CHAPTER-3

MATERIALS AND METHODS

3.1 Materials

3.1.1 Study Design

The present study was an age and sex matched population based case and control study. All the study subjects included in this study were people of Assam (primarily from southern part), Manipur, Nagaland and Mizoram. For the collection of samples, an informed written consent was taken from all patients, family members and controls. A questionnaire was made for retrieving the demographic information about subjects such as age, gender, family history of cancer as well as tobacco habits and food habits. Smokers and chewers were defined as having smoked or chewed at least for six months. Study population further divided into never, light and heavy smokers or chewers based on their frequency of tobacco consumption. Heavy smokers were those who smoked at least ≥ 20 bidis/cigarettes per day for ≥ 20 years, and light smokers where those who smoked less than 20 bidis/cigarettes per day for less than 20 years. Whereas, heavy chewers where those who chewed ≥ 10 doses per day for ≥ 20 years, and those who chewed less than 10 doses per day for less than 20 years were defined as light chewers (**Annexure-I**).

The present study is comprised of 180 head and neck cancer (HNC) patients that were pathologically confirmed, 300 first-degree relatives of patients (FDRs) and 300 healthy controls. The inclusion criteria for FDRs were those who had first-degree blood relation with HNC cases. The inclusion criteria for controls were those who had no blood relation with HNCs cases, non-malignant individual and residence of Northeast India. Sample of either blood or buccal swab or both collected from each subject included in the study. About 5 ml of venous blood were collected using sterilized needles and stored in EDTA vials. The buccal swabs of subjects were also collected using cotton buds and were stored in 1.5 ml tubes containing 500µl of TES buffer. Tumours tissue samples as well as adjacent normal tissue samples were collected at the time of diagnosis. All the tissue samples collected were diagnosed as squamous cell carcinoma (SCC) after histopathological examination using hematoxylin stain. Tissue samples collected mostly from Cachar Cancer Hospital and Research Centre (CCHRC) and Silchar Medical College and Hospital (SMC), Assam, India.

3.1.2 Ethical Clearance

Ethical clearance **(IEC/AUS/2013-07; dt-20/3/13)** for collection, consent form and analysis of samples was approved by the Institutional Ethical Committee (IEC), Assam University, Silchar, Assam, India.

3.1.3 Chemicals and Reagents

The chemical commonly used in present study were mostly of **molecular grade** and includes the following:

1. Tris base (hydroxymethyl amino methane)
2. Tris acetate
3. Concentrated HCl
4. Ethylene diamine tetrachloro acetic acid (EDTA)
5. Sodium chloride (NaCl)
6. Sodium dodecyl sulphate (SDS)
7. Proteinase K
8. Phenol: Chlorofom: Isomaylacohol (25:24:1)
9. Chloroform:Isoamyalcohol (24:1)
10. Sodium hydroxide pellets
11. Absolute ethanol
12. Ethidium bromide (EtBr)
13. Bromophenol blue
14. Agarose
15. PCR master mix
16. DNA polymerase
17. dNTPs master mix
18. Bisulfite DNA modification Kit

3.1.3.1 Stock Solution

I. **Tris (hydroxymethyl) aminomethane-HCl or Tris-HCl−10 ml (1M, pH−8.00)**

1.211 gm of Tris base was added to 7 ml of nuclease free water; pH was noted and adjusted to pH-8.00 by adding concentrated HCl. Once pH

becomes stable at pH-8.00, remaining amount of nuclease free water was added to make the final volume 10 ml and then autoclaved at 15 lb for 10 minutes and stored at room temperature.

II. **Ethylene diamine tetrachloro acetic acid or EDTA− 10 ml (0.5M, pH 8.00)**

1.862 gm of EDTA was added in 7 ml of distilled water; initial pH was noted and adjusted to pH 8.00 by adding NaOH. Once pH becomes stable at pH 8.00, remaining amount of water was added to make the final volume 10 ml and then autoclaved at 15 lb for 10 minutes and kept at 22°C.

III. **Sodium chloride or NaCl −10 ml (5M)**

2.422 gm of NaCl was added in 10 ml of distilled water and dissolved properly with an autoclaved stirrer and then further autoclaved at 15 lb for 10 mins and kept at 22°C.

IV. **Sodium Dodecyl Sulfate or SDS−10 ml (20%)**

2 gm of SDS powder dissolved in 10 ml of distilled water and store at room temperature.

V. **Ethidium Bromide or EtBr (10 mg/ml)**

10 mg of ethidium bromide dissolved in 1 ml of nuclease free water. Mixed properly and stored at room temperature after wrapping in aluminum foil.

VI. **Tris saturated phenol (pH 8.00)**

Solid phenol liquefied at 68°C in a water bath and equal volume of 1 M Tris HCl (pH 8.00) added. Discarded the upper aqueous phase and repeated the extraction procedure until the pH of the aqueous phase became 8.00. After final extraction, an equal volume of TE added to it and stored in a bottle wrapped with aluminum foil at 4°C.

VII. **Chloroform: Isoamyl alcohol (24:1)**

24 parts of chloroform added with 1 part of Isoamyl alcohol, mixed properly and stored at room temperature.

VIII. **Alcohol grades** − 100%, 90%, 70%

IX. **Proteinase K-** 200 µg of Proteinase K was dissolved in 1 ml of nuclease free water

X. Loading dye (6X)

Bromophenol Blue: 25 mg

Glycerol: 9 ml

Nuclease free water: 7 ml

3.1.3.2 Buffers

I. TES buffer (10 ml)

Tris-HCl: 50 mM (added 500 µl from stock 1 M Tris-HCl)

EDTA: 25 mM (added 500 µl from stock 0.5 M EDTA)

NaCl: 150 mM (added 300 µl from stock 5 M NaCl)

Nuclease free water: 8.7 ml

II. Lysis buffer (10 ml)

Tris-HCl: 10 mM (added 100 µl from stock 1 M Tris-HCl)

EDTA: 10 mM (added 200 µl from stock 0.5 M EDTA)

NaCl: 50 mM (added 100 µl from stock 5 M NaCl)

SDS: 20% (100 µl)

Nuclease free water: 8.6 ml

III. TAE buffer – 10 ml (50X)

Tris-HCl: 2.42 gm

Glacial acetic acid: 570 µl

EDTA-500mM: 1000 µl

Nuclease free water: 8.43 ml

IV. TE buffer

Tris-HCl : 10 mM

EDTA : 1 mM

3.1.2.3 Primers

In this study, PCR primers used were published or newly designed. Details of primers used for *GSTs* (Abdel-Rahman et al. 1996), *CYP1A1* (Chen et al. 2011), *XRCC1* (Yeh et al. 2005) , *XRCC2* (designed) genotyping, HPV detection (Winder et al. 2009) and for methylation specific PCR (Talukdar et al. 2014) are summarized in Table 3.1 and 3.2 respectively.

Table 3.1 List of primers used for amplification of different genes

Primer name	Primer sequence	Annealing Temp (°C)	Product size (bp)
GSTM1-F GSTM1-R	GAACTCCCTGAAAAGCTAAAGC GTTGGGCTCAAATATACGGTGG	59	215
GSTT1-F GSTT1-R	TTCCTTACTGGTCCTCACATTCTC TCACGGGATCATGGCCAGCA	59	480
CYP1A1-F* CYP1A1-R*	GAACTGCCACTTCAGCTGTCT GCTGCATTTGGAAGTGCTC	59	312
CYP1A1-F** CYP1A1-R**	TAGGAGTCTTGTCTCATGCCTT CAGTGAAGAGGTGTAGCCGCT	62	343
XRCC1-F XRCC1-R	TTGTGCTTTCTCTGTGTCCA TCCTCCAGCCTTTTCTGATA	58	615
XRCC2-F XRCC2-R	TCACCCATCTCTCTGCCTTTTG TTCTGATGAGCTCGAGGCTTTC	60	205
HPV My09-F HPV My11-R	CGT CCM ARR GGA WAC TGA TC GCM CAG GGW CAT AAY AAT GG	55	450
HPV GP5+ -F HPV GP6+ -R	TTTGTTACTGTGGTAGATACTAC GAAAAATAAACTGTAAATCATATTC	55	150

F-forward primer and R- reverse primer

*Primers of *CYP1A1* gene used as an internal control in multiplex PCR

** Primers used for *CYP1A1* (T3801C) genotyping

Table 3.2 List of primers used for methylation specific PCR (MSP)

Primer name	Primer sequence	Gene/loci	Annealing Temp (°C)
BRCA1-MF	TCGTGGTAACGGAAAAGCGC	***BRCA1***	66
BRCA1-MR	AACGAACTCACGCCGCGCAA		
BRCA-MF	TTGAGAGGTTGTTGTTTAGTGG		54
BRCA-MR	AACAAACTCACACCACACAA		
P16-MF	TTA TTA GAG GGT GGG GCG GATCGC	***P16***	60
P16-MR	GAC CCC GAA CCG CGA CCG TAA		
P16-UF	TTA TTA GAG GGT GGG GTG GATTGT		65
P16-UR	CAA CCC CAA ACC ACA ACC ATA A		
MLH1 MF	GATAGC GAT TTT TAA CGC	***MLH1***	56
MLH1 MR	TCT ATA AAT TAC TAA ATC TCT TCG		
MLH1 UF	AGAGTG GAT AGT GAT TTT TAA TGT		58
MLH1 UR	ACT CTA TAA ATT ACT AAA TCT CTT CA		
MINT1 MF	AAT TTT TTT ATA TAT ATT TTC GAA GC	***MINT1***	56
MINT1 MR	AAA AAC CTC AAC CCC GCG		
MINT1 UF	AAT TTT TTT ATA TAT ATT TTT GAA GTG T		55
MINT1 UR	AAC AAA AAA CCT CAA CCC CAC A		
MINT2 MF	TTG TTA AAG TGT TGAG TTC GTC	***MINT2***	58
MINT2 MR	AAT AAC GAC GAT TCC GTA CG		
MINT2 UF	GAT TTT GTT AAA GTG TTG AGT TTG TT		56
MINT2 UR	CAA AAT AAT AAC AAC AAT TCC ATA CA		
MINT31 MF	TGT TGG GGA AGT GTT TTT CGG C	***MINT31***	58
MINT31 MR	CGA AAA CGA AAC GCC GCG		
MINT31 UF	TAG ATG TTG GGG AAG TGT TTT TTG GT		59
MINT31 UR	TAA ATA CCC AAA AAC AAA ACA CCA CA		
ECAD-MF	TTA GGT TAG AGG GTT ATC GCG T	***ECAD***	57
ECAD-MR	TAA CTA AAA ATT CAC CTA CCG AC		
ECAD-UF	TAA TTT TAG GTT AGA GGG TTA TTG T		53
ECAD-UR	CAC AAC CAA TCA ACA ACA CA		
DAPK-MF	GGA TAG TCG GAT CGA GTT AAC GTC	***DAPK***	64
DAPK-MR	CCC TCC CAA ACG CCG A		
DAPK-UF	GGA GGA TAG TTG GAT TGA GTT AAT		64
DAPK-UR	GTT CAA ATC CCT CCC AAA CAC CAA		
RASSF1-MF	TTT TTC CAT TTC GCG TCT CT	***RASSF1***	60
RASSF1-MR	CGT TTT TGC CCT TTC TTC GC		
RASSF1-UF	TCACCCATTTTTCCATTTCTCT		60
RASSF1-UR	CTT TTT TTC CCT TTC TTC TCT T		
GSTP1-MF	GAA GTT GGC GAG GTC GAG TTT C	***GSTP1***	65
GSTP1-MR	ACC CGC CAC AAC CCG AAA AAA CG		
GSTP1-UF	GGG AAG TTG GTG AGG TTG AGT TTT		65
GSTP1-UR	CAA CCC ACC ACA ACC CAA AAA ACA		

UF=Unmethylated Forward; UM=Unmethylated Reverse; MF=Methylated Forward and MR=Methylated Reverse primer

3.2. Methods

3.2.1 Isolation and Purification of Genomic DNA from Tissue and Blood

For isolation of genomic DNA from study subject, both blood sample and tissue sample were preferred. Further steps of isolation were done following phenol-chlorofrom-isoamyl method (Green and Sambrook 2012).

3.2.1.1 Isolation of Genomic DNA from Tissue

1. The tissue sample was removed from alcohol and chopped with a sterilized blade as small as possible and then kept dry in − 80°C for 30 minutes
2. The tissue sample then added to a pre-cooled (dry ice) mortar, homogenized gently in 2 volume (w/v) cold TES buffer, homogenizer kept in ice previously. Adjusted the volume (500µl) with TES buffer
3. To the above sample, added 50 µl of 10% SDS followed by 5-10 µl of 20mg/ml of Proteinase-K, incubated at 56°C for 1-18 hours until the tissue is totally dissolve.
4. Equal volume of phenol: chloroform: isoamylalcohol (25:24:1) was added and mixed thoroughly for few minutes.
5. Centrifuged the samples for 10 minutes with 12000 rpm.
6. Transferred upper phase to new 1.5 ml tube and added equal volume of chloroform: isoamylalcohol and centrifuged with 12000 rpm for 10 minutes.
7. Transferred the upper aqueous layer in a fresh sterilized microcentrifuge tube and then adddouble the volume of chilled Absolute ethanol.
8. The above sample kept at -20° C for overnight for precipitation.
9. The above sample centrifuged at 10,000 rpm for 10 minutes.
10. Decanted the supernatant and retained the pellet
11. To the pellet, 500 µl of 70% ethanol was added and again centrifuged at 7000 rpm for 10 minutes and decanted the supernatant.
12. The pellet kept for air dry under laminar airflow.
13. The pellet was re-suspended in nuclease free water and stored at −20ºC for further use or − 86°C for long term preservation

3.2.1.2 Isolation of Genomic DNA from Blood

Blood samples are typically obtained as 1 ml of whole blood stored in EDTA vacutainer tubes frozen at 4° C.

1. The blood sample pipetted in a sterilized microcentrifuge tube containing 500 µl of Lysis buffer.
2. To the above sample, 50 µl of 10% SDS was added and mixed thoroughly for few minutes and then added 2-3 µl of Proteinase-K.
3. Mixed the sample thoroughly by inverting the microcentrifuge tube for few minutes and then incubated at 56° C for 30 minutes.
4. Added equal volume of phenol: chloroform: Isoamyl alcohol (25:24:1) mixed thoroughly for few minutes.
5. Centrifuged the sample for 10 minutes at 12,000 rpm centrifuge machine.
6. Carefully removed the aqueous layer to a new sterilized microcentrifuge tube, added 500 ml of Chloroform: Isoamyl alcohol (24:1), mixed thoroughly by repeated inverting the microcentrifuge tube for few minutes.
7. Centrifuged at 12,000 rpm for 10 minutes in a centrifuge machine.
8. Transferred the upper aqueous layer in a fresh sterilized microcentrifuge tube and added double the volume of Chilled absolute ethanol.
9. The above sample kept at -20° C for overnight for precipitation.
10. The above sample centrifuged at 10,000 rpm for 10 minutes.
11. Decanted the supernatant and retained the pellet.
12. To the pellet, added 500 µl of 70% ethanol and again centrifuged at 7000 rpm for 10 minutes and decanted the supernatant.
13. The pellet kept for air dry under laminar airflow.
14. The pellet was re-suspended in nuclease free water.
15. Then kept at 56ºC for 15 minutes and stored at -20 ºC for further use or −86◦ C for long-term preservation.

3.2.2 Purity Check and Yield of the Extracted DNA

The DNA isolated to be applied for downstream process such as PCR should be pure i.e. free from most of the associated proteins that keeps DNA coiled and should be in adequate quantity. Therefore, before undergoing PCR with the

extracted DNA, the purity and yield of the DNA checked by the processes as follows:

3.2.2.1 Spectrophotometric Analysis

Spectrophotometric determination of purity and yield of the extracted DNA may be explained stepwise as follows:

1. The isolated DNA stock solution taken at different dilution rate viz. 200, 50 and 10 times by adding nuclease free water.
2. Taken 50µl of nuclease free water in a cuvette and calibrated the spectrophotometer at 260 nm as well as 280nm.
3. Added 2 µl of each DNA sample to 48 µl of nuclease free water in a cuvette and mixed well.
4. Optical densities (OD) were measured at 260 (OD_{260}) and 280 (OD_{280}) in UV spectrophotometer (Biophotometer, Eppendrof) against nuclease free water as blank. The yield and purity of DNA samples estimated as follows:

    - Concentration of DNA stock solution (µg/ml) = OD_{260} X 100(dilution factor) X 50 µg/ml/1000
    - Purity of DNA stock solution = OD_{260}/OD_{280} (for pure DNA sample this ratio must be in the range of (1.75 – 1.80)
    - From the concentration of DNA stock solution, the total yield of DNA was calculated and recorded.

3.2.2.2 Agarose Gel Electrophoresis for Quantification and Quality Analysis

The stepwise procedure followed for agarose gel electrophoresis was as follows:

1. Agarose powder was poured in electrophoresis buffer (1X TAE) to the desired concentration and mixed properly and then heated in a microwave oven until completely melted
2. Ethidium bromide (10 mg/ml) was added to the gel (final concentration 0.5 ug/ml) at this point to facilitate visualization of DNA after electrophoresis.
3. After cooling the solution to about 60°C, it was poured into a casting tray containing a sample comb and allowed to solidify at room temperature.

4. After the gel solidified, the comb removed carefully without disturbing the wells formed
5. The gel, still in its plastic tray, was inserted horizontally into the electrophoresis chamber and covered with buffer
6. Samples containing DNA mixed with loading buffer were then pipette into the sample wells, the lid and power leads placed on the apparatus and current applied and flow confirmed through observing bubbles coming off the electrodes.
7. The distance DNA had migrated in the gel judged by visually monitoring migration of the tracking dyes, Bromophenol blue that migrate through agarose gels at roughly the same rate as double-stranded DNA fragments of 300 and 4000 bp.
8. When adequate migration occurred, DNA fragments visualized by staining with Ethidium Bromide
9. To visualize DNA, the gel placed on an ultraviolet transilluminator and photograph taken in Gel-DOC (BioRad).

The concentration of the DNA extracted was measured by calibrating the marker that was run parallel with the samples. The illumination intensity of the sample DNA in comparison to the marker DNA (whose value used for calibration) gives an idea regarding the concentration of the extracted DNA.

3.2.3 Genotyping of *GSTM1* and *GSTT1* by Multiplex Polymerase Chain Reaction (PCR)

Genotyping of the *GSTM1* and *GSTT1* was done by using multiplex PCR. The present and absent of *GSTM1* and *GSTT1* were determined simultaneously in a single PCR reaction using multiple sets of primers and each set of reaction included an internal control, here it is *CYP1A1* gene. Presence of 480, 315 and 215bp amplicons in agarose gel represents the *GSTT1*, *CYP1A1* and *GSTM1* genes respectively (Talukdar et al. 2013). All the PCR reactions were done in Veriti Mastercycler (Applied Biosystems Inc., CA, USA).

PCR Programme:

Each 10µl PCR reaction mixers contain:

2X Biomix (Bioline, UK):	5 µl
Forward primer for *GSTM1* (10 pmole/µl):	0.3 µl
Reverse primer for *GSTM1* (10 pmole/µl):	0.3 µl
Forward primer for *GSTT1* (10 pmole/µl):	0.3µl
Reverse primer for *GSTT1* (10 pmole/µl):	0.3 µl
Forward primer for *CYP1A1* (10 pmole/µl):	0.3 µl
Reverse primer for *CYP1A1* (10 pmole/µl):	0.3 µl
Genomic DNA (100-200 ng):	variable
Nuclease free water:	up to 10µl

PCR Cycling Condition:

Initial denaturation temperature:	94°C for 5 min
Denaturation:	94°C for 1 min for
Annealing:	59°C for 45 sec for
Extension:	72°C for 45 sec
Final extension:	72 °C for 10 min
Hold :	4°C

Gel-electrophoresis and Purification of PCR Product

The PCR-amplified products were analyzed in 1.5% low melting agarose gel containing Ethidium Bromide staining (10 mg/ml).

3.2.4 Genotyping of *CYP1A1*, *XRCC1* and *XRCC2*

3.2.4.1 Polymerase Chain Reaction-Restriction Fragment Length Polymorphism (PCR-RFLP)

Genotyping is the process of determining differences in the genetic make-up (genotype) of an individual by examining the individual's DNA sequence using biological assays and comparing it to another individual's sequence or a reference sequence. It reveals the alleles an individual has inherited from their parents. Traditionally genotyping is the use of DNA sequences to define biological

populations by use of molecular tools. It does not usually involve defining the genes of an individual. The polymorphisms of *CYP1A1* (T3801C), *XRCC1* (Arg399Gln) and *XRCC2* (Arg188His) genes were analyzed by this PCR-RFLP method (Chen et al. 2011,Yeh et al. 2005). The polymorphic site of *CYP1A1*, *XRCC1* and *XRCC2* was amplified using published forward and reverse primers in 20µl PCR reactions.

PCR programme:

Each 20µl PCR reaction mixers contain:

10X PCR buffer:	2 µl
dNTPs Mastermix (10mM):	2 µl
Forward primer (20 pmole/µl) :	0.5µl
Reverse primer (20 pmole/µl):	0.5 µl
High fidelity DNA polymerase (5unit/µl):	0.4 µl
Genomic DNA (100-200 ng) :	variable
Nuclease free water :	up to 20µl

PCR cycling condition:

Initial denaturation temperature:	94 °C for 5 min
Denaturation :	94 °C for 1 min for,
Annealing :	variable (°C)for 45 sec for
Extension:	72 °C for 45 sec
Final extension:	72 °C for 10 min
Hold :	4°C

The PCR products were visualized using a UV transilluminator after ethidium bromide staining.

Reaction Mixture for RFLP

PCR products of *CYP1A1*, *XRCC1* and *XRCC2* were digested separately with the restriction enzyme *Msp1*, *HpaII* and *HphI* enzymes (New England BioLabs, USA) respectively. The digested products were then resolved on 3% agarose gel to assess the size of the PCR–RFLP products. Material required for restriction

digestion includes 10x buffer, restriction enzyme, nuclease free water, 1.5 ml Eppendorf tubes, pipette, and water-bath for incubation.

Each 15 µl reaction of restriction enzyme contains:

10X PCR Buffer:	1.5 µl
Restriction enzymes (20U/ µl):	1 µl
PCR products:	7 µl
Nuclease free water:	upto 15µl

All the above contents were mixed in a 0.2 ml autoclaved PCR tube and incubated at 37°C overnight.

3.2.3.2 Sanger Sequencing

Sanger sequencing was performed for genotyping of *CYP1A1, XRCC1* and *XRCC2* genes to validate the PCR-RFLP results of polymorphisms. The amplified products were purified and sequenced bi-directionally using automated DNA Sequencer (ABI 3500 Genetic Analyzer; Applied Biosystem, Inc. USA).

Purification of PCR products by Gel purification

Purification of PCR products is required before sequencing the genes. Purification was done by GeneJET Gel Extraction Kit (Thermo Scientific, K0692) as per manufacturer protocol. The process is given below:

1. All the gel slices containing the amplified fragments of *CYP1A1, XRCC1,* and *XRCC2* using a clean scalpel was excised.

2. Then 1:1 volume of binding buffer was added to the each slice (volume : weight) (For example, 200µl of binding buffer was added for every 200mg of agarose gel slice)

3. The gel mixture was incubated at 56°C until the gel slices were completely dissolved. Then 200µl of Absolute ethanol was added to the solution.

4. All the samples were briefly vortexed and then the solution was transferred to GeneJET purification columns.

5. These were then centrifuged at 14000g for 1 minute. The flow through was discarded and the columns were placed back into the same collection tube.

6. 200µl of Pre-wash buffer (diluted with ethanol) was added to the GeneJET purification columns.

7. These were then centrifuged for 1 minute at 14000g. The flow through was discarded and the column was placed back into the same collection tube.

8. 700µl of Wash buffer (diluted with ethanol) was added to the GeneJET purification columns and then centrifuged for 1 minute at 14000g. Next, the flow through was discarded and the columns were placed back into the same collection tube. This step was carried out 2 times.

9. The empty GeneJET purification columns were centrifuged to an additional 1 minute at 14000g to completely remove the residual wash buffer.

10. These purification columns were then transferred into clean 1.5ml microcentrifuge tubes.

11. 10 µl of elution buffer was added to the centre of the each purification column membranes.

12. It was kept in room temperature for 5 minutes.

13. It was then centrifuged for 1 minute at 14000g.

14. The GeneJET purification columns were discarded and the purified PCR products were stored at -20°C for further downstream processes.

Sequencing Reaction

The sequencing reaction was performed using Ready reaction premix (2.5X), BigDye Terminator v3.1 dye Sequencing buffer (5X); 20 pmol of the primer (either forward or reverse), the purified PCR products as template (3-10ng) and nuclease free water upto 20 µl. The Chain termination reaction was carried out at 25 cycles of 1 min at 94°C, 5 sec at 50°C and 4 min at 60°C. The fragments were then purified

by Sodium Acetate/EDTA/Ethanol method prior to run in the 3500 Genetic Analyzer. The Sodium Acetate 3M, pH-5.2 and 125mM EDTA were added maintaining 1/10th of the total volume of the product followed by the addition of double the volume of absolute ethanol and consecutive

3.2.5 Bioinformatic Analysis

Different softwares such as ABI Sequence Scanner v1.0, ClustalX2, MEGA6, Bioedit and online servers such as NCBI Blast tool, Sequence manipulation suit, etc are used for the analysis of the sequences. The quality of the generated sequences (i.e. QV) were checked by Sequence Scanner v1 and SeqScape v2.7 (Applied Biosystems, Inc. USA) and further analyzed by nucleotide BLAST tool at the National Centre for Biotechnology Information (*http://www.ncbi.nlm.nih.gov/blast*) to check percentage similarity of the developed sequences with the database sequences. The resultant sequences were then annotated and multiple sequence alignment was performed to detect specific SNP using BioEdit software and MEGA6 software.

3.2.6 Methylation Analysis

Various methods used for methylation analysis are bisulfate sequencing, combined bisulfite restriction analysis (COBRA), methylation specific PCR (MSP), real time MSP or MethyLight, and pyrosequencing. All these methods are based on principles that differentially detect 5-methylcytosine (C^m) from cytosine (C). Bisulfite treatment of genomic DNA converts unmethylated cytosine (C) into uracil (U) while leaving methylated cytosine (C^m) unchanged.

3.2.6.1 Bisulfite Treatment of Genomic DNA

At first genomic DNA was modified using The Imprint® DNA Modification Kit (Sigma-Aldrich, USA; Catalog Number MOD50), the kit contains all of the reagents necessary for complete bisulfite conversion and subsequent purification of DNA samples. During modification process, DNA is chemically denatured to allow bisulfite reagent to react specifically with single-stranded DNA, thereby deaminating cytosine and creating a Uracil residue. DNA denaturation and bisulfite modification carried out simultaneously. The unique DNA protection reagents in

the modification buffer prevent chemical and thermophilic degradation of DNA in the bisulfite treatment. The Capture Solution enables the DNA to tightly bind to the column filter. This allows for the effective removal of residual sodium bisulfite and salts. The eluted modified DNA can be used immediately, or stored at –20 °C for up to 2 months. Converted DNA is suitable for a variety of downstream applications including Methylation-Specifc PCR, methylation sequencing, and pyrosequencing, as well as methylation microarray.

Kit Components (for 50 Reactions)

DNA Modification Powder	5 vials
DNA Modification Solution	6.5 ml
Balance Solution	0.5 ml
Capture Solution	20 ml
Cleaning Solution	3.5 ml
Elution Solution	1.5 ml
Spin Column	50 each
Capless Collection Tube, 2ml	50 each
Collection Tube, 1.5 ml	50 each

Reagent required (not provided with Kit)

1. Heating block or water baths for incubation at 37 °C or 99 °C, and 65 °C
2. Ethanol-diluted cleaning solution - Add 8.2ml of absolute ethanol to the bottle and mix.
3. 90% Ethanol solution - Add 500 µL of water to 4.5 ml of absolute ethanol.
4. Balance/Ethanol wash solution - Add 10 µL of Balance Solution to 1.1 ml of 90% ethanol.

5. The fragment of DNA to be amplified is intentionally small, to allow the assessment of methylation patterns in a limited region and to facilitate the application of this technique to samples, such as paraffin blocks, where amplication of larger fragments is not possible.

Procedure:

I. DNA Modification:

1. 1.1 ml of DNA modification solution was added to 1 vial of DNA modification powder. Vortex the vial for two minutes or until the solution is clear. The vial was examined for any particles that may not be dissolved. If the particles were remained in the vial, then incubate the vial at 65 ºC for two minutes and vortex briefly.

2. 40 µL of Balance solution was added vortex briefly.

3. Then 10 µL of DNA was added to 1.5 ml micro-centrifuge tube. 110µL of the solution from the Step-1was added into the tube. Vortex briefly and then incubated at 99 ºC for 6 minutes.

4. After the incubation in step-2, immediately follow with incubation at 65 ºC for 90 minutes and then proceed the post modification DNA clean up.

II. Post Modification Clean-up:

1. A spin column was placed into a cap less collection tube (2ml) for each sample that was modified.

2. 300 µL of capture solution was added to the spin column and allow the solution to sit on the column for one minute.

3. The modified DNA solution was added from Step-3 of the DNA modification procedure utilized on to the spin column already containing the Capture Solution. Then the column was centrifuged at 1200 x g for 20 seconds and the flow through was discarded.

4. 200 µL of ethanol-diluted cleanup solution was added to the spin column and centrifuged at 1200xg for 20 seconds.

5. 50 µL of the Balance/Ethanol wash solution was added to the bottom of the spin column. It was ensured that air bubbles are not impeding liquid flow to the column filter and incubated for 8 minutes at room temperature. After incubation, centrifuged for 20 seconds and discarded the flow through.

6. 200 µL of 90% ethanol solution was added to the spin column and centrifuged for 20 seconds and discarded the flow through.

7. Again, 200 mL of 90% Ethanol Solution was added to the Spin Column and centrifuge for 40 seconds. The 2 ml Cap-less Collection Tube was discarded, and the Spin Column was placed into the 1.5 ml Collection Tube.

8. In final step, 8-20 mL of Elution Solution was added to the bottom of the Spin Column. The solution was allowed to incubate for 1 minute and then centrifuged for 20 seconds. the Spin Column was removed and discarded. The eluted solution is the modified DNA. The modified DNA is now ready for downstream testing or it may store at –20 °C for up to 2 months.

Note:

1. Always cap the Spin Columns before placing them in a microcentrifuge.
2. All centrifugations are at 12,000 x g

3.2.6.2 PCR Amplification of the Modified Genomic DNA by Methylation Specific PCR (MSP)

This technique is very simple and can interrogate methylation status of several CpGs at primer sites by performing PCR with primers specific to unmethylated or methylated sequences. Methylation-specific PCR (MSP) is sensitive and specific for methylation of any sites of CpG in a CpG island. Primers were designed to distinguish methylated from unmethylated DNA in bisulfite-modified DNA. Unmodified DNA or DNA incompletely reacted with bisulfite can also be distinguished, since marked sequence differences exist between these DNAs. The frequency of CG sites in CpG islands renders this technique uniquely useful and extremely sensitive for such regions. Since the two strands of DNA are no longer complementary after bisulfite treatment, primer can be designed for either modified strand. The disadvantage of the technique is the high rate of false negative or

positive results and requires careful determination of the number of PCR cycles performed (Talukdar et al. 2013).

The reaction mixtures of the isolate modified DNA were made separately for amplification of each genes and carried out the amplification in the thermal cycler

Each 10µl PCR reaction mixers contain:

10X PCR buffer:	2 µl
dNTPs Mastermix (10mM):	2 µl
Forward primer (20 pmole/µl) :	0.5µl
Reverse primer (20 pmole/µl):	0.5 µl
High fidelity DNA polymerase (5unit/µl):	0.4 µl
Bisulfite modified DNA (100-200 ng) :	variable
Nuclease free water :	up to 10µl

PCR cycling condition:

Initial denaturation temperature:	94 °C for 5 min
Denaturation :	94 °C for 1 min for,
Annealing :	variable (°C) for 45 sec for
Extension:	72 °C for 45 sec
Final extension:	72 °C for 10 min
Hold :	4 °C

3.2.7 Statistical Analysis

Statistical analysis was performed using Statistical Package for the Social Sciences (SPSS) software version 18 and *p*-value less than 0.05 (two-tailed) was considered statistically significant. The differences in genotypes distribution of *GSTM1, GSTT1, CYP1A1, XRCC1* and *XRCC2* between patients having HNSCC and controls were correlated using Chi-square and a two-sided Fisher's exact test where appropriate. In cases of *XRCC1*, *XRCC2* and *CYP1A1,* the estimation of allele frequency for Hardy-Weinberg equilibrium was also analyzed by Chi-square test. We also used the Wilcoxon rank-sum test to compare promoter methylation levels of HNSCC tumour and normal samples, which permit the comparison of two groups of

independent samples. The significant values further adjusted for multiple testing by Bonferroni method (*p*-value multiplied by number of comparisons). Unsupervised hierarchical clustering was done using JMP version 12 software package of SAS, which identify subgroups among HNSCC patients based on promoter methylation frequency.

3.2.7.1 Logistic Regression (LR) Analysis

The association between the lifestyle habits, genetic factors and HNSCC risk was analyzed by calculating odds ratios (ORs), 95% confidence intervals (95% CI), and their corresponding *p*-values using logistic regression model. In addition, to strengthen the association between different genotypes and HNSCC risk, odds ratios were calculated after adjusting confounding factors such as age, gender, smoking, tobacco chewing status. Gene-gene and gene-environment interactions were estimated using the logistic regression model. Two-sided *p*-values less than 0.05 are considered statistically significant.

3.2.7.2 Multifactor Dimensionality Reduction (MDR)

The MDR is advanced, non-parametric and model-free statistical approach used to evaluate gene-gene and gene-environment interaction in cancer risk. The advantage of using MDR approach (*www.multifactordimensionalityreduction.org*) is to overcome sample size limitations encountered by conventional statistical methods (example: logistic regression analysis). Here, in present study MDR software package (MDR 3.0.2) was used to generate a best one-dimensional multi-factors model to classify and predict HNSCC susceptibility. We performed 10-fold cross-validation procedure and 10 times random seed numbers, to reduce the chance of observing false results due to the divisions of the data. The best model for each order of interaction was selected by maximum cross validation consistency (CVC) and testing balanced accuracy (TBA). Interaction models showing highest TBA and CVC was further tested by 1000 folds permutation tests and $\chi 2$ test at 0.05% significance levels.

3.2.7.3 Interaction Entropy Graphs

Interaction entropy graphs were constructed based on MDR results, for better verification and visualization of interactions between the variables. Interaction entropy graphs are consisting of a node for each variable with pair-wise connections between them. The percentage of entropy removed by each variable or by each pair-wise connection of variables is visualized for each node or connection. Therefore, the independent effect of each variable can be compared with the interaction effect of variables and determined whether interactions are synergistic or not (Talukdar et al. 2013).

3.2.7.4 False Positive Report Possibility (FPRP)

Results generated by multiple comparison analysis are often encountered by false positive discoveries, thus false positive report possibility (FPRP) is conducted to evaluate robustness of the findings from studies of gene-environment interaction using a Bayesian approach. In view of a small sample size of our epidemiological data, we set prior probabilities ranging from 0.25 to 10^{-5} with statistical power to detect OR of 1.5 and α level equal to the observed *p*-value. The FPRP cutoff point was kept at 0.5 (Talukdar et al. 2013).

3.2.7.5 Survival Analysis

Survival was calculated in months from the beginning of treatment to the month of death using Kaplan–Meier survival curves in SPSS software, version 18 (Windows). Deaths due to diseases/complication other than cancer were expelled from the study. The association between different characteristics (HPV, CIMP and cluster) and the event of death was analyzed using Log–rank (Mantel–Cox) tests.