CHAPTER-4

RESULTS

In Northeast (NE) India, prevalence of head and neck cancer is very high with oral cancer being the most frequent. However, very little is known about the environmental, clinical, genetic and epigenetic features of head and neck cancer from NE India. For that reason, we performed a comprehensive study on the environmental and viral risk factors, genetic alteration (polymorphisms), epigenetic alteration (promoter hypermethylation), their combined effect (gene-gene interaction and gene-environment interactions) and the survival status of HNC from NE India. The findings of the present study have been presented in five different sub-chapters of results:

Chapter 4.1 Prevalence of HPV and environmental risk factors associated with head and neck cancer

Chapter 4.3 Genetic polymorphism of DNA repair genes (*XRCC1* and *XRCC2*) and risk of HNSCC

Chapter 4.2 Genetic polymorphisms of carcinogen metabolizing genes in head and neck cancer

Chapter 4.4 Epigenetic alterations in head and neck cancer

Chapter 4.5 Correlation between HPV, environmental and genetic factors with promoter methylation profile of HNSCC patients

Chapter 4.1: Prevalence of HPV and Environmental Risk Factors Associated with Head and Neck Cancer

4.1.1 Demographic Characteristics of Study Subjects

The detailed demographic data of study population is depicted in **Table 4.1.1.** The percentage of male members in head and neck cancer (HNC) patients, first-degree relatives and healthy control groups were 67.1%, 64% and 62.7% respectively, whereas the frequencies of female members were 32.9%, 36% and 37.3% respectively. There was large majority of patients belongs to age group 45-65 years (55.2%) at the time of diagnosis. There was no significant difference observed in gender and age distribution among the study groups (p=0.631 and p=0.839 respectively), which indicate that the frequency matching of our study population was sufficient (**Figure 4.1.1**). We observed regular consumptions of fish, dry fish and meat among patients (82.9%, 54.8% and 60.6% respectively), first-degree relatives (78.7%, 66% and 65% respectively) and controls (81.7, 53.3 and 55.7% respectively). The consumption of dry fish among the study groups varies significantly (p=0.011), however no significant association was observed in fish and cooked meat consumption (p=0.465 and p=0.064 respectively) between patients, first-degree relatives and controls. The frequencies of smoking, betel-quid and tobacco chewing in HNC patients were 72.4%, 76.5% and 74.7% respectively. Whereas, in controls, the frequencies of tobacco smoking, betel-quid and tobacco chewing were 55%, 59.3% and 54.3% respectively. In first-degree relatives the frequencies of tobacco smoking, betel-quid and tobacco chewing were 48.3%, 55% and 61% respectively. Overall, there was significant differences observed in consumption of smoking, betel-quid and tobacco chewing among the study groups (p<0.01). The comparison between patients versus controls, first-degree relatives versus controls, and patients versus first-degree relatives in terms of tobacco habits, were summarized in **Table 4.1.2**. Analysis showed the prevalence of smoking, betel quid chewing and tobacco chewing were significantly higher in patients compared to first-degree relatives and controls (p<0.001). We also found that tobacco chewing habit is slightly higher in first-degree relatives compared to controls (p=0.098). However, in case of smoking and betel quid chewing, no significant variation was observed between first-degree relatives and controls (p=0.102 and p=0.283 respectively) (**Table 4.1.2**).

Table 4.1.1 Demographic characteristics of patients, first-degree relatives and controls

Characteristics	Patients No. (%)	First-degree relatives No. (%)	Controls No. (%)	*p*-value*
Gender				
Male	114 (67.1)	192 (64)	188 (62.7)	
Female	56 (32.9)	108 (36)	112 (37.3)	0.631
Age				
25-45	40 (23.6)	77 (25.7)	85 (28.3)	
45-65	94 (55.2)	159 (53)	153 (51)	
>65	36 (21.2)	64 (21.3)	62 (20.7)	0.839
Fish intake				
Never	29 (17.1)	64 (21.3)	55 (18.3)	
Ever	141 (82.9)	236 (78.7)	245 (81.7)	0.465
Dry fish intake				
Never	77 (45.2)	102 (34)	134 (44.7)	
Ever	93 (54.8)	198 (66)	166 (53.3)	0.011
Meat intake				
Never	67 (39.4)	105 (35)	133 (44.3)	
Ever	103 (60.6)	195 (65)	167 (55.7)	0.064
Smoking				
Non-smokers	47 (27.6)	155 (51.7)	135 (45)	
Smokers	123 (72.4)	145 (48.3)	165 (55)	<0.01
Betel-quid chewing				
Non-chewers	40 (23.5)	135 (45)	122 (40.7)	
Chewers	130 (76.5)	165 (55)	178 (59.3)	<0.01
Tobacco chewing				
Non-chewers	43 (25.3)	117 (39)	137 (45.7)	
Chewers	127 (74.7)	183 (61)	163 (54.3)	<0.01

*$p<0.05$ was considered as statistically significant

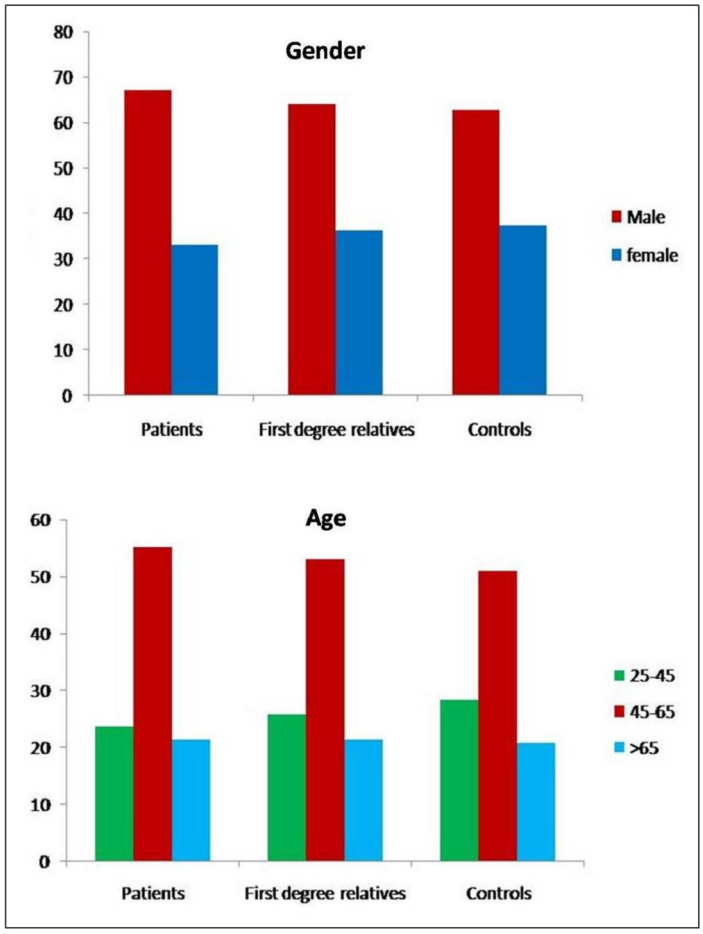

Figure 4.1.1: Age and gender distribution among patients, first-degree relatives and controls

4.1.2 The Association between Environmental Risk Factors and HNC

The association between head and neck cancer (HNC) and environmental risk factors such as tobacco consumption in various forms, was assessed and depicted in **Table 4.1.3**. The crude odds ratio (OR) was found higher for smoking (OR=2.14; 95% CI=1.43 to 3.21) and betel quid (OR=2.23; 95% CI=1.46 to 3.40), and for tobacco chewing (OR=2.48; 95% CI=1.64 to 3.75), also showed significant association (p<0.001). After adjusting for potential confounding factors (age, gender and tobacco habits), we found tobacco chewers had 2.35 fold (95% CI=1.54 to 3.59 and p<0.001) increase the risk of HNC compared to non-tobacco chewers. Whereas betel quid chewers and smokers showed 2.03 (95% CI=1.31 to 3.13 and p<0.001) and 1.96 fold (95% CI=1.29 to 2.99, p=0.002) increase the risk of HNC compared to non-chewer and non-smokers respectively.

Table 4.1.2 The comparison between patients, first-degree relatives and controls in terms of tobacco habits

Genotypes/habits	***p*-value***		
	Patients Vs. controls	First-degree relatives Vs. controls	Patients Vs. First-degree relatives
Smoking	<0.001	0.102	<0.001
Betel-quid chewing	<0.001	0.283	<0.001
Tobacco chewing	<0.001	0.098	0.003

**p*<0.05 was considered as statistically significant.

Table 4.1.3 Adjusted and crude odds ratio of environmental risk factors for assessment of HNC risk

Habits	**Patients/ Controls**	**Crude OR (95%CI)**	***p*-value**	**Adjusted OR* (95%CI)**	***p*-value**
Smoking					
Non-smokers	47/135	1 (Reference)		1 (Reference)	
smokers	123/165	2.14 (1.43 to 3.21)	<0.001	1.96 (1.29 to 2.99)	0.002
Betel-quid					
Non-chewers	40/122	1 (Reference)		1 (Reference)	
chewers	130/178	2.23 (1.46 to 3.40)	<0.001	2.03 (1.31 to 3.13)	<0.001
Tobacco					
Non-chewers	43/137	1 (Reference)		1 (Reference)	
chewers	127/163	2.48 (1.64 to 3.75)	<0.001	2.35 (1.54 to 3.59)	<0.001

OR=Odds ratio; CI=confidence interval

***** Adjusted for age, gender, smoking, betel quid chewing and tobacco chewing

p<0.05 was considered as statistically significant.

4.1.3 Prevalence of HPV in Head and Neck Cancer

4.1.3.1 PCR assay for detection of HPV

PCR of the oncogenic HPV DNA was carried out using consensus primers and 450 bp band confirmed HPV DNA presence (**Figure 4.1.2**). Overall, the prevalence of HPV in Northeast Indian population (cases-controls) was 30.7% (126 HPV positive samples out of 410). Based upon the most sensitive method of detection, the prevalence of HPV in head and neck cancer patients was 46.47% (79/170). The prevalence of HPV in healthy control samples was 19.59% (47/240). In the present study, HPV positive found to be associated with 3.43 fold (95% CI=2.21 to 5.32 and *p*-value less than 0.001) increased the risk of HNC in Northeast Indian population (**Table 4.1.4**). In males, frequency of HPV- positive samples was 66.7%, while in female 33.3%. The frequency of HPV-positive detected high among smokers, betel quid chewers and tobacco chewers (82.1%, 66.7% and 76.9% respectively) compared to non-smokers and non-chewers. However, there was no significant difference was observed between HPV-positive and HPV-negative HNC patients Among head and neck sites, HPV was most frequently detected in tumours of the oral cavity (68.4%), followed by the pharynx (12.6%), and larynx (11.3%) (**Table 4.1.5**)

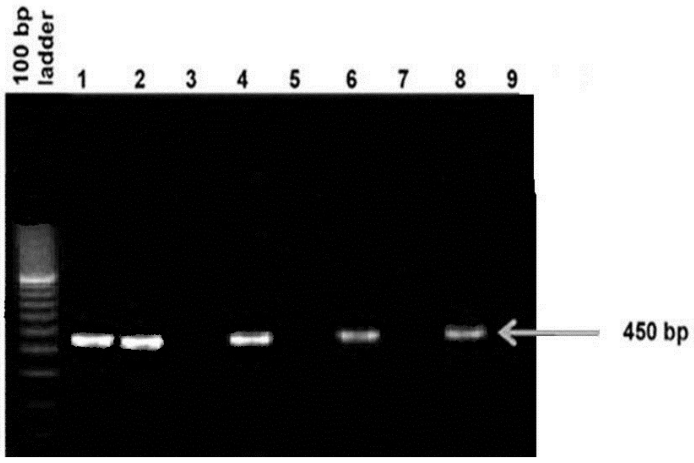

Figure 4.1.2: Agarose gel showing HPV presence or absence. 100 bp ladder (lane 1); **HPV positive** (lanes 1, 2, 4, 6 and 8); HPV-negative (lanes 3, 5, 7 and 9).

Table 4.1.4 HPV prevalence in HNSCC patients and healthy controls

HPV	Cases (170)	Controls (240)	OR	95% CI	*p*-value
Presence	79 (46.47)	47(19.59)	3.43	2.21 to 5.32	<0.001
Absence	91 (53.53)	193 (80.41)	1 (reference)		

OR=Odds ratio; CI=confidence interval
$p<0.05$ was considered as statistically significant.

Table 4.1.5 Characteristics of HPV-positive and HPV-negative samples

Characteristics	HPV-positive	HPV-negative	*p*-value
Gender			
Male	26 (66.7)	25 (78.1)	
Female	13 (33.3)	7 (21.9)	0.306
Smoking			
Yes	32 (82.1)	19 (59.3)	
No	7 (17.9)	13 (40.7)	0.062
Betel quid chewing			
Yes	26 (66.7)	21 (65.6)	
No	13 (33.3)	11 (34.4)	0.901
Tobacco chewing			
Yes	30 (76.9)	23 (71.9)	
No	9 (23.1)	9 (28.1)	0.785
Tumour sites			
Oral	54 (68.4)	55 60.44)	
Pharynx	10 (12.6)	11 (10.1)	
Larynx	9 (11.3)	15 (12.1)	
Other	6 (7.6)	10 (12.6)	

OR=Odds ratio; CI=confidence interval
$p<0.05$ was considered as statistically significant

4.1.3.2 Association of HPV and survival of HNSCC patients

Survival was examined with respect to HPV using Kaplan-Meier survival curves (**Figure. 4.1.3**). We had followed up data of 71 head and neck squamous cell carcinoma (HNSCC) patients with 47 deaths. Analysis revealed that the overall median survival time of 47 patients out of 71 was 15 months [95% CI=10.97-19.03], and the median survival time in HPV-positive and HPV-negative were 17 months [95% CI=12.13 to 21.86] and 13 months [95% CI=10.24 to 15.75]. Results revealed that HPV-positive HNSCC patients showed better survival compared to HPV-negative patients (p=0.041) (**Table 4.1.6**).

Table 4.1.6 Summary of survival data with respect to HPV

Variables	**Median (months)**	**95% Confidence Interval**		***p*-value**
		Lower	**Upper**	
Overall	15	10.97	19.03	
HPV-positive	17	12.13	21.86	0.041
HPV-negative	13	10.24	15.75	

p<0.05 was considered as statistically significant

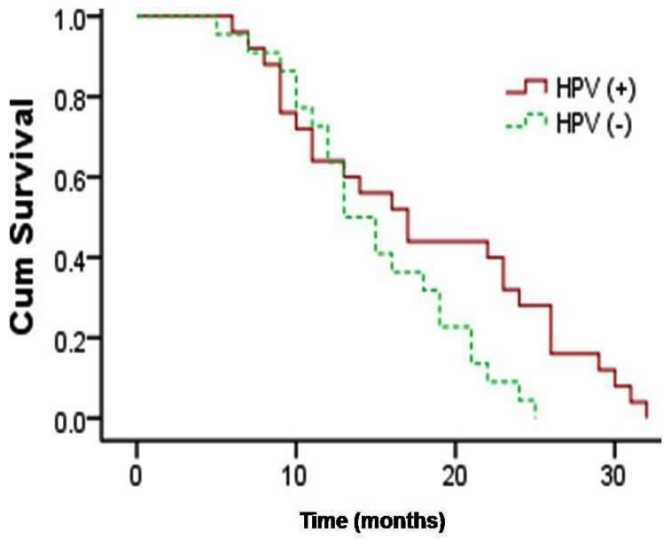

Figure 4.1.3 Kaplan-Meier survival plots for HPV. HPV-positive HNSCC tumours showing better survival compared to HPV-negative tumours.

Chapter 4.2:

Genetic Polymorphisms of Carcinogen Metabolizing Genes in Head and Neck Cancer

GSTM1 and *GSTT1* genes are Phage II carcinogen metabolizing polymorphic genes, and their null genotypes result in a complete loss of functional activity of GSTs enzymes. The *CYP1A1* gene is phase-I metabolizing polymorphic gene and most widely studied polymorphism of *CYP1A1* is ***m1*** (rs4646903), which is characterized by T→C transition at 3801 nucleotide position. This polymorphism may affect level of messenger RNA (mRNA) stability or gene expression and become play a crucial role in susceptibility to DNA adduct formation and increased cancer risk. The genotypes of *GSTT1* and *GSTM1* were detected by the observing present/absent of the desired band in 1.5% agarose gel (**Figure 4.2.1**). The allele types of *CYP1A1* T3801C gene were determined by three distinct banding patterns such as: a single 343 bp fragment indicate the wild-type TT allele, three fragments of 343, 200 and 143 bp for the TC heterozygous allele and two fragments of 200 and 143 bp indicate variant CC allele of *CYP1A1* (**Figure 4.2.2**).

4.2.1 Distribution of Genotypes of *GSTM1* and *GSTT1* in Cases, First-degree relatives and Controls

Table 4.2.1 depicted the genotype frequency distribution of *GSTM1* and *GSTT1* among HNSCC patients, their first-degree relatives and controls. The frequency of *GSTM1* null genotype found among patients, first-degree relatives (FDRs) and controls were 55.9%, 53% and 41.3% respectively. Whereas, frequency of *GSTT1* null genotype found among patients, first-degree relatives and controls were 35.9%, 34.3% and 26.7% respectively. For both *GSTM1* and *GSTT1* null genotype, the frequency distributions observed in patients, first- degree relatives and controls were 17.6%, 19.7% and 11% respectively (**Figure 4.2.3**). Analysis showed that *GSTM1* and *GSTT1* null genotype frequencies were significantly higher in patients compared to controls (p=0.002 and p=0.036 respectively). We also found a significant variation in frequencies of *GSTM1* and *GSTT1* null genotypes between first-degree relatives of patients and controls (p=0.004 and p=0.041 respectively). However, no significant variation was found in case of *GSTM1* and *GSTT1* null genotype distribution between patients and their first-degree relatives (p=0.547 and

p=0.735 respectively). Moreover, we found a highly significant difference in the frequency of both *GSTM1-GSTT1* null genotype in patients and first-degree relatives, when we compared the data with controls (*p*=0.001 and <0.001 respectively)

4.2.2 Association between Genetic Polymorphisms of *GSTM1* and *GSTT1* gene and HNC

The association between *GSTs* null genotype and HNC was evaluated using the logistic regression (LR) analysis (**Table 4.2.2**). We found significant association between *GSTM1* and *GSTT1* null genotype and HNC. The crude odds ratio (OR) was found higher for *GSTM1* (OR=1.81; 95%CI=1.23 to 2.63 and *p*=0.003) and *GSTT1* null genotype (OR=1.54; 95%CI=1.03 to 2.30 and *p*=0.046). After adjusting for potential confounding factors like age, gender and tobacco habits, individuals carrying *GSTM1* null genotype had 2.18 fold (95%CI=1.43 to 3.33 and *p*<0.001) increase the risk of HNC as compared to individuals having *GSTM1* genotype. Whereas, *GSTT1* null genotype (deletion) showed 1.61 fold (95%CI=1.04 to 2.52 and *p*=0.031) increased the risk of HNC compared to *GSTT1* wild genotype (presence).

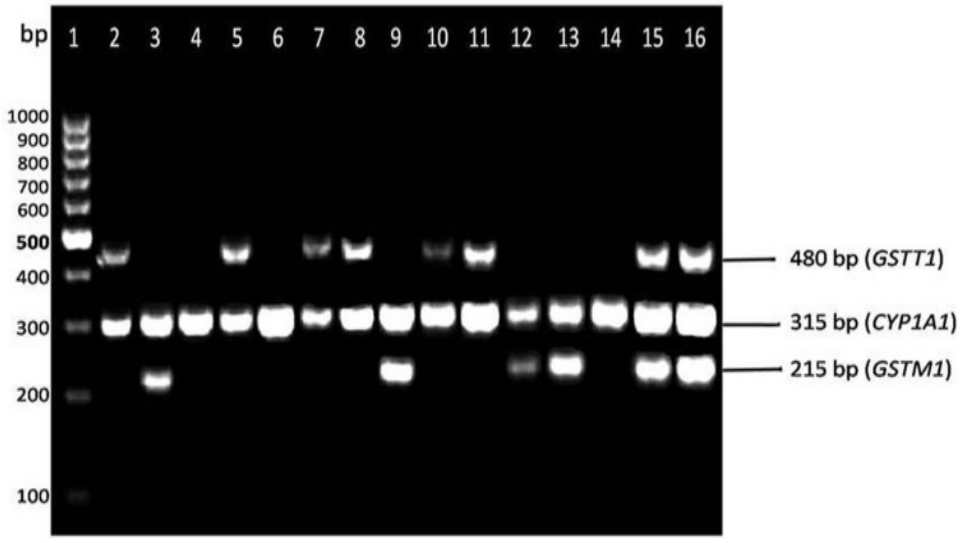

Figure 4.2.1: Polymorphism of *GSTM1* and *GSTT1* metabolic genes. Ethidium bromide stained gel showing 100 bp ladder (lane 1); *GSTM1* null genotype (lanes 2, 5, 7, 8, 10 and 11); *GSTT1* null genotype (lanes 3, 9, 12 and 13); both *GSTM1* & *GSTT1* null genotype (lanes 4, 6, and 14) and both *GSTM1* & *GSTT1* presence (lane 15 and 16).

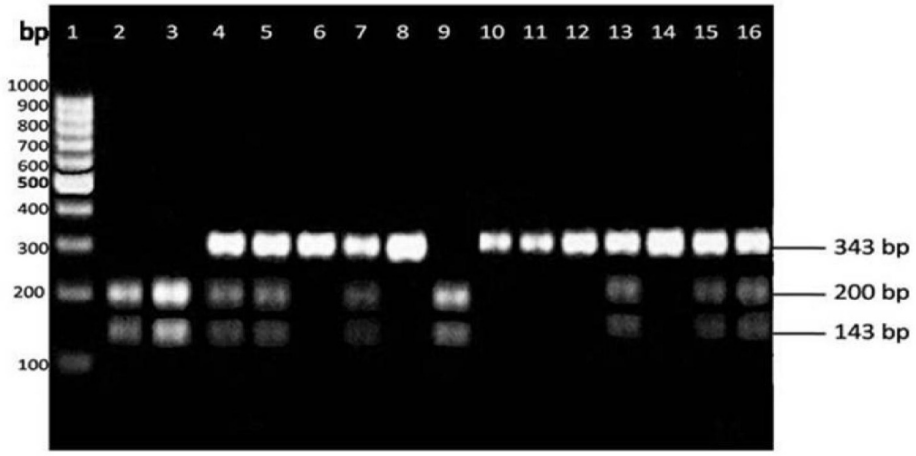

Figure 4.2.2: Ethidium bromide stained gel showing *CYP1A1* T3801C polymorphism. 100 bp ladder (lane 1); single band denotes *CYP1A1* **TT** wild type (lanes 6,8,10,11,12 and 14); three band indicates **TC** heterozygous genotype (lanes 4, 5, 7, 13, 15 and 16) and double band signify **CC** genotype (lanes 2, 3 and 9)

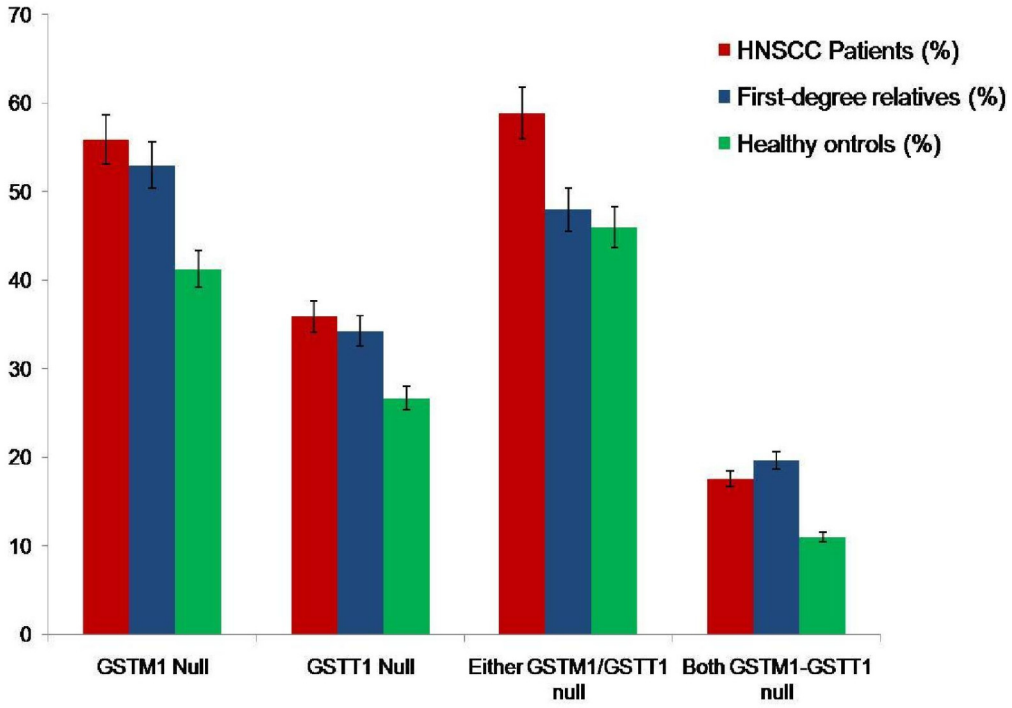

Figure 4.2.3: Frequency of genotype distribution of *GSTM1* and *GSTT1* in patients, first-degree relatives and controls

Table 4.2.1 Distribution of *GSTM1* and *GSTT1* genotypes among patients, first-degree relatives (FDRs) and controls

Genotypes	Patients No. (%)	FDRs No. (%)	Controls No. (%)	*p*-value		
				Patients Vs. controls	FDRs Vs. controls	Patients Vs. FDRs
GSTM1						
Present	75 (44.1)	141 (47)	176 (58.7)			
Null	95 (55.9)	159 (53)	124 (41.3)	**0.002**	**0.004**	0.547
GSTT1						
Present	109 (64.1)	197 (65.7)	220 (73.3)			
Null	61 (35.9)	103 (34.3)	80 (26.7)	**0.036**	**0.041**	0.735
GSTM1* and *GSTT1						
Both present	45 (26.4)	97 (32.3)	129 (43)			
Either null	95 (58.9)	144 (48)	138 (46)	**0.002**	0.068	0.115
Both null	30 (17.6)	59 (19.7)	33 (11)	**0.001**	**<0.001**	0.778

$p<0.05$ was considered as statistically significant.

Table 4.2.2 Adjusted and crude odds ratio of risk factors (*GSTM1* and *GSTT1* gene) for assessment of HNC risk

Genotypes	Patients/ Controls	Crude OR (95%CI)	*p*-value	Adjusted OR* (95%CI)	*p*-value
GSTM1					
Present	75/176	1 (Reference)		1 (Reference)	
Null	95/124	1.81 (1.23 to 2.63)	**0.003**	2.18 (1.43 to 3.33)	**<0.001**
GSTT1					
Present	109/220	1 (Reference)		1 (Reference)	
Null	61/80	1.54 (1.03 to 2.30)	**0.046**	1.61 (1.04 to 2.51)	**0.031**

*Odds ratio (OR) adjust for age, gender, smoking, tobacco-betel quid chewing as appropriate. $p<0.05$ was considered as statistically significant.

4.2.3 Genetic Polymorphisms in *CYP1A1* Gene and HNC Risk

The genotyping of *CYP1A1* T3801C was done in 180 HNC patients and 240 healthy controls. Among the patients, 68.9% were male and 31.1% female, whereas in control group, 61.7% were male and 38.3% female. Of the 180 cases; 63 (35%) had oral cancer, 24 (13.3%) had laryngeal malignancy, 70 (38.9%) had nasopharyngeal cancer, 10 (5.6) pharyngeal and 13 (7.6%) had other types of cancer in head and neck region. The three genotypes of *CYP1A1* T3801C viz. TT, TC and CC had frequency distributions of 44.4%, 40.6%, 15% and 54.2%, 35.4%, 10.4% in cases and controls respectively (**Table 4.2.3** and **Figure 4.2.4**). The Hardy–Weinberg Equilibrium analysis showed that the genotype frequency of *CYP1A1* gene is in equilibrium in both the cases and the controls ($\chi 2 = 2.25$, p=0.133 and $\chi 2 = 3.68$, p=0.06).

Table 4.2.3 Distribution of *CYP1A1* (T3801C) genotype among the study population

Genotype	Cases (%)	Controls (%)	Crude OR (95% CI)	*p*-value	Adjusted OR* (95% CI)	*p*-value
TT	80 (44.4)	130 (54.2)	1(reference)		1(reference)	
TC	73 (40.6)	85 (35.4)	1.40 (0.92-2.12)	0.135	1.45 (0.93-2.25)	0.098
CC	27 (15)	25 (10.4)	1.76 (0.96-3.22)	0.083	**1.97** (1.03-3.77)	**0.040**
TC+CC	100 (55.6)	110 (45.8)	1.48 (1.00-2.18)	0.061	**1.56** (1.03-2.36)	**0.033**
T-allele	233 (64.7)	345 (71.9)	1(reference)			
C-allele	127 (35.3)	135 (28.1)	1.39 (1.04-1.87)	**0.029**		

*Odds ratio (OR) adjust for age, gender, smoking and tobacco-betel quid chewing as appropriate
Bold values: Statistical significant

Logistic regression method was used to analyze the association between *CYP1A1* T3801C genotype and HNC risk. We found that CC (OR=1.97, 95% CI: 1.03-3.77; p=0.040) genotypes show significant risk association with HNC in the study population when compare to the TT wild genotype (**Table 4.2.3**). Combined TC and CC (TT + CC) genotypes also showed a risk of HNC (OR=1.56, 95% CI: 1.03-2.36, p=0.033). The risk associated with each allele was also investigated and we found significant risk of HNC in those individual carrying the C-allele as compared to T- allele (p=0.029).

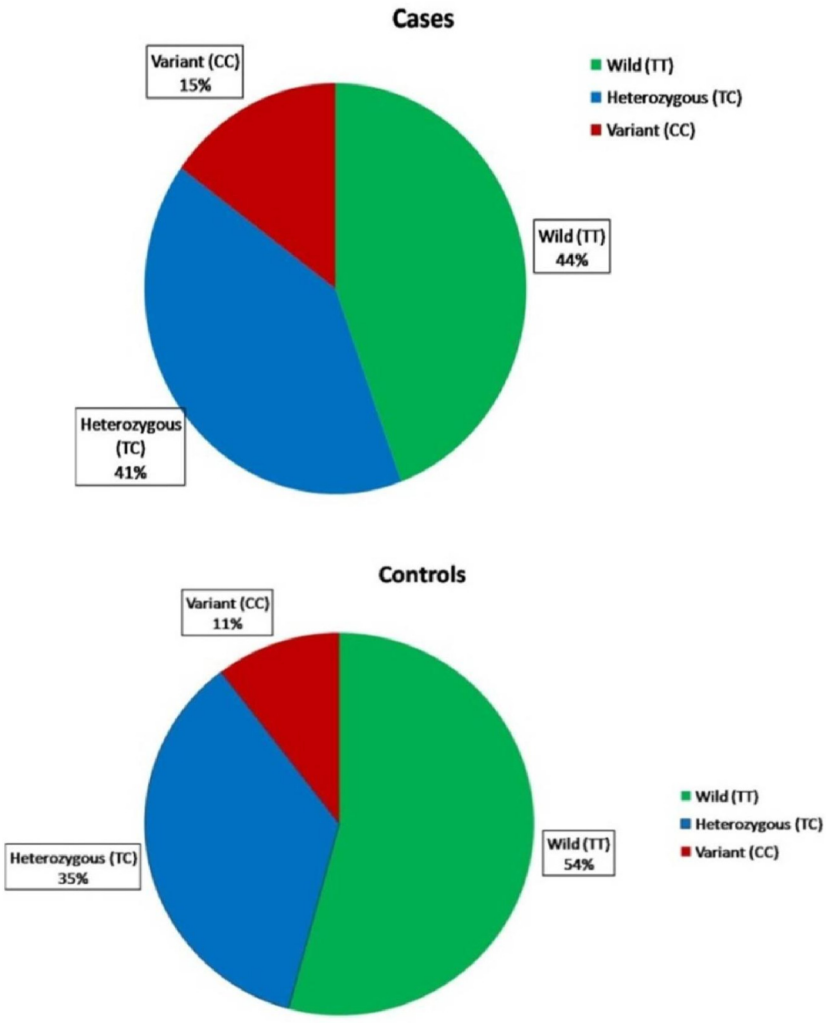

Figure 4.2.4: Frequency of genotype distribution of *CYP1A1* T3801C in patients and controls

4.2.4 Gene-Gene and Gene-Environment Interactions

We investigated gene-gene and gene-environment interaction taking into account that the impact of environmental risk factors likes smoking and tobacco-betel quid chewing and multiple genes association on head and neck cancer risk.

4.2.4.1 Gene-Gene Interaction of *CYP1A1* T3801C, *GSTM1* and *GSTT1* Genes

We analyzed the gene-gene interaction of *CYP1A1* T3801C and *GSTs* genes and their association with HNC using the combined low-risk wild genotypes of *CYP1A1*, *GSTM1* and *GSTT1* as a reference group (**Table 4.2.4**). Individuals carrying both *CYP1A1* TC/CC and *GSTM1* null genotypes, showed the highest risk of HNC (OR=3.52; 95%CI=1.90-6.51; $p<0.001$).

Table 4.2.4 Combined genotypes of *CYP1A1* T3801C, *GSTM1* and *GSTT1* genes and HNC risk

Combined genotypes	**Cases N (%)**	**Controls N (%)**	**OR (95% CI)**	***p*-value**
***CYP1A1* T3801C + *GSTM1* (Wild/Null)**				
TT + Wild type	37 (20.5)	66 (27.5)	1 (reference)	
TT + Null type	43 (23.9)	64 (26.7)	1.29 (0.72 to 2.31)	0.376
TC/CC+ Wild type	43 (23.9)	75 (31.2)	1.10 (0.62 to 1.96)	0.724
TT/CC + Null type	57 (31.7)	35 (14.6)	**3.52** (1.90 to 6.51)	**<0.001**
***CYP1A1* T3801C + *GSTT1* (Wild/Null)**				
TT + Wild type	50 (27.8)	90 (37.5)	1 (reference)	
TT + Null type	30 (16.7)	40 (16.7)	1.42 (0.78 to 2.62)	0.249
TC/CC+ Wild type	65 (36.1)	72 (30)	**1.75** (1.06 to 2.89)	**0.027**
TT/CC + Null type	35 (19.4)	38 (15.8)	**1.91** (1.05 to 3.47)	**0.033**

$p<0.05$ was considered as statistically significant.

OR=Odds Ratio and CI=Confidence Interval

However, *CYP1A1* TT & *GSTM1* null genotypes and *CYP1A1* TC/CC & *GSTM1* wild genotypes combination showed no significant association with risk of HNC (p=0.376 and 0.724 respectively).

On the other hand, a significant increased risk of HNC was observed in individuals carrying combined genotypes of *CYP1A1* TC/CC & *GSTT1* null, when compared with combined wild genotypes of *CYP1A1* and *GSTT1* (OR=1.91; 95%CI=1.05-3.47; p=0.033). Similarly, *CYP1A1* TC/CC and *GSTT1* wild genotypes combination also showed an increased risk of HNC (OR=1.75; 95%CI=1.06-2.89 and p=0.027) (**Table 4.2.4**).

4.2.4.2 Combined effects of Smoking, Betel quid and Tobacco Chewing and *GSTs* Polymorphisms

The interaction of smoking, betel quid and tobacco chewing and the frequencies of *GSTM1* and *GSTT1* null genotypes in cases and controls were summarized in **Table 4.2.5**. We found that smokers with *GSTM1* null genotype have more than three-fold (95% CI=1.99-6.18 and p<0.001) increased risk of developing HNC compared to non-smokers with *GSTM1* wild genotype. Whereas, interaction of null genotype of *GSTT1* with smoking habit resulted in a significant increase in the risk to HNC (OR=3.05; 95% CI=1.74-5.35; p<0.001) when compared to non-smokers with *GSTT1* genotype. Among betel quid chewers, individuals carrying *GSTM1* null genotype had significantly higher risk of HNC than non-betel quid chewer having wild-type *GSTM1* genotype (OR=4.09, 95% CI=2.16-7.74 p<0.001). Similarly, betel quid chewers carrying *GSTT1* null genotype have 3.74 fold (95% CI=2.04-6.83 and p<0.001) increases risk of HNC when compared to the reference group. The interaction between tobacco chewing and null genotype of *GSTM1* resulted in 4.35 fold (95% CI=2.38-7.95 and p<0.001) increase in the risk for HNSCC when compared to non-tobacco chewers with wild- type *GSTM1* genotype. Similarly, interaction of *GSTT1* null genotype with tobacco chewing resulted in a significant increase in HNSCC risk (OR=3.68; 95% CI=2.06-6.58; p<0.001) when compared to the reference group (**Table 4.2.5**).

Table 4.2.5 Combine effects of Smoking, betel quid and tobacco chewing and *GSTs* genes polymorphisms

	***GSTM1* present**			***GSTM1* null**		
	Pa/Co*	OR (95% CI)	*p*-value	Pa/Co	OR (95% CI)	*p*-value
Smoking						
Non-smokers	24/76	1 (reference)		23/59	1.23 (0.64-2.39)	0.61
Smokers	51/100	1.61 (0.92-2.85)	0.12	72/65	3.51(1.99-6.18)	**<0.001**
Betel quid						
Non-chewers	17/57	1 (reference)		23/65	1.19 (0.58-2.43)	0.71
Chewers	58/119	1.63 (0.88-3.05)	0.133	72/59	4.09 (2.16-7.74)	**<0.001**
Tobacco						
Non-chewers	21/65	1 (reference)		22/72	0.95 (0.48-1.87)	1.00
Chewers	54/111	1.51 (0.84-2.71)	0.193	73/52	4.35 (2.38-7.95)	**<0.001**
	***GSTT1* present**			***GSTT1* null**		
	Pa/Co	OR (95% CI)	*p*-value	Pa/Co	OR (95% CI)	*p*-value
Smoking						
Non-smokers	32/104	1 (reference)		15/31	1.57 (0.76-3.25)	0.245
Smokers	77/116	2.16 (1.32- 3.52)	0.002	46/49	3.05 (1.74-5.35)	**<0.001**
Betel quid						
Non-chewers	28/79	1 (reference)		12/43	0.79 (0.37-1.69)	0.571
Chewers	81/141	1.62 (0.97-2.69)	0.080	49/37	3.74 (2.04-6.83)	**<0.001**
Tobacco						
Non-chewers	28/103	1 (reference)		15/34	1.62 (0.78-3.37)	0.239
Chewers	81/117	2.55 (1.54-4.21)	0.001	46/46	3.68 (2.06-6.58)	**<0.001**

***Pa** = Patients and **Co** = Controls

$p<0.05$ was considered as statistically significant.

4.2.4.3 Interaction of Smoking and Combined Genotypes of *CYP1A1* T3801C, *GSTM1* and *GSTT1*

In present study, we also evaluated the interaction of smoking and the genotypes combinations of *CYP1A1* T3801C and *GSTM1* or *GSTT1* in HNC risk (**Table 4.2.6**). We found that smokers carrying both *CYP1A1* TC/CC and *GSTM1* null genotypes had the highest increase risk of developing HNC compared to never smokers with *CYP1A1* TT + *GSTM1* (wild) genotype (OR=6.42; 95%CI=2.70 to 15.26 & p<0.001). About 3.86 fold (95%CI=1.66 to 8.96 & p=0.003) increased the risk of HNC found in those who were smokers and carrying *CYP1A1* TC/CC + *GSTT1* null genotype. Again, smokers having *CYP1A1* TC/CC + *GSTT1* wild genotype also showed significant increased risk of HNC compared to reference group (OR= 3.58; 95% CI= 1.75-7.34 & p=0.001).

Table 4.2.6 Interaction of smoking and combined genotypes of *CYP1A1* T3801C, *GSTM1* and *GSTT1*

Smoking						
	Non-smokers			**Smokers**		
	Cases/ Controls	**OR (95% CI)**	***p*-value**	**Cases/ Control**	**OR (95% CI)**	***p*-value**
***CYP1A1* T3801C + *GSTM1* (Wild/Null)**						
TT + Wild	14/37	1 (reference)		23/29	2.10 (0.93 to 4.74)	0.101
TC/ CC + Wild	17/41	1.10 (0.48 to 2.51)	1.000	26/34	2.02 (0.92 to 4.46)	0.112
TT+ Null	19/33	1.52 (0.67 to 3.48)	0.400	24/31	2.05 (0.91 to 4.58)	0.106
TC/ CC + Null	23/21	2.89 (1.24 to 6.73)	0.020	34/14	**6.42** (2.70 to 15.26)	**<0.001**
***CYP1A1* T3801C + *GSTT1* (Wild/Null)**						
TT + Wild	17/50	1 (reference)		33/40	2.43 (1.19 to 4.95)	0.021
TC/ CC + Wild	25/40	1.84 (0.88 to 3.84)	0.135	39/32	**3.58** (1.75 to 7.34)	**0.001**
TT+ Null	16/20	2.35 (1.01 to 5.49)	0.075	14/20	2.06 (0.87 to 4.90)	0.116
TC/ CC + Null	15/22	2.01 (0.86 to 4.68)	0.125	21/16	**3.86** (1.66 to 8.96)	**0.003**

4.2.4.4 Interaction of Tobacco-Betel Quid and Combined Genotypes of *CYP1A1* T3801C, *GSTM1* and *GSTT1*

The interaction of tobacco-betel quid chewing and *CYP1A1* T3801C and *GSTM1* or *GSTT1* gene polymorphisms in HNC risk is summarized in **Table 4.2.7**. The results showed that tobacco-betel quid chewers carrying both *CYP1A1* TC/CC and *GSTM1* null genotypes had higher risk of developing HNC compared to never chewers with *CYP1A1* TT + *GSTM1* wild genotypes (OR= 9.31; 95%=3.33-25.98 and $p<0.001$) increase risk. Again, increased the risk of HNC was about 3.98 fold ($p=0.005$) in those who were chewers and carrying *CYP1A1* TT + *GSTM1* null genotypes. Individuals who were tobacco-betel quid chewers and carrying *CYP1A1* TT + *GSTT1* null genotype and *CYP1A1* TT/CC + *GSTT1* null genotypes also showed 4.77 fold (95% CI= 1.95-11.65; $p=0.001$) and 3.82 (95% CI= 1.57-9.30; $p=0.005$) increased risk of HNC respectively.

Table 4.2.7 Interaction of tobacco-betel quid chewing and combined genotypes of *CYP1A1* T3801C, *GSTM1* and *GSTT1*

	Tobacco-betel quid					
	Non-chewers			**Chewers**		
	Case/ Control	**OR (95% CI)**	***p*-value**	**Case/ Control**	**OR (95% CI)**	***p*-value**
***CYP1A1* T3801C + *GSTM1* (Wild/Null)**						
TT + Wild	7/24	1 (reference)		30/42	2.45 (0.95 to 6.33)	0.076
TC/ CC + Wild	16/29	1.89 (0.68 to 5.27)	0.311	27/46	2.01 (0.78 to 5.22)	0.176
TT+ Null	7/33	0.73 (0.23 to 2.31)	0.765	36/31	**3.98** (1.53 to 10.35)	**0.005**
TC/ CC + Null	19/21	3.10 (1.11 to 8.69)	0.046	38/14	**9.31** (3.33 to 25.98)	**<0.001**
***CYP1A1* T3801C + *GSTT1* (Wild/Null)**						
TT + Wild	11/35	1 (reference)		39/55	2.26 (1.03 to 4.94)	0.060
TC/ CC + Wild	23/32	2.29 (0.97 to 5.38)	0.090	41/40	3.26 (1.47 to 7.24)	0.005
TT+ Null	3/22	0.43 (0.11 to 1.69)	0.351	27/18	**4.77** (1.95 to 11.65)	**0.001**
TC/ CC + Null	12/18	3.89 (1.57 to 9.66)	0.004	24/20	**3.82** (1.57 to 9.30)	**0.005**

4.2.4.5 Multifactor Dimensionality Reduction (MDR)

Interactions between genetic and environmental factors for HNC risk were further evaluated by using MDR analysis and the best models to detect high-order interactions were summarized in **Table 4.2.8**. The MDR results showed that tobacco-betel quid chewing was the best one-factor model with TBA=0.55, CVC of 9/10 and $p<0.0003$. The combination of tobacco-betel quid chewing and *GSTM1* null was predicted as two factors model with 100% CVC (TBA=0.6118 and $P<0.0001$). However, among all the models, the four-factors model consists of smoking, tobacco-betel quid chewing, *CYP1A1* TC/CC and *GSTM1* null genotypes with CVC of 9/10 and highest TBA of 0.6292 and $P<0.0001$ was considered as the best model predicted for head and neck cancer.

4.2.4.6 Interaction Entropy Graphs

Using MDR results, we constructed interaction entropy graphs for HNC risk to determine synergistic or not-synergistic interactions (**Figure 4.2.5** and **4.2.6**). In interaction entropy graph, tobacco-betel quid chewing showed the highest independent effect (2.32%) and had synergistic interaction with *GSTM1* null genotype (1.45%) by removing 0.56% of entropy. On the other hand, *GSTM1* null genotype and *CYP1A1* TC/CC showed highest synergistic interaction removing 0.87% of entropy. Smoking (1.48%) and *CYP1A1* TC/CC (0.67%) might also explain considerable entropy independently.

4.2.4.7 False Positive Report Possibility (FPRP)

To strengthen our results we performed FPRP by testing the robustness and consistency of the gene-gene and gene-environment interaction obtained from both LR and MDR analysis. The FPRP values for all statistically significant result indicated that the interaction between smokers and tobacco-betel quid chewers with *CYP1A1* (TT/CC) + *GSTM1* null genotypes, and higher order predictor models obtained from multiple comparisons analysis showed excellent reliability even when assuming very low prior probabilities (upto 0.1 to 0.0001) when detecting ORs of 1.5 for an FPRP value of 0.5 (**Table 4.2.9**). The relatively greater FPRP values with very low prior probability assumptions (0.0001) might be attributed to the relatively small sample size of this study as well as moderate effects of selected SNP.

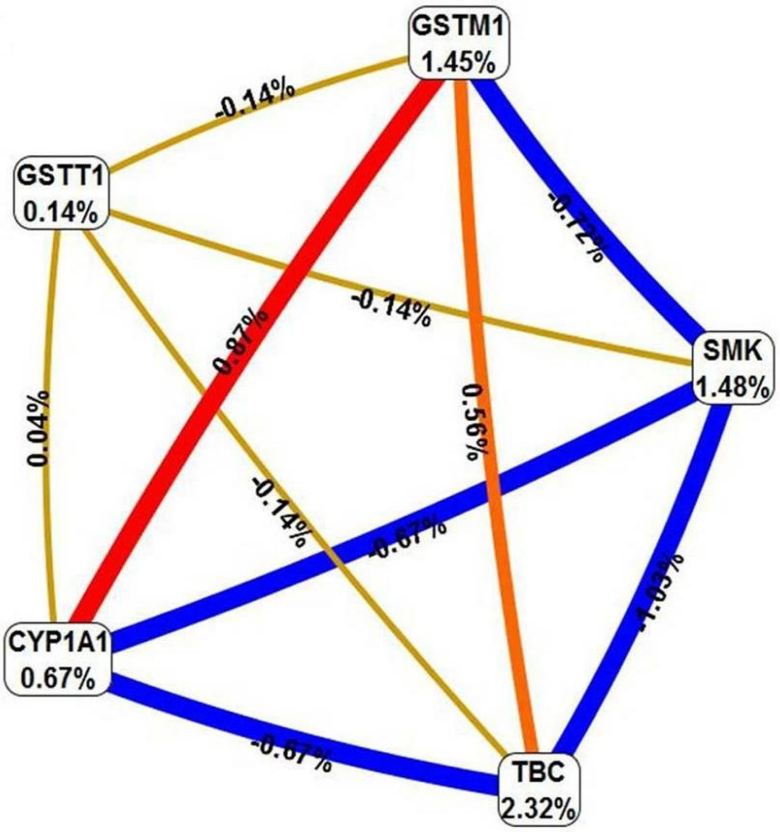

Figure 4.2.5: MDR analysis for high-order interaction. Interaction entropy graph of gene-environment interaction for head and neck cancer prediction. This graphical model explains the percent of the entropy in case-control removed by each factor (independent effect) and by each pair-wise combination of attributes (interaction effect). Positive percentage of entropy indicating synergistic interaction and negative values of entropy represent redundancy. The red colour indicating a high degree of synergistic interaction, orange a lesser degree whereas; gold represent midpoint; blue represents the highest level of redundancy followed by green. **TBC**-tobacco-betel quid chewing and **SMK**-smoking. The diagram indicate that, tobacco-betel quid chewing showed highest independent effect (2.32%), whereas interaction between *GSTM1* (null) and tobacco-betel quid chewing, and *CYP1A1* (variant) showed synergistic interactions towards HNC risk.

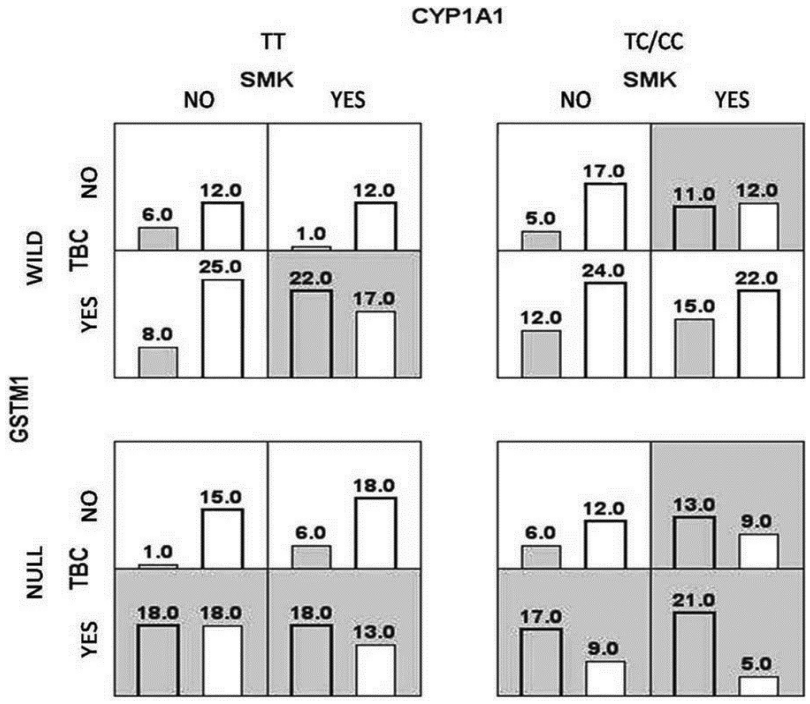

Figure 4.2.6: Summary of the four-factor model (*CYP1A1*, *GSTM1*, tobacco-betel quid chewing, and smoking) in MDR analysis. The distribution of high risk (dark shading) and low risk (light shading) combinations associated with HNC risk. The percentage of patients having HNC was represented by left column in each box, whereas right column in each box indicated percentage of controls.

Table 4.2.8 Summary of MDR analysis for HNC risk prediction

No. of Locus	Best model	TBA	TrBA	CVC	*p*-value
1st order	TBC	0.5500	0.5877	8/10	0.0003
2nd order	TBC, GSTM1	0.6118	0.6118	10/10	<0.0001
3rd order	SMK, TBC, GSTM1	0.5847	0.6359	6/10	<0.0001
4th order**	SMK, TBC, CYP1A1, GSTM1	0.6292	0.6620	9/10	<0.0001

CVC = cross-validation consistency; TBC = tobacco-betel quid chewing; SMK= smoking; TBA= testing balance accuracy; TrBA = training balance accuracy.

**Best model predicted for HNC risk with highest test balance and training balance accuracy and maximum CVC

Table 4.2.9 False Positive Reports Probability (FPRP) for odd ratios of the Logistic Regression (LR) and Multifactor Dimensionality Reduction (MDR) analysis

	Odds ratio	**OR= 1.5 (Prior Probability)**				
	OR (95% CI) ***P*-value**	**0.25**	**0.1**	**0.01**	**0.001**	**0.0001**
LR analysis						
SMK+*CYP1A1*(TT/CC)+ *GSTM1* (null)	6.42 (2.70-15.26) <0.001	**0.134**	**0.316**	0.836	0.981	0.998
SMK+*CYP1A1*(TT/CC)+ *GSTM1* (null)	3.86 (1.66-8.96) 0.003	**0.265**	0.519	0.922	0.992	0.999
TBC+*CYP1A1*(TT/CC)+ *GSTM1* (null)	9.31 (3.33-25.98) <0.001	**0.200**	**0.428**	0.892	0.988	0.999
TBC+*CYP1A1*(TT/CC)+ *GSTM1* (null)	3.82 (1.57-9.30) 0.005	**0.324**	0.590	0.941	0.994	0.999
MDR analysis		OR= 1.5 (Prior Probability)				
TBC	2.15 (1.41-3.26) 0.0003	**0.020**	**0.059**	**0.408**	0.874	0.986
TBC, *GSTM1*	3.02 (1.94-4.69) <0.0001	**0.003**	**0.008**	**0.085**	**0.484**	0.904
SMK, TBC, *GSTM1*	2.98 (2.01-4.46) <0.0001	**0.001**	**0.002**	**0.025**	**0.207**	0.724
SMK, TBC, *CYP1A1*, *GSTM1*	3.78 (2.51-5.69) <0.0001	**<0.001**	**<0.001**	**0.004**	**0.038**	**0.282**

Prior probability range = 0.25 – 0.00001 to detect OR = 1.5; α level = observed *p*-value;

Bold values represent noteworthy association at 0.5 FPRP

Chapter 4.3:

Genetic Polymorphisms of DNA Repair Genes (*XRCC1* and *XRCC2*) and Risk of HNSCC

Genotyping of DNA repair genes (*XRCC1* and *XRCC2*) were carried out on 110 HNSCC patients and 240 controls to evaluate the polymorphisms and the association with HNSCC risk. For tobacco habits, the study population was further divided into never, light and heavy smokers or chewers based on their frequency of consumption (number of doses/day). Heavy smokers were those who smoked at least 20 cigarettes or bidis per day and light smokers were those who smoked <20 cigarettes or bidis per day. Whereas, heavy chewers were defined as those who chewed more than 10 doses of tobacco or betel-quid/day and those who chewed ≤10 doses of tobacco or betel quid/day were classified as light chewers.

The allele types were determined by three distinct banding patterns such as: two fragments of 375 and 240 bp for the wild-type allele GG (Arg/Arg), three fragments of 615, 375 and 240 bp indicate heterozygous GA (Arg/Gln) and a single 615bp fragment for the variant allele AA (Gln/Gln) for *XRCC1* gene at Arg399Gln (**Figure 4.3.1**). Furthermore, to confirm the genotypes of *XRCC1*, PCR amplified products of *XRCC1* were sequenced (**Figure 4.3.2**).

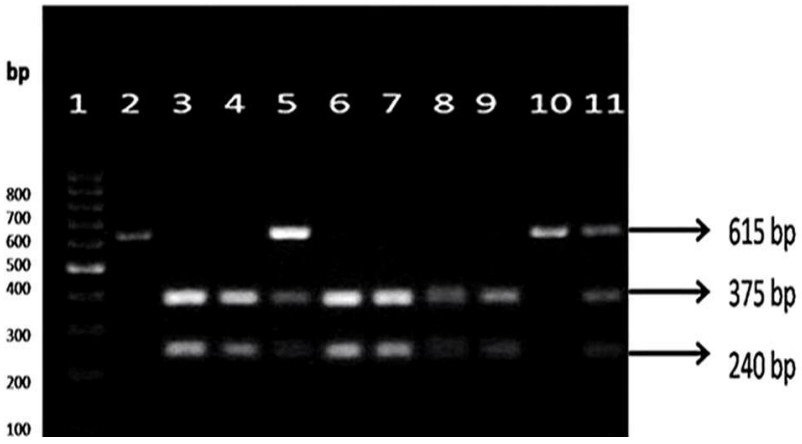

Figure 4.3.1: Ethidium bromide stained gel showing *XRCC1* Arg399Gln DNA repair gene polymorphism. 100 bp ladder (lane 1); two band denotes *XRCC1* Arg/Arg (GG) wild genotype (lanes 3,4,6,7,8 and 9); three band indicates *XRCC1* Arg/Gln (GA) heterozygous genotype (lanes 5 and 11) and single band denotes *XRCC1* Gln/Gln (AA) homozygous variant genotype (lanes 2 and 10).

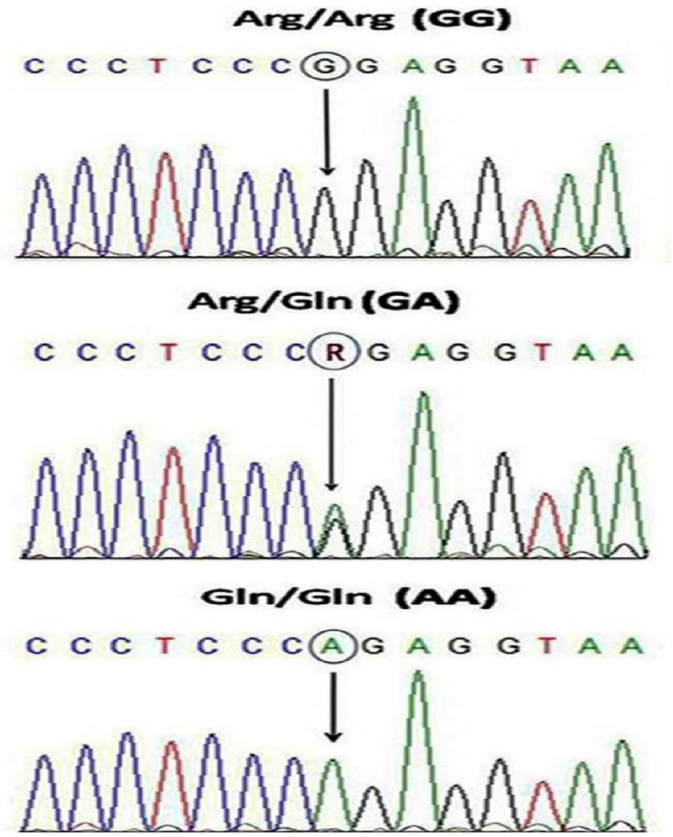

Figure 4.3.2: DNA sequencing results of *XRCC1* (Agr399Gln) gene. Diagram showing nucleotide variation (in black circle); black peak denotes guanine (G), **R** indicates heterozygous genotypes (G/A), whereas green peak denotes adenine (A).

Similarly, in case of *XRCC2* Arg188His, a single 205 bp fragment represents the wild-type allele GG (Arg/Arg), three fragments of 205, 137 and 68 bp denotes the heterozygous GA (Arg/His) and two fragments of 137 and 68bp represent the variant allele AA (His/His) of *XRCC2* at Arg188His (**Figure 4.3.3a**). However, we did not found any variant allele AA (His/His) of *XRCC2* in our study population. In addition, to confirm the genotype of *XRCC2* Arg188Gln, we sequenced the PCR products of *XRCC2*, which further validated the results of PCR-RFLP (**Figure 4.3.3b**).

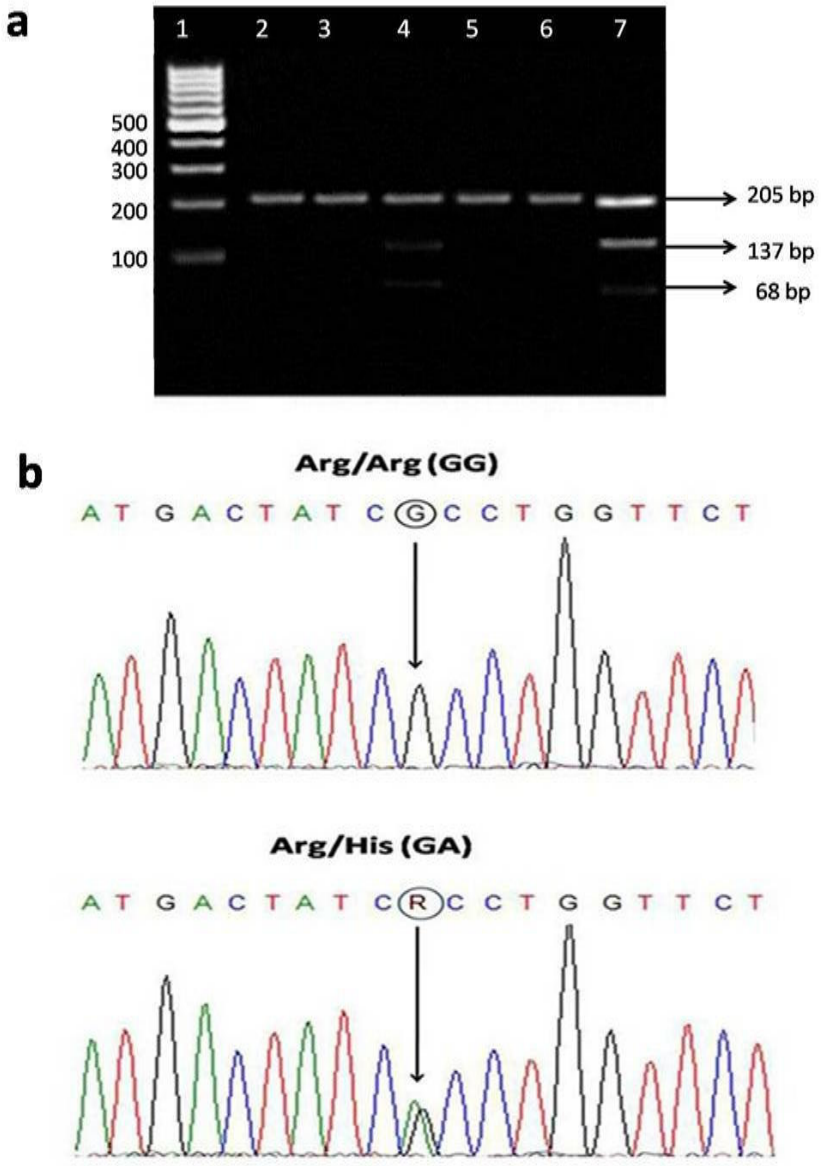

Figure 4.3.3: Polymorphism *XRCC2* Arg188His DNA repair genes.

(a) Ethidium bromide stained gel showing *XRCC2* Arg188His polymorphism: 100bp DNA ladder (lane 1); single band indicates *XRCC2* Arg/Arg(GG) genotype (lanes 2, 3, 5 and 6) and three band denotes *XRCC2* Arg/His(GA) genotype (lanes 4 and 7).

(b) DNA sequencing results showing nucleotide changes (mark by black circle). Nucleotide change marked as **R** indicates the heterozygous genotype, where green peak denotes adenine (A) while black peak denotes guanine (G).

4.3.1 Distribution of the Genotype Frequencies of *XRCC1* Arg399Gln and *XRCC2* Arg188His

The genotype frequencies of *XRCC1* (Arg399Gln) for GG (Arg/Arg), GA (Arg/Gln) and AA (Gln/Gln) were 32.7%, 50% and 17.3% in patients, whereas in controls, GG (Arg/Arg), GA (Arg/Gln) and AA (Gln/Gln) were 44.3%, 45% and 10.7% respectively. The genotypes frequencies of *XRCC2* (Arg188His) among patients were 67.3% and 32.7% for GG (Arg/Arg) and GA (Arg/His) genotype respectively, but in controls, GG (Arg/Arg) and GA (Arg/His) genotypes frequencies were 86.4% and 13.6% respectively (**Table 4.3.1**).

Table 4.3.1 Genotype frequency distribution of *XRCC1* Arg399Gln and *XRCC2* Arg188His polymorphisms and risk of HNSCC

Genotype	**Cases n=110 (%)**	**Controls n=140 (%)**	**OR* (95% CI)**	***p*-value**
***XRCC1* Arg399Gln** (G>A; exon 10; rs25487)				
GG (Arg/Arg)	36 (32.7)	62 (44.3)	1.00 (reference)	
GA (Arg/Gln)	55 (50)	63 (45)	1.53 (.883-2.66)	0.129
AA (Gln/Gln)	19 (17.3)	15 (10.7)	2.43 (1.085-5.45)	0.031
GA+AA (Arg/Gln+Gln/Gln)	74 (67.3)	78 (55.7)	1.68 (1.01-2.87)	0.048
G-allele	127 (57.7)	187 (66.8)	1.00 (reference)	
A-allele	93 (42.3)	93 (33.2)	1.47 (1.02 to 2.12)	0.041
***XRCC2* Arg188His** (G>A; exon 3; rs3218536)				
GG (Arg/Arg)	74 (67.3)	121 (86.4)	1.00 (reference)	
GA (Arg/His)	36 (32.7)	19 (13.6)	3.29 (1.74-6.21)	<0.01
G-allele	184 (83.6)	261 (93.2)	1.00 (reference)	
A-allele	36 (16.4)	19 (6.8)	2.69 (1.50 to 4.81)	<0.001

*Odd ratio (OR) adjusted for age, gender, smoking and tobacco-betel quid chewing
p <0.05 considered as statistically significance

4.3.2 Association between *XRCC1* Arg399Gln Polymorphisms and HNSCC

Using the logistic regression model, the association between *XRCC1* Arg399Gln polymorphism and HNSCC was analyzed. The analysis revealed that variant homozygote AA (Gln/Gln) genotype of the *XRCC1* Arg399Gln was associated with increased the risk of HNSCC that was statistically significant (OR= 2.43; 95% CI=1.08-5.45; p=0.031). In case of heterozygote GA (Arg/Gln) no statistically significant association with risk of HNSCC was found (OR=1.53; 95% CI=0.88-2.66; p=0.129). A significant increased risk of HNSCC was observed in case of variant-containing genotypes GA+AA (Arg/Gln+Gln/Gln) (OR=1.68; 95% CI= 1.01-2.87 and p=0.048) when compared with homozygous wild-type GG genotype. In addition, variant 'allele A' carrier genotypes were found to be associated with 1.47 folds risk of cancer (p= 0.041) when compared with 'G allele' (**Table 4.3.1**).

4.3.3 Association between *XRCC2* Arg188His Polymorphisms and HNSCC

The association between *XRCC2* Arg188His polymorphism and HNSCC was analysed (**Table 4.3.1**). The analysis shown that heterozygote genotype GA (Arg/His) of the *XRCC2* Arg188His polymorphism was associated with increased risk of HNSCC (OR=3.29; 95% CI=1.74-6.21 and p<0.001) and the association is highly significant. Furthermore, individuals carrying variant 'allele A' were found to be associated with 2.69 folds increased the risk of HNSCC (95%CI=1.50-4.81 and p<0.001) when compared with 'G allele'. It was notable that we did not find any individual carrying the variant homozygous AA (His/His) in case of *XRCC2* Arg188His polymorphism.

4.3.4 Gene-Gene Interactions of *XRCC1* and *XRCC2*

In our present study, we also investigated the gene-gene interaction of *XRCC1* Arg399Gln and *XRCC2* Arg188His and their association with HNSCC using the combined low-risk genotypes (*XRCC1* 399GG and *XRCC2* 188GG) as a reference group (**Table 4.3.2**). Individual carrying both *XRCC1* GA (Arg/Gln) and *XRCC2* GA (Arg/His) genotypes, showed the highest risk of HNSCC (OR=5.15;

95%CI=1.86-14.27; *p*=0.001). Similarly, *XRCC1* AA (Gln/Gln) and *XRCC2* GA (Arg/His) genotypes combination also showed an increased risk of HNSCC (OR=3.59 and *p*=0.021). On the other hand, genotypes combination of *XRCC1* GG(Arg/Arg)-*XRCC2* GA(Arg/His), *XRCC1* GA(Arg/Gln)-*XRCC2* GG(Arg/Arg) and *XRCC1* AA(Gln/Gln)-*XRCC2* GG(Arg/Arg) showed no significant association with risk of HNSCC (*p*=0.219, 00.442 and 0.556 respectively).

Table 4.3.2 Combined genotype analysis of *XRCC1* Arg399Gln and *XRCC2* Arg188His on risk of HNSCC

XRCC1 **Arg399Gln**	***XRCC2*** **Arg188His**	**Cases n=110**	**Controls n=140**	**OR (95% CI)**	***p*-value***
GG (Arg/Arg)	GG (Arg/Arg)	29	56	1.00 (reference)	
GA (Arg/Gln)	GG (Arg/Arg)	39	57	1.32 (0.72 to 2.41)	0.442
AA (Gln/Gln)	GG (Arg/Arg)	6	8	1.45 (0.48 to 4.41)	0.556
GG (Arg/Arg)	GA (Arg/His)	7	6	2.25 (0.72 to 7.03)	0.219
GA (Arg/Gln)	GA (Arg/His)	16	6	**5.15** (1.86 to 14.27)	**0.001**
AA (Gln/Gln)	GA (Arg/His)	13	7	**3.59** (1.32 to 9.75)	**0.021**

**p* <0.05 considered as statistically significance

4.3.5 Interaction of Tobacco with *XRCC1* and *XRCC2* Gene Polymorphisms

We evaluated the combined effect of tobacco smoking, tobacco-betel quid chewing and the genotypes of *XRCC1* and *XRCC2* in HNSCC risk (**Table 4.3.3** and **4.3.4**). The results suggest that heavy smokers carrying *XRCC1* GA (Arg/Gln) genotype were 2.79 fold (*p*=0.042) increase risk of developing HNSCC compared to never smokers with wild type *XRCC1* GG (Arg/Arg) genotype (**Figure 4.3.4**). Similarly, in case of *XRCC1* AA (Gln/Gln) genotype the risk increases 6.46 fold (*p*=0.017) in heavy smokers.

Among heavy tobacco-betel quid chewers, persons having *XRCC1* AA (Gln/Gln) genotype had significantly higher risk of developing HNSCC than non tobacco-betel quid chewer carrying *XRCC1* GG (Arg/Arg) genotype (OR=4.27, p=0.023). Whereas, individuals with *XRCC1* GA genotype and consumed heavy tobacco-betel quid had fold 2.04 (p=0.136) increased risk of HNSCC as compared to reference group. (**Table 4.3.3** and **Figure 4.3.4a**).

Table 4.3.3 Interaction of tobacco and *XRCC1* Arg399Gln polymorphisms

	***XRCC1* Arg399Gln polymorphism**								
	GG (Arg/Arg)			**GA (Arg/Gln)**			**AA (Gln/Gln)**		
	Pa/Co*	OR (95%CI)	*P*-value **	Pa/Co	OR (95% CI)	*P*-value	Pa/Co	OR (95% CI)	*P*-value
Smokers									
Never	13/28	1 (**Ref**)		20/29	1.49 (0.63-3.51)	0.390	5/9	1.20 (0.35-4.13)	1.000
Light	13/20	1.40 (0.54-3.61)	0.625	13/17	1.65 (0.63-4.31)	0.332	5/3	3.59 (0.81-15.9)	0.124
Heavy	10/14	1.54 (0.55-4.29)	0.435	22/17	**2.79** (1.13-6.87)	**0.042**	9/3	**6.46** (1.58-26.46)	**0.017**
Tobacco-betel quid chewers									
Never	15/32	1 (**Ref**)		15/23	1.39 (0.58-3.37)	0.501	2/4	1.07 (0.20-5.63)	1
Light	10/17	1.25 (0.47-3.34)	0.799	18/17	2.26 (0.93-5.51)	0.111	5/5	2.13 (0.57-8.01)	0.298
Heavy	11/13	1.81 (0.67-4.87)	0.302	22/23	2.04 (0.88-4.72)	0.136	12/6	**4.27** (1.38-13.20)	**0.023**

*Pa = Patients and Co=Controls; **Ref**-reference

**Fisher's exact test used to calculate p-value

$p < 0.05$ considered as statistically significance

Table 4.3.4 Interaction of tobacco and *XRCC2* Arg188His polymorphisms

	***XRCC2* Arg188His polymorphism**					
	GG (Arg/Arg)			**GA (Arg/His)**		
	Pa/Co*	OR (95% CI)	*p*-value**	Pa/Co	OR (95% CI)	*p*-value
Smoking						
Never smokers	25/58	1 (Reference)		13/8	3.77 (1.42-10.02)	**0.011**
Light smokers	18/32	1.31 (0.62-2.73)	0.567	13/8	3.77 (1.42-10.02)	**0.011**
Heavy smokers	31/31	2.32 (1.18-4.57)	**0.017**	10/3	**7.73** (2.06 to 29.07)	**0.003**
Tobacco-betel quid chewing						
Never chewers	21/50	1 (Reference)		11/9	2.91 (1.07-7.88)	0.061
Light chewers	23/33	1.66 (0.80-3.45)	0.193	10/6	3.97 (1.32-11.96)	**0.020**
Heavy chewers	30/38	1.88 (0.94-3.76)	0.082	15/4	8.93 (2.72-29.26)	**0.010**
Heavy chewers	30/38	1.88 (0.94-3.76)	0.082	15/4	**8.93** (2.72-29.26)	**0.010**

*Pa =Patients and Co=Controls

**Fisher's exact test used to calculate *p*-value and $p < 0.05$ considered as statistically significance

A highly statistically significant increased risk of HNSCC was observed in individuals carrying *XRCC2* GA (Arg/His) and with a heavy smoking habit (OR=7.73 and *p*=0.003), when we compared the data of never smoker individuals with *XRCC2* GG (Arg/Arg) genotype (**Table 4.3.4**). About 8.93 fold (*p*=0.01) increased the risk of HNSCC present in those who were heavy tobacco-betel quid chewers and also carrying *XRCC2* GA (Arg/His). Light smokers and tobacco-betel quid chewers carrying *XRCC2* GA (Arg/His) genotype also showed significant increased risk of HNSCC (*p*=0.011 and 0.020 respectively) (**Figure 4.3.4a**).

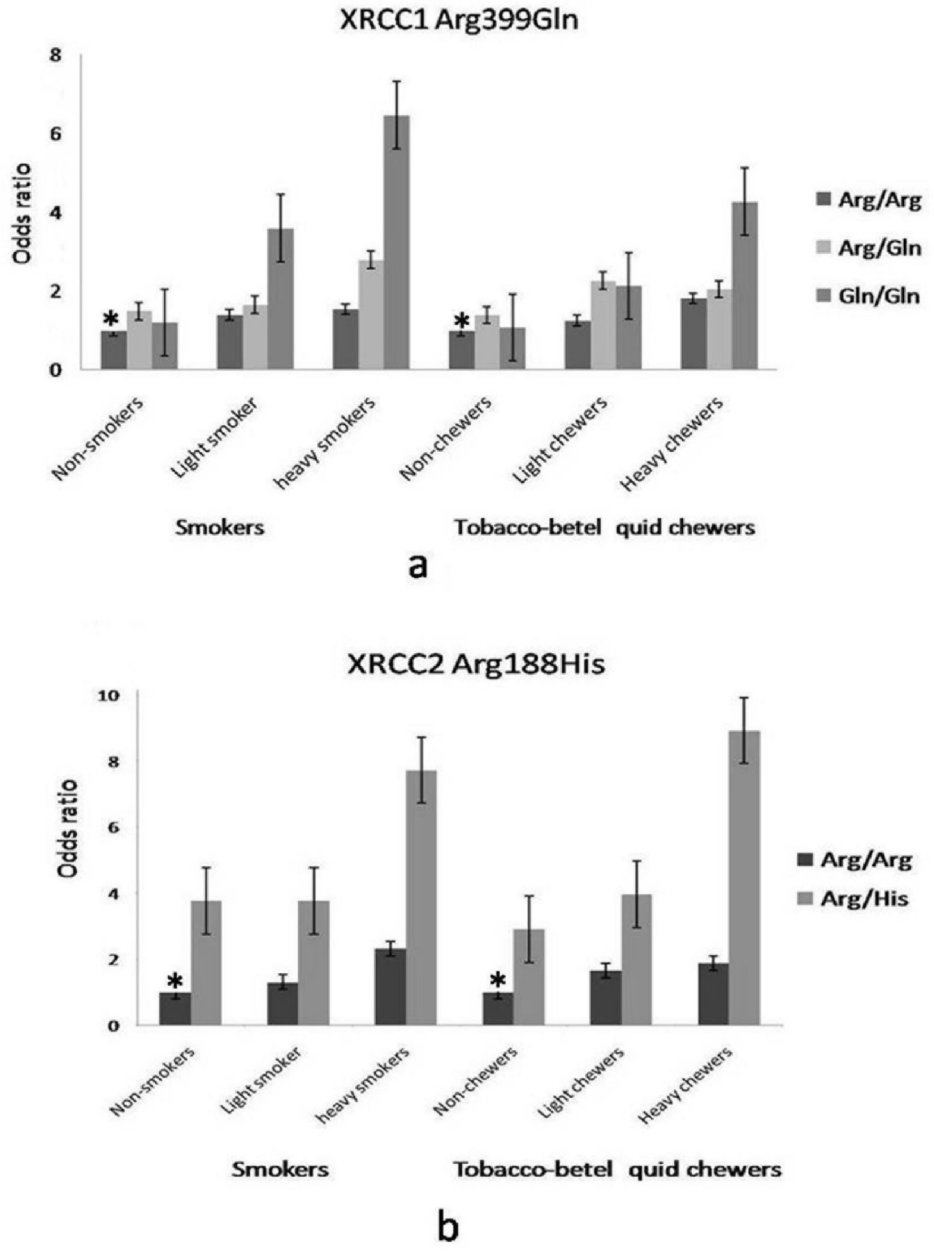

Figure 4.3.4: Combined effect of tobacco habits with *XRCC1* Arg399Gln and *XRCC2* Arg188His genotypes.

(a) Heavy smokers and tobacco-betel quid chewers carrying *XRCC1* AA (Gln/Gln) genotype were having 6.46 and 4.27 fold increase risk of developing HNSCC respectively.

(b) A highly significant risk of HNSCC was observed in individuals carrying *XRCC2* GA (Arg/His) genotype with heavy smoking (OR=7.73) or tobacco-betel quid chewing habit (OR=8.93).

Chapter 4.4: Epigenetic Alterations in Head and Neck Cancer

To understand the epigenetic alterations in head and neck cancer, we have analyzed the aberrant promoter methylation profile of HNSCC patients using seven important tumours-related pathway genes, including *DAPK*, *RASSF1* (apoptosis pathway), *BRCA1*, *MLH1* (DNA repair pathway), *p16* (cell-cycle pathway), *ECAD* (cell-cell adhesion), *GSTP1* (xenobiotic pathway) and three methylated loci (*MINT1*, *MINT2* and *MINT31*). We have evaluated promoter hypermethylation in 116 tissue samples (71 HNSCC tumours and 45 normal tissues) from the Northeast Indian population using methylation specific PCR (MSP) (**Figure 4.4.1**). The clinicopathological data of the 71 studied HNSCC tumour samples were summarized in **Table 4.4.1**. Of the 71 tumours samples; 38 (53.5%) oral cancer tissues (cheek, bottom of the tongue, tongue, gingivam, and buccal mucosa), 16 (22.6%) tissue samples of laryngeal cancer, 8 (11.2%) pharyngeal cancer tissues and 9 (12.7%) tissues of other cancer types in head and neck region. According to TMN classification, the majority of patients had advanced stage (III/IV) (67.3%).

Table 4.4.1 Clinicopathological data of HNSCC tumour samples

Parameters	**Cases (N=71)**	**Percentage (%)**
Tumours site		
Oral	38 (71)	53.5
Base of tongue	8 (71)	11.2
Tongue	4 (71)	5.6
Cheek	16 (71)	22.6
Gingivam	4 (71)	5.6
Buccal mucosa	6 (71)	8.4
Laryngeal	16 (71)	22.6
Pharyngeal	8 (71)	11.2
other	9 (71)	12.7
Stage at diagnosis		
Local (I/II)	17 (52)	32.7
Advanced (III/IV)	35 (52)	67.3
NA*	19 (71)	

***NA** = not available

4.4.1 Frequencies of Promoter Hypermethylation in Tumour and Normal Tissues

Promoter hypermethylation status of the *p16, DAPK, GSTP1, RASSF1, BRCA1, ECAD* and *MLH1* genes of 71 HNSCC tumour and 45 normal tissues samples was depicted in **Table 4.4.2**. Tumour tissues had much higher promoter hypermethylation frequency compared to normal tissue samples (32.4% vs. 13.3% for *p16* gene, 29.6% vs. 11.1% for *DAPK*, 18.3% vs. 8.9% for *BARC1*, 31% vs. 15.6% for *GSTP1*, 32.4% vs. 8.9% for *ECAD*, 50.7% vs. 22.2% for *RASSF1*, 5.6% vs. 2.2% for *MLH1*). However, significantly high level of hypermethylation were observed for *p16*, *DAPK*, *ECAD* and *RASSF1* genes (p=0.02, 0.02, 0.04 and 0.02 respectively), when HNSCC tissue was compared to normal tissue samples. The hypermethylation frequency of three MINT loci *MINT1*, *MINT2* and *MINT31* were 43.7%, 52.1% and 46.5% respectively in tumour tissue, whereas in normal tissues the hypermethylation frequency of *MINT1*, *MINT2* and *MINT31* were 13.3%, 11.1% and 17.8% respectively (**Table 4.4.3).** We found that *MINT1*, *MINT2* and *MINT31* promoter hypermethylation was significantly higher in tumour tissues compared to their normal counterpart (0.01, <0.01, and 0.01 respectively). Some of the genes or loci such as *RASSF1* and *MINT2* showed comparatively higher methylation frequency (more than 50%) in HNSCC tumour tissues. Whereas, *MLH1* gene showed lowest methylation frequency in tumour compared to normal tissue.

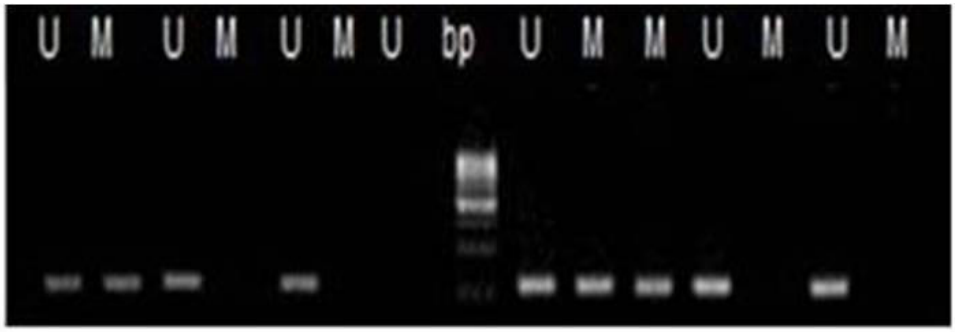

Figure 4.4.1: Agarose gel electrophoresis of MSP products. U=Unmethylated, M=Methylated

Table 4.4.2 Frequency of methylation of tumour suppressor genes in tumour and normal tissues

Tumour suppressor genes	**Frequency of methylation (%)**		
	Tumour tissue	**Normal tissue**	***p*-value**
p16			
Unmethylated	48 (67.6)	39 (86.7)	
Methylated	23 (32.4)	6 (13.3)	0.02
DAPK			
Unmethylated	50 (70.4)	40 (88.9)	
Methylated	21 (29.6)	5 (11.1)	0.02
BRCA1			
Unmethylated	58 (81.7)	41 (91.1)	
Methylated	13 (18.3)	4 (8.9)	0.16
GSTP1			
Unmethylated	49 (69)	38 (84.4)	
Methylated	22 (31)	7 (15.6)	0.06
ECAD			
Unmethylated	48 (67.6)	41 (91.1)	
Methylated	23 (32.4)	4 (8.9)	0.04*
RASSF1			
Unmethylated	35 (49.3)	35 (77.8)	
Methylated	36 (50.7)	10 (22.2)	0.02*
MLH1			
Unmethylated	67 (94.4)	44 (97.8)	
Methylated	4 (5.6)	1 (2.2)	0.38

Wilcoxon rank-sum test was used to calculate *p*-value

*Bonferroni correction of significance was applied

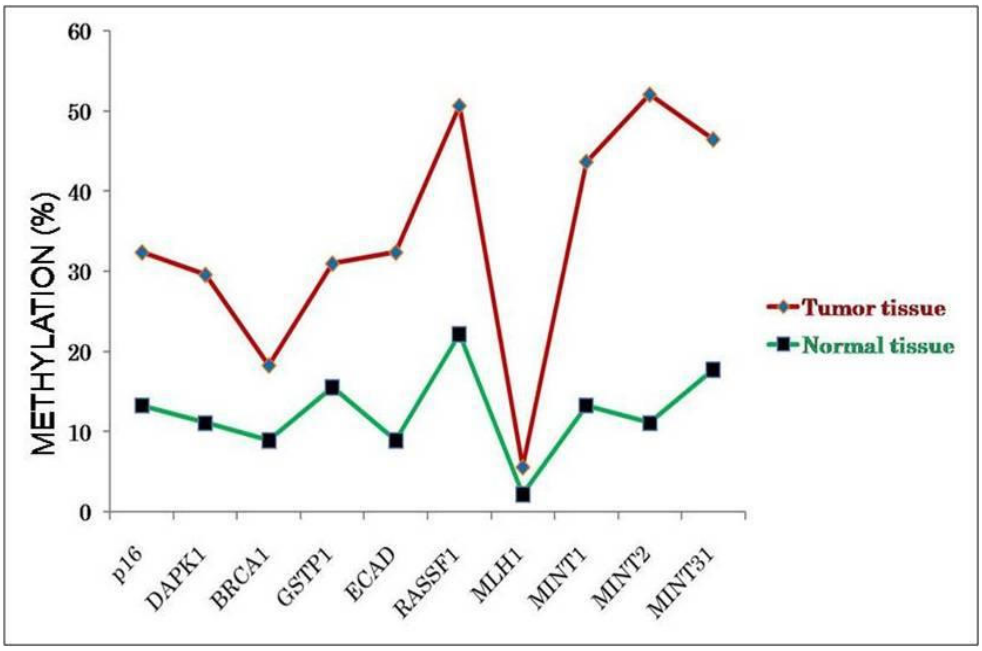

Figure 4.4.2: Frequencies of promoter hypermethylation in tumour and normal tissues. *P16, DAPK, ECAD, RASSF1* and *MINT1*, *MINT2* and *MINT31* genes/loci had significantly higher hypermethylation in tumours tissues compared to normal tissues

Table 4.4.3 Frequency of methylation of tumour-specific loci in tumour and normal tissues

	Frequency of methylation (%)		
Tumour–specific loci	**Tumour tissues**	**Normal tissues**	***p*-value**
MINT1			
Unmethylated	40 (56.3)	39 (86.7)	
Methylated	31 (43.7)	6 (13.3)	0.01*
MINT2			
Unmethylated	34 (47.9)	40 (88.9)	
Methylated	37 (52.1)	5 (11.1)	<0.01*
MINT31			
Unmethylated	38 (53.5)	37 (82.2)	
Methylated	33 (46.5)	8 (17.8)	0.01*

Wilcoxon rank-sum test was used to calculate *p*-value and *Bonferroni correction

4.4.2 CpG Island methylator Phenotype in HNSCC Tumour Tissues

In this study, CpG Island methylator Phenotype (CIMP) status was classified as CIMP-high (five or more methylated genes), CIMP-low (less than five methylated genes) and CIMP-negative (no methylated genes) based on the criteria previously used for CIMP status in several types of tumours (**Figure 4.4.3 and 4.4.4**). Among 71 HNSCC samples, 39.4% (28/71) were CIMP-high, 42.2% (30/71) were CIMP-low and 18.3% (13/71) were CIMP-negative. The methylation index (MI) (ratio of the number of methylated promoters and total number of promoters under study) ranged from 0 to 0.9 of the 71 patients. Out of the 71 HNSCC patients 14 (19.7%) had 0 (zero) MI, 33 (46.5%) had MI of 0.1-0.5 and 24 (33.8%) patients had MI of 0.6-0.9.

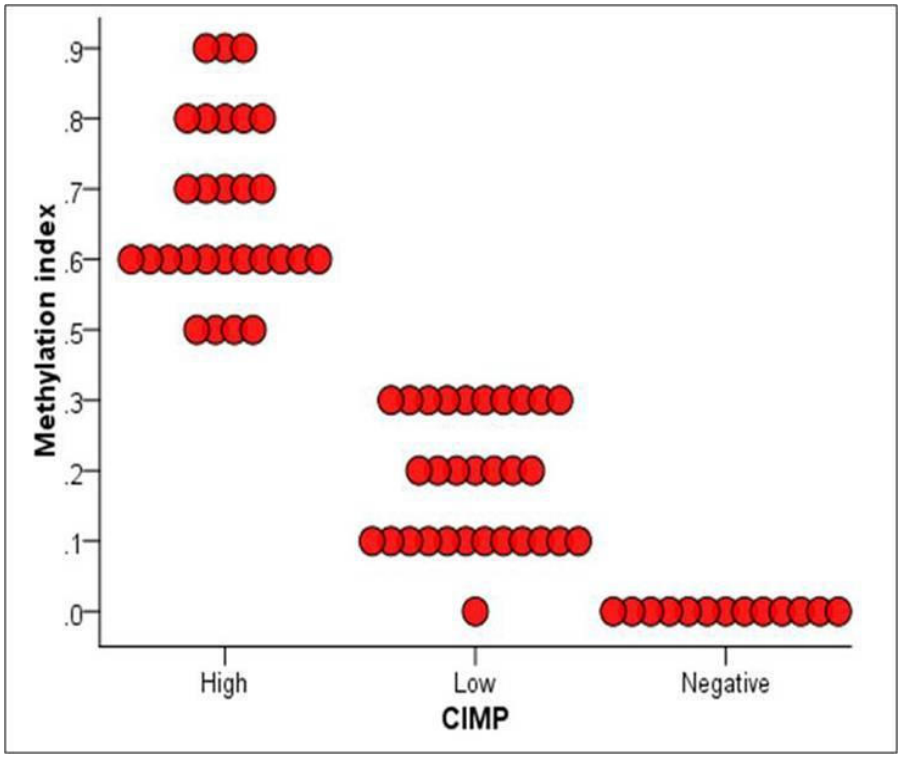

Figure 4.4.3: Methylation index (MI) in three CIMP-groups. Distribution of CIMP-high (0.5-0.9), CIMP-low (0.1-0.4) and CIMP-negative (zero) group of HNSCC tumours

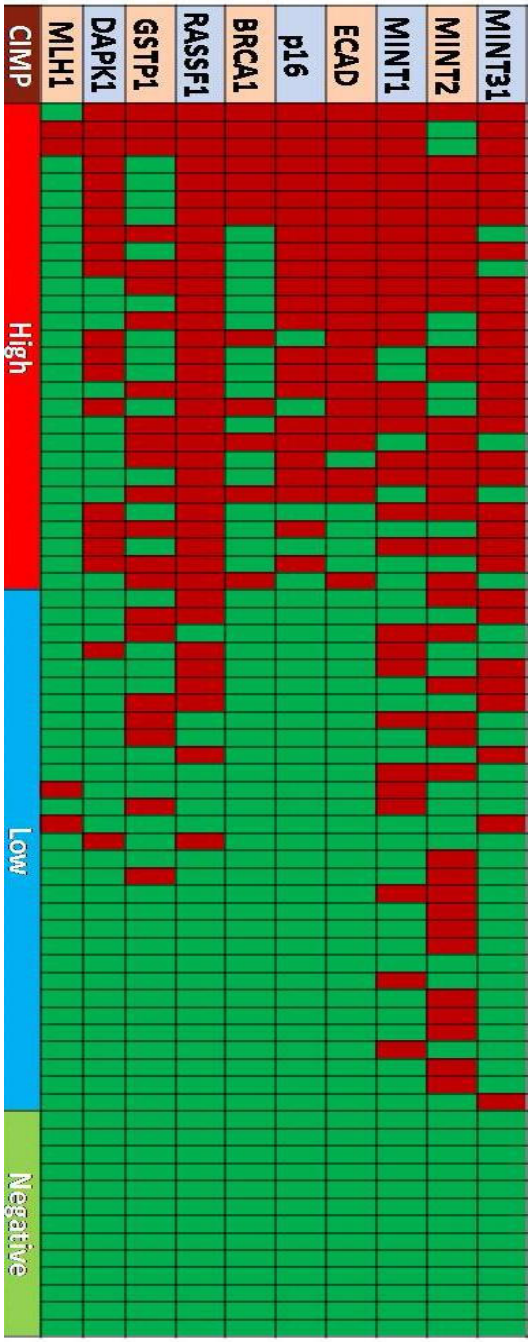

Figure 4.4.4: Promoter methylation status of 10 genes in the HNSCC tumour tissues. Each red rectangle represents methylated genes, while green rectangles represent unmethylated genes. Five or more methylated genes in a HNSCC tissue represent CIMP-high group, CIMP-low group denoted by less than five methylated genes) and CIMP-negative represent no methylated genes.

Chapter 4.5:

Correlation between HPV, Environmental and Genetic Factors with Promoter Methylation Profile of HNSCC patients

Epigenetic and genetic alteration plays a major role to the development of head and neck squamous cell carcinoma (HNSCC). Consumption of tobacco (smoking/chewing) and human papillomavirus (HPV) are also associated with an increase the risk of HNSCC. We correlated promoter methylation profile of patients with genetic (polymorphisms of *GSTM1*, *GSTT1*, *CYP1A1*, *XRCC1* and *XRCC2* genes) and environmental factors (smoking, betel quid and tobacco chewing) as well as with HPV and survival status. We also performed hierarchical cluster analysis to identify distinct subsets of HNSCC based on the promoter methylation profile. Furthermore, we investigated the outcome of patients, based on CpG island methylator phenotype (CIMP) and cluster (promoter methylation profiles) status.

4.5.1 Promoter Methylation Status in HPV-positive and HPV-negative HNSCC

In the study, HPV was detected in 37 out of 71 cases (52.11%) using consensus primers. The correlation between promoter methylation of tumour-related genes and HPV was summarized in **Table 4.5.1**. The overall frequency of promoter hypermethylation of selected genes/loci in HPV-positive samples ranged from 8.1% (*MLH1*) to 67.6% (*RASSF1*). Results shown that promoter methylation of *DAPK*, *RASSF1*, *p16* and *MINT31* were significantly higher in HPV-positive HNSCC patients compared to HPV-negative HNSCC (p=0.031, 0.013, 0.031 and 0.015 respectively) (**Figure 4.5.1)**. The promoter methylation frequency of *GSTP1* (43.2%), *ECAD* (43.2%) and *BRCA1* (27.1%) genes were found higher in HPV-positive samples, compared to HPV-negative. However, in cases of *MINT1*, *MINT2* and *MLH1* genes/loci, no much variation was found between HPV-positive and HPV-negative sample (p=0.522, 0.782 and 0.327 respectively). Moreover, HPV-positive tumour samples with CIMP-high showed highest methylation frequency (56.8%). We found a highly significant association between HPV-positive tumours and CIMP-high group (p=0.028). However, there was no correlation between CIMP-low and HPV-positive HNSCC (p=0.477), when compared with HPV-negative HNSCC tumours.

Table 4.5.1 Frequency of DNA methylation of genes/loci and CIMP status analyzed in HPV-positive and HPV-negative HNSCC

Gene/loci	HPV positive (%)	HPV negative (%)	*p*-value*
DAPK			
Methylated	16 (43.2)	5 (14.7)	
Unmethylated	21 (56.8)	29 (85.3)	**0.031**
GSTP1			
Methylated	16 (43.2)	6 (17.6)	
Unmethylated	21 (56.8)	28 (82.4)	0.057
RASSF1			
Methylated	25 (67.6)	11 (32.3)	
Unmethylated	12 (32.4)	23 (67.7)	**0.013**
BRCA1			
Methylated	10 (27.1)	3 (8.9)	
Unmethylated	27 (72.9)	31 (91.1)	0.066
p16			
Methylated	17 (45.9)	6 (17.6)	
Unmethylated	20 (54.1)	28 (82.4)	**0.031**
ECAD			
Methylated	16 (43.2)	7 (20.6)	
Unmethylated	21 (56.8)	27 (79.4)	0.087
MINT1			
Methylated	19 (51.3)	12 (35.3)	
Unmethylated	18 (48.7)	22 (64.7)	0.522
MINT2			
Methylated	20 (54.1)	17 (50)	
Unmethylated	17 (45.9)	17 (50)	0.782
MINT31			
Methylated	23 (62.2)	10 (29.4)	
Unmethylated	14 (37.8)	24 (70.6)	**0.015**
MLH1			
Methylated	3 (8.1)	1 (2.9)	
Unmethylated	34 (91.9)	33 (97.1)	0.327
CIMP status			
CIMP-negative	3 (8.1)	10 (29.4)	
CIMP-low	13 (35.1)	17 (50)	0.477
CIMP-high	21 (56.8)	7 (20.6)	**0.028**

*****Chi square test used to calculate *p*-value and all the *p*-value were adjusting for age, gender, smoking, betel-quid and tobacco chewing status

$p < 0.05$ considered as statistically significant

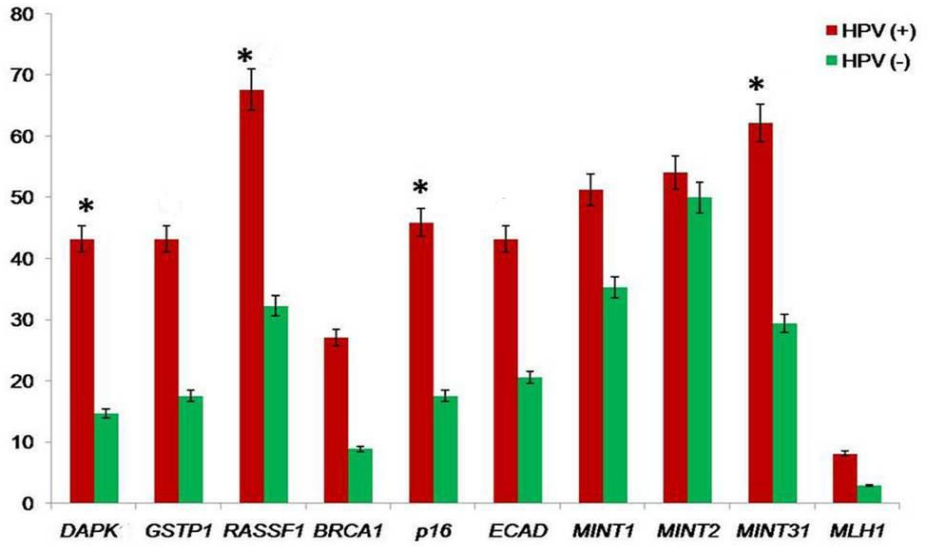

Figure 4.5.1: Frequency of promoter methylation of 10 tumour-related genes/loci in HPV (+) and HPV (-) HNSCC [*p<0.05]

4.5.2 Correlation between Environmental and Genetic Factors and CpG Island Methylator Phenotype (CIMP)

We analyzed the data on environmental factors such as smoking, betel quid chewing and with CIMP and summarized in **Table 4.5.2**. Smokers, betel-quid chewers and tobacco chewers had 89.3%, 71.4% and 89.3% CIMP-high (CIMP-H) respectively. Whereas, non-smokers, non-betel-quid chewers and non-tobacco chewers had show 38.5%, 53.8% and 46.2% CIMP-H characteristic respectively (**Figure 4.5.3**). Smoking and tobacco chewing had strong correlation with CIMP-H (p=0.008 and 0.034 respectively) compared to CIMP-N. However, betel quid chewing had no significant correlation with CIMP-markers. We also had not found any significant variation between CIMP-low versus CIMP-negative and CIMP-high versus CIMP-low in terms of tobacco smoking or chewing (**Table 4.5.2**).

We also correlated the genetic alteration data of carcinogen metabolizing (*GSTM1*, *GSTT1* and *CYP1A1*) and DNA repair (*XRCC1* and *XRCC2*) genes with CIMP panel data (**Table 4.5.3**). The frequency of *GSTM1* and *GSTT1* null genotype was 78.6%, 35.7% respectively in CIMP-high, whereas, 30.8% and 30.8% in CIMP-negative respectively (**Figure 4.5.3)**. The CIMP-high tumours had significantly higher frequency of *GSTM1* null, when compared with the CIMP-low and CIMP-

negative tumours (p=0.004 and 0.023 respectively). However, there was no significant correlation found between *GSTT1* (null), *XRCC1* (Arg/Gln) and *XRCC2* (Arg/His) variant genotypes and different CIMP markers (**Figure 4.5.3**). HNSCC patients with different habits and genetic profiles were found to exhibit differential methylation index (MI). Environmental factors such as smoking and tobacco chewing and genetic variant of *GSTM1*, *CYP1A1* and *XRCC2* showed higher MI. Methylation index found to be high in HPV-positive compared to HPV-negative. HNSCC patients with different habits and genetic profiles were found to exhibit differential methylation index (MI). Mean MI of tobacco chewers/smokers was higher than non chewers/smokers. Similarly, patients with *GSTM1* null, *CYP1A1* CC genotype and HPV-positiv also shown higher methylation index (**Figure 4.5.2**).

Table 4.5.2 Correlation between environmental factors and CIMP-status

Characteristics	**CIMP-H (N=28)**	**CIMP-L (N=30)**	**CIMP-N (N=13)**	***p*-values***		
				CIMP – H Vs CIMP-N	**CIMP – L Vs CIMP-N**	**CIMP – H Vs CIMP-L**
Smoking						
Yes	25 (89.3)	21 (70)	5 (38.5)			
No	3 (10.7)	9 (30)	8 (61.5)	**0.008**	**0.044**	0.214
Betel quid chewing						
Yes	20 (71.4)	20 (66.7)	7 (53.8)			
No	9 (28.6)	10 33.3)	6 (46.2)	0.303	0.249	0.968
Tobacco chewing						
Yes	25 (89.3)	22 (73.3)	6 (46.2)			
No	3 (10.7)	8 (26.7)	7 (53.8)	**0.034**	0.207	0.170

CIMP-H=CIMP-high; **CIMP-L**=CIMP-low and **CIMP-N**= CIMP-negative
**p*-values (Chi-square) were adjusted for age, gender, smoking, HPV, betel-quid and tobacco chewing as appropriate

Table 4.5.3 Correlation between genetic factors and CIMP-status

Characteristic	CIMP-H (N=28)	CIMP-L (N=30)	CIMP-N (N=13)	*p*-values*		
				CIMP – H Vs CIMP-N	CIMP – L Vs CIMP-N	CIMP – H Vs CIMP-L
GSTM1						
Null	22 (78.6)	14 (46.7)	4 (30.8)			
Present	6 (21.4)	16 (53.3)	9 (69.2)	**0.023**	0.731	**0.004**
GSTT1						
Null	10 (35.7)	15 (50)	4 (30.8)			
Present	18 (64.3)	15 (50)	9 (69.2)	0.697	0.762	0.281
***CYP1A1* (T3801C)**						
Wild (TT)	8 (28.6)	13 (43.3)	8 (61.5)			
Heterozygous (TC)	10 (35.7)	11 (36.7)	4 (30.8)	0.997	0.931	0.910
Variant (CC)	10 (35.7)	6 (20)	1 (7.7)	0.180	0.456	0.256
***XRCC1* (Arg399Gln)**						
Arg/Arg	7 (25)	12 (40)	6 (46.2)			
Arg/Gln	14 (50)	11 (36.7)	6 (46.2)	0.108	0.624	0.281
Gln/Gin	7 (25)	7 (23.3)	1 (7.7)	0.056	0.230	0.434
***XRCC2* (Arg188His)**						
Arg/Arg	14 (50)	22 (73.3)	10 (76.9)			
Arg/His	14 (50)	8 (26.7)	3 (23.1)	0.402	0.994	0.186

CIMP-H=CIMP-high; **CIMP-L**=CIMP-low and **CIMP-N**= CIMP-negative
**p*-value was calculated by Chi-square test and further *p*-value was adjusting for age, gender, HPV, smoking, betel-quid and tobacco chewing status (as appropriate)

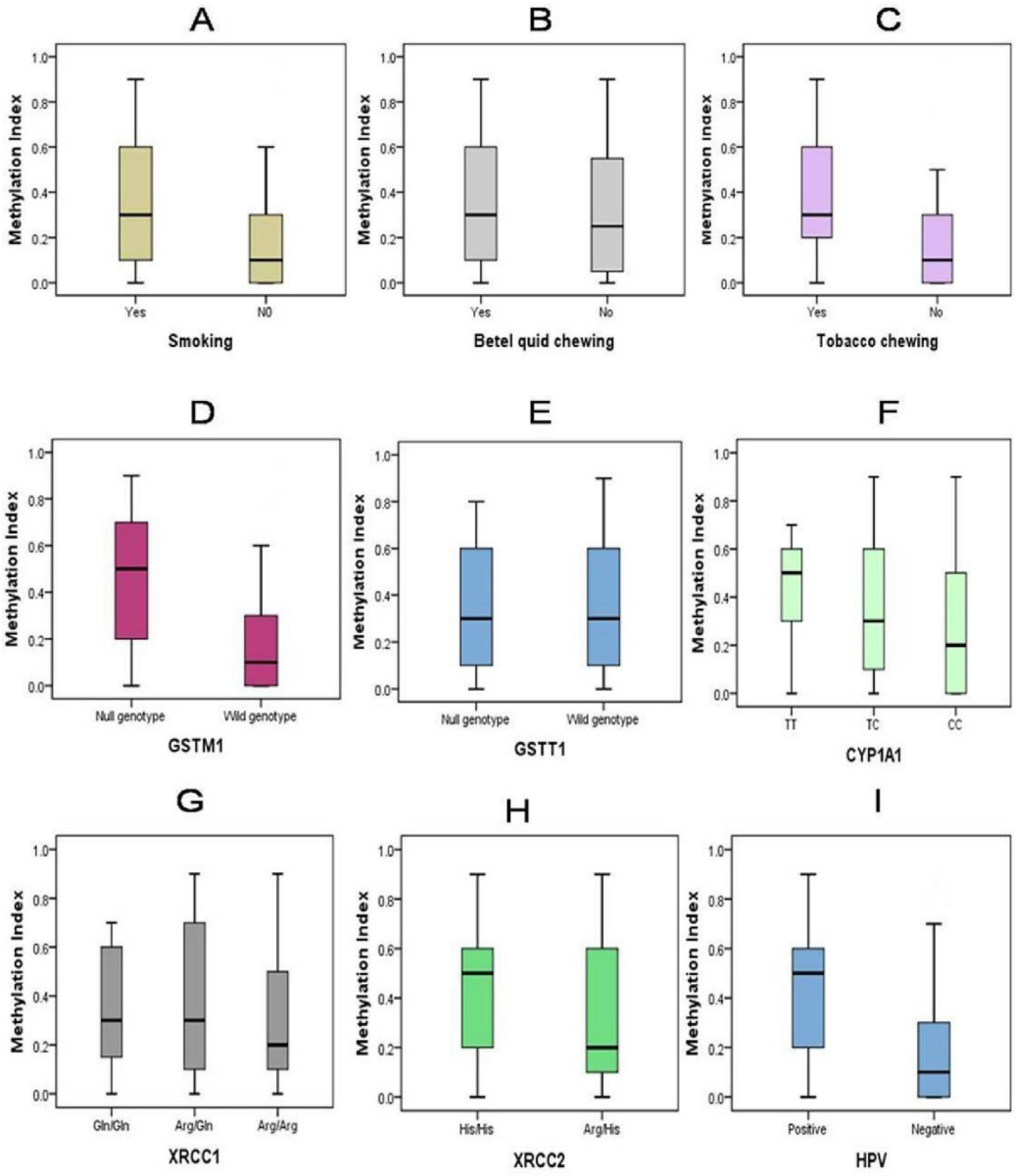

Figure 4.5.2: Methylation index (MI) stratified by genetic and habit related risk factors in HNSCC. Each boxplot represents differential methylation index (MI) among smokers/chewers and non smokers/chewers or wild genotype vs. variant genotype or HPV-positive vs. HPV-negative HNSCC patients**.**

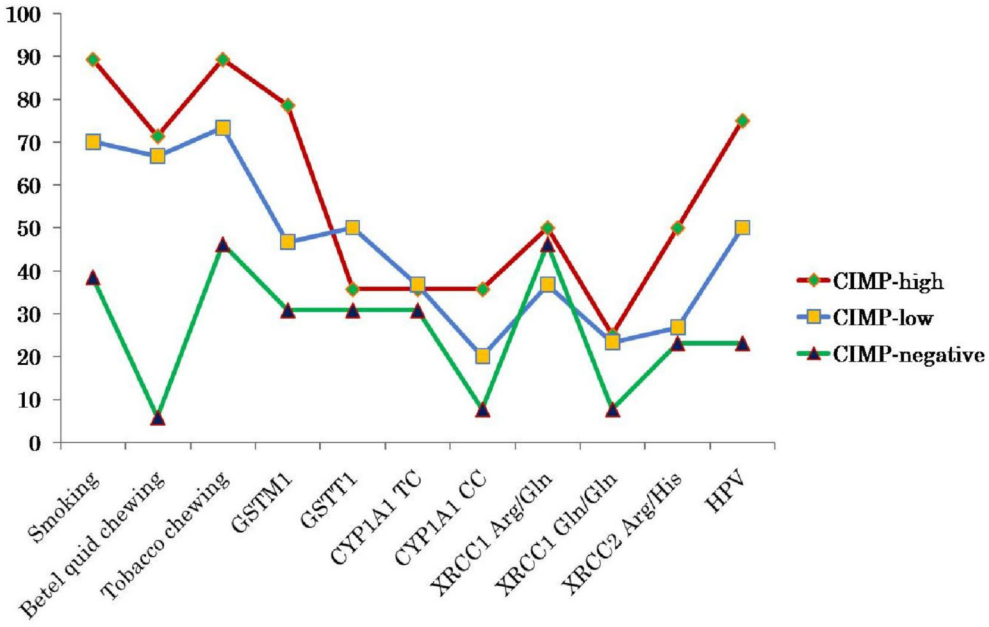

Figure 4.5.3: Correlation between environmental, genetic factors, HPV and CIMP-status. CIMP-high showed higher frequency in terms of smoking, betel-quid and tobacco chewing, HPV, *GSTM1* null, *CYP1A1* and *XRCC2* variant genotype

4.5.3 Identified Tumour Clusters and Correlation with Environmental, Genetic and CIMP characteristics

In this study, we performed unsupervised hierarchical clustering and identified two different classes or sub-groups based on promoter hypermethylation data on tumour samples (**Figure 4.5.4**). The two clusters viz. Cluster-1 and Cluster-2 as identified consist of 29 (40.8%) and 42 (59.2%) HNSCC cases respectively. We had constructed hierarchical clusters using seven genes/loci, as after adjustment seven gene/loci out of ten found to be significantly hypermethylated. The two identified clusters had distinct environmental, genetic and epigenetic features. The frequency of smoking (86.2%) betel quid chewing (72.4%) and tobacco chewing (89.7%) was higher in Cluster-1 compared to Cluster-2 (**Table 4.5.4**). However, only smoking and tobacco chewing shown statistically significant variation among the clusters (p=0.048 and 0.034 respectively).

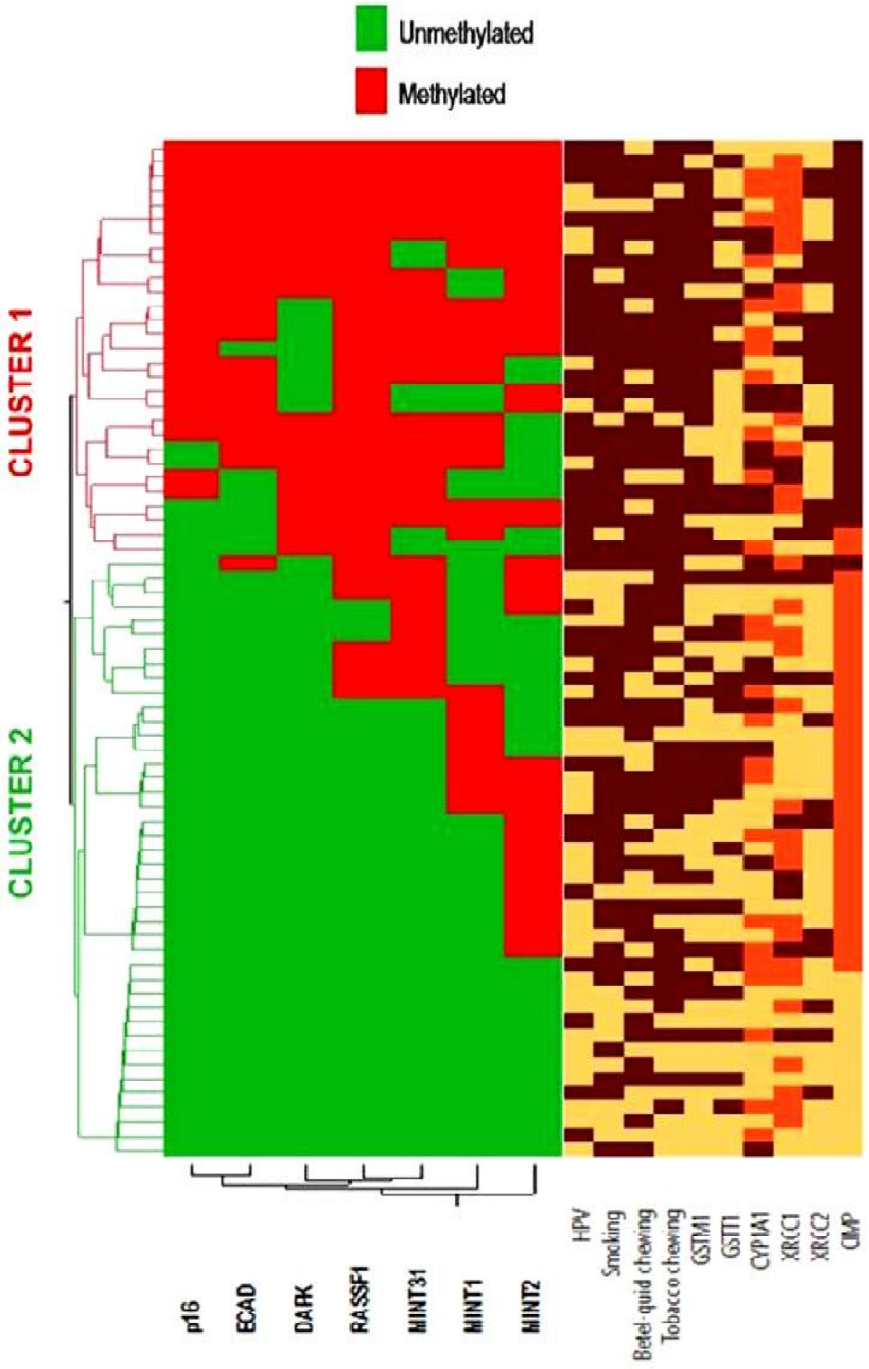

Figure 4.5.4: Hierarchical clustering and heatmap was constructed based on promoter methylation data of HNSCC in 7 tumour-related genes/loci. The different factors in heatmap were represented by colour variation: tobacco consumers, HPV presence and *GSTM1* null, *GSTT1* null (pink colour); tobacco non-consumers and HPV absence *GSTM1* present and *GSTT1* present (green colour). For *CYP1A1*, *XRCC1* and *XRCC2* status: wild type (green); heterozygous (blue) and homozygous variant allele (pink) and for CIMP status: CIMP-high (pink), CIMP-low (blue) and CIMP-negative (green).

Table 4.5.4 Environmental characteristics of the two identified clusters

Characteristics	**Cluster 1 (N=29) (%)**	**Cluster 2 (N=42) (%)**	***p*-values***
Smoking			
Yes	25 (86.2)	26 (61.9)	**0.048**
No	4 (13.8)	16 (38.1)	
Betel quid chewing			
Yes	21 (72.4)	26 (61.9)	0.360
No	8 (27.6)	16 (38.1)	
Tobacco chewing			
Yes	26 (89.7)	27 (64.3)	**0.034**
No	3 (10.3)	15 (35.7)	

**p*-values (Chi-square) were adjusted for age, gender, smoking, HPV, betel-quid and tobacco chewing as appropriate
$P < 0.05$ considered as statistically significant

Genetic, HPV and CIMP characteristics of the Cluster-1 and Cluster-2 were summarized in **Table 4.5.5**. The frequency of *GSTM1* null (82.8%) and *CYP1A1* (31.05%), *XRCC1* (27.6%) and *XRCC2* (48.3%) variant genotypes was higher in Cluster-1 compared to Cluster-2. However, only *GSTM1* null genotype had shown statistically significant variation among the clusters (p= 0.002). In addition, the frequency of HPV positive HNSCC tumours (75.9%) was significantly higher in Cluster-1 as compared to Cluster-2 (35.7%) (p=0.009). Moreover, CIMP-high group (93.1%) is significantly higher in of Cluster-1 (p<0.001) compared to Cluster-2. Whereas, Cluster-2 characterized by CIMP-low (66.7%) and CIMP-negative groups (30.9%), also showed statistical significant differences between Cluster-1 and Cluster-2 (p<0.001).

Table 4.5.5 Genetic, HPV and CIMP characteristics of the two identified clusters

Characteristics	Cluster 1 (N=29) (%)	Cluster 2 (N=42) (%)	*p*-value
GSTM1			
Null	24 (82.8)	16 (38.1)	**0.002**
Present	5 (17.2)	26 (61.9)	
GSTT1			
Null	11 (37.9)	18 (42.9)	0.474
Present	18 (62.1)	24 (57.1)	
CYP1A1 **(T3801)**			
Wild (TT)	9 (31.05)	20(47.6)	
Heterozygous (TC)	11 (37.9)	14 (33.3)	0.997
Variant (CC)	9 (31.05)	8 (19.1)	**0.371**
XRCC1 **(Arg399Gln)**			
Arg/Arg	8 (27.6)	17 (40.4)	
Arg/Gln	13 (44.8)	18 (42.9)	0.428
Gln/Gin	8 (27.6)	7 (16.7)	0.169
XRCC2 **(Arg188His)**			
Arg/Arg	15 (51.7)	31 (73.8)	
Arg/His	14 (48.3)	11 (26.2)	0.481
His/His	0	0	
HPV status			
Present	22 (75.9)	15 (35.7)	**0.009**
Absent	7 (24.1)	27 (64.3)	
CIMP status			
Positive	27 (93.1)	1 (2.4)	
Low	2 (6.9)	28 (66.7)	**<0.001**
Negative	0	13 (30.9)	**<0.001**

$p < 0.05$ considered as statistically significant

4.5.4 Correlation between Survival data and promoter methylation profiles

Survival was examined with respect to CpG island methylator phenotype (CIMP) markers and identified Clusters, using Kaplan-Meier survival curves. Analysis revealed that the overall median survival time of 47 patients out of 71 was 15 months [95% CI=10.97-19.03], and the median survival time in CIMP-high, CIMP-low and CIMP-negative were 11 months [95% CI=7.71 to 14.28], 18 months [95% CI=13.73 to 22.26], and 19 months [95% CI=5.14 to 32.85], respectively (P_{trend}=0.011). Again median survival time for Cluster-1 and Cluster-2 characteristic were 13 and 18 months, respectively (p=0.026) (**Table 4.5.6**). The CIMP-high showed less cumulative survival as compared to CIMP-low and CIMP-negative (**Figure 4.5.5**), moreover Cluster-1 also showed poor survival compared to Cluster-2 (**Figure 4.5.6**). Further analysis showed that CIMP-high and Cluster-2 were significantly associated with poor survival in patients with HNSCC.

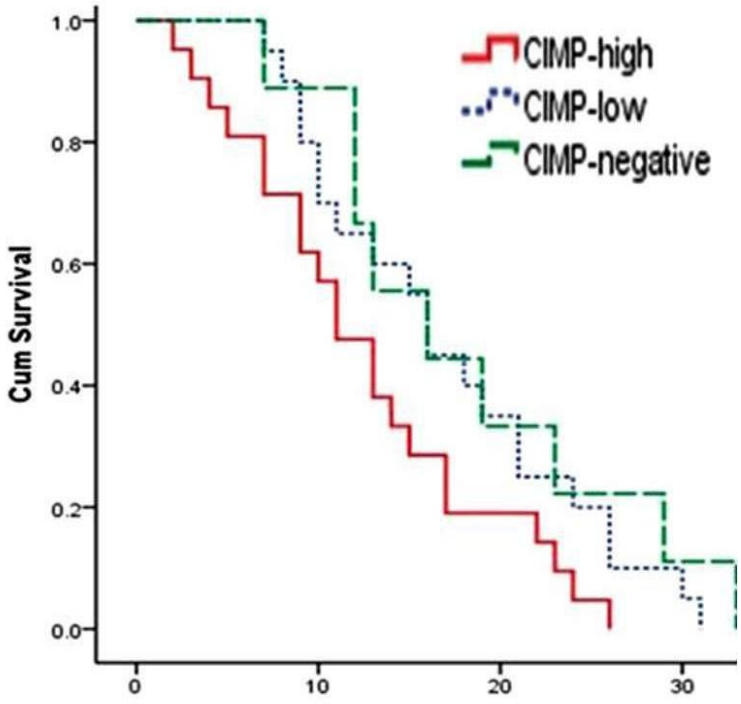

Figure 4.5.5: Kaplan-Meier survival plot for CIMP-status. CIMP-high group of HNSCC tumours showing poorer survival compared to CIMP-low and CIMP-negative group.

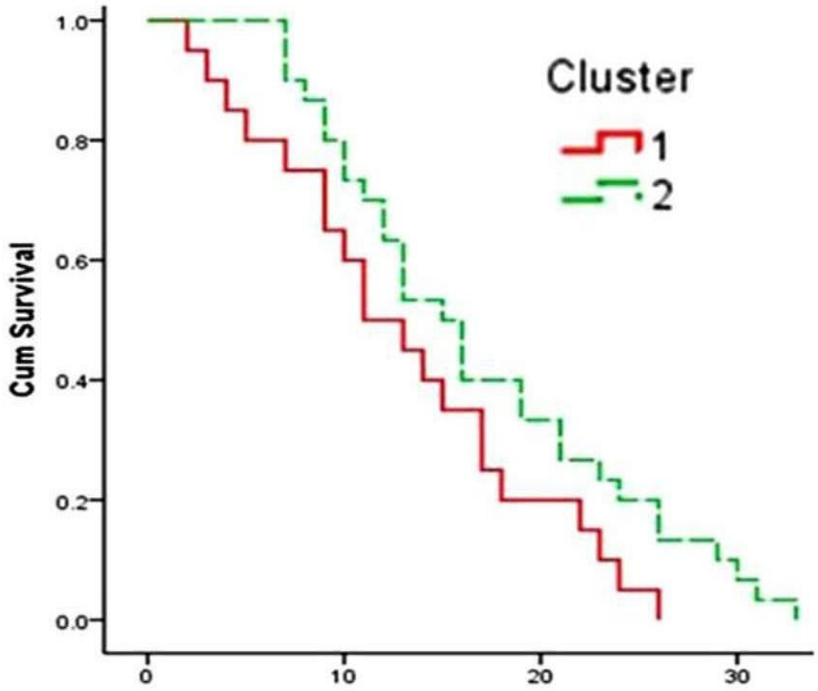

Figure 4.5.6: Kaplan-Meier survival plots for identified Clusters. Two epigenetic cluster also showed differential survival with Cluster-1 had a poor survival.

Table 4.5.6 Summary of survival data of HNSCC patients in Clusters and CIMP group

Variables	**Median (months)**	**95% Confidence Interval**		***p*-value**
		Lower	**Upper**	
Overall	15	10.97	19.03	
Cluster-1	13	9.83	16.16	**0.026**
Cluster-2	18	13.85	22.14	
CIMP-High	11	7.71	14.28	**0.011*** ($*P_{trend}$)
CIMP-Low	18	13.73	22.26	
CIMP-Negative	19	5.14	32.85	