 2013; Hussein et al., 2014; Yang et al., 2015) However, report for association of CYP polymorphism and NPC is very rare to our knowledge. Glutathione S-transferases (GSTs) are phase II metabolic isozymes which catalyse converting glutathione (GSH) to xenobiotic substrates for the purpose of detoxification. Absence of GSTM1 and/or GSTT1 was associated with increased risk of NPC (Jiang et al., 2011; Murthy et al., 2013; Wei et al., 2013).

The present case-control study is to evaluate the risk factors associated with NPC development among the NPC high-risk population of India. EBV load, genetic polymorphisms of CYP1A1, CYP2A6, CYP2A6, GSTM1, and GSTT1, and etiological environmental factors including cigarette smoking, tobacco consumption, alcohol consumption, smoked meat consumption, fermented food consumption, dust exposure, smoke exposure and poor ventilated house were compared between 105 NPC patients and 115 healthy control individuals who are age, sex and ethnic matched.

Materials and Methods

Blood specimens

The participants for this study were recruited from the Regional Institute of Medical Sciences (RIMS), Manipur. The study was approved both by the Institutional Human Ethics Committees (IHECs) at the RIMS and at the Manipur University, Manipur, India. A total of 105 patients with biopsy-proved NPC were enrolled between April 2011 and November 2013. The patient's cancer stage was defined according to the 1992 American Joint Committee on Cancer (AJCC TNM staging system). In addition, 115 healthy volunteers whose age, sex and ethnic matched with the 105 cases were recruited consecutively for the case-control study. Informed consent was obtained from each individual. In addition to intravenous 3ml blood sample collection from all individuals and endoscopic biopsy done from primary site biopsy in case group, all subjects were requested to respond to a personal interview that elicited detailed information on potential risk factors for nasopharyngeal carcinoma, including socio-demographic characteristics, cigarette smoking, alcohol consumption, dietary history, occupational history, and a family history of cancer.

Cell culture

Burkitt's lymphoma B-cell line (Raji) was obtained from the National Centre for Cell Sciences (NCCS), Pune. The cells was grown and maintained in a humidified incubator at 37°C and in 5% CO₂ atmosphere. RPMI-1640 medium was supplemented with 10% FBS (Fetal Bovine Serum), 100 units/ml penicillin, and 100 µg/ml streptomycin. All the culture reagents were obtained from

Invitrogen Gibco, USA.

DNA Extraction

Blood samples (3 ml) from 105 NPC patients (Case) and 115 healthy donors (Control) were collected in an EDTA coated tube and DNA were isolated from whole blood samples and the Burkitt's lymphoma B-cell (Raji) culture cells using QIAamp DNA Blood Mini Isolation Kit (Qiagen, USA) according to the manufacture's manual. Extracts were aliquoted in single use volumes to prevent freeze-thaw cycles and stored at -80°C prior to testing.

EBV gene targets, Albumin and PCR controls

Blood DNA sample and the DNA isolated from Raji cell line as positive control were subjected to amplification of the EBNA1 gene by conventional PCR as described previously (Tune et al., 1999) in parallel with the albumin gene as a 'housekeeping gene'. The primers for amplifying the EBNA1 were 5'-TGAATACCACCAAGAAGGTG-3' and 5'-AGTTCCTTCGTCGGTAGTC-3' and that of the albumin gene were 5'-GCCCTCTGCTAACAAAGTCCTAC-3' and 5'-GCCCTAAAAAGAAAATCGCCAATC-3' respectively.

Quantitative analysis of EBV DNA by Real Time PCR

All the DNA samples were accurately quantified on a Qubit 2.0 Fluorometer (Life Technologies, USA) using a Quant-iT dsDNA BR Assay Kit (Invitrogen, USA) in a Qubit Fluorometer (Invitrogen). Quantification of EBV DNA load of the blood sample was performed on a Step OnePlus Real-time PCR system (Applied Biosystems, USA). PCR primers targeting unique EBNA1 (Accession number V01555) for Real Time PCR were designed using Primer Express 3.0 software (ABI) and synthesized by Applied Biosystems, USA. The Real time PCR primers sequences for EBNA1 gene were Forward 5'-CGAGGAAGTGGCCCTTGCTAT-3' and Reverse 5'-CCAAAGGGGAGACGACTCAA-3' which give an amplicon size of 61 bp.

A calibration curve was generated using 10-fold serial dilutions using Raji cell DNA, varying from 250 to 250,000 copies, as a standard. Each DNA samples was then diluted to the same starting concentration of 5 ng/μl. PCR reactions were set up in 96-well plates (Applied Biosystems, USA) to a final volume of 25 μl comprising, 12.5 μl of 2X SybrGreen Master Mix (Invitrogen, USA), 200 nM each of forward and reverse primers and 15 ng of target DNA. Thermal cycling was initiated with a 2 min incubation at 50°C for the uracil N-glycosylase to act, followed by an initial denaturation step of 3 min at 95°C, and then 40 cycles of 95°C for 30 s and 61°C for 30 secs. A melt curve analysis immediately followed at between 60°C to 99°C as a check for amplicon purity.

Viral load calculation and result interpretation

Raji is a diploid cell line containing EBV viral genome 50 copies/cell (Nonoyama and Pagano, 1973). A conversion factor of 6.6 pg of DNA per diploid cell (Saiki et al., 1988) was calculated for the amount of EBV. The CT values from blood DNA were plotted on the standard curve, and the copy number was calculated using a

software package (StepOne Plus software v2.1, Applied Biosystems, USA) for data analysis. Each sample was performed in triplicates in three independent experiments, and quantitative results were averaged. EBV DNA load, expressed as viral copy number per mL of blood, was determined using the equation below as described previously (Lo et al., 1999):-

$$C = Q \times V_{\text{DNA}} / V_{\text{PCR}} \times 1 / V_{\text{EXT}}$$

Where,

C=EBV target concentration in blood (copies/ml)

Q=Target quantity (copies) determined by a sequence detector in PCR.

V_{DNA} =Total volume of DNA obtained after extraction (typically 50 μl of DNA extracted)

V_{PCR} =Volume of DNA used in PCR.

V_{EXT} =Volume of Blood extracted.

Genotyping

The blood DNA samples of NPC patients (Case) and healthy donors (Control) were analyzed for genetic polymorphisms of GSTM1, GSTT1 by PCR whereas the polymorphism analysis of CYP1A1, CYP2E1 and CYP2A6 was done by PCR-RFLP (Restriction Fragment Length Polymorphism) as described previously (Hayashi et al., 1991; Arand et al., 1996; Ariyoshi et al., 2000; Jiang et al., 2014). All the restriction enzymes were obtained from New England Biolabs (USA). The resulted PCR amplicon or the restriction enzyme fragments were analyzed on 0.8-2% agarose gel electrophoresis and bands were visualized with an ultraviolet transilluminator after ethidium bromide staining.

CYP2A6 genotyping

CYP2A6 genotypes, including homozygous for the wild-type (*1A/*1A), conversion-type (*1B/*1B), deletion-type (*4C/*4C) and other three heterozygous genotypes (*1A/*1B, *1A/*4C and *1B/*4C) were determined by PCR-RFLP assay as previously described (Ariyoshi et al., 2000). The PCR products of 1322 bp length were either digested overnight at 37°C with the enzymes BstUI or Bsu361. The 1322 bp PCR products digested with BstUI generated 1322 or 1004 bp fragment; 1322 bp fragment for *1A allele and 1004 bp fragment for *1B or *4C allele. PCR products digested with Bsu361 generated 800 or 750 bp fragment; 800 bp fragment for *1A or *1B allele and 750 bp fragment for *4C allele.

Statistical analysis

Differences in the distributions of demographic characteristics, selected variables, and frequencies of genotypes of the GSTM1, GSTT1, CYP2E1, CYP1A1 and CYP2A6 polymorphism between the cases and controls were evaluated by using the chi-square test and odds ratios (ORs) with 95% confidence interval (95%CI).

EBV load levels among patients with advanced stage and early-stage NPC were compared using the Mann-Whitney rank-sum test. The concentrations of blood EBV DNA between case and controls were compared with the Mann-Whitney rank sum test for binary categories or the Kruskal-Wallis test for more than two categories. All

statistical tests were two-sided, and a p-value of less than 0.05 was considered to indicate statistical significance. Analyses were performed with the use of GraphPad Prism (version 6), USA.

Results

Study characteristics

Table 1 presents data on the characteristics of the study subjects. The mean age of 105 cases at NPC diagnosis was 48.9 (range, 18-80) years and that of 115 controls at enrollment was 44 (range, 22-73) years ($p>0.05$). Among cases the number of male (68.5%) is more than female (31.4%) by 2.2 times. The results indicate that NPC incidence is higher among male as compared to female and the incidence increases 3 folds after approximately 50 years of age (median age 47.47). Tumor histological types were available for all the 105 case samples. Using the World Health Organization (WHO) classification for NPC (1991) criteria, 5.71% ($n=6$) of NPC patients had Type I (keratinizing squamous cell carcinoma), 15.2% ($n=16$) NPC cases had Type II (non-keratinizing carcinoma), and 79.0% ($n=83$) NPC cases had Type III (undifferentiated carcinoma). Stage wise, 10.2% ($n=10$) of cases was in stage I, 24.4% ($n=24$) in stage II, 42.8% ($n=42$) in stage III and 22.4% ($n=22$) in stage IV indicating 35% ($n=34$) of the cases was in the early stage while 65% ($n=64$) was in the advanced stage.

Association with NPC Risk

Analyzing of some previously reported epidemiological factors associated with NPC incidence, the following factors, in high order, are significantly higher in the cases than controls group; (1) smoked meat consumption ($p<0.00001$), (2) smoke exposure ($p=0.0007$), (3) living in poor ventilated house ($p=0.003$), (4) alcohol intake ($p<0.05$), as shown in Table 1. There are no statistically significant differences between cases and controls regarding association other factors like cigarette smoking, tobacco consumption, dust exposure and fermented food consumption distribution.

Table 2 summarizes data for the overall distribution of cases and controls by GSTM1, GSTT1, CYP2E1 Rsa1, CYP1A1 Msp1 and CYP2A6 gene polymorphism. Among controls, the frequencies of the GSTM1 and GSTT1 null genotypes are 62.6% and 31.3% respectively. Similarly, the heterozygous variant allele frequencies of the CYP2E1 Rsa1, CYP1A1 Msp1 and CYP2A6 *1A/*1B among the controls are 16.9%, 41.7% and 42.6% respectively. None of the GST genotypes were significantly associated with NPC ($p>0.05$). Further we also found that there was also no significant association between the polymorphisms of CYP2E1 ($p=0.88$), CYP1A1 ($p=0.49$) and NPC susceptibility.

PCR-RFLP analysis was used to evaluate CYP2A6 polymorphism (Figure 1). As summarized in Table 2, there is significant differences in the polymorphism variants of CYP2A6 ($p<0.05$) between cases and control, but not in other genes investigated. In the healthy control group, percentage of wild type allele, *1A/*1A is 33, but it is increased to 51.4 in the NPC patient group, whereas

Table 1. Characteristics of Study Subjects and Etiological Environmental Factors

Variables	Case	Controls	p-value
Sex			
a) Male	72 (68.5)	71 (61.7)	0.28
b) Female	33 (31.4)	44 (38.2)	
Age			
≤ 30	15 (14.2)	24 (20.8)	0.58
31 - 40	21 (20)	22 (19.1)	
41 - 50	25 (23.8)	22 (19.1)	
> 50	44 (41.9)	47 (40.8)	
Age			
Mean	48.9	44	0.54
S.D.	15	10.1	
Range	18-80	22-73	
WHO pathological classification			
Type I	6 (5.7)		
Type II	16 (15.2)		
Type III	83 (79.0)		
Tumour Stage			
T1 or T2	34 (34.2)		
T3 or T4	64 (65.7)		
Nodal Stage			
N0 or N1	32 (30.4)		
N2	53 (50.4)		
N3	20 (19.0)		
Etiological Environmental Factors			
Cigarette Smoking			
Never	49 (46.6)	61 (53.0)	0.34
Ever	56	54	
a) 1- 10 yrs	23 (21.9)	19 (16.5)	
b) > 10 yrs	33 (31.4)	35 (30.4)	
Alcohol Intake			
Never	46 (43.8)	70 (60.8)	0.01
Ever	59	45	
a) 1-10 yrs	16 (15.2)	17 (14.7)	
b) > 10 yrs	43 (40.9)	28 (24.3)	
Tobacco Consumption			
Never	68 (64.7)	77 (66.9)	0.73
Ever	37	38	
a) 1-10 yrs	16 (15.2)	17 (14.7)	
b) > 10 yrs	21 (20)	21 (18.2)	
Smokes Meat Consumption			
Never	22 (20.9)	78 (67.8)	$p<0.00001$
Ever	83	37	
a) 1-10 yrs	34 (32.3)	11 (9.5)	
b) > 10 yrs	49 (46.6)	26 (22.6)	
Dust exposure (cultivated or roadside)			
Never	46 (43.8)	60 (52.1)	0.21
Ever	59	55	
a) 1-10 yrs	27 (25.7)	30 (26.0)	
b) > 10 yrs	32 (30.4)	25 (21.7)	
Smoke Exposure			
Never	27 (25.7)	55 (47.8)	0.0007
Ever	78	60	
a) 1-10 yrs	46 (43.8)	40 (34.7)	
b) > 10 yrs	32 (31.2)	20 (17.3)	
Poor ventilated house			
Never	43 (40.9)	70 (60.8)	0.0032
Ever	62	45	
a) 1 - 10 yrs	18 (17.1)	19 (16.5)	
b) > 10 yrs	32 (30.4)	20 (17.3)	
Fermented Food Consumption			
Never	39 (37.1)	47 (40.8)	0.607
Ever	66	69	
a) 1-10 yrs	25 (23.8)	22 (19.1)	
b) > 10 yrs	41 (39.0)	47 (40.8)	

*p-values are calculated by Chi square test of significance

Table 2. Distribution of CYP2A6, CYP2E1, CYP1A1, GSTM1 and GSTT1 Gene Polymorphism between NPC Patients and Controls

	Case (%)	Controls (%)	OR s (95% CI)	p-value
CYP2A6				
*1A/*1A	54 (51.4)	38 (33.0)	1	$\chi^2 = 11.58, p=0.003$
*1A/*1B	41 (39.0)	49 (42.6)	0.58 (0.31 - 1.10)	
*1B/*1B	10 (9.52)	28 (24.3)	0.25 (0.10 - 0.62)	
GSTM1				
Present	39 (37.1)	43 (37.3)	1	0.97
Null	66 (62.8)	72 (62.6)	1.01 (0.56 - 1.81)	
GSTT1				
Present	72 (68.5)	79 (68.6)	1	0.98
Null	33 (31.4)	36 (31.3)	1.01 (0.55 - 1.85)	
CYP2E1 RsaI				
w/w	86 (81.9)	95 (82.6)	1	0.88
w/h	19 (18.0)	20 (16.9)	1.05 (0.49 - 2.23)	
h/h	0 (0.00)	0 (0.00)		
CYP1A1 MspI				
w/w	27 (25.7)	28 (24.3)	1	0.49
w/h	50 (47.6)	48 (41.7)	1.08 (0.45 - 1.89)	
h/h	28 (26.6)	39 (33.9)	0.74 (0.62 - 2.93)	

*Significance set at $P < 0.05$; Abbreviations: 95% CI: 95 % confidence interval; OR: odds ratio; Pearson's chi-square test was used to evaluate the frequency distribution of genes among case and controls; significant at $p < 0.05$; *1A=wild type, *1B=conversion type; (w=wild type, h= heterozygosity)

Table 3. Analysis for Genotype Combinations (GSTM1, GSTT1 and CYP1A1) between Case and Controls

Genotypes	Case (%)	Control (%)	P-value
GSTM1 ^{wild} GSTT1 ^{wild}	21 (20)	28 (24.3)	0.62
GSTM1 ^{null} GSTT1 ^{wild}	50 (47.6)	50 (44.3)	
GSTM1 ^{wild} GSTT1 ^{null}	18 (17.1)	15 (13.0)	
GSTM1 ^{null} GSTT1 ^{null}	16 (15.2)	22 (19.1)	
GSTM1 ^{wild} CYP1A1 ^{wild}	15 (14.2)	12 (10.4)	0.66
GSTM1 ^{null} CYP1A1 ^{wild}	12 (11.4)	17 (14.7)	
GSTM1 ^{wild} CYP1A1 ^{variant}	26 (24.7)	33 (28.6)	
GSTM1 ^{null} CYP1A1 ^{variant}	52 (49.5)	53 (46.0)	
GSTT1 ^{wild} CYP1A1 ^{wild}	19 (18.0)	23 (20)	0.48
GSTT1 ^{null} CYP1A1 ^{wild}	9 (8.57)	6 (5.21)	
GSTT1 ^{wild} CYP1A1 ^{variant}	56 (53.3)	55 (47.8)	
GSTT1 ^{null} CYP1A1 ^{variant}	21 (20)	31 (26.9)	
GSTM1 ^{wild} GSTT1 ^{wild} CYP1A1 ^{wild}	8 (7.61)	9 (7.82)	0.69
GSTM1 ^{or GSTT1} either null CYP1A1 ^{wild}	17 (16.1)	16 (13.9)	
GSTM1 ^{null} GSTT1 ^{null} CYP1A1 ^{wild}	2 (1.90)	3 (2.6)	
GSTM1 ^{wild} GSTT1 ^{wild} CYP1A1 ^{variant}	14 (13.3)	19 (16.5)	
GSTM1 ^{or GSTT1} either null CYP1A1 ^{variant}	54 (51.4)	50 (43.4)	
GSTM1 ^{null} GSTT1 ^{null} CYP1A1 ^{variant}	10 (9.52)	18 (15.6)	

* $p < 0.05$ is considered statistically significant

percentages of heterozygous *1A/*1B and mutant homozygous *1B/*1B decrease from 42.6 to 39 and 24.3 to 9.52 respectively.

Interestingly, the frequency of cases with homozygotes of the CYP2A6*1A/*1A in the NPC patients (51.4%) was higher than that in the controls (33%). In contrast, the frequency of cases (9.52%) with homozygotes CYP2A6 *1B/*1B was lower than that in the controls (24.3%). The χ^2 test indicated a significant difference in the distribution of the CYP2A6 genotypes between case and controls ($\chi^2=11.58, p=0.003$), suggesting that there was a relationship between genetic polymorphism of CYP2A6 and NPC. The odds ratio (OR) for the cases versus controls harbouring CYP2A6 *1B/*1B was 0.25 (95% CI=0.10 - 0.62). This data implies that the NPC risk of having the CYP2A6 *1B/*1B genotype could be four

times lower than that of the CYP2A6 *1A/*1A wild type genotype and therefore, suggests that deficient CYP2A6 activity due to genetic polymorphism reduces NPC risk. CYP2A6 *4C allele was not found either in healthy control and NPC patient groups and therefore are not included in the analysis. Overall, a significant association between CYP2A6 polymorphism and NPC development was observed ($P < 0.05$). Therefore, our results indicate that genetic polymorphism of CYP2A6 is associated with NPC but not in GSTM1, GSTT1, CYP2E1 and CYP1A1 genes.

We have also analysed whether any combined genetic polymorphisms between GSTM1/GSTT1, GSTM1/CYP1A1, GSTT1/CYP1A1 and GSTM1/GSTT1/CYP1A1 is associated with NPC. However, our results suggest that either combined genetic polymorphisms are also not associated with this disease (Table 3).

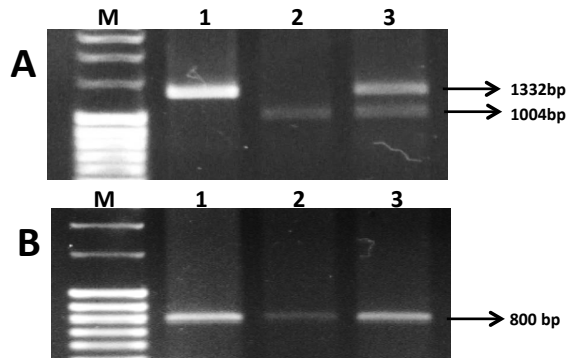


Figure 1. Representative Agarose gel Electrophoresis for PCR-RFLP analysis of CYP2A6 Polymorphism in NPC Patients. PCR products of 1322 bp in length digested with BstUI enzyme (A) and Bsu361 (B) generated patterns of DNA fragments; 1332 bp with BstUI and 800 bp with Bsu361 (*1A/*1A genotype, Lane 1); 1004 bp with BstUI and 800 bp with Bsu361 (*1B/*1B genotype, Lane 2); both 1332 bp and 1004 bp with BstUI and 800 bp with Bsu361 (*1A/*1B, Lane 3). M denotes 1kb marker

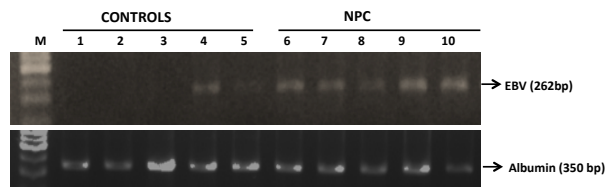


Figure 2. Detection of Epstein-Barr Virus (EBV) DNA by Polymerase Chain Reaction (PCR). DNA samples of case and controls were checked for the presence of EBV by conventional PCR. EBNA1 gene was amplified from whole blood samples of case (representative; lane 6-10) and controls (representative; lane 1-5). The amplicon was analysed on 0.8% Agarose. Among cases EBV amplicon was obtained in all samples while among control, EBV amplicon was obtained in few samples (lane 4 and 5). Albumin serves as internal control

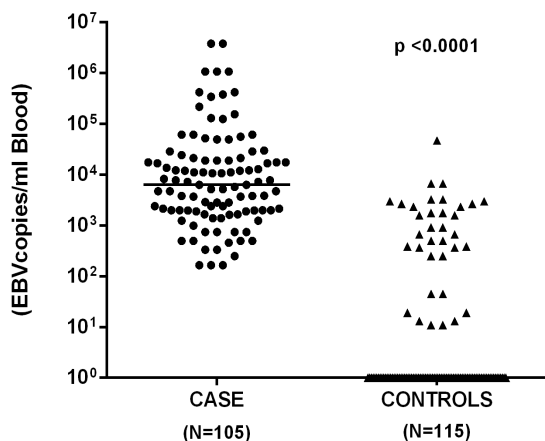


Figure 3. Scatter Plot of EBV DNA Load (copies/ml) in Blood Samples of NPC and Healthy Controls. EBV load of whole blood was measured by EBNA1 real-time PCR in NPC patients (n=105) and healthy controls (n=115). EBV DNA copy numbers are expressed on the logarithmic scale on Y-axis. The categories (Case and controls) are plotted on the X-axis. The P-values were calculated comparing the EBV load between patients and control groups by means of Mann-Whitney test. The horizontal lines represent median values. The difference was statistically significant ($P < 0.0001$)

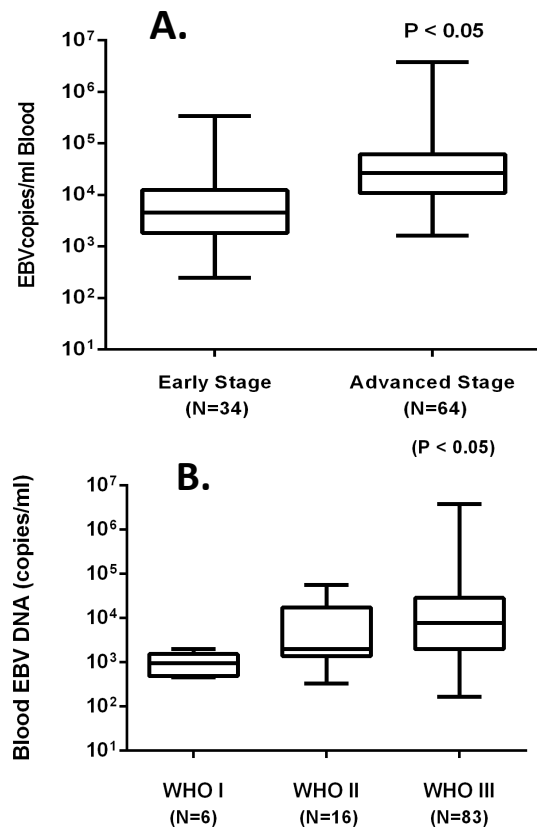


Figure 4. EBV DNA load Directly Correlates with the Stages of disease and WHO Classification in NPC Patients. A). Correlation of Blood EBV loads according to disease staging in NPC patients. Peripheral blood samples were obtained from NPC patients with early stage disease (n=34) and advanced stages disease (n=64). EBV DNA copy numbers according to disease stage were expressed on the logarithmic scale. Bars show the median EBV loads and standard deviations for each group. The difference in the clinical stage were significant ($p < 0.05$) calculated by Mann Whitney Rank Sum Test. B). EBV load according to WHO pathological classification types. The median load and standard deviation of different groups of classification (WHO I, II and III) are given and the difference in the groups were significant ($p < 0.05$) as calculated by Kruskal Wallis Test.

EBV load measurement among NPC cases and healthy community controls

EBV loads were measured from unfractionated whole blood from the blood of cases and controls by real-time PCR because previous studies have demonstrated that EBV load measurement from unfractionated whole blood (UWB) is strongly preferred to serum because it combines all blood compartments that may harbor EBV and best reflects the absolute viral burden in the patient's circulation (Stevens et al., 2001a). Prior to EBV load measurement by RTQ-PCR, each DNA sample was subjected to detection of the EBNA1 gene by conventional PCR in parallel with the albumin gene as a 'house-keeping' gene. EBNA1 gene was detected in 105 out of 105 (100%) cases whereas in 24 out of 115 controls (21%) by conventional PCR. Albumin gene was detected in all cases and controls. Figure 2 shows a few representative results.

The EBV loads of cases and controls as well as distribution of clinical stages in cases is shown in Figure

3 and 4. The median values of the EBV load of cases and controls were 6308 and 0.00 copies/ml respectively. EBV load of 81 controls out of 115 were very low (less than 100 copies/ml). There is a very high significant difference of EBV viral load between cases and controls ($p < 0.0001$). The median viral loads in early-stage (stages I and II) and advanced-stage (stages III and IV) patients are 4523.5 and 26518 copies/ml respectively. The median viral load distribution of cases analyzed based on the WHO histopathological type, WHO I, WHO II and WHO III were 956.5, 2004 and 7819 copies/ml respectively with their significant p -value is ($p < 0.05$) as shown in Figure 4B. As expected, patients with stage III or IV disease have significantly higher viral load compared to patients with stage I or II disease ($p < 0.05$), Mann-Whitney rank-sum test).

Discussion

The aetiology of NPC is majorly attributed to three risk factors namely infection with Epstein-Barr virus (EBV), genetic predisposition, and environmental pollutants. (Mathur, 2003) Therefore, in this study, we sought to investigate the association of three main risk factors with NPC incidence among population of Nagaland and Manipur, states of northeastern India where NPC incidence is high.

Our data suggest that whole blood EBV load measurement may be used as diagnostic and disease staging biomarker for NPC. There is a very high significantly difference of whole blood EBV load among the patients of NPC (median value; 6308 copies/ml) and control community (median value; 0.00 copies/ml) was observed which indicates that NPC incidence is highly associated with EBV infection. There was also correlation between EBV load and disease stages as well as WHO classification. The patients in advanced-stage disease has significantly higher (5.8 folds, $p < 0.05$) viral load than patients in early-stage disease. Moreover, the median viral load of patients who are in the WHO III classification is significantly higher than patients in the WHO II and similar pattern is found when compared the patients in WHO I and WHO II.

A number of etiological environmental factors like cigarette smoking, tobacco consumption, alcohol consumption, smoked meat consumption, fermented food consumption, dust exposure, firewood and petrochemical related smoke exposure and living in poor ventilated house were also investigated whether these factors which have been considered as possible risk factors for NPC development. Our study suggests that smoked meat consumption, dwelling in poor ventilated house, alcohol consumption and exposure to smoke may be the etiological environmental factors for NPC development among the population of Manipur. Out of these factors, smoked meat consumption and exposure to smokes are the high significant factors, although dwelling in poor ventilated house and alcohol consumption are significant but lesser than the above two factors. When our manuscript was under preparation, similar findings were reported from other group (Lakhanpal et al., 2014) from these region

underlying the importance of possible risk factors and confirming the role of EBV in the aetiology of NPC in Northeastern India.

Polymorphisms in glutathione-S-transferase genes (GSTM1 and GSTT1) and cytochrome p450 genes (CYP2E1, CYP1A1 and CYP2A6) are candidate cancer susceptibility genes (Gattas et al., 2006; Huang et al., 2013; Islam et al., 2013; Murthy et al., 2013; Russo et al., 2013; Jiang et al., 2014; Yang et al., 2015). Their association with NPC has been reported in some ethnic groups (Tiwawech et al., 2006; Murthy et al., 2013; Russo et al., 2013; Wei et al., 2013; Yang et al., 2015) [. However, non-association has also been reported in many studies (Cheng et al., 2003; Rossini et al., 2007; Boccia et al., 2008; Moraes et al., 2012). The levels of expression and catalytic activities of and GSTM1 and GSTT1 enzymes in NPC, and their metabolic balance, may be an important determinant host factor underlying NPC. Our study shows relationships between cytochrome P450 2A6 (CYP2A6) gene polymorphism and NPC development. The significant association between the CYP2A6 genetic polymorphism and NPC risk was found ($\chi^2 = 11.58$, $p = 0.003$). However, association of NPC development with polymorphism of GSTM1, GSTT1 and CYP1A1 or combination of these three genes is not found. It is thus suggested that CYP2A6 polymorphism may play a crucial role in NPC susceptibility and it may be used as a risk marker for NPC among risk group of Nagaland and Manipur. Although association of CYP2A6 polymorphism has been studied well in other cancer like lung cancer, oral cancer and breast cancer, the association of CYP2A6 polymorphism with NPC has not been well studied except one study on Thailand population (Tiwawech et al., 2006). CYP enzyme family metabolized a wide variety of xenobiotic compounds including carcinogens which are metabolically activated to generate ultimate carcinogens. CYP2A6 is known to metabolically activate nitrosamines (Guengerich and Shimada, 1991) and CYP2A6 deletion or variants reduced risk of lung, oral and bladder cancer in Asian population studies (Topcu et al., 2002; Song et al., 2009; Tamaki et al., 2011; Liu et al., 2013a; Liu et al., 2013b).

Taking together the results, the present study lead to two important findings; firstly EBV load may be useful as biomarker for NPC detection and prognosis for the high risk-group in India. Secondly, the etiological environmental factors like, consumption of smoked meat and alcohol, exposure to firewood/petrochemicals related smokes and living in poorly ventilated house are main EBV activator in the individual who has CYP2A6 *1B/*1B genetic polymorphism among the high risk-area of NPC, Nagaland and Manipur.

Acknowledgements

We are grateful to Ksh Anand Singh (Department of Statistics, Manipur University) for assisting statistical analysis. The support from the Department of Biotechnology (DBT), Ministry of Science and Technology, Government of India (Grant No. BT/57/NE/TBP to LSS) is gratefully acknowledged.

References

- Adham M, Greijer AE, Verkuijlen SA, et al (2013). Epstein-barr virus DNA load in nasopharyngeal brushings and whole blood in nasopharyngeal carcinoma patients before and after treatment. *Clin Cancer Res*, **19**, 2175-86.
- Arand M, Muhlbauer R, Hengstler J, et al (1996). A multiplex polymerase chain reaction protocol for the simultaneous analysis of the glutathione S-transferase GSTM1 and GSTT1 polymorphisms. *Anal Biochem*, **236**, 184-6.
- Ariyoshi N, Takahashi Y, Miyamoto M, et al (2000). Structural characterization of a new variant of the CYP2A6 gene (CYP2A6*1B) apparently diagnosed as heterozygotes of CYP2A6*1A and CYP2A6*4C. *Pharmacogenetics*, **10**, 687-93.
- Bhatia P, Singh L (1981). Evaluation of contrast radiography in nasopharyngeal malignancy. *Indian J Cancer*, **18**, 141.
- Boccia S, Cadoni G, Sayed-Tabatabaei FA, et al (2008). CYP1A1, CYP2E1, GSTM1, GSTT1, EPHX1 exons 3 and 4, and NAT2 polymorphisms, smoking, consumption of alcohol and fruit and vegetables and risk of head and neck cancer. *J Cancer Res Clin Oncol*, **134**, 93-100.
- Bozina N, Bradamante V, Lovric M (2009). Genetic polymorphism of metabolic enzymes P450 (CYP) as a susceptibility factor for drug response, toxicity, and cancer risk. *Arh Hig Rada Toksikol*, **60**, 217-42.
- Burt RD, Vaughan TL, McKnight B (1992). Descriptive epidemiology and survival analysis of nasopharyngeal carcinoma in the United States. *Int J Cancer*, **52**, 549-56.
- Chang ET, Adami HO (2006). The enigmatic epidemiology of nasopharyngeal carcinoma. *Cancer Epidemiol Biomarkers Prev*, **15**, 1765-77.
- Cheng YJ, Chien YC, Hildesheim A, et al (2003). No association between genetic polymorphisms of CYP1A1, GSTM1, GSTT1, GSTP1, NAT2, and nasopharyngeal carcinoma in Taiwan. *Cancer Epidemiol Biomarkers Prev*, **12**, 179-80.
- Gattas GJ, de Carvalho MB, Siraque MS, et al (2006). Genetic polymorphisms of CYP1A1, CYP2E1, GSTM1, and GSTT1 associated with head and neck cancer. *Head Neck*, **28**, 819-26.
- Ghosh SK, Singh AS, Mondal R, et al (2014). Dysfunction of mitochondria due to environmental carcinogens in nasopharyngeal carcinoma in the ethnic group of Northeast Indian population. *Tumour Biol*, **35**, 6715-24.
- Go RE, Hwang KA, Choi KC (2015). Cytochrome P450 1 family and cancers. *J Steroid Biochem Mol Biol*, **147**, 24-30.
- Guengerich FP, Shimada T (1991). Oxidation of toxic and carcinogenic chemicals by human cytochrome P-450 enzymes. *Chem Res Toxicol*, **4**, 391-407.
- Hayashi S, Watanabe J, Kawajiri K (1991). Genetic polymorphisms in the 5'-flanking region change transcriptional regulation of the human cytochrome P450IIE1 gene. *J Biochem*, **110**, 559-65.
- Ho HC (1976). Epidemiology of nasopharyngeal carcinoma. *Gann Monograph Cancer Res*, **18**, 49-61.
- Huang FM, Chen HC, Khan MA, et al (2013). CYP2A6, CYP1A1, and CYP2D6 polymorphisms in lung cancer patients from central south China. *Med Oncol*, **30**, 521.
- Hussein AG, Pasha HF, El-Shahat HM, et al (2014). CYP1A1 gene polymorphisms and smoking status as modifier factors for lung cancer risk. *Gene*, **541**, 26-30.
- Islam MS, Ahmed MU, Sayeed MS, et al (2013). Lung cancer risk in relation to nicotinic acetylcholine receptor, CYP2A6 and CYP1A1 genotypes in the Bangladeshi population. *Clin Chim Acta*, **416**, 11-9.
- Ji MF, Huang QH, Yu X, et al (2014). Evaluation of plasma Epstein-Barr virus DNA load to distinguish nasopharyngeal carcinoma patients from healthy high-risk populations in Southern China. *Cancer*, **120**, 1353-60.
- Jiang XY, Chang FH, Bai TY, et al (2014). Susceptibility of lung cancer with polymorphisms of CYP1A1, GSTM1, GSTM3, GSTT1 and GSTP1 genotypes in the population of Inner Mongolia region. *Asian Pac J Cancer Prev*, **15**, 5207-14.
- Jiang Y, Li N, Dong P, et al (2011). Polymorphisms in GSTM1, GSTT1 and GSTP1 and nasopharyngeal cancer in the east of China: a case-control study. *Asian Pac J Cancer Prev*, **12**, 3097-100.
- Kataki AC, Simons MJ, Das AK, et al (2011). Nasopharyngeal carcinoma in the Northeastern states of India. *Chin J Cancer*, **30**, 106-13.
- Lakhanpal M, Singh LC, Rahman T, et al (2014). Contribution of susceptibility locus at HLA class I region and environmental factors to occurrence of nasopharyngeal cancer in Northeast India. *Tumour Biol*, [DETAIL].
- Lin JC, Chen KY, Wang WY, et al (2001). Detection of Epstein-Barr virus DNA in the peripheral-blood cells of patients with nasopharyngeal carcinoma: relationship to distant metastasis and survival. *J Clin Oncol*, **19**, 2607-15.
- Lin JC, Wang WY, Chen KY, et al (2004). Quantification of plasma Epstein-Barr virus DNA in patients with advanced nasopharyngeal carcinoma. *N Engl J Med*, **350**, 2461-70.
- Liu YL, Xu Y, Li F, et al (2013a). CYP2A6 deletion polymorphism is associated with decreased susceptibility of lung cancer in Asian smokers: a meta-analysis. *Tumour Biol*, **34**, 2651-7.
- Liu ZB, Shu J, Wang LP, et al (2013b). Cytochrome P450 2A6 deletion polymorphism and risk of lung cancer: a meta-analysis. *Mol Biol Rep*, **40**, 5255-9.
- Lo YM, Chan LY, Lo KW, et al (1999). Quantitative analysis of cell-free Epstein-Barr virus DNA in plasma of patients with nasopharyngeal carcinoma. *Cancer Res*, **59**, 1188-91.
- Lo YM, Leung SF, Chan LY, et al (2000). Plasma cell-free Epstein-Barr virus DNA quantitation in patients with nasopharyngeal carcinoma. Correlation with clinical staging. *Ann NY Acad Sci*, **906**, 99-101.
- Lung ML, Cheung AK, Ko JM, et al (2014). The interplay of host genetic factors and Epstein-Barr virus in the development of nasopharyngeal carcinoma. *Chin J Cancer*, **33**, 556-68.
- Mathur S (2003). Epidemiological and etiological factors associated with nasopharyngeal carcinoma. ICMR Bulletin. Indian Press New Delhi, 33.
- Moraes LN, Borges MF, Sousa PA, et al (2012). Lack of association of CYP1A1-MspI SNP and GSTM1 null genotypes with cancer in a Brazilian family with unusually high cancer incidence. *Genet Mol Res*, **11**, 1610-7.
- Murthy AK, Kumar V, Suresh KP (2013). Meta-analysis of GSTM1 and GSTT1 polymorphisms and risk of nasopharyngeal cancer. *Asian Pac J Cancer Prev*, **14**, 1697-701.
- Nguyen QM, Nguyen HC, Parkin DM (1998). Cancer incidence in Ho Chi Minh City, Viet Nam, 1995-1996. *Int J Cancer*, **76**, 472-9.
- Nonoyama M, Pagano JS (1973). Homology between Epstein-Barr virus DNA and viral DNA from Burkitt's lymphoma and nasopharyngeal carcinoma determined by DNA-DNA reassociation kinetics. *Nature*, **242**, 44-7.
- Nor Hashim NA, Ramzi NH, Velapasamy S, et al (2012). Identification of genetic and non-genetic risk factors for nasopharyngeal carcinoma in a Southeast Asian population. *Asian Pac J Cancer Prev*, **13**, 6005-10.
- Parkin D, Whelan S, Ferlay J, et al (2002). Cancer incidence in five continents Vol. VIII. IARC scientific publications, 155.
- Rossini A, Rapozo DC, Soares Lima SC, et al (2007). Polymorphisms of GSTP1 and GSTT1, but not of CYP2A6, CYP2E1 or GSTM1, modify the risk for esophageal cancer

- in a western population. *Carcinogenesis*, **28**, 2537-42.
- Russo A, Francelin PR, Galbiatti AL, et al (2013). Association between GSTP1, GSTM1 and GSTT1 polymorphisms involved in xenobiotic metabolism and head and neck cancer development. *Mol Biol Rep*, **40**, 4181-8.
- Saiki RK, Gelfand DH, Stoffel S, et al (1988). Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science*, **239**, 487-91.
- Sharma TD, Singh TT, Laishram RS, et al (2011). Nasopharyngeal carcinoma--a clinico-pathological study in a regional cancer centre of northeastern India. *Asian Pac J Cancer Prev*, **12**, 1583-7.
- Shotelersuk K, Khorprasert C, Sakdikul S, et al (2000). Epstein-Barr virus DNA in serum/plasma as a tumor marker for nasopharyngeal cancer. *Clin Cancer Res*, **6**, 1046-51.
- Song DK, Xing DL, Zhang LR, et al (2009). Association of NAT2, GSTM1, GSTT1, CYP2A6, and CYP2A13 gene polymorphisms with susceptibility and clinicopathologic characteristics of bladder cancer in Central China. *Cancer Detect Prev*, **32**, 416-23.
- Stevens SJ, Pronk I, Middeldorp JM (2001a). Toward standardization of Epstein-Barr virus DNA load monitoring: unfractionated whole blood as preferred clinical specimen. *J Clin Microbiol*, **39**, 1211-6.
- Stevens SJ, Verschuuren EA, Pronk I, et al (2001b). Frequent monitoring of Epstein-Barr virus DNA load in unfractionated whole blood is essential for early detection of posttransplant lymphoproliferative disease in high-risk patients. *Blood*, **97**, 1165-71.
- Stevens SJ, Vervoort MB, van den Brule AJ, et al (1999). Monitoring of epstein-barr virus DNA load in peripheral blood by quantitative competitive PCR. *J Clin Microbiol*, **37**, 2852-7.
- Tamaki Y, Arai T, Sugimura H, et al (2011). Association between cancer risk and drug-metabolizing enzyme gene (CYP2A6, CYP2A13, CYP4B1, SULT1A1, GSTM1, and GSTT1) polymorphisms in cases of lung cancer in Japan. *Drug Metab Pharmacokinet*, **26**, 516-22.
- Tiwawech D, Srivatanakul P, Karalak A, et al (2006). Cytochrome P450 2A6 polymorphism in nasopharyngeal carcinoma. *Cancer Lett*, **241**, 135-41.
- Topcu Z, Chiba I, Fujieda M, et al (2002). CYP2A6 gene deletion reduces oral cancer risk in betel quid chewers in Sri Lanka. *Carcinogenesis*, **23**, 595-8.
- Tsang CM, Deng W, Yip YL, et al (2014). Epstein-Barr virus infection and persistence in nasopharyngeal epithelial cells. *Chin J Cancer*, **33**, 549-55.
- Tsao SW, Tsang CM, To KF, et al (2015). The role of Epstein-Barr virus in epithelial malignancies. *J Pathol*, **235**, 323-33.
- Tsao SW, Yip YL, Tsang CM, et al (2014). Etiological factors of nasopharyngeal carcinoma. *Oral Oncol*, **50**, 330-8.
- Tune CE, Liavaag PG, Freeman JL, et al (1999). Nasopharyngeal brush biopsies and detection of nasopharyngeal cancer in a high-risk population. *J Natl Cancer Inst*, **91**, 796-800.
- Wei Y, Zhou T, Lin H, et al (2013). Significant associations between GSTM1/GSTT1 polymorphisms and nasopharyngeal cancer risk. *Tumour Biol*, **34**, 887-94.
- Yang J, Li L, Yin X, et al (2015). The association between gene polymorphisms and risk of nasopharyngeal carcinoma. *Med Oncol*, **32**, 398.
- Yip TT, Ngan RK, Fong AH, et al (2014). Application of circulating plasma/serum EBV DNA in the clinical management of nasopharyngeal carcinoma. *Oral Oncol*, **50**, 527-38.