ANALYSIS OF MUTATIONS IN KEY TRIGGER GENES ASSOCIATED WITH GASTRIC CANCER FROM SUBTYPING TO EARLY PROGNOSTIC ASSESSMENT

Thesis submitted in partial fulfillment of the requirement for the degree of Doctor of Philosophy in Biotechnology

By

Souvik Ghatak

Ph.D. Registration No: MZU/PhD/615 of 31.10.2013

Under the supervision of

Dr. N. Senthil Kumar
Professor, Department of Biotechnology
School of Life Sciences
Mizoram University, Mizoram

2016

Department of Biotechnology
School of Life Sciences
MIZORAM UNIVERSITY
(*A Central University*)
Aizawl - 796 004, Mizoram, India

Dr. N. Senthil Kumar Ph.D
Professor & Head

Phone: 09436352574
Email: nskmzu@gmail.com

CERTIFICATE

This is to certify that **Mr. Souvik Ghatak**, Ph.D. Scholar, Registration No. MZU/Ph.D./667 of 31.10.2013 has work on the Thesis entitled **"Analysis of mutations in key trigger genes associated with gastric cancer from subtyping to early prognostic assessment".** He has fulfilled all criteria prescribed by the UGC (Minimum Standard and Procedure governing Ph.D. Regulations). He has fulfilled the mandatory publication (Publication enclosed) and completed Ph.D. course work.

This also certifies that the scholar has been admitted in the Department through an entrance test, followed by an interview as per clause 9(i) and (ii) of the UGC Regulation 2009.

Date:

(N. Senthil Kumar)
Supervisor

MIZORAM UNIVERSITY

AIZAWL-796004

17th October, 2016

DECLARATION

I, **Souvik Ghatak,** hereby declare that the subject matter of this thesis is the record of work done by me, that the contents of this thesis did not form basis of the award of any previous degree to me or to do the best of my knowledge to anybody else, and that the thesis has not been submitted by me for any research degree in any other University/Institute.

This is being submitted to the Mizoram University for the degree of Doctor of Philosophy in Biotechnology.

Date:

(Souvik Ghatak**)**
Candidate

(Prof. N. Senthil Kumar**)**
Head

(Prof. N. Senthil Kumar**)**
Supervisor

Acknowledgement

First and foremost, I am grateful to **God**, the Creator and **my parents**, and to whom I owe my very existence, and for providing me the opportunity to step in the excellent world of science. I am thankful to him for giving me the grace and privilege to pursue this program and successfully complete it in spite of many challenges faced. The journey has been quite remarkable, and it is a unique stepping stone to many exploits ahead. I have also been supported and supervised by many people to whom I would like to express my deepest gratitude. I would also like to thank my **grandfather**for his blessing.

I would like to acknowledge my indebtedness and render my warmest gratitude to my supervisor **Prof. N. Senthil Kumar,Head,** Department of Biotechnology, Mizoram University,who is captivating, honest, and the true embodiment of a mentor. His hard working nature, unconventional approach, openness and flexibility at accepting new ideas and creativity provided me with a perfect environment in which I was able to improve myself both at the personal and academic levels to achieve the very best and to explore new frontiers with devotion and confidence. His sacrifice of time and patience in listening to my ideas and problems, untiring suggestions and thoughtful criticisms are things that I would always cherish in the days to come. Learning from him was enormous, his intuition has made him as a constant oasis of ideas and passions in science to provide me unflinching encouragement to transform me a budding researcher and I believe his wishes are immensely valuable for my career ahead. I would also like to pay my respects to all the faculty member**,** Department of Biotechnology, Mizoram University for extending all required facilities to complete my research work. All the faculty members from Department of Biotechnology, Mizoram University were kind enough for their help and encouragement at various phases of research, whenever I approached them, and I do hereby acknowledge all of them. I am thankful to the office and technical staff of the Department for their constant help and support.

I owe my gratitude to **Dr. Jeremy L. Pautu**, Oncologist, Mizoram State Ccancer Institute, Zemabawk, Aizawl, Mizoram; **Dr. John Zohmingthanga**, Head pathologist, Civil Hospital, Aizawl, Mizoramfor providing me samples and scientific suggestions**. Dr. Arup Bhoumik**, Pathologist, Agartala Govt. Medical College, Agartala, Tripura, **Dr. Ravi**

Kannan,Oncologist, Cachar Cancer Hospital and Research Centre, Silchar, Assam and **Dr. Anirban Mukhopadhyay**, Assistant professor, Department of Zoology, University of Kalyani, West Bengalfor providing research facilities in their lab and scientific suggestions. I would like to express my deepest thank to **Dr. Madhusudhan Das,** Professor, Department of Zoology, Calcutta University, Kolkata, West Bengal for his inspiration, constant support and for giving new ideas during research.I would like to express my deepest thank to **Prof. Partha P Majumder,** Director and **Dr. Arindam Maitra,** Associate Professor, National Institute of Biomedical Genomics, Kalyani, West-Bengalfor his inspiration, constant support and for giving new ideas during research.I owe specialmention to my friends, **Subhajit Mukherjee, Surajit De Mandal, Ravi Prakash Yadav, Payel Chakraborty, S. Sarath Babu,Brinda Senthil Kumar,Freda Lalrohlui,Christeen Sailo,Rebecca Lalmuanpuii,Lalhma Chhuani, Ruth Lalfelpuii, Ajit Kumar Passari, Vineet Kumar Mishra**, **Vincent Vineeth Leo, Thangjam Premabati, David K. Zorinsanga, Dayna Rajkumari, Loknath Samanta, Ravi Kumar Mishra**whoworked in lab and helped me throughout the completion of my research as well as for sharingthe joyous, sad feelings and giving me the golden moments during this period.

Specially, I also thankful to **Department of Biotechnology (DBT),** Government of India, New Delhifor providing the fellowship to complete my research works**under DBT STB-HUB (BT/04/NE/ 2009 dt. 29.08.2014)**andDBT-Twinning Project on Gastric Cancer (**BT/360/NE/TBP/2012**)**.** I am also thankful to the Department of Biotechnology, New Delhifor the establishment of the **DBT-BIF Centre**(**BT/BI/12/060/2012 (NERBIF-MUA**) in the Department of Biotechnology, Mizoram University.

At last but not the least, I acknowledge all those who provide their samples and consent for collecting samples and those who knowingly and unknowingly contributed in making my work easier and a real success.

Date: **Souvik Ghatak**

Department of Biotechnology
School of Life Sciences
MIZORAM UNIVERSITY
(*A Central University*)
Aizawl - 796 004, Mizoram, India

Dr. N. Senthil Kumar Ph.D
Professor & Head

Phone: 09436352574
Email: nskmzu@gmail.com

TO WHOM IT MAY CONCERN

This is to certify that **Mr. Souvik Ghatak**, Ph.D scholar (Reg. No. MZU/PhD/615 of 31.10.2013) completed his Research work on the topic "**Analysis of mutations in key trigger genes Associated with gastric cancer from subtyping to early prognostic assessment**". He has fulfilled all criteria prescribed by the UGC (Minimum Standard and Procedure governing Ph. D regulations).

This also certifies that the scholar has been admitted to Ph. D programme in the department through an entrance test, followed by an interview as per clause 9(i) & (ii) of the UGC regulation, 2009. He has completed Ph. D course work (Mark sheet enclosed) and mandatory publication (publication enclosed) for thesis submission.

Abbreviations

SNP	Single nucleotide polymorphism
GC content	Guanine and Cytosine content
DNA	Deoxyribonucleic Acid
PCR	Polymerase chain reaction
PMFS	Phenylmethylsulfonyl fluoride
RNA	Ribonucleic Acid
PCR	Polymerase Chain Reaction
16S rRNA	16S-Ribosomal Ribonucleic Acid
RFLP	Restriction Fragment Length Polymorphism
RT-PCR	Reverse transcriptase - Polymerase Chain Reaction
pH	Negative Ion Of Hydrogen Ion Concentration
mM	Milli Molar
M	Molar
mm	Millimeter
h	Hour
d	Days
ng	Nanogram
pmol	Pico-Mol
µM	Micromolar
dNTPs	Deoxynucleotides
U/ µl	Unit Per Micro Liter
TAE	Tris Base, Acetic Acid and EDTA
TBE	Tris Base, Boric Acid and EDTA
UV	Ultra-Violet
BLAST	Basic Local Alignment Search Tool
EMBL	European Molecular Biology Laboratory
UCSC	University of California, Santa Cruz

IARC	International Agency for Research on Cancer
dbSNP	Database for SNP
bp	Base Pair
DMSO	Dimethyl Sulfoxide
ANOVA	Analysis Of Variance
$MgSO_4$	Magnesium Sulfate
$MgCl_2$	Magnesium Chloride
CAS	Chrome Azurol S
KCN	Potassium Cyanide
g	Gram
mg/ml	Milligram Per Ml
SD	Standard Deviation
NaCl	Sodium Chloride
mtDNA	Mitochondrial DNA
TP53	Tumor Protein 53
CDH1	Cadherin-1
MDM2	Mouse Double Minute 2
GC	Gastric Cancer
HC	Healthy Control
cagA	cytotoxin-associated gene A
HVR	Hyper Variable Region
hg	Haplogroup
rCRS	Revised Cambridge Reference Sequence
BSPs	Bayesian Skyline Plots
GTR	General Time-Reversible
MCMC	Markov Chain Monte Carlo
TMRCA	Time of Most Recent Common Ancestor
YBP	Year Before present
CNV	Copy Number Variation
MSI	Microsatellite Instability
D-loop	Displacement Loop

MT-CO1/MT-COXI	Mitochondrial Cytochrome C Oxidase Subunit I
ROS	Reactive Oxygen Species
RBCs	Red Blood Cells
rpm	Rotation per Minute
DEPC	Diethylpyrocarbonate
tRNA	transfer RNA
MSS	Microsatellite Stable
EBV	*Epstein–Barr virus*
H. pylori	*Helicobacter pylori*
PAGE	Polyacrylamide Gel Electrophoresis
SDS	Sodium Dodecyl Sulfate
NADH	Nicotinamide adenine dinucleotide
ND1/NDI	NADH-ubiquinone Oxidoreductase chain 1
ND5/NDV	NADH-ubiquinone Oxidoreductase chain 5
ND2/NDII	NADH-ubiquinone Oxidoreductase chain 2
DCPIP	2,6-dichlorophenolidophenol
DB	Decylubiquinone
DBH2	Decylubiquinone
ORs	Odds Ratios
CIs	95% Confidence Intervals
df	Degree of Freedom
SPSS	Statistical Package of the Social Sciences
LR	Logistic Regression
MDR	Multifactor Dimensionality Reduction
ATP	Adenosine Tri-Phosphate
MFI	Mean Fluorescence Intensity
del	Deletion
int	Insertion
Vh	Volt-hours
SSCP	Single-Strand Conformation Polymorphism
Arg	Arginine

Pro	Proline
GSTs	Glutathione-S-Transferases
Ile	Isoleucine
Val	Valine
NCRP	National Cancer Registry Program
AAR	Age-Adjusted Rates
CVC	Cross Validation Consistency

CONTENTS

		Page No.
	Certificate	**i**
	Declaration	**ii**
	Acknowledgement	**iii-iv**
	Abbreviation	**v-viii**
	Contents	**ix-xi**
	List of Figures	**xii-xvii**
	List of Tables	**xviii-xix**
	Chapter I : Introduction and Review of literature	**1-7**
	Chapter II: Aim and Objectives	**8-9**
	Chapter III: Mitochondrial haplogroup analysis from Mizo tribal population of Northeast India	**10-27**
1.	Introduction	**11-12**
2.	Materials and Methods	**12-18**
3.	Results and Discussion	**18-27**
	Chapter IV: Study of Mitochondrial control region, Cytochrome Oxidase C subunit I sequence variation and Mitochondrial Microsatellite Instability associated with gastric cancer	**28-75**

1.	Introduction	**29-30**
2.	Materials and Methods	**30-38**
3.	Results	**38-60**
4.	Discussion	**61-75**
	Chapter V: Characterization of novel Missense and splice site mutations in TP53 gene and their genetic and functional analysis in gastric tumorigenesis	**76-101**
1.	Introduction	**77-78**
2.	Materials and Methods	**78-85**
3.	Results	**85-95**
4.	Discussion	**95-101**
	Chapter VI: Novel mutations and inactivation of the E-Cadherin Gene in diffuse Type Gastric Cancer	**102-123**
1.	Introduction	**103-105**
2.	Materials and Methods	**105-111**
3.	Results	**111-121**
4.	Discussion	**121-123**
	Chapter VII: Xenobiotic pathway gene polymorphisms associated with gastric cancer in high risk Mizo-Mongoloid population, Northeast India	**124-146**
1.	Introduction	**125-127**
2.	Materials and Methods	**127-132**

3. Results **133-142**

4. Discussion **143-146**

Chapter VIII: Summary and Conclusion **147-156**

Bibliography **157-188**

Conference/Seminars/Workshops Attended **189-191**

List of publications/ Communicated paper **192-195**

List of Figures

Page No.

Chapter III: Mitochondrial haplogroup analysis from Mizo tribal population of Northeast India

1. **Figure 3.1.** Average base quality for the sample MVC and MGC. (A) R1 reads of MVC_WGS sample, (B) R2 reads of MVC_WGS sample, (C) R1 reads of MGC_WGS sample, (D) R2 reads of MGC_WGS sample, (E) Mapping quality distribution of all aligned reads. **13**

2. **Figure 3.2**. F1 haplogroup distribution throughout Asian continent. **15**

3. **Figure 3.3.** Mitochondrial genome map. Map of genes in the Mizo-Mongoloid mitochondrial genome (16569 bp). It represents one of the two possible master circles, Blast1 for MVC and Blast2 for MGC. Blue lines indicate CDS, red lines indicate ORF. The genes with blue text color outside of the circle are transcribed in clockwise direction; those with red text color outside the circle are transcribed in counter clockwise direction. **16**

4. **Figure 3.4.** Complete mitochondrial genome of Mizo F1d2 and their associated SNPs. Each dot in the picture represents one SNP. **21**

5. **Figure 3.5.** Novel variation of Mizo F1d2 mitochondrial genome**.** **21**

6. **Figure 3.6.** Median joining phylogenetic tree of haplogroup F1 complete mitochondrial genomes. A haplogroup sequence R sequence was chosen as the root of the tree. Mutations are reported according to the Revised Cambridge Reference Sequence. ‘‘d’’ represents deletions, ‘‘i’’ represents insertions, “underline” represents Mutations that are reversions to an ancestral state and “<” represents Mutation that are parallel to an ancestral state. Novel haplogroup highlighted with yellow box. **23**

7. **Figure 3.7.** Time calibrated phylogenetic tree of 41 human F1 haplogroup complete mitochondrial genome sequences. A BEAST phylogeny with posterior probability density estimates of the time of most recent common ancestor (TMRCA) of F1 haplogroup sequences. Different color shown different sub haplogroup of hg F1. **24**

8. **Figure 3.8.** Network representation of F1 Complete mitochondrial sequences in modern East and Southeast Asian populations. Each haplotype is represented by a circle. The haplotypes are colour-coded according to their division. Each section of the circle represents one individual sampled from a single population. Mutations are all substitutions and are reported according to the reconstructed phylo-tree. **25**

9. **Figure 3.9.** Bayesian skyline plots. The y-axis for each plot is the product of the effective population size and the generation time. (A) Bayesian Skyline plots of haplogroup F1. The bold blue line represents the median posterior effective population size through time. The blue field delimit the 95% highest posterior density for effective population size, accounting for uncertainty in there constructed phylogeny and substitution model parameters. (B) plots the sum of the separate lineage population estimates through time. Effective population size is plotted on a log scale and assumes a generation time of 30000 years ago. These estimates of effective population size have an inverse relationship with the evolutionary rate of mtDNA used for the calibration, such that they will be lower for faster rates and higher for slower rates. **26**

Chapter IV: Study of Mitochondrial control region, Cytochrome Oxidase C subunit I sequence variations and Mitochondrial Microsatellite Instability associated with gastric cancer

10. **Figure 4.1.** (A) Mitoseek circos plot of all the somatic mutation in mitochondrial control region and MT-CO1 region of gastric cancer samples. Gray dots represent position and frequency of the mutations. (B) Mutation position in the different domain of the mitochondrial control region. **43**

11. **Figure 4.2.** (A) 3-D structure of the mutation in gastric cancer samples. Both the wild-type and mutant side chain are shown in green and red, respectively. The rest of the protein is shown in grey. (A) Mutation of Isoleucine to Asparagine at position 162 due to 6388T>A, (B) Mutation of Phenylalanine to Leucine at position 164 due to 6395C>G, (C) Mutation of Tryptophan to Glycine at position 236 due to 6609T>G, (D) Mutation of Tryptophan to Arginine at position 396 due to 7089T>C. **48**

(E) Mutation of Serine to Tryptophan at position 478 due to 7336C>G, (F) Mutation of Histidine to Glutamine at position 429 due to 7190C>A, (G) Wild and mutant type COXI protein stand, helix, transmembrane region and protein binding region.

12. **Figure 4.3.** (A) Cell cycle histogram of the adjacent normal representative sample, (B) Cell cycle histogram of the representative same Tumor samples showed aneuploidy. Comparison of mitochondrial respiratory enzymes activity. (A) Complex I activity. (B) Complex II activity. (C) Complex III activity. (D) Complex IV activity. The data represent the means SD for three separate experiments. a, ab and b represented a significant decrease compared with the wild-type, using Duncan test (Anova) for data analysis. The P value for (A) is 0.004, (B) is 0.072, (C) is 0.017 and (D) is 0.006. **50**

13. **Figure 4.4.** (A) Representative example of mitochondrial microsatellites instability (mtMSI) at the (C)n, (CA)n for D-loop and $(C)_6$ for ND1 and $(C)_6$ for ND5 SSCP gel pictures. Black arrows indicating mobility-shift band in gastric tumor tissue. (B) Representative Electropherograms obtained by direct sequencing of the mitochondrial D-loop C_7TC_4 and C_6TC_4, COXI$(A)_7$,ND1$(C)_6$, ND5$(C)_5$, $tRNA^{ser}(A)_6$ and $tRNA^{ser}(C)_6$. **53**

14. **Figure 4.5.** (A) Circular association plot between mitochondrial microsatellites instability and demographic factors in gastric cancer, created using Circos with custom settings. The circos plot depicts frequencies different factor and mtMSI status in the inner ring and association of MSI status and different factors in the outer ring. Each factor and MSI status has been assigned different color. The area of each colored ribbon depicts the frequency of the different factor associated with particular mtMSI status; (B) Interaction entropy graph to find the gene-environment interaction with Gastric cancer risk. The percent of the entropy for and mtDNA mutations and potential confounder's interactions are represented in the graph. Bar wide shows frequencies of sample. Positive percentage of entropy denotes synergistic interaction while negative percentage denotes redundancy. Here, the red color denotes the high degree of synergistic interaction, golden color denotes the moderate degree of synergistic interaction, orange color represents minimum degree of synergistic interaction, green color represents minimum redundancy interaction while blue color **57-58**

denotes the highest redundancy.

15. **Figure 4.6.** (a) Microscopic view of gastric adenocarcinoma of gastric tumor cells and adjacent normal cell. (A) Positive high immunoexpression of anti-MTCOI antibody in normal cells (B) Positive moderate immunoexpression of anti- MTCOI antibody in cancer cells (C) moderate immunoexpression of anti-MTCOI antibody in adjacent normal cells (D) negative or less immunoexpression of anti-MTCOI antibody in same tumor cell (from moderate immunoexpression adjacent normal cell), represented by the brownish colour in the cytoplasm and membrane. (b) Representative samples of mitochondrial COXI gene and expression study, *MT-COXI* (455 bp) expression in 10% polyacrylamide gel, β-Actin (392 bp) as control and COXI protein (57kda); (c) Heat map representation of MTCOXI gene differential gene and protein expression and cluster analysis. **59**

Chapter V: Characterization of novel Missense and splice site mutations in TP53 gene and their genetic and functional analysis in gastric tumorigenesis

16. **Figure 5.1.** The c.693-2A>G mutation in the mutant induces a fully penetrant splicing defect. (A) cDNA sequence of exon 5 and 6; (B) cDNA sequence of the exon 5 and 6 after 24 nucleotide deletion; (C) amplified product for tumor and normal samples (WT- Wild type having normal 258 bp allele, MT- Mutant type having heterozygyous 234 bp allele); (D) genomic DNA sequence and the potential splice site deep intronic mutation (c.693-2A>G) (E) Overview about the localization of the mutation and its consequence upon transcript and protein domain processing. **91**

17. **Figure 5.2.** DNA content analysis in gastric cancer samples measured by the flowcytometry. Histogram of A) Diploid Normal sample with G0/G1 peak (red), S-phase (shaded peak) and G2/M peak (green) B) Diploid gastric cancer tumour sample C) Aneuploid gastric cancer tumour sample with aneuploidy peak (yellow). **93**

18. **Figure 5.3.** (A) Positive high immunoexpression of anti-TP53 in Normal cell (B) Positive high immunoexpression of anti-TP53 antibody in tumor cell (C) Positive moderate immunoexpression of anti-TP53 antibody in tumor cell (D) Positive low **93**

immunoexpression of anti-TP53 antibody in tumor cell (E) Negative immunoexpression of anti–TP53 antibody in tumor cell, represented by the brownish colour in the nucleus. (F) Positive moderate immunoexpression of anti-TP53 antibody in adjacent normal cell (from negative immunoexpression cancer cell).

19. **Figure 5.4.** Representative tumor samples of TP53 (121 bp) expression in 10% polyacrylamide gel. Densitometric quantitative band intensity for TP53 and β-actin (392 bp) amplified products. **94**

20. **Figure 5.5.** (A) Positive high immunoexpression of anti-MDM2 in Normal cell (B) Positive high immunoexpression of anti-TP53 antibody in tumor cell (C) Positive moderate immunoexpression of anti-TP53 antibody in tumor cell. **94**

Chapter VI: Novel mutations and inactivation of the E-Cadherin Gene in diffuse Type Gastric Cancer

21. **Figure 6.1.** (A) Positive high immunoexpression of anti-CDH1 in Normal cell; (B) Positive high immunoexpression of anti-CDH1 antibody in tumor cell (Signet ring cell); (C) Positive moderate immunoexpression of anti-CDH1 antibody in tumor cell; (D) Negative immunoexpression of anti–CDH1 antibody in tumor cell represented by the brownish colour in the nucleus (Signet ring cell); (E) Positive low immunoexpression of anti-CDH1 antibody in adjacent normal cell (from negative immunoexpression cancer cell). **113**

22. **Figure 6.2.** Representative tumor samples of CDH1 (262 bp) expression in 10% polyacrylamide gel. Densitometric quantitative band intensity for CDH1 and β-actin (392 bp) amplified products. T1-T4: Representative tumor samples; N1 - N3: Representative adjacent normal samples. **114**

Chapter VII: Xenobiotic pathway gene polymorphisms associated with gastric cancer in high risk Mizo-Mongoloid population, Northeast India

23. **Figure 7.1.** (A) Multiplex PCR amplified products of GSTT1 (459 bp), GSTM1 (219 bp) and internal control gene albumin (349 bp); (B) 16SrRNA (109 bp) and CagA (400 bp) region of *H. pylori* amplified by Multiplex PCR; (C) GSTP1 (433 bp) **134**

expression in 10% polyacrylamide gel, β-Actin (392 bp) as control.

24. **Figure 7.2.** Recevier operating Characteristic curve based on **(A)** multiplex analyses of GSTM1, GSTT1 and GSTP1 polymorphisms; **(B)** GST polymorphism associated with food habits; **(C)** GST polymorphism associated with *H. pylori*. **137**

25. **Figure 7.3.** (A) Interaction entropy graph to find the gene-environment interaction with Gastric cancer risk. The percent of the entropy for independent factors as well as their interactions are represented in the graph. (B) Summary of the two SNPs GSTP1 rs1695: A>G (A/G and G/G) and GSTP1 rs1138272: C>T (C/T and T/T) relation with GSTP1 gene expression in MDR analysis. Black bar height shows GSTP1 expression. **140**

26. **Figure 7.4.** Heat map and graphical representation of GSTP1 gene expression relation with two different SNPs and their genotyping. **142**

List of Tables

Page No.

Chapter III: **Mitochondrial haplogroup analysis from Mizo tribal population of Northeast India**

1. **Table 3.1.** Nucleotide variations in Mizo population using the rCRS **20**
2. **Table 3.2.** SNPs used to detect novel variations in mitochondrial genes of Mizo ethnic population **20**

Chapter IV: **Study of Mitochondrial control region, Cytochrome Oxidase C subunit I sequence variation and Mitochondrial Microsatellite Instability associated with gastric cancer**

3. **Table 4.1.** Primers used in mtDNA amplification and pathogen identification. **33**
4. **Table 4.2.** List of primers for amplification of each repeat microsatellites sequences **33**
5. **Table 4.3.** Demographic and clinical characteristics of the samples **40**
6. **Table 4.4.** Demographical and Clinicopathological features of Gastric Cancer Patients in Mizoram (Stratified by Age) **40**
7. **Table 4.5.** List of mutations with frequency in D-loop region of gastric tumor samples **41-42**
8. **Table 4.6.** List of mutations and their pathogenicity in cytochrome c oxidase subunit I region of gastric tumor samples **46-47**
9. **Table 4.7.** ROS level and cell cycle in tumor, gene mutation group and control group **55**
10. **Table 4.8.** Association tables between mitochondrial microsatellites instability and demographical and Clinicopathological features of gastric cancer patients and healthy control in Mizoram. **55**
11. **Table 4.9.** Association tables between mitochondrial microsatellites instability and gender of gastric cancer patients and healthy control in Mizoram. **56**

Chapter V: Characterization of novel Missense and splice site mutations in TP53 gene and their genetic and functional analysis in gastric tumorigenesis

12. **Table 5.1.** Primers designed in the present study for amplification of exons and splice sites in TP53 gene **81**

13. **Table 5.2.** Mutations of TP53 gene and their pathogenicity in gastric tumor samples. **88-90**

Chapter VI: Novel mutations and inactivation of the E-Cadherin Gene in diffuse Type Gastric Cancer

14. **Table 6.1.** Primers designed in the present study for amplification of exons and splice sites in CDH1 gene **108**

15. **Table 6.2.** E-cadherin expression in primary gastric cancer correlated with various histopathological classifications. Staining intensity E-cadherin-positive cells was evaluated in comparison to the adjacent normal, N= number of tumors of each type investigated. According to the WHO system: papillary, tubular, and mucinous tumor types were combined under the term "Adenocarcinoma". **115**

16. **Table 6.3.** Mutations of CDH1 gene and their pathogenicity in gastric tumor samples. **118-120**

Chapter VII: Xenobiotic pathway gene polymorphisms associated with gastric cancer in high risk Mizo-Mongoloid population, Northeast India

17. **Table 7.1.** Primers tested for polymorphism pattern in the present study **129**

18. **Table 7.2.** Genotype pattern for GST gene polymorphism in the samples **135**

19. **Table 7.3.** Mediation effect of *H. pylori* infection and their association with GST gene polymorphism **138**

Chapter I

Introduction and Review of literature

Stomach cancer is the second-most common cancer among men and third-most among women in Asia and worldwide (Ferlayet al.,2010). There are reports that the gastric cancer has the worldwide variation incidence. A high incidence of gastric cancer has been reported from Southeast Asia, mainly from Japan, China and South Korea (Albertset al.,2003). The literature on incidence of gastric cancer in India and its comparison with the global incidences is in scanty. The age-adjusted rate [AAR] of gastric cancer among urban registries in India is [3.0–13.2] compared to the worldwide AAR [4.1–95.5] (Yeole, 2008; Satyanarayanaet al.,2008; Rastogiet al.,2008). Worldwide and more so in the developed world, there has been a decline in the incidence of gastric cancer and this has been attributed to improved food hygiene, sanitation, and food preservation techniques. However, this declining trend has not been seen in certain parts of India (Pavithranet al.,2002). Differences in some dietary pattern and use of tobacco and alcohol have been considered as potential risk factors. In a case–control study from Trivandrum, a high consumption of rice and chili and consumption of high-temperature food found to be independent risk factors for gastric cancer in multivariate analysis (Mathew et al.,2000). The decrease in the incidence of stomach cancer is associated with the improving standard of living, which results in changes in dietary habits. There is an opinion that adequate intake of certain vitamins decreases the probability of contracting stomach cancer (Yeole, 2008). North-Eastern India is known as one of the most diverse regions in Asia ethnically and culturally and inhabits more than 166 tribes. The year 2006-2008 report of National Cancer Registry Program, ICMR reveals that in India the highest incidence of gastric cancer is in Aizawl, Mizoram which is far

more than the other districts of the Mizoram state, the report says the state of Sikkim ranks second in terms of the gastric cancer incidences (NCRP-ICMR, 2007 and 2010).

The Mizo populations share their ancestral origin in China (Vumson, 1986) and have the very high tobacco smoking rate (Chaturvediet al.,1998). A peculiar habit of using "tuibur" [tobacco smoke–infused water] has also been observed in Mizoram. The habit of chewing betel quid, containing fresh betel nut, slaked lime wrapped in betel leaf is also consumed in Mizoram (Kaushalet al.,2010). "Khaini" is a locally used tobacco product that contains powdered leaf blended and mixed with fine lime paste is also consumed extensively and many scientific reports say that this kind of tobacco use including smoking is highly associated with stomach cancer (Phukanet al.,2005). An international comparison of Age adjusted rate [AAR] with that of Population Based Cancer Registries [PBCR] in India shows that, the major cities like Delhi, Mumbai, Chennai, Pune, Kolkata comes within the top 20 high incidence area. District wise comparison of Minimum age adjusted rate [MAAR] with that of PBCR's under National Cancer Registry Programme [NCRP] shows that Aizawl ranks highest in cancer incidence in India with a rate of 217.9 per 100,000. According to PBCR 2006-2008, there are high incidence of stomach, lung and esophagus cancer among males and females are effected with breast, cervix uteri, stomach and lung cancer (http://www.ncrpindia.org/ Annual_Reports.aspx).

Cancer is a multifactorial disorder and most of the cancer cases seem to be a result of sporadic mutations occurring due to spontaneous genetic events, environmental events or interaction between genetic and environmental factors. Very less research has been done on the dietary habits of the Mizo people, Stomach cancer in Mizoram has been shown to have a positive association with their food habits i.e. consumption of smoked, dried and salted fish and meat

(Phukanet al.,2006). Gastric cancer is the second most common cancer in the world (Parkinet al.,2003). A focused investigation till date has not been done on genetic susceptibility Mizo population with reference molecular genetics and cancer biology.

Various mutations and genetic polymorphisms have been observed in the mitochondrial genome and nuclear genome among cancer patients. Mitochondrial DNA mutations, insertions and deletions have been observed in many types of human cancer (Chatterjee et al.,2006). The human mitochondrial genome, being one of the smallest mitochondrial genomes found in eukaryotes, consists of 16,569 base pairs. The human mtDNA is a super coiled, double-stranded circular molecule containing 37 genes coding for 13 polypeptides of the mitochondrial electron respiratory chain, 22 tRNAs and 2 rRNAs. A non-coding mtDNA control region contains three hypervariable regions and a displacement [D-loop] region (Pentaet al.,2001). Many mitochondrial DNA mutations have been correlated to various types of cancer like gastric cancer (Zhao et al.,2005), breast cancer (Ye et al.,2010), prostate cancer (Jessie et al.,2001).An increase in reactive oxygen species [ROS] has been observed in cancer cells, which may be due to leakage of electrons from electron transport complexes during mitochondrial respiration. These electrons react with molecular oxygen forming superoxide radicals which are then converted to ROS. A direct correlation between mtDNA mutation and ROS has not been observed so far (Carew and Huang, 2002).

Mitochondrial DNA mutations/polymorphisms, deletions specifically in the non-coding Dloop region and have been found to be associated with gastric cancer (Zhao et al.,2005). Mitochondrial DNA D-loop polymorphism, somatic mutation, and 4977 bp deletion are

correlated with Gastric cancer (Chatterjee et al.,2006). Somatic mutations of the mitochondrial DNA [mtDNA] are common in many human cancers, which maybe a reflection of altered DNA repair mechanisms and also predisposition of mtDNA to free radical damage (Wallace et al.,1999; Hochhauser, 2000; Singh, 2006). mtDNA mutations thus, can serve as an important biomarker for human cancers (Jakupciaket al.,2005; Parr et al.,2006). Although, mtDNA mutation has been frequently observed in human cancer, its use as a biomarker has not been explored much. The mutations can occur anywhere in the mitochondrial genome which makes it different from nuclear genome (Chatterjee et al.,2006; Kagan et al.,2005). Eight somatic mutations in Mitochondrial genome were found in gastric cancer patients, among them five point mutations [G3697A, G4996A, G9986A, C12405T and T13015C] are homoplasmic and three mutations [5895delC, 7472insC and 12418insA] are heteroplasmic (Hung et al., 2010) and a polyplasmic 260 bp tandem duplication and triplication mutation in the mtDNA D-loop is also observed in gastric cancer cases (Hung et al., 2008). Mutations occurring in the D-loop region might interfere with transcription of the entire mtDNA genome and may lead to severe alterations in mitochondrial function. Altered mitochondrial physiology has been described in tumor cells and large deletions of the mtDNA genome have been reported in different solid tumor types. The expression of CO I and ND4 mRNA are significantly higher in gastric cancer tissues than in normal gastric tissues. Expression of mitochondrial CO I and ND4 were associated with gastric cancer dedifferentiation and suggested that the detection of mitochondrial CO I and ND4 expression should be evaluated as biomarkers for the diagnosis of gastric cancer. (Jie-Ta et al.,2012).The transcriptional expressions of the ND4 and COI varied remarkably under hypoxia, and the ND4 and COI transcripts were reduced to one fourth and one half, respectively after hypoxia for 24h (Chengboet al.,2009).

Recent research also revealed the association of gastric cancer withgenetic mutations/polymorphisms in the nuclear genes Alcohol Dehydrogenase, Cadherin1 [CDH1] (Christofori and Semb, 1999; Corsoet al.,2013), Tumor protein53 [TP53] (Tolbert et al.,1999), Tumor Necrosis Factor-α [TNF-α] (Yu et al.,2013), Cytochrome P450 2e1 [CYP2E1], Gluthione s-transferase [GST] (Gonzalez et al.,2012), XRCC1 [DNA Repair protein gene] and Methylenetetrahydrofolate reductase [MTHFR] (Ottet al., 2011). Genetic variation is found in alcohol dehydrogenase [ADH1A, ADH1B, ADH1C, and ADH7] and aldehyde dehydrogenase [ALDH2] (Duellet al., 2012). CYP1A1, CYP2E1, and GSTM1 gene polymorphisms (Darazyet al.,2011) and increases the risk for gastric cancer. In gastric cancer patients, MTHFR variant [rs1801131] could serve as a potential prognostic marker (Ottet al., 2011). Fitzgerald *et al.* (2010) said that for both men and women, CDH1 mutation carriers have a cumulative risk of gastric carcinoma by 80 years of age of 80%, with a mean age at diagnosis of 40 years. Additionally, women carrying a CDH1 mutation have a 60% lifetime risk for developing lobular stomach and breast cancer. The human E-cadherin function is to suppress cell invasion; in fact its deregulation is correlated with the infiltrative and metastatic ability of the tumor (Takeichi, 1993), with the consequent loss of cell adhesion and concomitant increase in cell motility. Ihsanet al.(2011) worked on the role of p53 codon 72 polymorphism related with this cancer on the population of NEI. Gopal *et al*. (2012) reported the down regulation of Endo-sulfatasesulf 1 protein expression in gastric cancer. Mir et al.(2012) stated the association between p16, hMLH1 and E-cadherin genes promoter hypermthylation and intake of local hot salted tea and sun-dried foods in Kashmiris with gastric tumors. Tripathiet al. (2011) reported that gastric cancer is associated with reduced GST activity.

A study on the mutations in the mitochondrial and nuclear DNA related with gastric cancer in the Mizo population was investigated for the better understanding of the genetic makeup of the population leading to early diagnosis, proper treatment and counseling to the patients well in time.The proposed work is a part of an effort to determine whether frequent mutations in mitochondrial and nuclear DNA exist in gastric cancer specimens or not with the long-term goal of determining whether the quantity or nature of the mutations can be used to determine the effective therapy of the patients.

Chapter II

Aim and Objectives

The following aims are set forth to carry out the proposed work in this study:

- Study of Mitochondrial control region, Cytochrome Oxidase C subunit I sequence variations and Mitochondrial Microsatellite Instability associated with gastric cancer.

- To study the Germ line and somatic mutations and polymorphic variation of E-cadherin [CDH1] and Tp53 genes associated with gastric cancer.

- To study the relevance of GSTM1, GSTT1 and GSTP1 gene polymorphisms to gastric cancer susceptibility and phenotype.

Chapter III

Mitochondrial haplogroup analysis from Mizo tribal population of Northeast India

1. Introduction

It has been studied that Eastern and Southern Asia differ strongly in their human mtDNA with only limited zones of admixture (Tolk et al., 2001). This haplogroup was first defined as group A by Ballinger et al. (1992), and later renamed as F by Torroni et al. (1994). Mitochondrial haplogroup F is thought to have diverged from haplogroup R around 50,000 years ago in East Asia, spreading throughout East and South East Asia (Soares et al., 2009; Kivisild et al., 2007).

Most of the hg F diversity is found today in native populations of China, Japan, Tibet and Nepal. Phylogenetic analyses of complete mtDNA genome revealed four major sub-clades of hg F, termed as F1 (China, Japan, Tibet, Nepal), F2 (China, Japan), F3 (Island South East Asia), and F4 (Taiwan, Philippines, Indonesia, China, Uzbekistan) (Hill et al., 2006). Further, hg F1 is sub divided into 7 different groups: F1a (South East Asia), F1b (China, Mongolia, Kazakhstan), F1c (China, Japan, Tibet), F1d (China, Japan, Nepal), F1e (China, Japan), F1f (Japan) and F1g (China, Tibet) (Hill et al., 2006). The hypervariable region I and II (HVR-I and HVR-II) motifs: G16129A, T16172C, T16304C, T16519C (HVR-I; according to the Revised Cambridge Reference Sequence,rCRS) (www.Phylotree.org) and A73G, C146T, C152T, C195T, A249d and A263G (HVR-II) denoted the basal F1 haplotype.

Mizoram is one of the eight states of North Eastern Indian Region, covering an area of 21,087sq km with a small population of 1,09,0000 as per 2011 census. Mizo is a collectivenoun, consisting of eight distinct tribes like Lushai, Ralte, Hmar, Paite, Pawih, Lakher (Mara) with many other sub-tribes(Singha, 2012).Like other tribes in the North Eastern India, the origin of the Mizo's is not fully understood. It is commonly accepted that Mizos are the part of a great Mongoloid wave of migration from China and later moved into India to their present habitat. The firstMizo group migrated to India were known as Kukis,

the second immigrants were called New Kukis and the last Mizo tribe immigrants were Lushais (Singha, 2012).

In this study, we analysed the mtDNA genome from the Mizo sample and compared with all currently available hg F1 lineages. This was performed to determine the population specific polymorphisms that can be later used to eliminate the common variations (found in Mizo population) during Cancer mutation analysis.

2. Materials and methods

2.1. Sample collection and DNA extraction

Our initial study consisted of 50 subjects (22 – Male and 28 – Female) with self-identified and Mizo-defined ethnic group from Aizawl, Mizoram, India. Subjects were all healthy individuals and were chosen between 18-50 years old to minimize the variability due to growth and aging. Proper ethical approval from the institutional review board were obtained and all individuals signed a consent form. Genomic DNA was isolated from the peripheral blood using standard method and stored in -20^0 C deep freezer until further use (Ghatak et al., 2013).

2.2. Massively parallel sequencing of mitochondrial genome and sequence analysis

Two Mizo ethnic individuals were selected who had Mizo grandparents for complete mitochondrial genome sequencing using illumine platform. Illumina Paired-End (Illumina Inc., San Diego, CA) library preparation kit was used for preparation of DNA libraries. The next generation tool Fastqc was used to check the read quality, raw sequencing data and was re-confirmed in an internal Galaxy serveras well (Figure 3.1) (Giardine et al., 2005). Sequencing reads were mapped to the hg19 reference genome (NCBI GRCh37;AF347015.1 [16571 bp]/NC_012920.1 [16569 bp]) using the Bowtie

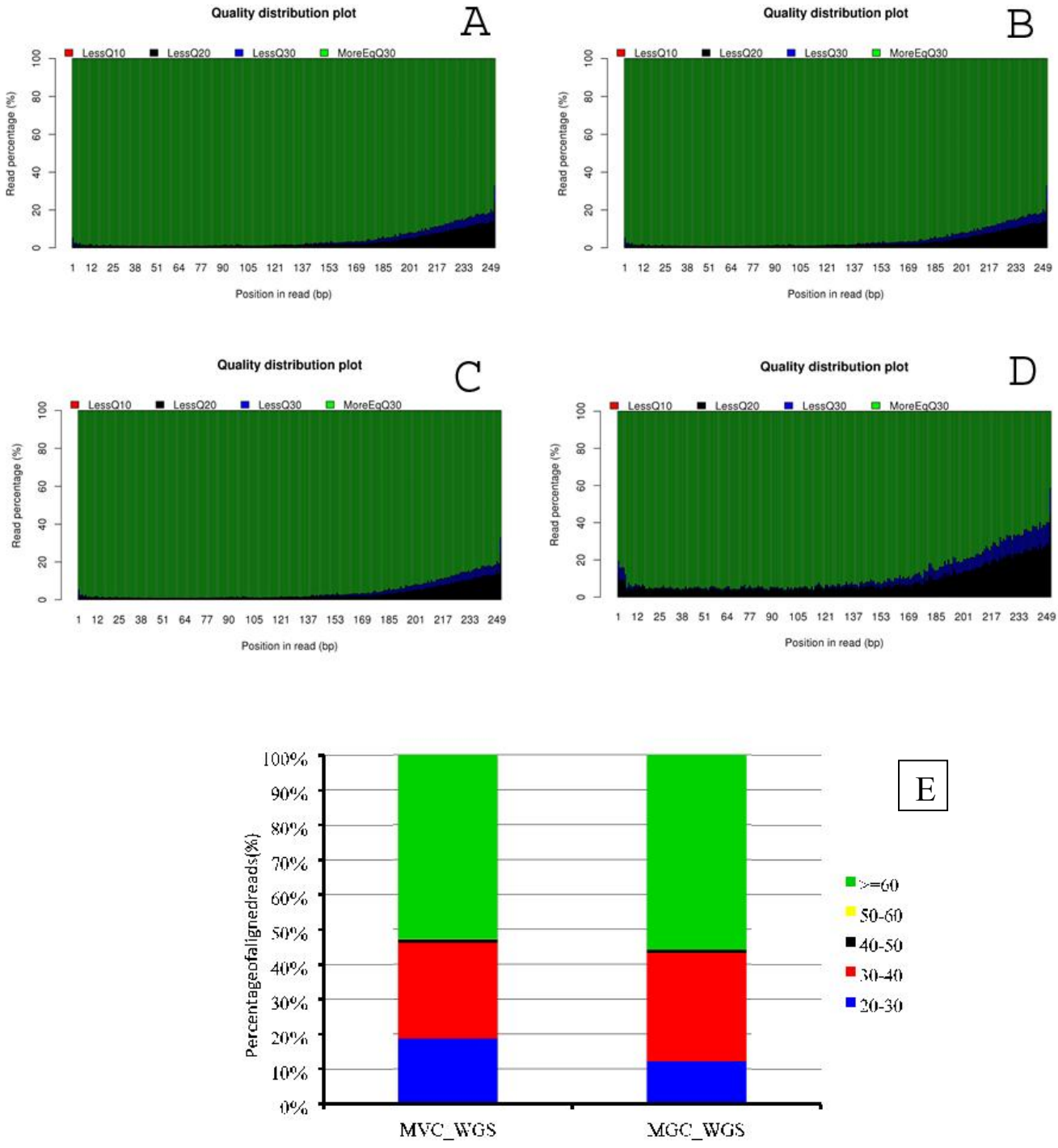


Figure 3.1. Average base quality for the sample MVC and MGC. (A) R1 reads of MVC_WGS sample, (B) R2 reads of MVC_WGS sample, (C) R1 reads of MGC_WGS sample, (D) R2 reads of MGC_WGS sample, (E) Mapping quality distribution of all aligned reads.

alignment package (0.12.7). The Integrative Genomic Viewer was used to visualize and analyse the BAM files and "Filter pileup" tools were used to call Single Nucleotide Polymorphisms (SNPs) from the generated pileup files (Robinson et al., 2011). A quality threshold Q ≥ 30 was set for the potential variant and the SNP were called if it was present at least in one read. In addition, the MITOBAM annotator tool was used to detect any variant present in at least one read and with a Q ≥ 30 for annotation. Potential variants were checked with Mitomap mtDNA Sequence Polymorphism database (http://www.mitomap.org/MITOMAP)(Ruiz-Pesini et al., 2007). All sequence data has been submitted to NCBI Gen Bank (Accession No - SRR2075910, SRR2075911).

2.3. Mitochondrial haplogroup identification

Complete genome sequences were used for the mitochondrial haplotype estimation by using Haplogrep server to detect mitochondrial DNA profiles based on Phylotree build 16 (VanOven and Kayser, 2009). The estimated haplogroup were verified using manual checking of Phylotree build 16 from www.Phylotree.org database. We compiled a dataset of 40 complete hg F1 Asian mtDNA sequences representing 40 different haplotypes from the reported literatures (www.phylotree.org)(Figure 3.2) (VanOven and Kayser, 2009). All hg F1 sub-haplotype sequences were sampled from database and were aligned in clustalW using MEGA 6 software. Each haplogroup was treated as a separate population. The revised Cambridge Reference Sequence (rCRS) (haplogroup H and commonly found in European population) was used to identify the variations from all the sequences.

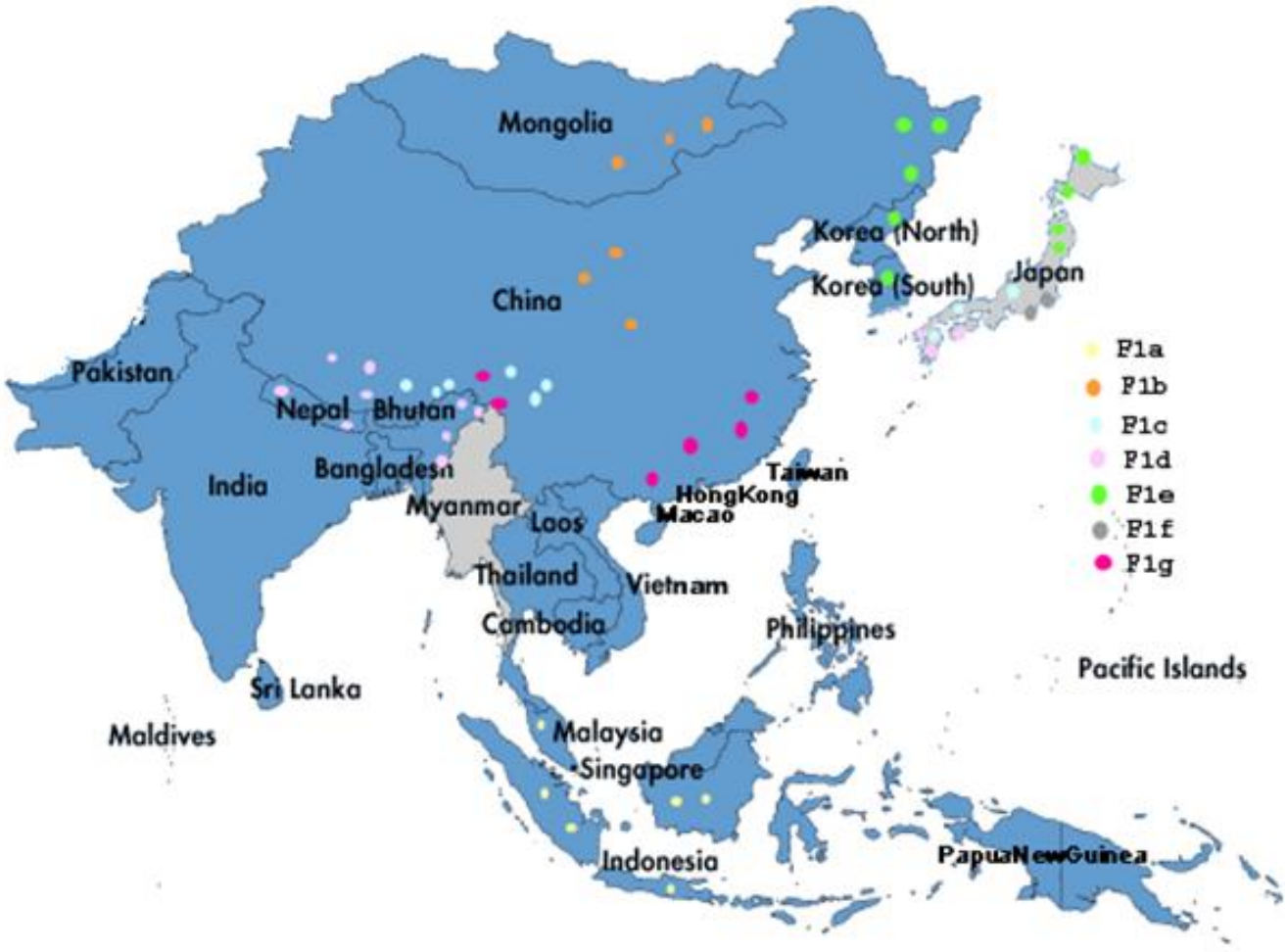


Figure 3.2.F1 haplogroup distribution throughout Asian continent.

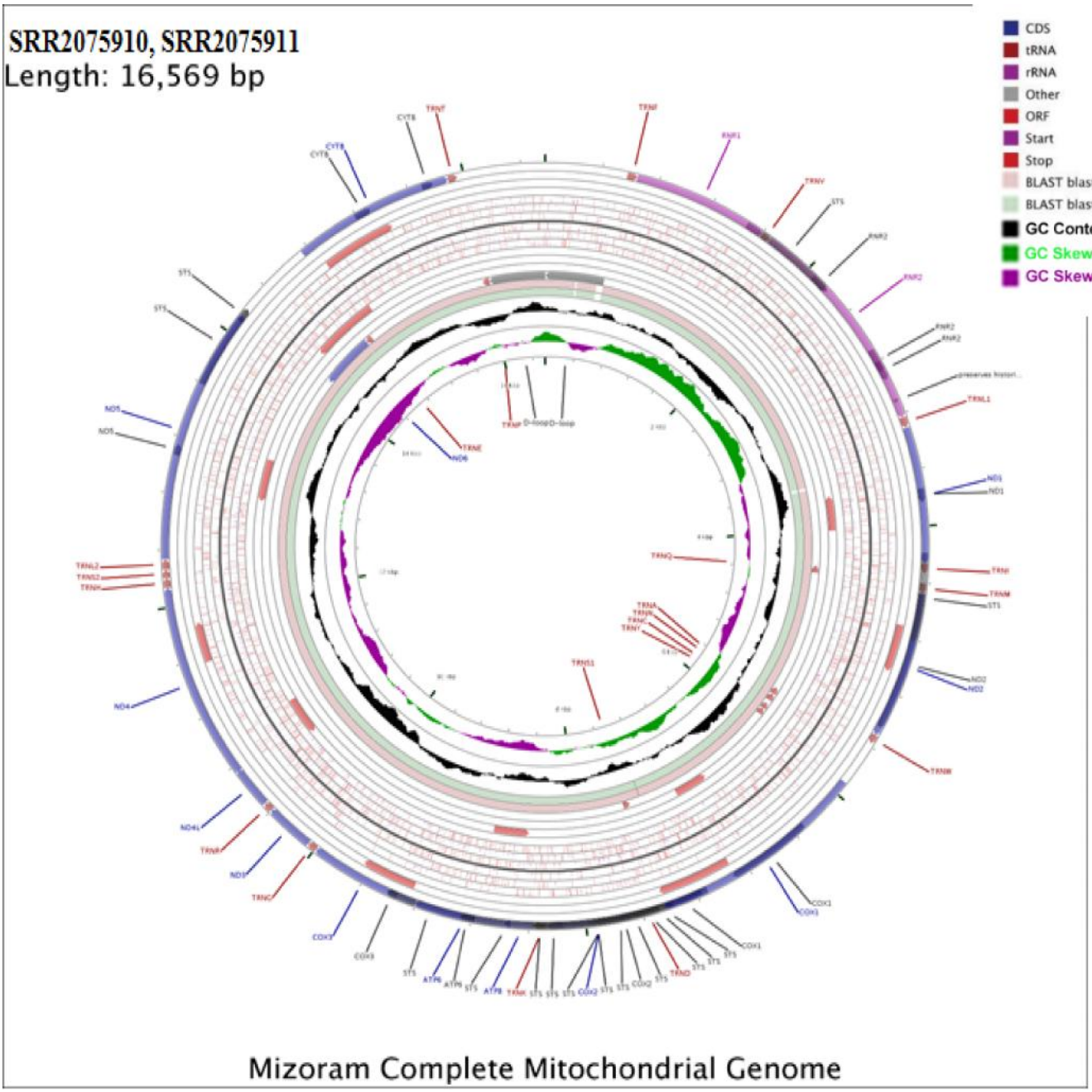


Figure 3.3 Mitochondrial genome map. Map of genes in the Mizo-Mongoloid mitochondrial genome (16569 bp). It represents one of the twopossible master circles, Blast1 for MVC and Blast2 for MGC. Blue lines indicate CDS, red lines indicate ORF. The genes with blue text coloroutside of the circle are transcribed in clockwise direction; those with red text color outside the circle are transcribed in counter clockwisedirection.

Translated query vs. translated database BLAST searches were used to identify the boundaries of protein coding genes and gene homologies based on the open reading frames with >80 bp in length. The mitochondrial sequences were analyzed by MitoOffice software for separating sequence based on the specific region. The sequence position and final sequence annotation analyses were conducted with the program Artemis v8.0 and the mitochondrial genome and AT GC content were visualized using the program CGView for comparing the genome with Revised Cambridge reference sequence (Figure3.3).

2.4. Parsimony network construction

We reconstructed a median joining statistical parsimony network (by Network 4.613) with the two completely sequenced mitochondrial genomes from this study (MVC and MGC) and all the hg F1 complete mitochondrial genomes. Mutations 309.1C, 315.1C, A16182c, A16183c, 16193.1C and C16519T were scientifically ignored, as these positions were known to characterize mutational hotspots and/or recurrent sequencing artefacts.

2.5. Bayesian skyline plots and maximum parsimony phylogenetic tree

For each of the F1 haplogroups, MCMC sampling algorithm(Hastings, 1970) was used to construct the Bayesian skyline plots (BSPs)(Drummond et al., 2005) for effective population size by using BEAST v. 1.4 (Drummond and Rambaut, 2007). To estimate total Asian hg F1 population size through time, we used constructed BSPs from the sets of 42 sequences from phylotree dataset. The general time-reversible (GTR) substitution model was used with site-specific rates for all the three codon positions to construct the ancestral gene tree for each haplogroup. The MCMC sampling algorithm was based on a run of 40000000 generations, for every 4000 sample, with burn-in discarded for first 4000000 generations. A

new nomenclature based on the van Oven and Kayser with several new modifications was used. The most-parsimonious tree was reconstructed manually for the complete mtDNA sequences and was verified by using mtPhyl 2.8.0.0 software (http://eltsov.org), which is designed to reconstruct maximum parsimony phylogenetic trees of hg F1 variation list.

2.6. Conformation of novel variation by Sanger sequencing

To confirm the novel whole mitochondrial genome variation (listed in Table 3.1), we performed Sanger sequencing with the fifty DNA samples from Mizo ethnic group. The novel variations were analysed using primers specific for six novel variations (Table 3.2). PCR was performed using a thermal cycler as follows: 35 cycles of 40 s at 94°C for denaturation, 40 s at specific annealing temperatures, and 40 s at 72°C for extension (Table 3.2). Final extension was performed at 72°C for 10 min. The PCR products were electrophoresed on an agarose gel and stained with ethidium bromide to confirm the size of the bands. The amplified product was purified and sequenced by gold standard Sanger sequencer. All products were sequenced from the opposite directions to ensure reading accuracy. Sequences and chromatograms obtained were examined by DNA baser and chromas software version 2.13, and aligned by BLAST [www.ncbi.nlm.nih.gov/blast]. All mitochondrial sequences were compared with latest version of Revised Cambridge Reference Sequence [rCRS] of the Human Mitochondrial DNA [NC_012920] and analysis the variation of sequences by Mito Tool Programming and novel variations were checked manually with www.phylotree.org.

3. Results and discussions

The complete mt genome study has identified six novel variations in the Mizo ethnic group: two each in the protein coding and control regions and one in non-coding region (Table 3.1; Figure 3.4). These

variants returned no hits in Mitomap or mtDB database together for any of the haplogroups currently available, thus these positions are deemed to be novel(Ruiz-Pesini et al., 2007). All the novel variations were confirmed by Sanger sequencing (Table 3.2; Figure 3.5). The novel variants were noted in both the individuals in complete mitochondrial genome and in 78% of other Mizo samples (by sanger sequencing) and are thus probably ancestral variations. T152C was observed at the position of replication origin region of H-strand of hypervariable segment2 which can transfer from one generation to other.

A total of 41 haplogroups were observed in our analysis including the novel F1d2. Supplementary Figure 3.5 presents the SNPs and haplogroup frequencies in control region of the studied population. The mt sequence variation of the hg F1d2 individuals based on the rCRS were identified (Figure 3.4). The study identified 8 defining markers of mitochondrial haplogroup F1 which were fixed in all hg F1 individuals.

Six novel variations were detected for the two Mizo F1 individuals and its further substructure (F1d2). The novel sequence motifs were validated by calculation of Θ and π diversity statistics by DNAsp and Mega6 software. The hg F1d group was found to exhibit low haplotype diversity ($\Theta = 0.00052$). Studied hg F1 sub haplogroup revealed high diversity levels signifying the past population expansion based on substantial negative values for Tajima's D neutrality tests.

Sample IDs	Variations	Haplogroup
1. MVC **2. MGC**	A73G, T146C, 249d, A263G, A750G, A1438G, C1734T, A2706G, C3970T, A4769G, T5628C, T6392C, G6962A, C7028T, T7738C, A8860G, G10310A, T10609C, G11719A, G12406A, C12882T, G13928C, C14766T, A15326G, C15402T, T16304C	F1d (pre-defined)
	Additional Variations in Mizo ethnic group: T152C, G5460A, A5894G, A7445G, T16311C, C16189T!!	F1d2 (New Haplogroup)

"!!" double back mutation as per PhyloTree.org - mtDNA tree Build 16

"d" deletion

Table 3.1. Nucleotide variations in Mizo population using the rCRS

Novel SNPs	Primer Sequence	Amplified product Size (bp)	Annealing Temp. (^{0}C)	Location in Mitochondria
T152C	F: GTATGCACGCGATAGCATTGC R: GATGTCTGTGTGGAAAGTGGCTG	210	63	D-Loop (hypervariable segment II, h strand replication origin region)
G5460A	F: CTACTCCCCATATCTAACAACG R: CTGATTTGCGTTCAGTTGATG	253	61	ND2 (Complex I, First nucleotide of codon-331; Ala>Thr)
A5894G	F: CAGGTTTGAAGCTGCTTCTTCG R: CCTAGGACTCCAGCTCATGCGC	223	54	Non coding position between tRNA-Tyr and COXI
A7445G	F: GAGAAGCCTTCGCTTCGAAGCG R: GACCTACTTGCGCTGCATGTGC	290	57	COXI (Complex IV, Third nucleotide in the stop codon, AGA>AGG)
T16311C and C16189T!!	F: GCCACCATGAATATTGTACGG R: GTGGTCAAGGGACCCCTATCTG	294	55	D-Loop (hypervariable segment I)

"!!" double back mutation as per PhyloTree.org - mtDNA tree Build 16

Table 3.2.SNPs used to detect novel variations in mitochondrial genes of Mizo ethnic population

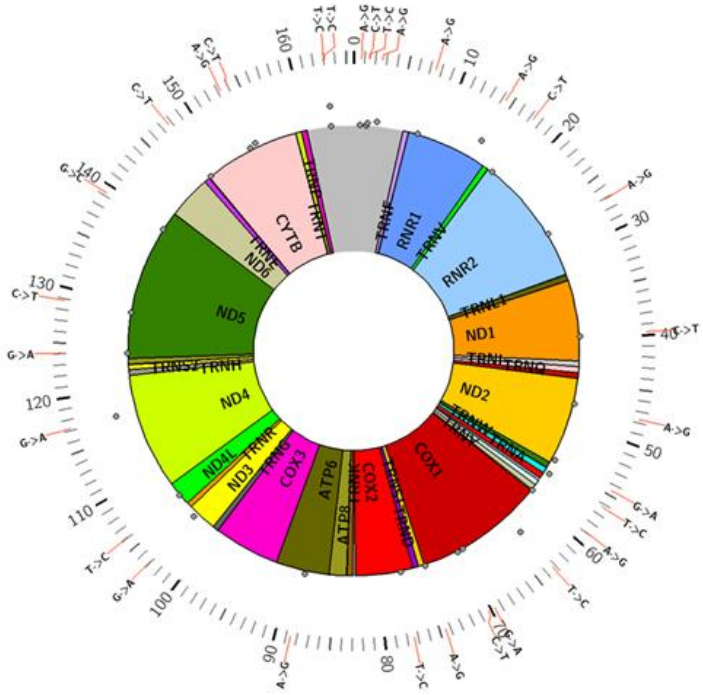


Figure 3.4. Complete mitochondrial genome of Mizo F1d2 and their associated SNPs.Each dot in the picture represents one SNP.

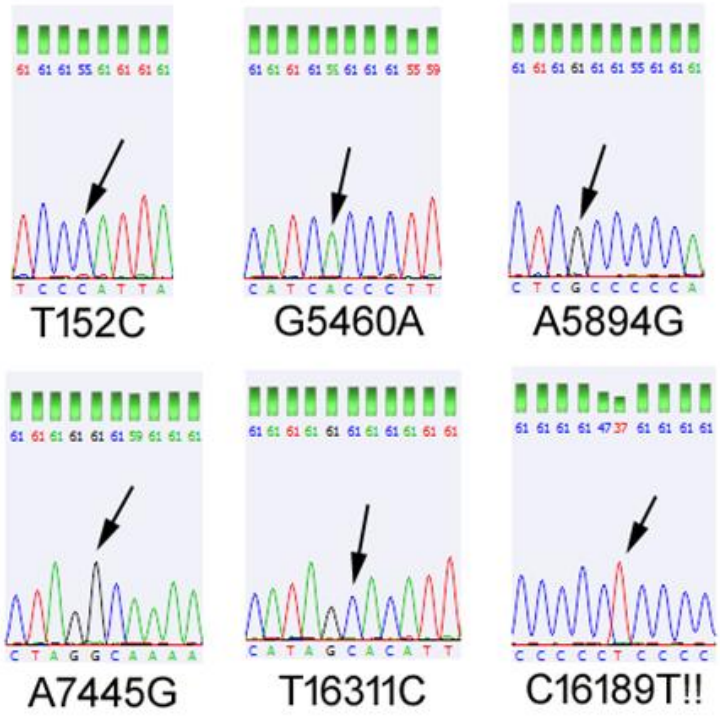


Figure 3.5. Novel variation ofMizo F1d2 mitochondrial genome**.**

Phylogenetic analysis of the F1 sequences in the software mtPhyl confirmed that all Mizo individuals belongs to haplogroup F1d2 (Figure 3.6). A detailed phylogeny was reconstructed with 42 complete sequences to investigate the sub-structure of haplogroup F1 (Figure 3.7). Supplementary Figure 3.7 shows that Mizo hg F1 group clustered in F1d2 position with Japan (F1dand F1d1) mt sequences (Tanaka et al., 2004). This novel haplotype, F1d2, is not currently included in an updated Phylotree (http://www. phylotree.org/) (VanOven and Kayser, 2009). Map of the statistical median joining network of F1 haplotype was showing separate haplotype for F1d2 (represented by pink ellipses) (Figure 3.8). Two haplotypes are joined by a line if they are separated by a median vector; additional median vectors are indicated by small red circle. The six small red circle represents the separation of hg F1d2 by six motifs as a different SNP from the other F1d haplogroups. A large number of haplogroup-specific and haplogroup-associated polymorphisms relative to the complete mtDNA of hg F1 was drawn in this network. The F1a- and F1b-haplotypes are clustered distantly from the F1c and F1d-haplotype sequences, hence the haplogroups F1c and F1d were might the most recent common ancestors of the Asian.

A BEAST phylogeny with posterior probability density estimates of the time of most recent common ancestor (TMRCA) of F1 haplogroup sequences is shown in Figure 3.8. The tree separated from the most common ancestor with two major clades (F1a-F1c-F1f and F1b-F1d-F1e). The hg F1a-F1c-F1f together separated from F1 haplogroup, whereas hg F1b-F1d and F1e together separated from the hg F1 haplogroup. Hg F1b and F1d are the most recent haplogroup in hg F1 tree than any other hg F1 sub haplogroup and both the haplogroup separated in same time from the major haplogroup F1.

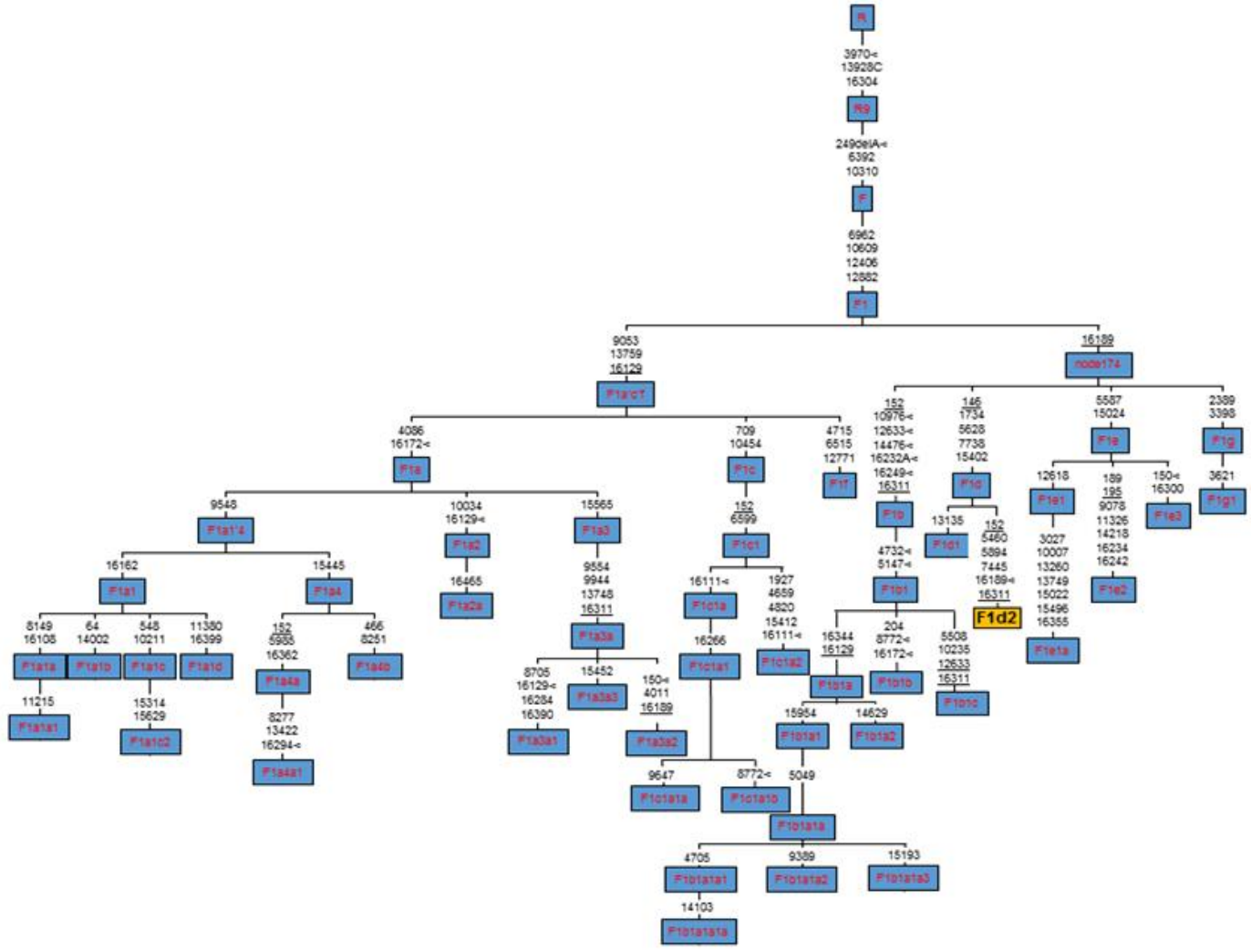


Figure 3.6. Median joining phylogenetic tree of haplogroup F1 complete mitochondrial genomes.A haplogroup sequence R sequence was chosen as the root of the tree. Mutations are reported according to the Revised Cambridge Reference Sequence. ‘‘d’’ represents deletions, ‘‘i’’ represents insertions, “underline” represents Mutations that are reversions to an ancestral state and “<” represents Mutation that are parallel to an ancestral state. Novel haplogroup highlighted with yellow box.

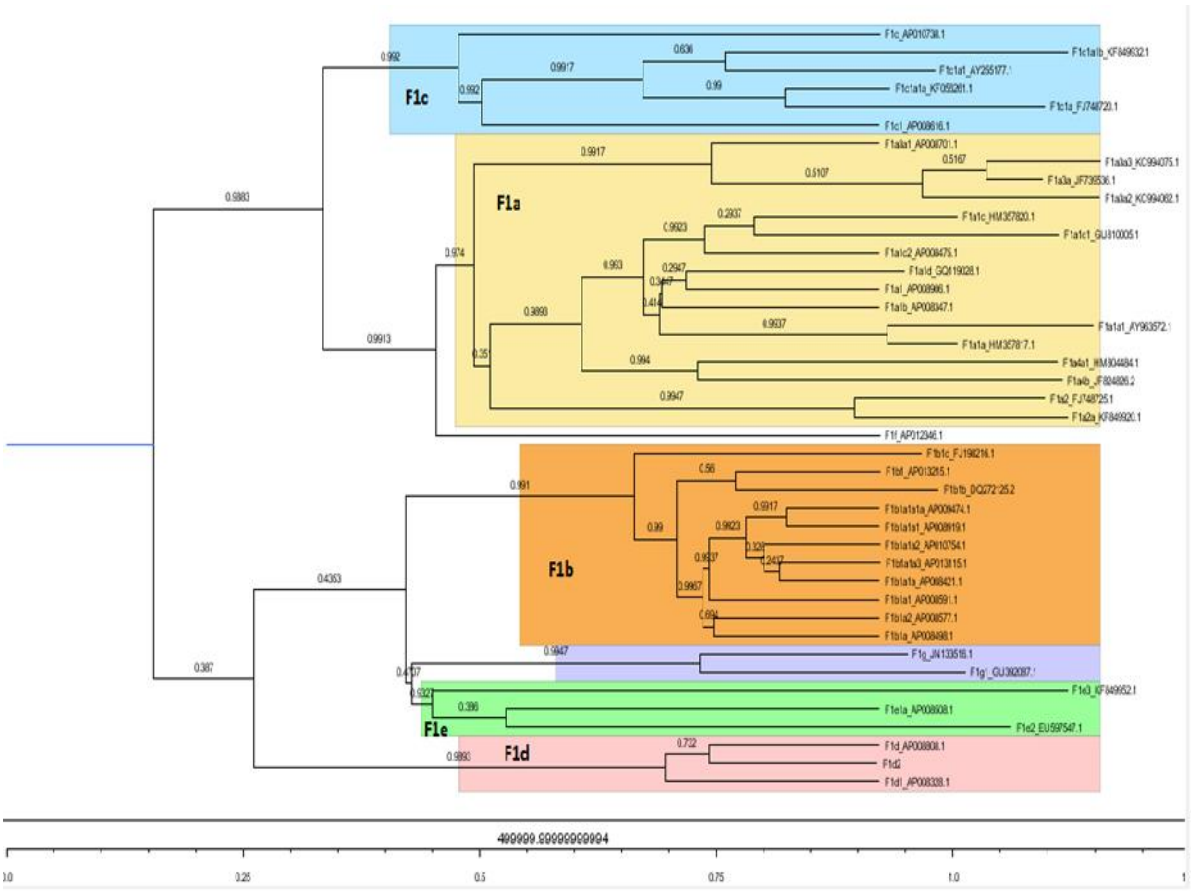


Figure 3.7. Time calibrated phylogenetic tree of 41 human F1 haplogroup complete mitochondrial genome sequences. A BEAST phylogeny with posterior probability density estimates of the time of most recent common ancestor (TMRCA) of F1 haplogroup sequences. Different colour shown different sub haplogroup of hg F1.

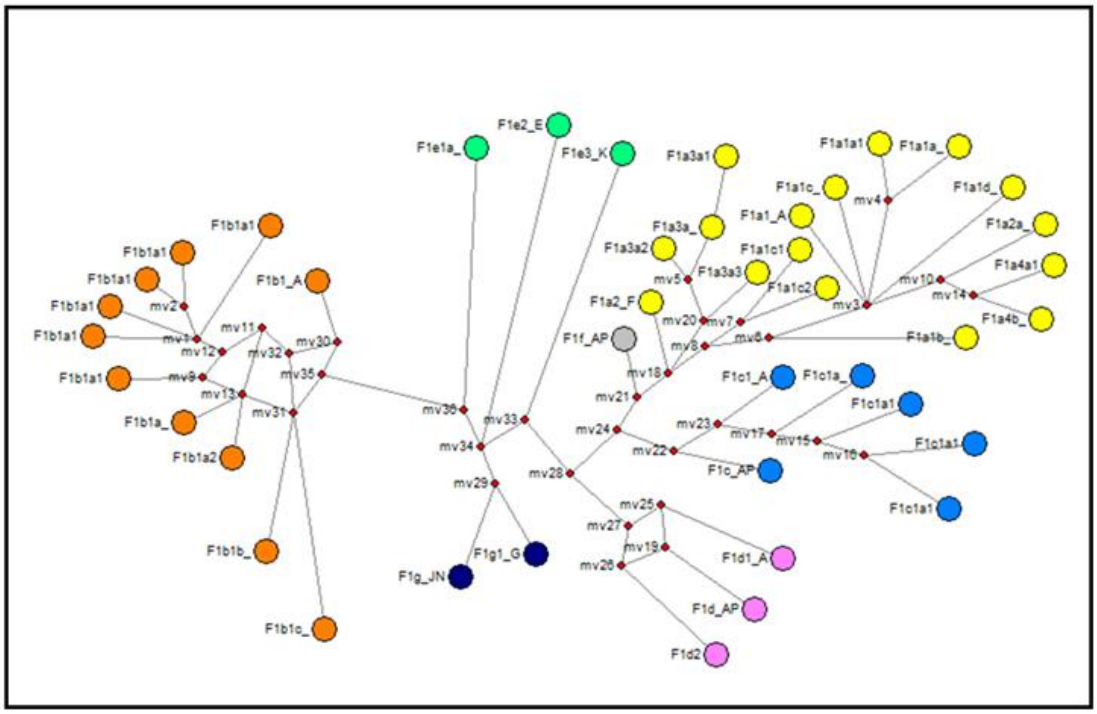


Figure 3.8. Network representation of F1 Complete mitochondrial sequences in modern East and SoutheastAsian populations. Each haplotype is represented by a circle. The haplotypes are colour-coded accordingto their division. Each section of the circle represents one individual sampled from a single population. Mutations are all substitutions and are reported according to the reconstructedphylo-tree.

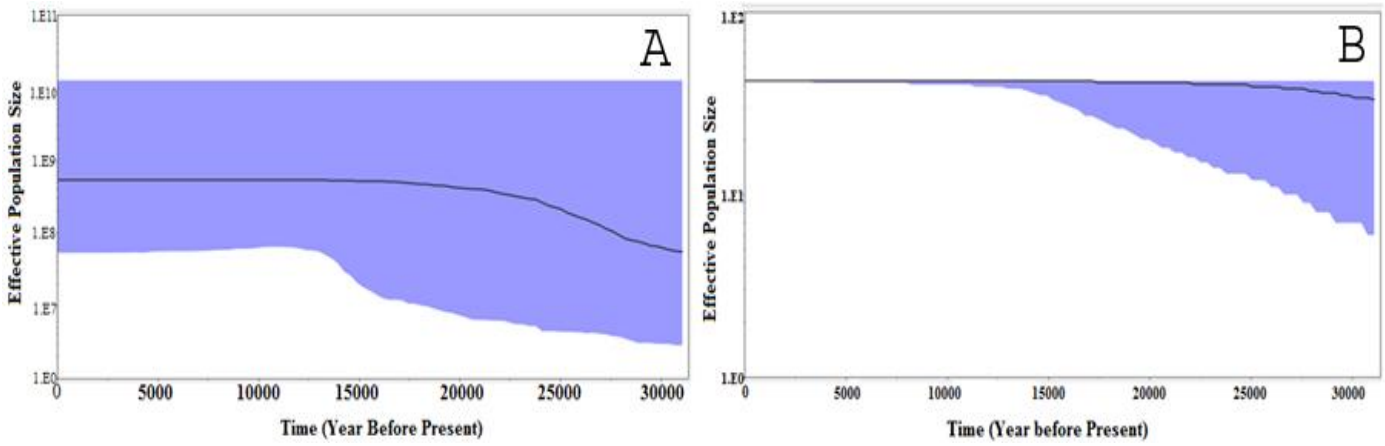


Figure 3.9. Bayesian skyline plots.The y-axis for each plot is the product of the effective population size and the generation time. (A) Bayesian Skyline plots of haplogroup F1. The bold blue line represents the median posterior effective population size through time. The blue field delimit the 95% highest posterior density for effective population size, accounting for uncertainty in thereconstructed phylogeny and substitution model parameters. (B)plots the sum of the separate lineagepopulation estimates through time. Effective population sizeis plotted on a log scale and assumes a generation time of 30000 years ago. These estimates of effective population size havean inverse relationship with the evolutionary rate of mtDNAused for the calibration, such that they will be lower for fasterrates and higher for slower rates.

Hg F1 datasets of skyline plot reveals that the F1-type individuals were 30,000 Year Before present (YBP) and continuing to the present(Hill et al., 2006). The expansion of population was between 15,000 and 20,000 YBP for hg F1 with a median posterior effective population size through time (Figure 3.9). The hg F1 might be primarily present in independent differentiations in different regions or diverse populations, because the skyline plot did not show any population history before 30,000 YBP for all sub types of F1 haplogroup.

This study has identified six novel mtDNA variations: two in the protein coding, two in control region and one variation in noncoding region (Table 3.1 and Figure 3.5). These variants returned no hits in Mitomap or mtDB database, thus these positions are deemed to be novel(Ruiz-Pesini et al., 2007). The novel variants were noted in both the individuals and are thus probably ancestral variations.

Chapter IV

Study of Mitochondrial control region, Cytochrome Oxidase C subunit I sequence variations and Mitochondrial Microsatellite Instability associated with gastric cancer

1. Introduction

Gastric cancer is one of the frequent causes of death worldwide. Earlier studies on gastric cancer patient samples observed large-scale deletion and somatic mutations in mtDNA (Han et al., 2003; Maximo et al., 2001). Mitochondrial D-loop region mutations are frequent in several type of cancerand have been documented as a hot spot region for human cancer (Lee et al., 2004; Yin et al., 2004). The D-loop is involved in the control of replication and transcription of mtDNA, mutations in this particular region might cause a decrease in the copy number and alteration in different protein coding gene expression in the mitochondrial genome (Shadel and Clayton, 1997).However, the relationship between clinicopathologic features of gastric carcinoma and mutation or depletion of mtDNA has remained unclear.

Damage in mitochondrial respiratory complex function not only decreases the supply of energy, which may stop energy-dependent apoptosis, but also increases ROS production that may induce mutation and oxidative damage to mitochondrial DNA (mtDNA) (Cuezva et al., 2002; Isidoro et al., 2004). These findings suggest that there are common mechanisms by which the bioenergetic function of mitochondria is altered in cancer cells. However, only fewresearch have been documented to explain how qualitative and quantitative changes in mtDNA influence the mitochondrial oxidative phosphorylation in cancer tissues.

Microsatellite instability (MSI) is a major manifestation of genomic instability. Nuclear MSI has been reported in many common malignancies in different mismatch repair system. However, any genetically instability in mitochondrial genome in tumor cells (Richard et al., 2000; Bianchi et al., 2001) is less

characterized, and the responsible mechanismforROS generation still remains unclear. To extend these findings, we analyzed the occurrence of mtMSI in 12 different regions of mtDNA in gastric carcinomas to determine: 1)whether mtMSIassociation with demographic factors is responsible for alterations in mtDNA in gastric tumor?, 2)Does mitochondrial MSI in the D-loop and protein coding regions affect mitochondrial- protein expression, function and complex enzyme alterations?, 3) To prove the speculation that some DNA repair mechanisms are potentially involved in maintenance of the normal mitochondrial genome.

In the present study, the mutation of mitochondrial control region (D-Loop) and cytochrome c oxidase subunit I (MT-COXI) associated with gastric cancer has been studied to clarify the influence of these mutations on gene expression, ROS generation and cell cycle progression. We also analyzed the mtMSI in gastric cancer and its premalignant lesions to elucidate whether mtMSI led to the progression of gastric cancer association with demographic factors. The purpose of this study was to investigate the influence of genetic instability on cell carcinogenesis and progression of gastric cancer.

2. Materials and Methods

2.1.Subjects and sample collection

This study consisted of 80 random patientsamples with primary gastric tumour and adjacent non-tumour sites (aged from 37 to 79 years) collected during surgical resection from Civil Hospital, Aizawl, Mizoram and 80 unrelated healthy controls (aged from 31 to 73 years) belonging to the same ethnic group. None of these patients received preoperative chemotherapy or radiotherapy. The samples were all confirmed histologically and all subjects provided informed consent for obtaining the study specimens. The study design and data collection methods have been described previously(Ghatak et al., 2014). Patients with gastric neoplasms including adenocarcinoma (MALT lymphoma, stromal, or carcinoid tumors) were

included. Patients with secondary or recurrent GC, previous history of other malignancies, or refusal to participate were excluded from the study. Medical charts were reviewed using a standard protocol to obtain information on cancer treatment, clinical stage, dietary habits, previous disease history, physical activity, tobacco, and alcohol use by an in-person interview using a structured questionnaire (Ghatak et al., 2014). The study protocol was approved by the Institutional Review Board of all institutes involved in the study.

2.2. DNA extraction from the blood sample

The lymphocytes from whole blood were separated by lysing the RBCs using a hypotonic buffer (ammonium bicarbonate and ammonium chloride, Hi-media) with minimal lysing effect on lymphocytes. Three volumes of RBC lysis buffer were added to the blood sample, mixed by vortexing and inverting thoroughly for 5 min and centrifuged (Eppendorf 5415R, Germany) at 2,000 X g for 10 min. The lymphocytes was used for DNA extraction by modified protocol of Ghatak et al. (2013).

2.3.DNA and RNA extraction from the tissue samples

Deparaffinization was carried out by adding 1 ml of xylene to the tissue section in each microfuge tube followed by vigorous vortexing for 10 minutes. The mixture was centrifuged at 12000 rpm for 10 minutes. The supernatant was discarded and the deparaffinization steps were repeated once again, followed by rehydration through subsequent washings with 100%, 90% and 70% absolute ethanol diluted in RNase free DEPC treated water, respectively. The deparaffinised tissue and surgical resection specimen was used for the DNA and RNAco-extraction by the modified protocol of Ghatak et al. (2015).

2.4.PCR amplification of the mtDNA D-loop and COXI region

The mtDNA D-loop region (1030 bp) was amplified by polymerase chain reaction (PCR) using primers Mt–F (5'-CACCATTAGCACCCAAAGCT-3') and Mt-R (5'-CTGTTAAAAGTGCATACCGCCA-3').

The mtDNA COXI region (700bp) was amplified using Universal barcode primer as well as complete COXI gene (1650 bp) was amplified with Hmt-COF (5'-ATCACCTCGGAGCTGGTAAA-3') and Hmt-COR (5'-CGCTGCATGTGCCATTAAGATA-3') primer set (Table 4.1). The PCR amplification products (3µl) were subjected to electrophoresis in a 1.2 % agarose gel in 1X TBE buffer at 80 Vh for 30 min, stained with Ethidium Bromide, and images were obtained in GBOX gel documentation systems (UK). PCR products were purified with a Qiagen gel extraction kit (Qiaquick columns; Qiagen, Chatsworth, CA) and stored at -20^0 C until for sequencing.

2.5.MT-MSI detection

PCR-single strand conformation polymorphism (PCR-SSCP) was performed to amplify the microsatellite sequence of mtDNA using published primers (Ling et al., 2004). The primer consisted of three D-loop regions, six coding regions and one tRNA (Table 4.2). The reaction conditions and procedures were similar to those reported by Ling et al., (2004) and Hebano et al., (1998). PCR products that showed mobility shifts were directly sequenced using an appropriate internal primer and analyzed using the 3500 automated DNA sequencer (Applied Bio system, USA). All experiments were repeated at least twice to rule out any artifacts. A tumor will be defined as: High-MSI (MSI-H) - if more than 2 of the 10 examined loci showed unequivocal instabilities (at least 20% unstable loci); Microsatellite stable (MSS) - if no microsatellite instability was found; Low microsatellite instable (MSI-L) if less than 2 marker showed instability (20% unstable loci).

Gene	Primer (5' to 3')	Annealing Temperature/ Time	Amplified products/ Alleles
D-loop	F: CACCATTAGCACCCAAAGCT R: CTGTTAAAAGTGCATACCGCCA	63°C - 60 sec	1030 bp
COXI (Barcoding Region)	F: TTATACGACTCACTATAGGGGGTCAACAAATCATAAAGATATTGG R: ATTAACCCTCACTAAAGTAAACTTCAGGGTGACCAAAAAATCA	54°C - 60 sec	650 bp
COXI (Complete ORF)	F: ACAGGATTTGGTACTAGCCT R: AGTGCCTTCACATAGTCATCCTTG	50°C - 40 sec	1730 bp
H. pylori (16srRNA)	F: CTGGAGAGACTAAGCCCTCC R: ATTACTGACGCTGATTGTGC	60°C - 45 sec	109 bp
H. pylori (CagA)	F: AATACACCAACGCCTCCAAG R: TTGTTGCCGCTTTTGCTCTC		400 bp
EBNA3C	F: AGAAGGGGAGCGTGTGTTGT R: GGCTCGTTTTTGACGTCGGC	56°C - 45 sec	153 bp for type 1 EBV 246 bp for type 2 EBV
COXI For expression	F: CTAGCAGGTGTCTCCTCTATCT R: GCTCGTGTGTCTACGTCTATTC		455 bp band
β-Actin Control	F: ACCATGGATGATGATATCGC R: ACATGGCTGGGGTGTTGAAG	54°C- 45 sec	392 bp band

Table 4.1. Primers used in mtDNA amplification and pathogen identification.

Repeat sequence	mtDNA region (nucleotide position)	Gene	Primer sequence (5' – 3')
$(PolyC)_n$	267 - 423	D-loop (1)	F: TCCACACAGACATCATAACA R: AAAGTGCATACCGCCAAAAG
$(CA)_n$	453 - 637	D-loop (2)	F: CCTCCCACTCCCATACTACTAA R: GTGATGTGAGCCCGTCTAAACA
$(PolyC)_n$	16112 -16379	D-loop (3)	F: CACCATGAATATTGTACG R: CAAGGGACCCCTATCTGAGG
$(C)_6$	3397 - 3617	NDI	F: ATACAACTACGCAAAGGCCCCA R: AATAGGAGGCCTAGGTTGAGGT
$(A)_7$	4555 -4644	NDII	F: CCTGAGTAGGCCTAGAAATAAA R: ACTTGATGGCAGCTTCTGTG
$(A)_7$	6644 - 6733	COXI	F: CCTACCAGGCTTCGGAATAA R: ATAGCTCAGACCATACCTATG
$(T)_7$	9431 - 9526	COXIII	F: CCAAAAAGGCCTTCGATACG R: GCTAGGCTGGAGTGGTAAAA
$(C)_6$ and $(A)_8$	12360 - 12465	NDV (1)	F: CACCCTAACCCTGACTTCC R: GGTGGATGCGACAATGGATT
$(CCT)_3 – (AGC)_3$	12940 – 13032	NDV (2)	F: GCCCTTCTAAACGCTAATCC R: TCAGGGGTGGAGACCTAATT
$(A)_6$ and $(C)_6$	7359 - 7606	$tRNA^{Ser}$	F: GAAGAACCCTCCATAAACCTGG R: TACTTGCGCTGCATGTGCCATT

Table 4.2. List of primers for amplification of each repeat microsatellites sequences

2.6.Immunohistochemical analysis

Immunohistochemical study was performed by rabbit polyclonal anti-MT-COXI antibody (ab14705) (Abcam, Japan), which detects the wild-type MT-COXI protein and is recommended for detection of the respective protein in human tissue. As positive control, slides with histological sections previously demonstrated as being positive for these antibodies were used. A similar slide was used as a negative control, subtracting the primary antibody from the reaction (Grace et al., 2002). Staining was recorded as either present or absentwith regard to staining intensity. Extent of staining was graded as: 0, 0-10% of cells positive; 1, 10-50% of cells positive; 2, greater than 50% of cells positive for COX1. Staining was considered positive if the extent of staining was graded as 2. Staining was considered reduced if the extent was graded as 1 and graded as 0.

2.7.Sequence analysis

The samples exhibiting mutation or polymorphism and instability after SSCP analysis was taken for further sequencing and mutation analysis. All products were sequenced from opposite directions to ensure reading accuracy. Sequences and chromatograms obtained were examined by chromas software version 2.13, DNA baser and align by BLAST [www.ncbi.nlm.nih.gov/blast]. All mitochondrial sequences was compared with latest version of Revised Cambridge Reference Sequence [rCRS] of the Human Mitochondrial DNA [NC_012920] and analysis the variation of sequences by Mito Tool Programing. The results of the DNA sequence analysis was compared with the published Cambridge Sequence using Mutation Surveyor Version 1.4 DNA mutation analysis software [Softgenetics, State College, PA]. Sequence differences between malignant and matched tissue and blood was recorded as mtDNA polymorphisms or mutation.

2.8.PCR-Based Pathogen Detection

The presence of *H. pylori* infection was determined in patients with GC by multiplex PCR amplification of 16S rRNA and CagA genes using Hp1-Hp2 and CagAF-CagAR primers, respectively (Table 4.1) (Ghatak et al., 2016). The PCR products were electrophoresed with 1.5% agarose gel. *H. pylori* infection was defined by the presence of an intact band of 109 bp (16S rRNA) and 400 bp (CagA gene). The presence of EBV type1/type 2 infections was carried out by a standard PCR assay crossways type-specific regions of EBNA3C gene using previously described primer sets (Table 4.1) (Kingma et al., 2997). The PCR amplification was carried out for 35 cycles at 95 °C for 45 seconds, 56 °C for 45 seconds, 72 °C for 1 minutes followed by a final extension at 72 °C for 10 minutes. The PCR products were analyzed by 12% polyacrylamide gel, and the presence of 153 and 246 bp was considered as type 1 and type 2 EBV, respectively.

2.9.Reverse Transcriptase PCR (RT-PCR) for COXI Expression

First-strand synthesis of cDNA was performed using Superscript II reverse-transcriptase (Life Technologies Inc., USA). Five μg aliquot of total cellular RNA was used for each reverse transcription reaction, and one-tenth of this reaction was used for PCR. Primers targeting MT-COXI region and exons 1 and 3 of house-keeping gene, β-actin, were used (DeBruin et al., 2000) (Table 1). The PCR for COXI was run for 35 cycles consisting of 94°C for 30 seconds, 1 minutes at gradually decreasing temperatures using a touchdown protocol (4 cycles each at 62° C and 60°C; and 30 cycles at 58°C), and 74 °C for 5 minutes. RT-PCR products were electrophoresed in 10% polyacrylamide gel and visualized by ethidium bromide staining.

2.10. Western blot analysis

Western blot analysis was carried out on protein lysates from gastric cancer tumor and adjacent normal cell. Briefly, equal amounts of total protein (30 μg) from each sample were run on a 4-20% gradient SDS-PAGE gel under reducing conditions and proteins were transferred to PVDF membranes (BioRad, Hercules, CA). Membranes were probed with a 1:500 dilution of anti-COXI polyclonal mouse antibody (ab14705) (Abcam, Japan. Immunoreactive proteins were visualized by incubating membranes with horseradish peroxidase-conjugated anti-mouse or anti-rabbit IgG antibody (1:10000) followed by reaction with ECL Plus (Amersham, St. Louis MO). Membranes were subsequently probed with a mouse monoclonal β-actin antibody (Sigma, St. Louis MO) as an internal protein loading control. Experiments were repeated three times. All data are expressed as the mean ± SD of three experiments. Statistical significance was assessed by the unpaired two-tail T-test.

2.11.Mitochondrial respiratory enzymes activity assay

Activities of mitochondrial enzymes were determined spectrophotometrically according to previous methods with slight modifications (Singer, 1994; Trounce et al., 1996; Zhou et al., 2003). Cell membranes were permeabilized with 0.5% Triton X-100. Samples were kept on ice for subsequent determination of mitochondrial enzyme activities. Specifically, Complex I (NADH dehydrogenase) activity was measured by monitoring the reduction of ferricyanide at 420 nm (Singer, 1994). The activities are expressed as nmol reduced ferricyanide/min/mg protein using an extinction coefficient of 1.0 per mM/cm. Complex II (succinate dehydrogenase) activity was measured as the rate of reduction of 2,6-dichlorophenolidophenol (DCPIP) at 600 nm when coupled to the complex II-catalyzed reduction of decylubiquinone (DB) (Trounce et al., 1996). Complex III (ubiquinol–cytochrome c oxidoreductase) activity was measured by monitoring reduction of cytochrome c at 550 nm catalyzed by complex III in the presence of reduced decylubiquinone (DBH2) (Trounce et al., 1996). Complex IV (cytochrome c

oxidase) activity was measured by monitoring the oxidation of reduced cytochrome c at 550 nm (Trounce et al., 1996).

2.12.Measurement of ROS level

DCFH-DA (from Sigma Company) was dissolved in 95% ethanol to a concentration of 5 mmol/L and stored at 4 □ in the dark, diluted to 5 μmol/L with PBS before use. The intensity of DCF green fluorescence was measured after DCFH-DA reaction with molecular device florescence microplate reader (spectra-max) with an excitation wavelength of 485 nm and an emission wavelength of 535 nm. The results were expressed as mean fluorescence intensity (MFI) (Kang et al., 2016).

2.13. Cell cycle estimation

0.1 g of grossly gastric tumor and adjacent normal gastric mucosa tissue were used for cell cycle analysis. Cells were harvested by mechanical dis-aggregation and fine-needle aspiration. The data obtained was analyzed using the ModFit LT software (DNA Modeling System) version 2.0 (Verity Software House, Inc.) and single parameter histograms were obtained (Blanco et al., 2013).

2.14.Statistical Analysis

D-loop and MT-COXI gene mutations among case–control subjects was tested for Hardy–Weinberg equilibrium by a chi-square test with one degree of freedom (df). The mutations in each group were estimated using odds ratios (ORs) and 95% confidence intervals (CIs). Additionally, logistic regression analyses were conducted to compute the potential confounder's influence of both genetic and environmental factors for gastric cancer. Then, the independent effect of risk factors was investigated in a multivariate model (introducing all variables and terms of interactions) keeping only those statistically significant or showing a confounding effect on the studied factors. The likelihood ratio test was used to decide whether to retain each covariate in the model. Sex, cigarette smoking, family history of gastric

cancer, tuibur consumption, alcohol intake, smoked meat/vegetable intake, high consumption of salt and pickle, high consumption of fermented fat, *H. pylori* status and genotyping, EBV status, and genotyping were all considered in the regression model as potential confounders to evaluate the association of risk factors and susceptibility to gastric cancer. For all tests, a two-sided p-value <.05 was considered statistically significant. All statistical analyses were performed using SPSS 20.0 program (SPSS Iberica, Madrid, Spain) and conformed by R-console statistical package ver3.3.0 (The R Foundation for Statistical Computing). Heat map analysis was performed by Hemi IBP software.Circos plots (Krzywinski et al., 2009) were generated to visualize variant distributions in MT-COXI gene and their association with gastric cancer stages, ploidy level, MT-COXI gene and protein expression and demographic factors.

2.15.Multifactor dimensionality reduction (MDR) analysis and Interaction entropy graph

MDR analysis (MDR 3.0.2) was performed to assess the risk of gastric cancer by studying the effect of demographic factors (Hahn et al., 2003). It is a non-parametric, model-free statistical approach which is used to generate one-dimensional model to predict gastric cancer susceptibility. The test is used to overcome the limitations encountered by parametric methods like logistic regression (LR) due to small sample size (Manuguerra et al., 2007). The potential confounding factors used in the multivariate model were considered for interaction entropy estimation.

3. RESULTS

3.1. Demographical and Clinicopathological characteristics of gastric cancer patients

Gastric cancer was more prevalent in males (70%) among the collected samples from Mizoram population. Among the various demographic factors in the bivariate analysis, *Helicobacter pylori*infection (OR: 1.168; 95% CI: 0.365 – 3.733; P=0.044) with cagA genotyping (OR: 8.298; 95% CI:

0.365 – 6.578; P=0.005), Smoked Meat/Vegetable Intake (OR: 6.181; 95% CI: 2.528 – 12.346; P=0.001) and High fermented fat intake (OR: 3.387; 95%CI: 1.902 – 6.031; P=0.0003) were found to be significant risk factors for gastric cancer (Table 4.3). High salt intake also exhibited potential factors for gastric cancer for this population (OR: 2.687; 95%CI: 0.244 – 3.936; P=0.078). No significance was found with consumption of alcohol, and cigarette smoking.The patient characteristics of two groups are shown in Table 4.

The median age in the young group was 36 years (range16–45), and this group contained a lesser proportion of patients (35%) than the older aged group (65%). As shown in Table 4.3, the older patient group had a higher proportion of men (40%) and women (25%) (P <0.326). Tumour size and subtotal gastrostomy were more in case of older age than in younger patients. Most of the gastric cancer patients are operated with stage II tumor. The symptoms at presentation in both groups are shown in Table 4.4. The most common symptom was abdominal pain (62.50%) followed by weight loss (42.50%; p = 0.05). Hemorrhage was more in young patients group rather than older patients group and Vomiting was more in older age group rather than young patients group.

3.2. Somatic Mutation in D-Loop

We detected nucleotide changes in gastric cancers by comparison with the sequence in the noncancerous part of the same patient and healthy control sequences. In studying the 80 patients with gastric cancer, we found that in an average 15% of these patients' harbored somatic mutations in their gastric cancers (Table 4.5). Total 49 mutations are found in hyper variable region I, 6 are in non-coding region and 17 are in

Demographic factor	[a]HC (n=80)	[b]GC (n =80)	[c]ORs (95% CI)[d]	P value
Age years ± SD (range)	52.18 ± 12.35	58.7 ± 9.76	-	-
Male	40 (50%)	56 (70%)	-	-
Female	40 (50%)	24 (30%)		
Cigarette smoking	52 (65%)	56 (70%)	0.595 (0.373 – 1.947)	0.029
H. pylori Positive	32 (40%)	56 (70%)	1.168 (0.365 – 3.733)	0.094
CagA Positive	8 (10%)	36 (45%)	8.298 (0.365 – 6.578)	0.005
EBV Positive	6 (7.5%)	18 (22.5%)	1.332 (0.255 – 6.957)	0.084
Family history of gastric cancer	10 (12.5%)	12 (15%)	0.708 (0.674 – 2.030)	0.477
High Salt and pickle Intake	36 (45 %)	62 (77.5%)	2.687 (0.244 – 3.936)	0.078
Tuibur consumption	28 (35%)	34 (42.5%)	0.828 (0.459 – 1.492)	0.049
Smoked Meat/Vegetable Intake	50 (62.5%)	72 (90%)	6.181 (2.528 – 12.346)	0.001
High fermented fat intake	28 (35%)	72 (90%)	3.387 (1.902 – 6.031)	0.0003
Alcohol Intake	34 (42.5)	40 (50%)	0.543 (0.312 – 1.944)	0.031

[a] HC – Healthy Control; [b] GC – Gastric Cancer; [c]OR – Odd ratio; [d] 95% CI – 95 % Confidence Interval

Table 4.3. Demographic and clinical characteristics of the samples

Parameters	Young group (Age ≤ 45 years)	Elderly group (Age 46 - 79 years)	P-value
Tumor location			
Upper	16 (20%)	26 (32.5%)	0.869
Middle	04 (5%)	10 (12.5%)	
Lower	08 (10%)	12 (15%)	
Whole	0	04 (5%)	
Tumor size (cm),(mean ± SD)	4.6 ± 2.8	4.9 ± 3.1	0.922
Type of gastrectomy			0.186
Total	20 (25%)	24 (30%)	
Subtotal	08 (10%)	28 (35%)	
Stage			
Stage I	4 (5%)	5 (6.25%)	
Stage II	12 (17.5%)	20 (30%)	0.828
Stage III	8 (12.5%)	15 (20%)	
Stage IV	4 (5%)	12 (15%)	
Abdominal pain	18 (22.5%)	32 (40%)	0.161
Weight loss	10 (12.5%)	24 (30%)	0.05
Hemorrhage	14 (17.5%)	6 (7.5%)	0.205
Dysphagia	12 (15%)	10 (12.5%)	0.76
Early satiety	6 (7.5%)	6 (7.5%)	1.00
Vomiting	8 (10%)	24 (30%)	0.045
Increased Abdominal girth	2 (2.5%)	0	0.317
BMI(Mean ± SD)	21.4 kg/m^2 ± 3.6	22.1 kg/m^2 ± 2.9	0.006

Table 4.4. Demographical and Clinicopathological features of Gastric Cancer Patients in Mizoram

(Stratified by Age)

Position	Substitution	Frequency (%)	Region
16051	A>G	10	HV1
16086	T>C	2.5	
16092	T>C	2.5	
16093[a]	T>C	12.5	
16129	T>C	2.5	
16136[a]	T>C	10	
16145	G>A	5	
16147[a]	C>T	5	
16158	A>G	5	
16183	A>C	5	
16184	C>A	5	
16184-16193	Poly (C) change	55	
16209	T>C	2.5	MT-5, HV1
16213	G>A	2.5	
16215	A>G	2.5	
16217[a]	T>C	7.5	
16218	C>T	2.5	
16223	T>C	30	HV1
16227	A>G	5	
16234	C>T	5	
16235	A>G	5	
16239	C>T	5	
16246	A>T	2.5	
16258	Adel	5	
16260	C>T	5	
16261	C>T	2.5	
16264	C>T	2.5	
16266	C>T	10	
16272	A>G	5	
16275	A>G	2.5	
16278	C>T	12.5	
16287	C>T	2.5	
16290[a]	C>T	12.5	
16298*	T>C	7.5	
16304	T>C	2.5	
16311	T>C	10	
16316[a]	A>G	2.5	
16318	A>G	5	
16319	G>A	15	
16324	T>C	12.5	
16325	T>C	7.5	
16355	C>T	10	
16362	T>C	42.5	
16381*	T>C	2.5	
16391	G>A	12.5	Non-coding
16400	C>T	5	
16444[a]	C>T	2.5	
16501	C>T	2.5	
16519	T>C	60	

Position	Substitution	Frequency (%)	Region
64*	C>T	2.5	
146	T>C	30	HV2
150	C>T	7.5	
151	C>T	12.5	
152	C>T	22.5	
153	A>G	7.5	
185	G>A	5	
195	T>C	2.5	
199	T>C	2.5	
200	A>G	12.5	
204	T>C	7.5	
207	G>A	2.5	
227*	A>G	5	MT-CSB1, HV2
228*	G>A	2.5	
235*[a]	A>G	12.5	MT-TFX, MT-CSB1, HV2
248	A del	12.5	MT-TFX, HV2
253	C>T	2.5	
262	C>T	2.5	
303-309	Poly (C) change	63	HV2

'*' – represents novel mutations; 'a' - represents heterozygous mutations.

Table 4.5. List of mutations with frequency in D-loop region of gastric tumor samples.

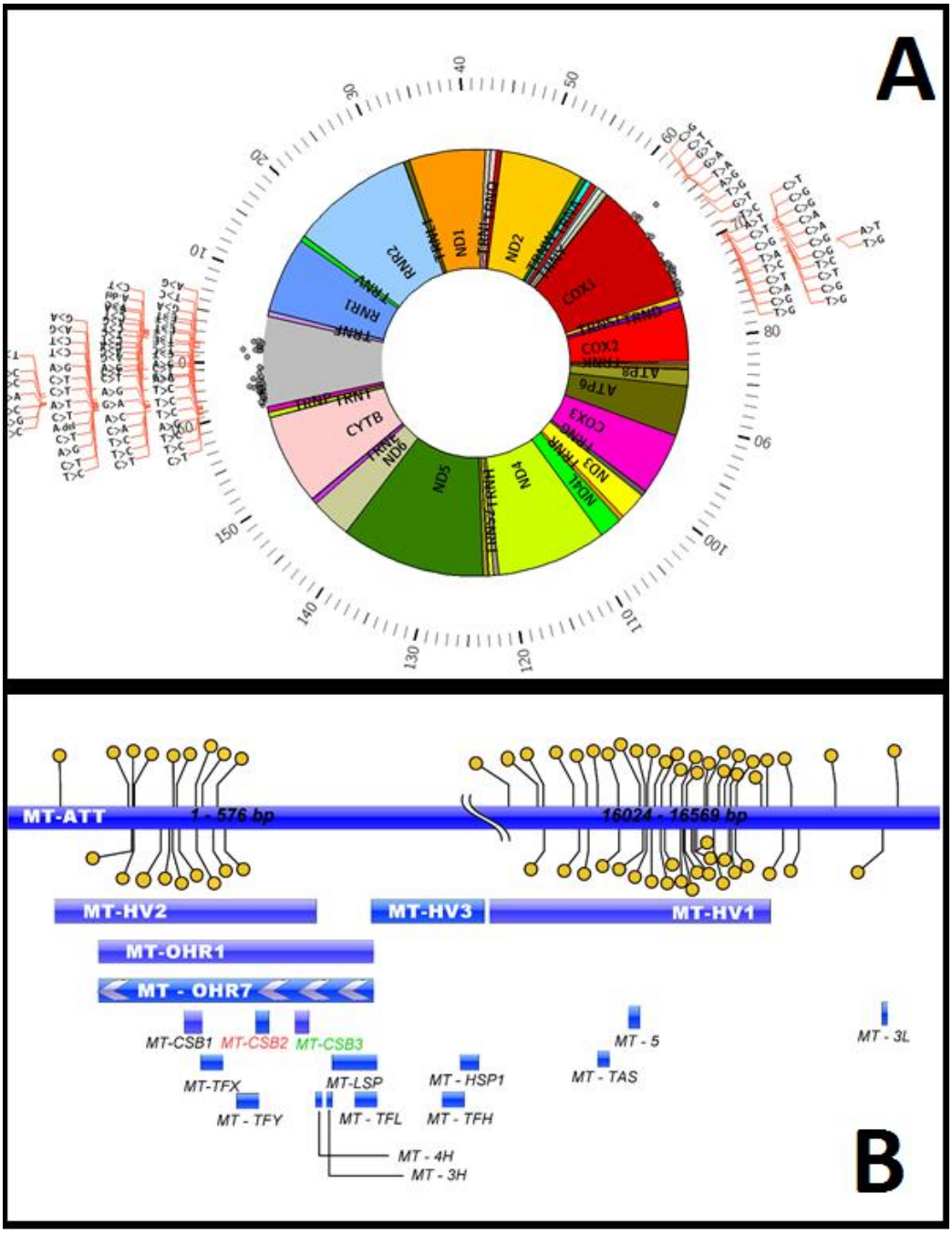


Figure 4.1.(A) Mitoseek circos plot of all the somatic mutation in mitochondrial control region and MT-COXI region of gastric cancer samples. Gray dots are represent position and frequency of the mutations. (B) Mutation position in the different domain of the mitochondrial control region.

hyper variable region II (HVII). Three of the gastric cancer patients, we found three distinct point mutations (228G>A, 235A>G and 248Adel) in the D-loop CSB1 and transcription factor X binding site region of the mtDNA (Figure 4.1 and 4.2; Table 4.5). The all three mutations are novel mutations. Among the 72 mutations, 17 (23.61%) were heteroplasmic. Two novel adenine deletion were found in HVI and HVII region (16258Adel and 248Adel). It is important to note that 55% of the gastric cancers carried a change (insertion or deletion) in the nucleotide position16184-16193 poly-C mononucleotide repeat of the mtDNA. The results indicate that somatic mutations occurred in the D-loop region of the mtDNA in gastric carcinomas.

3.3. Mutation in the COXI region of mtDNA

A total of 27 MT-COXImutationwere identified in gastric cancer tumor samples with respect to the adjacent normal, blood and rCRSsequences (Table 4.6; Figure 4.1). Among the 27 COXI mutations, 22 (14, non-synonymous and 8, synonymouswere novel mutation. All the 14non-synonymous COXI mutations were not previously reported in the literature or the public mtDNA mutation databases (mtDBase: http://www.genpat. uu.se/mtDB/index; MITOMAP: http://mitomap.org/MITO MAP). Two heteroplasmicnonsynonymous mutations were found at position 6395C>G and 7336C>G in 60% and 26.25%of gastriccancer patients, respectively (Table 4.6). All 14 mutations were identified as disease causing by mutation taster and probably damaging by polyphen 2. 6395C>G (164F>L) mutation location is annotated in UniPort database as transmembrane domain and the mutant amino residues is smaller than the wild type amino acids so, it's might not be interact with the lipid membrane.Due to the mutation the protein electron carrier activity and oxidoreductase activity might be lost. Both of the mutations in the COXI region were potentially harmful missense/nonsense mutations because they occurred at the evolutionarily highly conserved amino acid residue with a non-conserved amino acid change. The 6395C>G changes phenylalanine at position 164 to lysine, and C7336G changes a serine residue at 478

to tryptophan in COXI are highly conserved throughout evolution, from sea urchin and fruit flies to the mouse and human. This suggests that 6395C>G may have a functional consequence. The arrangement of stand and helix are changing between transmembrane α helices V and VI and between VII, VIII and IX due to the frequent mutation in this region in gastric cancer tumor samples. Due to the mutation of 7103C<G (400F>L) and 7190C<A (429H>Q), the single protein binding region might be totally lost predicted by ISIS.

3.4. ROS level and cell cycle in mutation group and control group

Cell cycle and apoptosis could be detected by flowcytometry synchronously. In cell cycle, DNA was synthesized in synthesis (S) phase. As a result, the percentage of cells in synthesis phase could reflect cell proliferation. Insufficient mitochondrial activity might, in turn, decrease ATP levels and subsequently increase the AMP/ATP ratio, which is known to activate AMPK and lead to G1 > S arrest. The mitochondrial complex function might be lost and G1 might be arrested with high S phase DNA quantity (Figure 4.3A and B) (p value = 0.039). Most of the gastric cancer samples showed diploid, except in six samples where aneuploidy (containing 6395C>G change) resulted in less DNA content in G2/M phase, whereas high DNA content in S phase. For most of the gastric tumor and adjacent normal tissues were shows high level of ROS generation (MFI) than endoscopic control group. As shown in Table 4.7, level of ROS, rate of cell apoptosis and proliferation in mutated tumor tissues were higher than adjacent normal and control group. There was a significant ROS (MFI) observed between tumor tissues and healthy control samples. Percentage of G0/G1 phase significantly reduced and S phase significantly increased for the D-loop, COXI and mtMSI associated tumor cells than healthy control samples. There was no significant change of G2/M phase between mutated tumor cells and control cells. These results establish that increased ROS production causes mutant mitochondrial genome in tumor tissue to be blocked in G1-S transition and that relieving this stress allows the cells to re-enter the cell cycle.

Mutation	Frequency (%)	Amino acids change	Mutation Taster/Polyphen2/ SNP&GO	Effect of the Mutations by HOPE
6388T>A *	15	162I>N	Disease causing / 0.976(Probably damaging)/ Disease(RI-6)	• The hydrophobicity of the wild- type and mutant residue differs. • The mutation might cause loss of hydrophobic interactions with other molecules on the surface of the protein. • The wild-type residue is more hydrophobic than the mutant residue. This differences in hydrophobicity can affect the hydrophobic interactions with the membrane lipids.
6395C>G *[a]	60	164F>L	Disease causing/ 0.998(Probably damaging)/ Disease (RI-7)	• The mutant residue is smaller than the wild-type residue. The size difference can affect the contact with the lipid-membrane.
6609T>G *	15	236W>G	Disease causing / 0.999(Probably damaging)/ Disease(RI-7)	• The hydrophobicity of the wild- type and mutant residue differs. • The mutation will cause loss of hydrophobic interactions in the core of the protein • The mutation introduces a glycine at this position. Glycines are very flexible and can disturb the required rigidity of the protein at this position.
6651G>T *	10	250G>C	Disease causing/ 1.000(Probably damaging)/ Disease(RI-4)	• The wild-type residue was buried in the core of the protein. The mutant residue is bigger and probably will not fit. • The torsion angles for this residue are unusual. Only glycine is flexible enough to make these torsion angles, mutation into another residue will force the local backbone into an incorrect conformation and will disturb the local structure. • The wild-type residue is a glycine, the most flexible of all residues. The flexibility might be necessary for the proteins function. Mutation of this glycine can abolish this function.
7018A>T*	27.5	372Y>F	Disease causing/ 0.999(Probably damaging)/ Neutral(RI-6)	• The hydrophobicity of the wild- type and mutant residue differs. • The mutation will cause loss of hydrogen bonds in the core of the protein and as a result disturb correct folding. • The mutation is located with a topological domain annotated in the uniprot databases as mitochondrial matrix. The mutation might disrupt this topological domain.
7027C>T *[a]	23.75	375A>V	Disease causing/ 1.000(Probably damaging)/ Disease(RI-8)	• The mutant residue is bigger than the wild-type residue. The size difference can affect the contact with the lipid-membrane. • The wild-type residue was buried in the core of the protein. The mutant residue is bigger and probably will not fit.
7063T>A*	15	387F>Y	Disease causing/ 0.995(Probably damaging)/ Disease(RI-6)	• The hydrophobicity of the wild- type and mutant residue differs. • The wild-type residue is more hydrophobic than the mutant residue. This differences in hydrophobicity can affect the hydrophobic interactions with the membrane lipids.
7089T>C*	10	396W>R	Disease causing/	• The hydrophobicity of the wild- type and mutant residue differs.

			1.000(Probably damaging)/ Disease(RI-9)	• The wild-type residue is more hydrophobic than the mutant residue. This differences in hydrophobicity can affect the hydrophobic interactions with the membrane lipids
7103C>G *	7.5	400F>L	Disease causing/ 0.998(Probably damaging)/ Disease(RI-1)	• The mutant residue is smaller than the wild-type residue. • The mutation will cause an empty space in the core of the protein. • The mutation is located with a topological domain annotated in the uniprot databases as mitochondrial matrix. The mutation might disrupt this topological domain.
7190C>A*	20	429H>Q	Disease causing/ 0.993(Probably damaging)/ Disease(RI-8)	• The mutant residue is smaller than the wild-type residue. The size difference can affect the contact with the lipid-membrane. • The mutation will cause an empty space in the core of the protein.
7194C>G*	31.25	431L>V	Disease causing/ 0.919Probably damaging)/ Neutral(RI-1)	• The mutant residue is smaller than the wild-type residue. The size difference can affect the contact with the lipid-membrane. • The mutation will cause an empty space in the core of the protein.
7336C>G*[a]	26.25	478S>W	Disease causing/ 1.000Probably damaging)/ Disease(RI-1)	• The residue is located on the surface of the protein, mutation of this residue can disturb interactions with other molecules or other parts of the proteins. • The hydrophobicity of the wild- type and mutant residue differs. • The mutation is located with a topological domain annotated in the uniprot databases as mitochondrial matrix. The mutation might disrupt this topological domain.
7383T>G*	10	494W>G	Disease causing/ 0.999Probably damaging)/ Disease(RI-6)	• The hydrophobicity of the wild- type and mutant residue differs. • The mutation will cause loss of hydrophobic interactions in the core of the protein. • The mutation introduces a glycine at this position. Glycines are very flexible and can disturb the required rigidity of the protein at this position.
7413A>T*	5	504T>S	Disease causing/ 0.955Probably damaging)/ Neutral(RI-9)	• The mutant residue is smaller than the wild-type residue. • The mutation will cause an empty space in the core of the protein. • The mutation is located with a topological domain annotated in the uniprot databases as mitochondrial matrix. The mutation might disrupt this topological domain

Table 4.6. List of mutations and their pathogenicity in cytochrome c oxidase subunit I region of gastric tumor samples.

'*' - representing novel mutations; 'a' - representing heterozygous mutations.

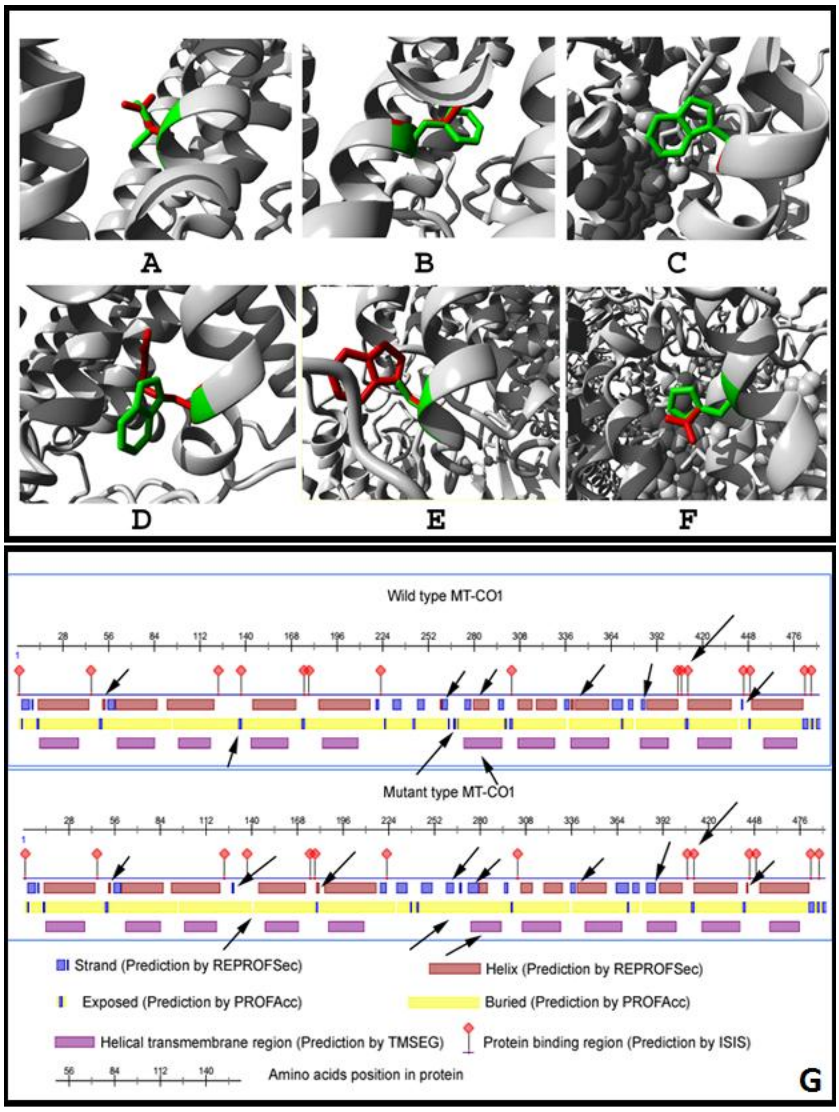


Figure 4.2. (A) 3-D structure of the mutation in gastric cancer samples. Both the wild-type and mutant side chain are shown in green and red, respectively. The rest of the protein is shown in grey. (A) Mutation of Isoleucine to Asparagine at position 162 due to 6388T>A, (B) Mutation of Phenylalanine to Leucine at position 164 due to 6395C>G, (C) Mutation of Tryptophan to Glycine at position 236 due to 6609T>G, (D) Mutation of Tryptophan to Arginineat position 396 due to 7089T>C. (E) Mutation of Serine to Tryptophan at position 478 due to 7336C>G, (F) Mutation of Histidine to Glutamine at position 429 due to 7190C>A, (G) Wild and mutant type COXI protein stand, helix, transmembrane region and protein binding region.

3.5. Mitochondrial respiratory enzymes activity

In the present study, we determined whether tumor cell and mitochondrial DNA mutations could lead to altered mitochondrial enzyme activity. We compared activities of the mitochondrial complexes in tumor and mutated cells. As shown in Figure 4.3, there were no differences in activities of complexes II between the gastric tumor, mutated and adjacent normal cell. However, complex I, II and IV activity in tumor and mutated cells was significantly lower compared with adjacent normal (Figure 4.3). This study suggests that mitochondrial mutation in gastric tumor cell inactivation leads to reduced complex I, II and IV activity. The complex II and IV enzymes activity were significantly reduced for tumor, D-loop, COXI mutated and MSI cells than the adjacent normal cells. COX activity in total cell homogenate from the tumor and adjacent normal cells was measured to determine the effectof COXI mutation in the activity of complex IV. Figure 4.3 shows thatcells harboring the 6395C>G, 7103C>G and 7194C>G mutations cause more than50% decreasein the activity of complex IV. This COX activity in the adjacent normal control cells are normal, indicating that these mutationsin gastric cancer influences the activity of a different complexin the respiratory chain. Six samples with complex IV enzymes activity lower were exhibited aneuploidy cell in cell cycle analysis and most of the samples were having very high S phase and proliferation of the tumor cells.

3.6. mtDNAmicrosatellite instability in the D-loop region

Three distinct microsatellite sequences within the mitochondrial D-loop region were examined for mutation and instability status in 80 gastric tumor samples. 303-bp and 16185-bp positions (C)n tract and 517-bp position (CA)n were examined using the polymerase-single stranded conformation polymorphism (PCR-SSCP) and direct sequencing methods. The 303 and 16185 poly (C)n tract exhibited

C_7TC_6 and C_7TC_4 for Mizo population, respectively. Seventeen cancer patients (Male -13; Female -5) and three healthy control

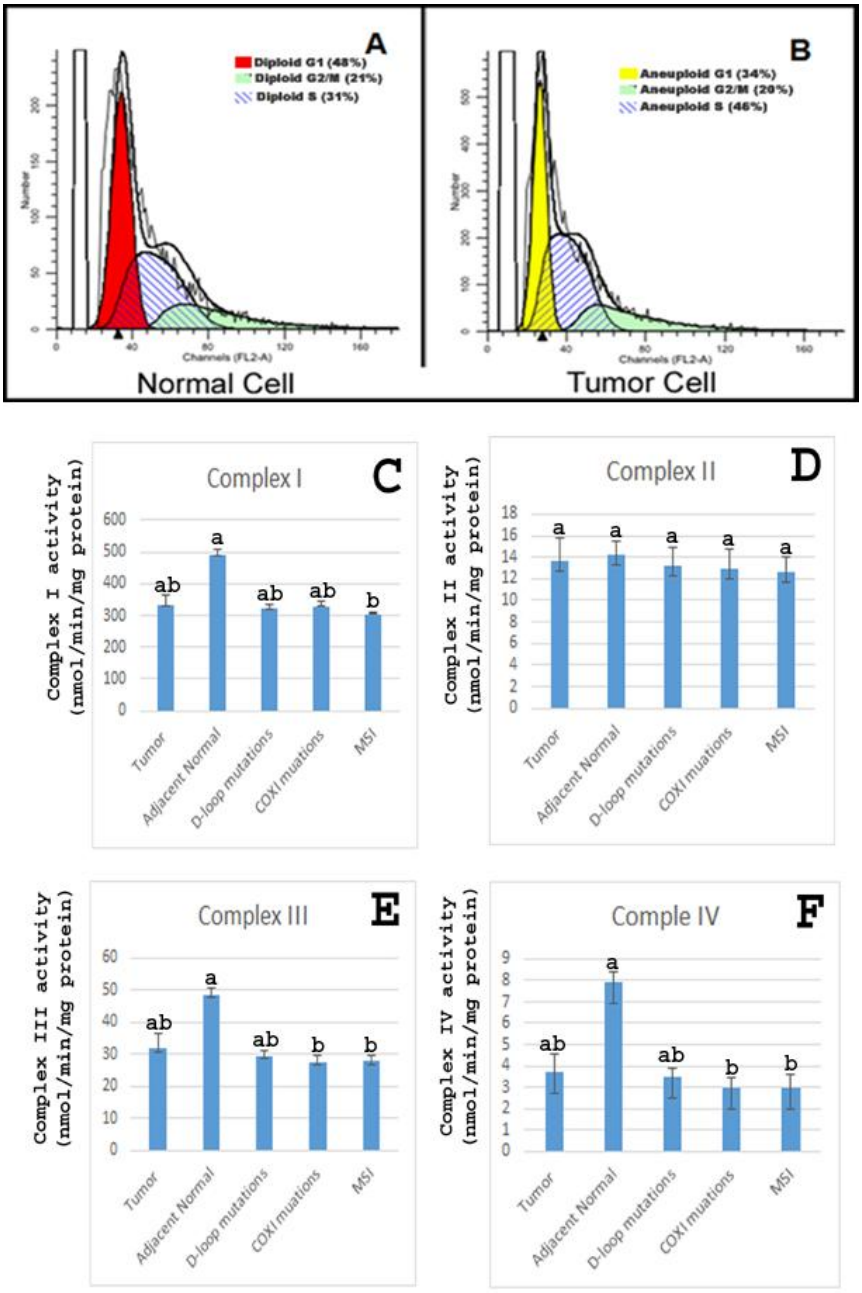


Figure 4.3. (A) Cell cycle histogram of the adjacent normal representative sample, (B) Cell cycle histogram of the representative same Tumor samples showed aneuploidy. Comparison of mitochondrial respiratory enzymes activity. (A) Complex I activity. (B) Complex II activity. (C) Complex III activity. (D) Complex IV activity. The data represent the means SD for three separate experiments. a, ab and

brepresented a significant decrease compared with the wild-type, using Duncan test (Anova) for data analysis. The P value for (A) is 0.004, (B) is 0.072, (C) is 0.017 and (D) is 0.006.

(Male – 2; Female -1) exhibited abnormal bands with different mobilities for the 303 poly (C)n tract (Figure 4.5; Table 4.7). Five cancer patients (Male -3; Female -2) and two healthy control (Male – 1; Female -1) exhibited abnormal bands with different mobilities for the 16185 poly (C)n tract. Direct DNA sequencing revealed that most alterations consisted of deletion or insertion of cytidine residues in the (C)n tract (Figue 5). Six tumor samples revealed both the 303 and 16185 poly (C)n tract instability. None of the individuals examined had the same (C)nrepeats in germline cells (blood and adjacent normal cell), suggesting that this variation reflected a natural polymorphism with respect to mitochondrial function. Some carcinoma (12.5%) tissue contained two distinct mtDNA subpopulations with respect to (C)n status (heteroplasmy). In some cases for 517-bp (CA)n, the mtDNA subpopulation that appeared to be specific to carcinoma tissue in the SSCP analysis was also detectable in the normal counterpart using direct sequencing. Such cases revealed a different proportion of mtDNA subpopulations between normal and carcinoma tissue. Whether the alterations resulted from sequence changes in the proportion of mtDNA subpopulation during mitochondrial segregation could not be determined. Therefore, these alterations were comprehensively characterized as changes in mitochondrial heteroplasmy (CMH) that were never detected in normal counterparts. The frequency of the (CA)n tract were very less for the gastric tumor samples. Thus, microsatellite alteration in the mitochondrial D-loop region (MSI) occurred commonly in the (C)n tract, but not in the (CA)n tract, in gastric carcinoma. The 303 poly(C)n tract is significantly associated with the gender (Male) with high Odd ratios and 95% confidence interval (OR: 7.45; 95% CI: 2.09-26.45; P=0.020).

3.7. mtDNA microsatellite instability in the coding region

In order to determine whether microsatellite alteration occurred in the mitochondrial coding regions, seven independent repeat sequences were examined using PCR-SSCP and direct sequencing. The coding regions are NDI$(C)_6$, NDII$(A)_7$, COXI$(A)_7$, COXIII$(T)_7$, NDV$(C)_6$ and $(A)_8$, NDV(CCT)3 and (AGC)3 and $tRNA^{ser}(A)_6$ and $(C)_6$ (Table 4.2). As shown in Table 4.9, four cases (male -3, female -1) reveled microsatellite instability in NDI region. Four samples (male -3, female -1) exhibited a cytidine deletion in the $(C)_6$ tract within the NADH dehydrogenase 5 (NDV) gene (Figure 4.5). This frame shift mutation created a stop codon (TAA) downstream and was predicted to encode a truncated ND5 protein (a change from 603 to 35 amino acids). The MSI mtDNA was homoplasmic for all the four tumor tissues. Two samples (male -2) and four samples (male -3, female -1) exhibited a adenine and cytidine insertion in the $(A)_8$and $(C)_6$ for NDV and NDI coding regions, respectively. These samples also exhibited frameshift mutations inthe $(A)_8$and $(C)_6$ tracts and created mitochondria specific stop codons AGA and AGG, respectively. The samples exhibited NDI and NDV coding region MSI and also D-loop 303 (C)n and 16185 (C)n tract instability. $(C)_7$tract in the cytochrome c oxidase subunit I also exhibited single cytidine deletion and insertion and generated frameshift mutation in mtDNA COXI coding region. Due to the insertion of single cytidine in the $(A)_7$ tract of COXI gene stop codon was generating at 266 position of the COXI protein and the deletion of a cytidine in the $(A)_7$ tract stop codon was generating at the position at 271 amino acid position. Thus, microsatellite alteration in the mitochondrial coding region was observed in gastric carcinoma, and appears to be associated with alteration in the D-loop region. Two (male – 1, female -1)of the gastric cancer tumor samples exhibited $tRNA^{ser}(A)_6$tract microsatellite instability (A_6>A_7). The MSI may affect the function of the mitochondrial encoded tRNASer 2 (AGU/C) (MT-TS2).

3.8. Decreased level of COX I gene and protein in mutated tumor cell

We investigated whether the alteration in complex III and IV activities were due to lower levels of cytochrome c oxidase subunit I protein encoded by the mitochondrial genome. Therefore, we examined the expression of COXI mRNA by RT–PCR. For RT-PCR, β-actin was used as the control. The results show significant differences in the levels of COXI mRNA COXI mutated tumor and adjacent normal cells

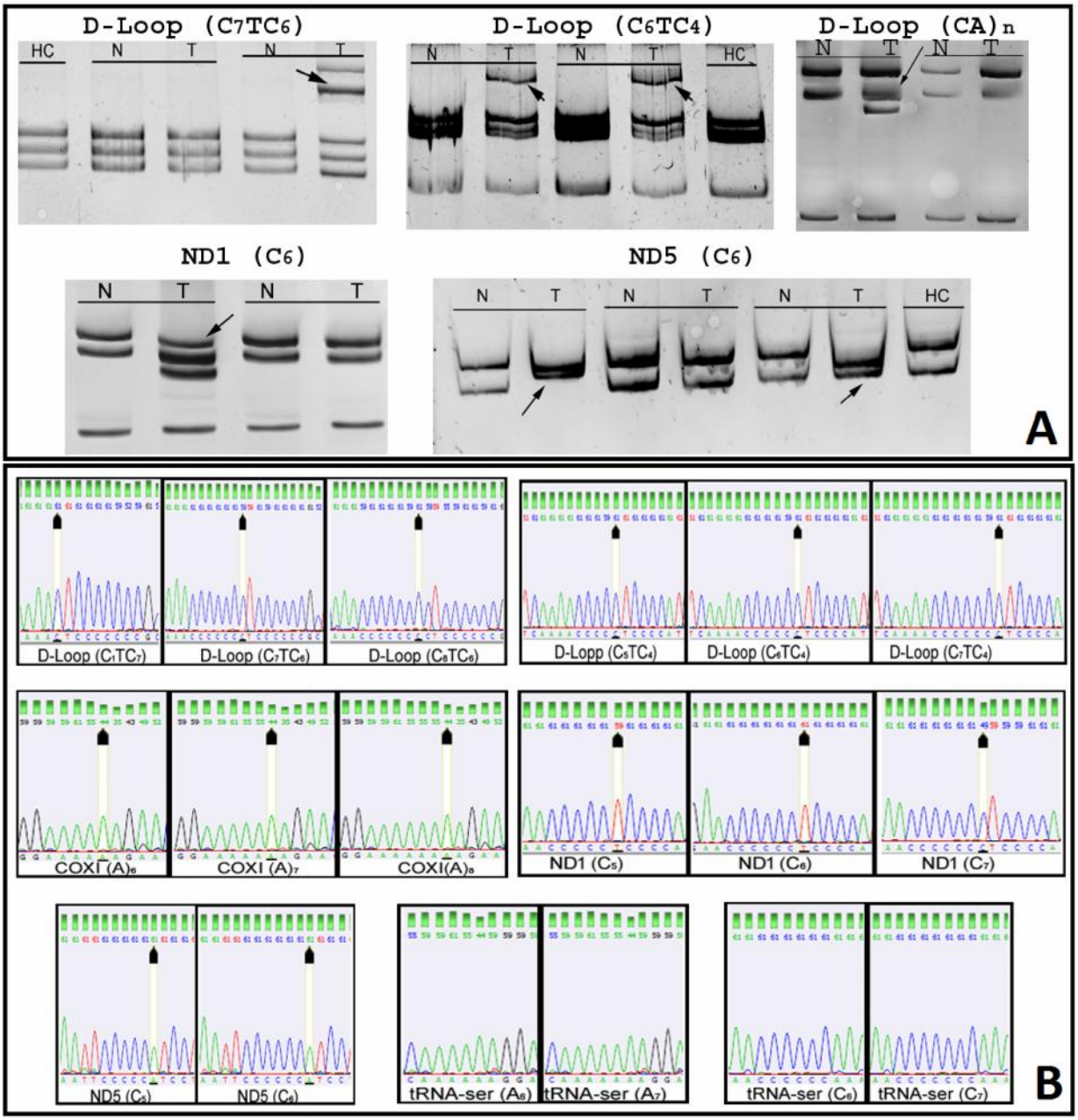


Figure 4.4. (A) Representative example of mitochondrial microsatellites instability (mtMSI) at the (C)n, (CA)n for D-loop and $(C)_6$ for ND1 and $(C)_6$ for ND5 SSCP gel pictures. Black arrows indicating mobility-shift band in gastric tumor tissue. (B) Representative Electropherograms obtained by direct sequencing of the mitochondrial D-loop C_7TC_4 and C_6TC_4, COXI$(A)_7$,ND1$(C)_6$, ND5$(C)_5$, $tRNA^{ser}(A)_6$ and $tRNA^{ser}(C)_6$.

(Figure 5.5B). We next examined the protein levels of COXI and by western blot analysis using antibodies against COXI. The same blot was probed with an antibody against β-actin as a control for variations in loading. The results show that the COXI protein level was significantly reduced in COXI mutated tumor cell (Figure 5.5B an 5.5C). Taken together, the discoordinate expression of COXI mRNA and COXI protein suggest post-transcriptional regulation of COXI by COXI gene mutations. For some gastric cancer tumor samples, moderate levels of COXI protein expression was observed although the mRNA levels were very less (Figure 5.5C). Given all the ways, protein levels are controlled post-translationally and given that mRNA translation is regulated by microRNAs in most of the cases. Might be microRNAs regulate the less expressed mRNA to moderate level of protein translation.

3.9. Association of demographic factors, mitochondrial mutations and microsatellite instability

Various factors associated with the occurrence of gastric cancer includes diet and family history affects the MSI status. A significant relationship was observed with presence of *H. pylori* infection in gastric tumor tissue and MSI (OR: 13.20; 95% CI: 5.18 – 33.62; P= 0.035) along with cagA genotyping (OR: 8.21; 95% CI: 3.77 – 17.87; P= 0.035). High consumption of smoked meat and smoked vegetables also significantly affect the mitochondrial MSI (OR: 5.90; 95% CI: 1.97 – 17.68; P=0.028) as well as high fermented fat intake associated with mt-MSI (OR: 6.20; 95% CI: 2.67 – 14.39; P=0.069). Alcohol

consumption, tuibur consumption and cigarette smoking were not associated with gastric cancer and mt-MSI (Figure 4.5). The rate of mtMSI was significantly higher in samples with *H. pylori* infection than in those without the infection ($P < 0.035$). MDR analysis was used to further validate our logistic regression based study on gene-gene and gene-environment interaction by using a model free approach. A four order interaction model was chosen along with their cross validation consistency (CVC) and testing balance accuracy (TBA). The best model was selected based on the highest CVC and TBA value among all the models. Among the entire dataset,

Samples	ROS (MFI)	Apoptosis (%)	G_0/G_1 (%)	S (%)	G_2/M (%)
Tumor samples	183.72 ± 20.13^{a}	16.54 ± 1.48^{a}	38.74 ± 5.06^{ab}	46.85 ± 2.77^{ab}	14.41 ± 3.45^{a}
Control Samples	94.53 ± 6.72^{b}	5.94 ± 0.83^{a}	52.81 ± 3.92^{a}	30.59 ± 2.24^{b}	16.6 ± 2.47^{a}
D-loop mutation	165.73 ± 2.14^{a}	16.62 ± 1.40^{a}	36.96 ± 2.63^{ab}	45.73 ± 1.65^{ab}	17.31 ± 0.55^{a}
COXI mutation	169.5 ± 4.61^{a}	15.98 ± 0.47^{a}	38.82 ± 4.98^{ab}	46.92 ± 1.73^{ab}	14.26 ± 3.30^{a}
mtMSI	184.84 ± 19.01^{a}	15.4 ± 0.34^{a}	34.39 ± 0.71^{b}	48.63 ± 0.99^{a}	16.98 ± 1.69^{a}

Table 4.7. ROS level and cell cycle in tumor, gene mutation group and control group (mean±SD). a, ab and b Duncan test (Anova) significant associations; ROS – Reactive oxygen species; MFI - mean fluorescence intensity

Factors	mtMSI		mtMSS		ORs (95% CI)	P value
	Normal samples (%)	Tumor Samples (%)	Normal Samples (%)	Tumor Samples (%)		
Male Female	6 2	32 14	34 38	24 10	1.95 (0.75 - 5.37)	0.0310
Cigarette smoking	3	35	49	21	1.22 (0.60 – 2.48)	0.907
H. pylori Positive	4	44	28	12	13.20 (5.18 – 33.62)	0.035
*CagA*Positive	1	29	7	7	8.21 (3.77 – 17.87)	0.090
EBV Positive	0	17	6	1	6.49 (2.49 – 16.93)	0.374
High Salt and pickle Intake	6	34	50	28	1.02 (0.48 – 2.16)	0.849
Tuibur consumption	4	12	24	22	0.54 (0.27 – 1.10)	0.709

Smoked Meat/Vegetable Intake	8	42	42	30	5.90 (1.97 – 17.68)	0.028
High fermented fat intake	6	40	19	32	6.20 (2.67 – 14.39)	0.069
Alcohol Intake	5	11	29	29	0.34 (0.17 – 0.70)	0.779

Table4.8. Association tables between mitochondrial microsatellites instability and demographical and Clinicopathological features of gastric cancer patients and healthy control in Mizoram.

	D-loop (1) $(PolyC)_n$	**D-loop $(CA)_n$**	**D-loop (2) $(PolyC)_n$**	**NDI $(C)_6$**	**NDII $(A)_7$**	**COXI $(A)_7$**	**COXIII $(T)_7$**	**NDV $(C)_6$**	**NDV $(A)_8$**	**NDV $(CCT)_3$**	**NDV $(AGC)_3$**	**$tRNA^{Ser}$ $(A)_6$**	**$tRNA^{Ser}$ $(C)_6$**
Cancer													
Male	13	3	3	3	Nil	3	Nil	3	2	Nil	Nil	1	1
Female	5	1	2	1		2		1	Nil			1	Nil
Healthy													
Male	2	1	1	Nil	Nil	1	Nil	Nil	Nil	Nil	Nil	1	Nil
Female	1	Nil	1			Nil						NIl	
Odd ratios	7.45	4.15	2.60	Nil	Nil	5.26	Nil	Nil	Nil	Nil	Nil	3.07	Nil
(Confidence interval)	(2.09 – 26.45)	(0.41 – 38.04)	(0.48 – 13.81)			(0.60 – 46.13)						(0.31 – 30.23)	
P value	0.020	0.223	0.502	Nil	Nil	0.620	Nil	Nil	Nil	Nil	Nil	0.094	Nil

Table 4.9. Association tables between mitochondrial microsatellites instability and gender of gastric cancer patients and healthy control in Mizoram.

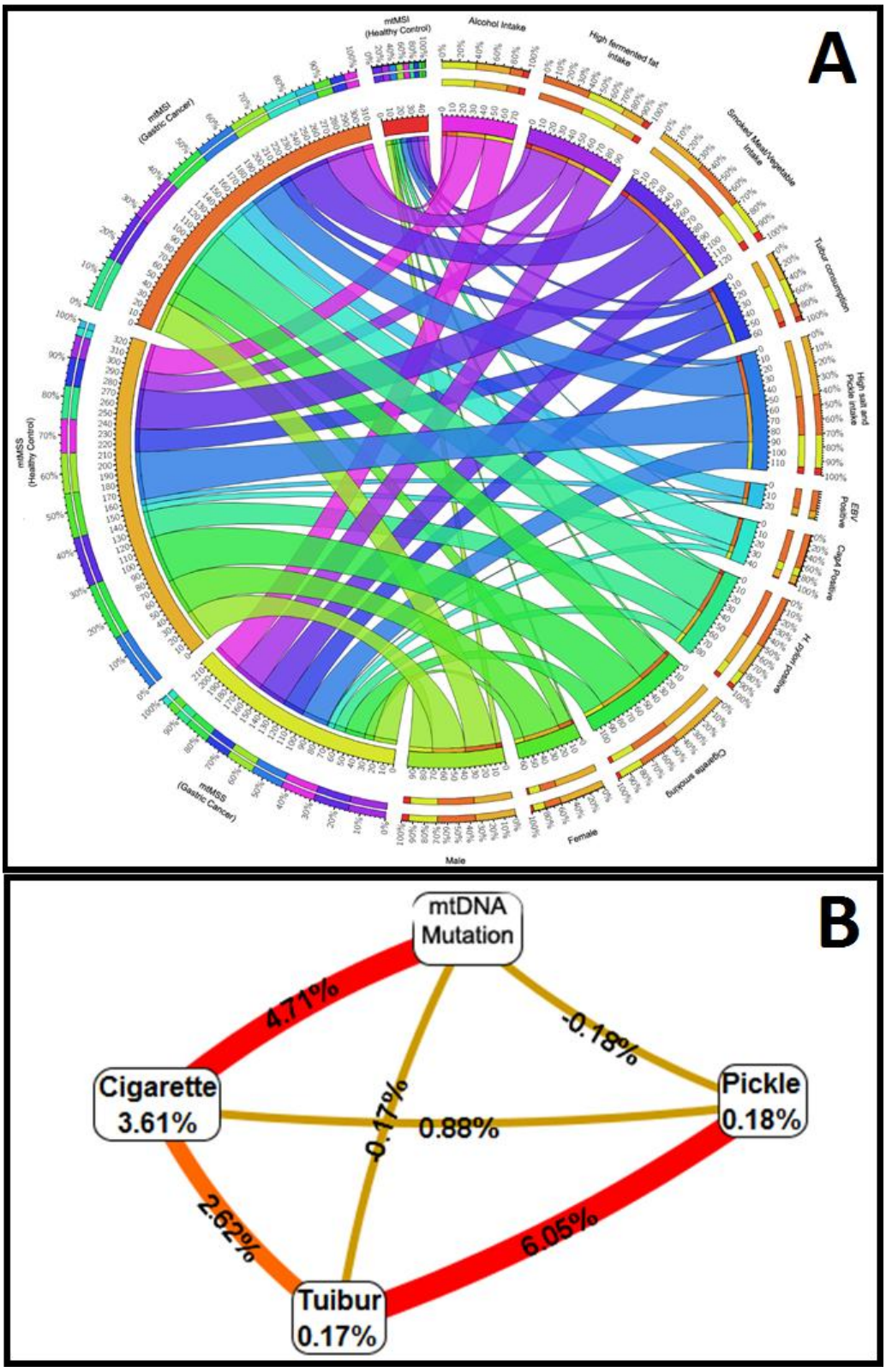


Figure 4.5. (A) Circular association plot between mitochondrial microsatellites instability and demographic factors in gastric cancer, created using Circos with custom settings. The circos plot

depicts frequencies different factor and mtMSI status in the inner ring and association of MSI status and different factors in the outer ring. Each factor and MSI status has been assigned different color. The area of each colored ribbon depicts the frequency of the different factor associated with particular mtMSI status; (B) Interaction entropy graph to find the gene-environment interaction with Gastric cancer risk. The percent of the entropy for and mtDNA mutations and potential confounder's interactions are represented in the graph. Bar wide shows frequencies of sample. Positive percentage of entropy denotes synergistic interaction while negative percentage denotes redundancy. Here, the red color denotes the high degree of synergistic interaction, golden color denotes the moderate degree of synergistic interaction, orange color represents minimum degree of synergistic interaction, green color represents minimum redundancy interaction while blue color denotes the highest redundancy.

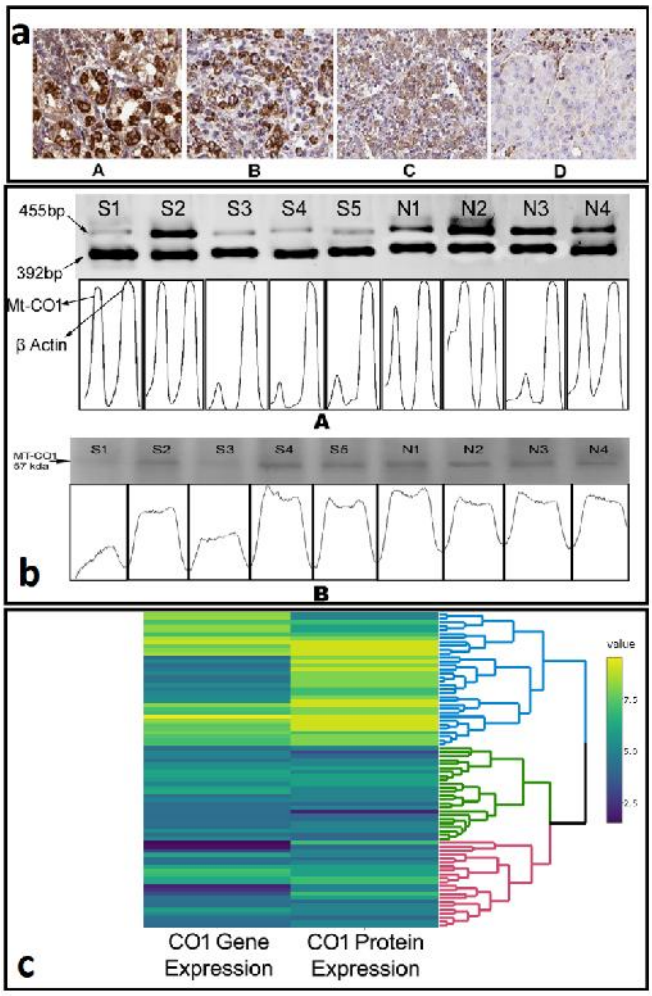


Figure 4.6.(a) Microscopic view of gastric adenocarcinoma of gastric tumor cells and adjacent normal cell. (A) Positive high immunoexpression of anti-MTCOXI antibody in normal cells (B) Positive moderate immunoexpression of anti- MTCOXI antibody in cancer cells (C) moderate immunoexpression of anti-MTCOXI antibody in adjacent normal cells (D) negative or less immunoexpression of anti-MTCOXI antibody in same tumor cell (from moderate immunoexpression adjacent normal cell), represented by the brownish colour in the cytoplasm and membrane. (b) Representative samples of mitochondrial COXI gene and expression study, *MT-COXI*(455bp) expression in 10% polyacrylamide gel, β-Actin (392 bp) as control and COXI protein (57kda); (c) Heat map representation of MTCOXI gene differential gene and protein expression and cluster analysis.

mitochondrial mutations are synergistically influenced by the cigarette smoking and tuibur consumption and are the best two factor model that was found to be statistically significant ($p < 0.002$) with a CVC of 10/10 and testing accuracy of 0.63. In our study, cigarette smoking is highly (3.61%) influencing for the mtDNA alteration followed by tuibur consumption (Figure 4.5).

3.10. Immunostaining of COX I in gastric cancer tumor and adjacent normal cells

Rabbit polyclonal anti-MT-COXI antibody (ab52223) (Abcam, Japan) specificity was reported for endogenous levels of total MT-COXI protein and it's expressed in a variety of tissues (http://www.abcam.com/MT-COXI-antibody-ab52223.html).Figures4.6A and B are representative images of stomach mucosa from healthy control specimen showing normal levels of cytochrome c oxidase subunit I immunostaining. In Figure 4.6A, the intense positive stain for COX I is found throughout the cells with deep staining at the cytoplasm and the membrane. Figures4.6C and D are representative images of section from the stomach tumor mucosa of a patient with gastric tumor. The cells with total absence of staining (Figure 4.6D) is clearly demarcated for all the tumor cells. Figure 4.6C are representative image of section of adjacent normal cells of the sample tumor sample of figure 4.6D. The adjacent normal cells are moderately stained and develop moderate amount of color intensity for the COX I immunostaining. The overall decreased immunostaining is particularly evident of the patient with gastric tumor shown in Figure 4.6D. Most of the negative COX I immunostaining was observed in cases of tumor having 6395C>G, 7103C>G mutations together and at stage-III only. Eight samples (10%) were negative, 46 (57.5%) were with moderate expression and 26 (32.5%) were positive for COXI protein expression in gastric adenocarcinoma tissues. After performing

the Fisher exact test, the MT-COXI expression was significantly correlated with the Stages III of gastric cancer tissues ($p = 0.038$).

4. Discussion

In the current study, we attempted to gather evidence to identify the relationship between dietary factor and gastric cancer risk in association with mitochondrial pathogenicity. Gastric cancer has been shown to have a positive association with consumption of smoked and salted meat and vegetables. Smoke-drying and preservation leads to formation of N-nitroso compounds. Nitrite reacts with amines and amides found in meats and other proteins to form N-nitroso compounds, which are highly carcinogens (Correa, 1985; Ghatak et al., 2016). The presence of nitrite in processed smoked meatsindirectly contributes to the development of chronic atrophic gastritis, leading to the development of stomach cancer (Nomura, 1990). Studies in the past have also shown positive associations of high intake of smoked meats as potential confounder for gastric cancer (Ward et al., 1999; Kneller et al., 1991). High salt intake showed significantly positive association with diffuse type gastric cancer. Intragastric high salt concentration destroys the mucosal barrier through the increase of surface mucous cell mucin and decrease of gland mucous cell mucin (Tamura et al., 1994), leading to inflammation and damage such as diffuse erosion, atrophic gastritis and decreased acidity of the stomach (Tsuganeand Sasazuki, 2007). It creates a condition favoring *H. pylori* infection (Kato et al., 2006). Salt may also directly damage gastric mucus, improve inflammatory responses of the gastric epithelium leading to increased epithelial cell proliferation as part of the repair process(Campos et al., 2006). This potentiates the action of carcinogens, and increases the probability of endogenous mutations (Kim et al., 2010). Infection with *H. pylori* and *CagA*+ strains were more frequent in gastric cancer patients than in controls. However, very less differences in the prevalence of *Epstein–Barr virus* (*EBV*) strains were

observed between patients and controls (22.5 vs. 7.5%, respectively). Frequent consumption of sa-um (fermented fat) was found to be associated with the risk of developing stomach cancer. This is a food material uniquely consumed in Mizoram (Phukan et al., 2006). Dietary intakes of total or saturated fat have been shown to be associated with stomach cancer (Gonzálezand Lopez-Carrillo, 2010; Kneller et al., 1992). Boiled pork fat, in addition to being a rich source of saturated fat, may form carcinogenic compounds during long storage, as in other stored meats.

A recent meta-analysis including 203 relevant studies (assessing 225 polymorphisms across 95 genes) found a total of 37 polymorphisms across 27 genes to be significantly associated with gastric cancer in Asians, and 12 polymorphisms across 11 genes in Caucasians (Camargo et al., 2006). In terms of salt consumption, studies *in vitro* have shown that several *H. pylori* genes associated with virulence (including *cagA*) were upregulated when the bacterium was cultured in a medium with high salt concentration. As a result, increased expression of CagA was observed leading to alteration of gastric epithelial cell morphology and function (Loh et al., 2009).

In the current study, we found more than half of the gastric carcinoma samples carried somatic mutation in the D-loop region and most of them were heteroplasmic. The D-loop region of mtDNA is a crucial position for replication and expression of the mitochondrial genome because it possesses essential transcriptional promoters and is the leading-strand origin of replication (Fliss et al., 2000). Furthermore, the D-loop region is hypervariable and susceptible to somatic mutations because of its distinctive triple-stranded DNA structure. Mitochondrial mutations may modify the function of normal oxidative phosphorylation chain which operates as a metabolic caretaker to prevent unexpected alterations to the glycolytic metabolic phenotype and also serves

as a gatekeeper to avoid improper production of genotoxic reactive oxygen species (Zhou et al., 2007). A total 49 mutations in hyper variable region I (HVI), 6 in non-coding region and 17 in hyper variable region II (HVII) were detected in the gastric cancer patients indicating that the D-loop region has a high mutation rate and might be in HVI region. The mtDNA D-loop contains the initial site of heavy chain replication and the promoters for heavy and light chain transcription. In the past few years, somatic mtDNA mutations have been identified in several types of human tumours (Polyak et al, 1998; Fliss et al, 2000; Maximo et al, 2000; Richard et al, 2000; Yeh et al, 2000; Hibi et al, 2001; Kirches et al, 2001; Liu et al, 2001; SanchezCespedes et al, 2001; Nomoto et al, 2002; Lievre et al, 2005), including gastric cancer (Habano et al., 2000). In the present study, for mtDNA mutations, we report a 15% frequency of D-Loop somatic mutations. Wu et al. (2005) reported that 48% (15/31) gastric carcinomas displayed somatic mutations in the D-loop region. Interestingly, in the same study, they showed that 67% cancers with the somatic mutations in the D-loop had insertion or deletion mutations in nucleotide position (np) 303–309 of the mononucleotide repeat region (poly-C track). In a separate study of the mtDNA D-loop region, three of eight gastric tumors showed similar D-loop alterations (Alonsoet al., 1997).Burgartet al. (1995) analyzed 77 gastric adenocarcinomas and found a 50 bp deletion in the mitochondrial D-loop region in 12.5% tumor samples. The deletion included the CSB2 region and was flanked by 9-bp direct repeats in the 303-309 base pair. In our study, we also found 303-309 poly-C repeat in higher frequency (63%) in Mizo population. During tumor clone expansion, continuous replication of mtDNA needs repeated melting and reannealing, which would make this region unstable and mutation formation (Pham et al., 2006). An alternative explanation is D-loop mutations might alter mtDNA transcription and lead to a respiratory chain alteration which is responsible for release of high reactive oxygen species

levels and therefore contribute to nuclear genome damage and also to cancer initiation and promotion (Bandy and Davison, 1990). The difference in the frequency of heteroplasmy across different age groups was statistically significant in GC (P=0.027), which suggests that heteroplasmy increase with age (Chinnery et al., 2005). And the different frequencies of alterations in the HVI and HVII distinct mtDNA D-Loop regions may reflect the importance of mutations at different nucleotide positions in tumor transformation. Consequently, the high frequency of mtDNA D-loop alterations presence in GC could perhaps be exploited as clinical markers for early cancer detection (Verma and Kumar, 2007).

The current study provides the firstconvincing evidence that COXI mutations play an important role in the etiology of gastric cancer. Gastric cancers have a significantly increased frequency of functionally important COXI mutations, and the introduction of a mtDNA mutation into gastric cancer cells (which inhibits OXPHOS and increases ROS production) increased their in vivo growth. The COXI mutations were significantly more frequent in gastric cancer tumor samples than in adjacent normal controls or in the general population. The COXI mutations significantly altered the highly conserved amino acids, which included both new heteroplasmic somatic and recurrent homoplasmic germ-line mutations. Hence, mtDNA COXI mutations appears to be a causal factor in the etiology of gastric cancer and this finding may explain why COXI mutations are so common in the gastric cancer and other cancers. Gastric cancer kills middle aged or older males and females, but the mtDNA is exclusively maternally inherited. Hence, deleterious COXI mutations that cause gastric cancer would have minimal effect on the genetic fitness of the mutant mtDNA. This study observed two novel heterozygous mutation, 6395C>G changes phenylalanine at position 164 to lysine, and C7336G changes a serine residue at 478 to

tryptophan in COXI which are highly conserved throughout evolution and is found in the tumor, adjacent normal and blood samples of the patients. This means that the mutations are maternally inherited and novel germline mutation in this population are associated with gastric cancer. In fact, the phenylalanine 164 is located in the transmembrane α helices IV of the COX I which are critical for the interaction between subunits I and III of the cytochrome c oxidase (Tsukihara et al., 1996). The substitution of threonine (which has the capacity of forming hydrogen bonds) by alanine with a small hydrophobic chain is expected to impair the contact between subunits I and III. Percentage of COXI helix and strand regions are altered due to the mutations (Figure 4.2). Due to the mutation in 6395C>G domain (annotated for HemeA and iron binding site of COXI protein), the protein transmembrane and substrate transporter activities might be lost. The mutation 7336C>G (478S>W) which is located in the topological domain (domain with certain spatial position) of mitochondrial matrix, might disrupt the topology and the resultant protein might be partially disordered (43%) as predicted by HOPE server.

We assessed mtDNA instability in the noncoding D-loop region in 80 gastric carcinomasamples. In the present study, mtDNA microsatellite instability within the D-loop region were frequently detected in gastric carcinomas (46%), most of which involved insertion or deletion of cytidine in the (C)n sequence. Burgart et al. (1995) demonstrated that mtDNA alteration in the D-loop region are found in gastric cancers (12.5%) and all the series had deletions in the 50 bp portion including the (C)n tract.We assume that the CMH in the mitochondrial genome plays a role in gastric carcinogenesis, because many carcinomas that showed heteroplasmy in the normal counterparts retained exactly the same proportion. But in our study, we have taken the adjacent normal tissue with the absence of tumor cell after confirmation by histochemistry. Hence, we can

get the real status of CMH in gastric tumor. The 303 (C)n tract is located in the conserved sequence block 2 (CSB2) that is an essential element for mtDNA replication (Jemt et al., 2015). Six tumor samples revealed both the 303 and 16185 poly (C)n tract instability and all the samples are taken from male gastric cancer patients. Thus, the poly (C)n tract instability frequently happened for male patients and its might be gender specific. From all the mtMSI gastric cancer samples, one sampleexhibited abnormal MSI for 303 poly(C)nin D-loop region. After sequencing the 303 nucleotide position, it was found that C_1TC_7 nucleotide tract was present instead of `C_7TC_6. Instability of these region may decrease in the copy number and alter the gene expression in the mitochondrial genome, because the D-loop is involved in the control of replication and transcription of the mtDNA (Lee et al., 2005).

On the other hand, mtMSI in the coding region were also detected for gastric cancer. All the coding region MSI were not detected in the normal counterpart, these mutations reflected generation of new mutant mtDNA during cancer development. The cases with NDV(C)6 and NDV(A)8 microsatellites instability showed frameshift mutation in the NDV gene and predictably led to truncated proteins that lacked a large portion of the C-terminus. Remarkably all these MSI are homoplasmic throughout all the tumor tissue. Hence, unlikely that only a less component of mutant mtDNA has accumulated in each carcinoma tissue without any selection because the number of mitochondrial genomes per cell is quite large (Shay and Werbin, 1987). But, an insertion or deletion of only one nucleotide in the ND5 sequence most likely did not result in a replicative advantage over normal mitochondrial genomes. Therefore, the truncated ND5 proteins probably conferred some specific tumor cells growth advantage. In other case, poly (C)6 tract in NDI gene was also observed for male and female patients, and predictably led

to frameshift mutations that produces a truncated protein. Only one tissue from female gastric cancer patients exhibited poly(C)5 microsatellite instability for NDI, but three male patients exhibited poly(C)7 MSI in tumor tissue. Hence the NDI poly(C)7 MSI may also be important in male gastric carcinogenesis for this population. Interestingly, all samples exhibited NDI and NDV coding region MSI, also exhibited D-loop 303 (C)n and 16185 (C)n tract instability. Therefore, ND1 and ND5 gene may be one of the target genes for mtMSI in gastric tumor. Both NDV and NDI gene products are subunits of the NADH dehydrogenase complex. Horton et al., (1996) detected mtDNA deletions in the NDI region in renal cell carcinomas suggesting that functional changes in a NADH dehydrogenase component may significantly influence carcinogenesis, although these alterations most likely do not accelerate tumor cell growth directly. Bandy and Davison (1990) proposed an alternative hypothesis that damage to the mitochondrial gene may result in impaired function of normal respiration and release abnormally high level of ROS which amplify damage to the nuclear oncogenes and tumor suppressor genes. Ozawa (1995) also suggested that production of ROS can be synergistically enhanced by the malignant cycle and result in a mtDNA mutation and pathogenicity. These hypotheses provide a possible mechanism by which specific mitochondrial mutations promotes carcinogenesis. Two samples for male and one sample for female exhibited cytidine insertion in the $(A)_7$ tract of COXI gene and stop codon was generating at 266 position of the COXI protein and one male and one female gastric cancer tumor sample exhibited cytidine deletion in the $(A)_7$ tract generating stop codon at the 271 amino acid position. Due to the COXI $(A)_7$ tract alteration, complex III and IV enzyme activities might be altered and it will reflect the mitochondrial pathogenicity in gastric cancer. Two samples exhibited $tRNA^{ser}(A)_6$microsatellite instability (A_6>A_7) which may affect the function of the mitochondrial encoded tRNASer 2 (AGU/C) (MT-TS2) in

mitochondrial protein synthesis and that almost half of the proteins synthesized by mitochondrial ribosomes are structural subunits of complex I (Fox, 2012).

Thus, D-loop is responsible for the regulation of mtDNA replication and transcription, its mutation leading to mutations in coding region and change of protein synthesis, and finally affects the function of respiration chain which hampers the energy supply of cells and produces volume of ROS. ROS results in injury to the genome and then induces cancer. High level of ROS is toxic through activating cell apoptosis and causing injuries to the genome. ROS might regulate cell apoptosis in the following ways. ROS is the message molecule of some transcription factors (such as Apaf-1) and can activate some useful components of cell apoptosis (Lovat et al., 2003). The increase of ROS is often accompanied with the decrease of intracellular anti-oxidant, resulting in imbalance between oxidant and reductive, which is just the common central step of cell apoptosis (Bianchi et al., 2001). Most people believe that ROS is necessary for cell apoptosis. High level of ROS inspires cell necrosis or drives cells to the way from apoptosis to necrosis. ROS not only participates in the process of cell apoptosis but also is a kinetin for cell division that promotes nuclear DNA mutation, cell mitosis and selective growth of tumor cells. ROS is relatively stable and easy to diffuse within cells and exist universally in various cell types. The formation and elimination of them are under strong cellular regulation. All the above properties make ROS extraordinary proper for secondary messengers. The level of intracellular ROS increases under extracellular stimulation signals such as cytokine and growth factor and then they take part in cellular signal transduction. Mitochondrial complex I of the electron transport has been recently shown to induce cell-cycle arrest through a pathway involving mitochondrial reactive oxygen species (ROS) generation, followed by activation of the homologs

of the c-Jun N-terminal kinase (JNK), the FOXO transcription factor, and the cell-cycle regulator p27 (Finkel and Hwang, 2000). One major source of intracellular ROS is the NADPH oxidases. NADPH oxidases catalyze the production of superoxide from O2 and NADPH (Sullivan and Chandel, 2014). In this study, we also observed that ROS are generated in more amount in the tumor sample which was having less enzyme activity of complex I and heteroplasmic mutations in D-loop, ND1 and ND5 genes. Heteroplasmic mutations have been observed to be enriched in tumors relative to normal tissue and have been implicated to confer a selective advantage in tumorigenesis (Larman et al., 2012). Heteroplasmic mutations in complex I have been shown to increase mROS, increase colony formation in soft agar, and increase tumor formation *in vivo* (Park et al., 2009). As a result, the percentage of cells in synthesis phase could reflect cell proliferation. Due to the high level of ROS production, the mitochondrial complex function might be lost and G1 might be arrested with high S phase DNA quantity (Figure 4.3A and B) (p value = 0.039). In this study, six gastric tumor samples where aneuploidy (containing COXI 6395C>G change, NDI and NDV microsatellites instability) resulted in less DNA content in G2/M phase, whereas high DNA content in S phase. Direct defects in mitochondrial complexes can also elicit such cell cycle defects. Indeed, mitochondrial dysfunction due to loss-of-function mutants in the ETC complexes I or IV induce G1 > S arrest. For most of the gastric tumor and adjacent normal tissues, high level of ROS generation (MFI) was found than endoscopic control group. Defects in respiratory enzymes due to mtDNA mutations or directly due to oxidative damage of the enzyme proteins and associated lipids can lead to enhanced ROS production. Cytochromec c oxidase subunit I may also contribute to decrease the level of ROS due to its antioxidant effects. Thus, a defect in COXI can cause enhanced ROS levels and a subsequent increase in mtDNA damage (Cavelier et al., 1995), which was also observed in the

present study. All the tumor samples containing COXI 6395C>Gproduces high level of ROS. This might be major cause of less activityof complex IV enzymes. The production of reactive oxygen species (ROS) is a consequence of monovalent reduction of molecular oxygen, and depends on the electron flow rate. The nitric oxide (NO) is a powerful modulator of oxygen uptake by reversible binding to COX and thus, it reduces O_2 utilization and increases ROS production. On exposure to NO from different sources, either NOS or NO-donors, cells stop growth at G1 or G2 phase, or high proliferation of cell in S phase (Arciuch et al., 2012). In our study, the COX activity are very less for some gastric tumor samples, which are subsequently producing high amount of ROS and high amount of proliferative S phase in cell cycle.

COX activity is described to be lower in colonic adenocarcinoma and hepatocellular carcinoma than in the normal cells (Sun and Cederbaum, 1980; Sun et al., 1981; Heerdt et al., 1990). COX deficiency is also one of the most frequent causes of respiratory chain defects in humans. Patients with COX deficiency present heterogeneous clinical phenotypes, including hepatic failure and encephalomyopathy (Barrientos et al., 2002). Although our analysis showed a decrease in COXI subunit protein, and we have found corresponding decrease in COXI mRNA gastric tumor cells. For some gastric cancer tumor samples, moderate levels of COXI protein expression was observed although the mRNA levels are very less (Figure 4.6C). The Phosphorylation status modulates mechanisms that control cell division and differentiation, dramatically affecting protein levels to incising up to optimal level. It is possible that the stability of COXI mRNA was affected, which may affect the COXI protein level. Together, our study suggests that COXI is post-transcriptionally regulated and modified by nuclear protein and that the lower level of COXI contributes to the reduced COX activity in tumor cells. Consistent with

our results in gastric cancer cells, COXI is also regulated at the post-transcriptional level (Pinkham et al., 1994). Damage to mtDNA is often highlighted as a factor affecting expression of mtDNA genes and thus contributing to secondary consequences of OXPHOS deficiencies. Considering its structure, maintenance, and close localization to sources of oxidative damage, this may well be a significant factor in gastric cancer progression.

The present study, indicates a statistically significant correlation between decreased cytochrome c oxidase subunit I expression in mRNA and protein level (western blotting and immunohistochemistry) and gene mutations in the stomach mucosa of patients with gastric cancer. This provides support for a possible cause-and effect relationship between these two biological processes. It has been shown that expression of mitochondrial proteins or loss of histochemical activity within stomach mucosa (Taylor et al., 2003) reflect stem cell mutations. It has been well documented that damaged mitochondria generate excessive reactive oxygen speciescausing lipid peroxidation (Kirkland et al., 2002) lose their membranepotential resulting in the opening of the mitochondrialpermeability transition pore. This initiates theprocess of autophagy (Rodriguez-Enriquez et al., 2009), which is the main degradative process in the cell responsible for the removal ofdamaged mitochondria and other organelles. They may include the release of proautophagic signals through the mitochondrial permeability transition pore and/or the presentation of abnormal aggregates of damaged, mis-folded proteins in mitochondria (Rodriquez-Enriquez et al., 2009). Cytochrome c oxidase subunit I is one of the 13 proteins encoded by mtDNA (complex IV), we hypothesize that the loss of mitochondrial-encoded cytochrome c oxidase subunit I protein expression in the stomach tumor mucosa of patients with gastric cancer or at high risk for gastric cancer may be due to specific mitochondrial mutations.

Consider a scenario in which the COXI gene is mutated in a mtDNA molecule within a given mitochondrion. The mutant mtDNA replicates along with wild-type mtDNA and upon mitochondrial division, copies segregate to daughter mitochondria. After a succession of mitochondrial divisions, the mitochondrion whichby chance has a higher proportion of the mutant mtDNA will exhibit a reduction in the cytochrome c oxidase subunit I protein (Payne et al., 2005). A mutational defect in any of the 13 mtDNA-encoded respiratory subunits can lead to respiratory chain dysfunction (Joseph et al., 2004; Suliman et al., 2004). If the cytochrome c oxidase subunit I mutation and loss of cytochrome c oxidase subunit I proteindecreases reactive oxygen species formation by the respiratory chain, there will be less lipid peroxidation, and the permeability transition pore will remain closed. This leads to the failure to elicit autophagy (Rodriquez-Enriquez et al., 2009) resulting in a replicative advantage to the mutant mitochondria and the accumulation over time of homoplastic mutant mitochondria within the cell (Brunk et al., 2002). We have also found similar kind of result for our study. Most of the novel COXI mutations and coding region MSI were homoplasmic in nature. It has been shown through cell fusion experiments and PCR amplification experiments in tissue from patients with mitochondrial disease, that mutant mtDNA can be functionally dominant over wild-type mitochondrial genomes (Shoubridge et al., 1990), can have a replicative advantage (Yoneda et al., 1992), and can be clonally expanded (Nekhaeva et al., 2002).

The state of Mizoram being one of the high risk/incidence regions of stomach cancer in India (Phukan et al., 2005; Ghatak et al., 2016), few epidemiological studies have been carried out earlier to find out various potential risk factors for stomach cancer. It has been pointed out that the consumption of Tuibur (smoke-infued tobacco water), cigarette smoking and smoked meet

consumption may be one of the important risk factors for the high prevalence of stomach cancer among the populace in Mizoram (Lalmuanpuii et al., 2015; Ghatak et al., 2016) which is consistent with the present study. Oxidative damage caused by smoking has also been shown to inhibit mitochondrial enzyme activity in platelets and cause mitochondrial dysfunction in alveolar macrophages (Smith et al., 1993). In addition, an increase in mtDNA content and decline in mitochondrial function also occurs in response to DNA damaging agents including tobacco related product (Lewis et al., 2002). Although, a causal relationship exists between *H. pylori* infection and the development of gastric carcinoma, the molecular mechanisms underlying this relationship have remained elusive. In this study, we show that *H. pylori* infection may correlate with mitochondrial gene mutations and mtMSI in the gastric mucosa of patients with various gastric pathologies. In this study, a huge frequency of *H. pylori* positive samples were observed with gastric cancer and some of the samples are infected with cagA genotyping. After association study between *H. pylori* with or without cagA genotyping and mt-MSI status, we observed a significant associations. Machado et al.(2009) showed that gastric epithelial cells are closely related to the genetic instability of mitochondrial DNAdue to *H. pylori* infection. In our study, it was found that as carcinogenesis progresses, the level of mtMSI increases, and the rate of mtMSI was found to be significantly higher in the *H. pylori*-positive groups than the *H. pylori*-negative groups, implying that mtMSI might play a role in the occurrence of those gastric cancers that are *H. pylori* positive. Reactiveoxygen species (ROS) are commonly released in gastricmucosae, which was found in the current study, that are inflamed as a result of infection with*H. pylori*, especially with CagA+ strains, and could be responsible formtMSI-positive gastric cancer (Bagchi et al., 2002;Shimoyama et al., 2002). The mitochondrial genome is particularly susceptible to oxidative damage and mutation because of the high rate of ROS

generation in this organelle and its inefficient DNA repair system (Akhmedov et al., 2015; Szczepanowska et al., 2015). Increased damage caused by ROS and defective DNA repair are the two causes that have been proposed to explain mtMSI in *H. pylori*associated gastric cancer [Akhmedov et al., 2015]. In the current study, we did not find an obvious relationship between mtMSI and tumor size, the depth of invasion, node metastasis or clinical stage, indicating a limited role of mtMSI in predicting the prognosis of gastric carcinoma. However, a marked difference in mtMSI was noted in gastric cancers distinguished by histological type. mtMSI was significantly more frequent in intestinal-type gastric cancers than diffuse-type gastric cancers, suggesting that mtMSI is a predisposing event in intestinal gastric cancer.

To our knowledge, the present study is a novel finding in terms of the possible role of mtDNA COXI and D-loop mutations in gastric carcinogenesis. In summary, we have identified novel mutations in mitochondrial control region and cytochrome c oxidase subunit I genes, which is frequently less expressed in gastric tumor cells. Mitochondrial microsatellites instability is also a major factors for progression of gastric tumor and it was significantly associated with *H. pylori* infection with cagA positive or negative and smoked meat consumption for Mizo population. All the tumor samples containing COXI 6395C>G heterozygous mutation, are producing high level of ROS. The COXI mutations altered significantly more conserved amino acids, and they included both new heteroplasmic somatic and recurrent homoplasmic germ-line mutations. This was might be major cause of less activity of complex IV enzymes. Due to the COXI $(A)_7$ tract alteration the complex III and IV enzyme activity might be altered and it will reflect to the mitochondrial pathogenicity in gastric cancer. Mitochondrial gene alterations may attribute for gastric cancer risk in Mizoram, Northeast Indian population. The analysis of mtMSI and D-loop

alteration might help to identify patients at high risk for gastric cancer diagnosis. The mutations of D-loop and COXI takes part in the carcinogenesis and progression of gastric cancer through the effect of increased ROS and altering the mitochondrial ETC complex enzyme activity and cell cycle. The mutations in COXI markedly decreased overall immunostaining of cytochrome c oxidase subunit I may prove to be reliable biomarkers of gastric cancer risk, because they occur at elevated frequency in the stomach mucosa of patients with, and at high risk for, cancer. We have also demonstrated that mtMSI is an early and important event in the progression of gastric carcinogenesis, especially in intestinal-type gastric cancer. *H. pylori* infection contributes to mtMSI in gastric cancer development. Our results support a role for mtMSI in different mechanisms of gastric carcinogenesis. Because the majority of patients with *H. pylori* infection will not progress to cancer and only a subset of these patients harbor mtMSI, it is conceivable that patients with *H. pylori* infection displaying mtMSI are at greater risk of developing gastric cancer than those without instability. Further, prospective studies with larger sample size analyzing gene–environment relations, mitochondrial mutations and microsatellites instability in different geographic areas, and ethnic groups are needed to assess the significance of each factor for gastric cancer.

Chapter V

Characterization of novel Missense and splice site mutations in TP53 gene and their genetic and functional analysis in gastric tumorigenesis

1. Introduction

Gastric cancer (GC) is second most common malignancy in the world, and is a leading cause of cancer mortalityin East Asia (Northeast China and India, Korea, Japan), Eastern Europe, and the Andean countries of South America (Ferlay et al., 2012). According to the Indian council of Medical Research for 2013, gastric cancer is the second most common cancer among men and the third most among women in northeast region of India. Intestinal type of tumors are the most common in high risk gastric cancer population (Correa, 2000) which is developed in the antrumpart of the stomach and lead to atrophic multifocal gastritisand intestinal metaplasia due to the infection of *Helicobacter pylori*(Nomura et al., 1990). The etiology of gastric cancer is complex, involving genetic susceptibility, familial predisposition and combinations of dietary and environmental factors that drive the accumulation of genetic alterations associated with increasing genetic instability. Other major risk factors for gastric cancer are smoking, tobacco consumption, high dietary intake of salt and nitrates, smoked meat consumption and low intake of antioxidant micronutrients [Ghatak et al., 2016].

The p53 gene, located on the short arm of chromosome 17, encodes a protein which plays a critical role tumor suppression, DNA transcription and cell cycle regulation (Fearon and Vogelstein, 1990). It has been suggested that the frequency of TP53 mutations increases with the progression of GC from normal gastric mucosa, however, the reported frequency of mutations varies widely between studies (Shiao et al., 2000). TP53 mutations have been shown to be a late event in the progression of colorectal cancer (Fearon and Vogelstein, 1999); however this remains to be systematically explored in GC.In addition to direct

mutations of the TP53 gene, a loss of p53 function can occur through elevation in the levels of one or more of its negative regulators. Among these, Mdm2 (known collectively as Mdm proteins) are the most powerful regulators of TP53. They are frequently altered in human cancer, and their expression is often positively correlated with that of wild-type p53 (Wade and Wahl, 2013). Mutations in TP53 usually result in the accumulation of mutant p53 protein, which disrupts the auto-regulatory loops with the Mdm proteins.Recent studies have suggested that elevated levels of Mdm2 are detected in GC compared to adjacent normal tissue and correlate with poor prognosis (Ye et al., 2013). Recent studies have demonstrated that polymorphisms of apoptosis-related genes, such as TP53 codon 72 and the MDM2 promoter, are associated with the susceptibility or prognosis of solid tumors (Xu et al., 2005; Ohmiya et al., 2006).The p53specific genotype of the codon 72 polymorphism could be a risk factor for certain tumors and by making an environment favorable for tumor formation (Hiyama et al., 2002).

We hypothesize that gastric cancers arising in with different immunoreactivitypatterns of p53 protein might exhibit different molecular abnormalities in addition to their differing epidemiological backgrounds.The aim of this study was to determine whether p53 protein expression and/or TP53 mutation status could be used as a biomarker to allow targeted screening of high risk groups with gastric cancer.

2. Materials and methods

2.1.Subjects and sample collection

This study consisted of 80 random patients with primary gastric cancer (aged from 37 to 79 years) and the paired biopsy specimens of gastric tumour and adjacent non-tumour sites were collected during surgical resection from Civil Hospital, Aizawl, Mizoram. Eighty unrelated healthy controls (aged from

31 to 73 years) belonging to the same ethnic group were also collected. None of these patients received pre-operative chemotherapy or radiotherapy. The samples were all confirmed histologically and all subjects provided informed consent for obtaining the study specimens. Patients with gastric neoplasms including adenocarcinoma (MALT lymphoma, stromal, or carcinoid tumors) were included. Patients with secondary or recurrent GC, previous history of other malignancies, or refusal to participate were excluded from the study. Medical and demographic charts were reviewed using a standard protocol to obtain information on cancer treatment, clinical stage, dietary habits, previous disease history, physical activity, tobacco, and alcohol use by an in-person interview using a structured questionnaire (Ghatak et al., 2014). The study protocol was approved by the Institutional Review Board of all institutes involved in the study.

2.2. DNA extraction from the blood sample

The lymphocytes from whole blood were separated by lysing the RBCs using a hypotonic buffer (ammonium bicarbonate and ammonium chloride, Hi-media) with minimal lysing effect on lymphocytes. Three volumes of RBC lysis buffer were added to the blood sample, mixed by vortexing and inverting thoroughly for 5 min and centrifuged (Eppendorf 5415R, Germany) at 2,000 X g for 10 min. The lymphocytes was used for DNA extraction by modified protocol of Ghatak et al. (2013).

2.3.DNA and RNA extraction from the tissue samples

Deparaffinization was carried out by adding 1 ml of xylene to the tissue section in each microfuge tube, followed by vigorous vortexing for 10 minutes. The mixture was centrifuged at 12000 rpm for 10 minutes. The supernatant was discarded and the deparaffinization steps were repeated once again, followed by rehydration through subsequent washings with 100%, 90% and 70% absolute ethanol diluted

in RNase free DEPC treated water, respectively. The deparaffinised tissue and surgical resection specimen was used for the DNA and RNA co-extraction by the modified protocol of Ghatak et al. (2015).

2.4.PCR amplification of the TP53 exon region with splicing site

The Tp53 different exon region with splicing site was amplified by polymerase chain reaction (PCR) using eight different primers pair (Table 5.1). PCR was carried out in 25 µl total reaction volumes, each containing 100 ng of template DNA, 0.2 pM of each primer, 1X PCR buffer, 1.5 mM $MgCl_2$, 200 mMdNTPs, and 1 unit of Taq DNA polymerase (Fermentas, Germany). The reaction mixture was heated to 94^0 C for 5 min, followed by 30 cycles each consisting of 1 min denaturation at 94^0 C, 1 min annealing at respective temperature (Table 5.1), 1 min of extension at 72^0 C and a final 5 min extension at 72^0 C. The PCR amplification products (3 µl) were subjected to electrophoresis in a 1.2 % agarose gel in 1X TBE buffer at 80 Vh for 30 min, stained with Ethidium Bromide, and images were obtained in GBOX gel documentation systems (UK). PCR products were purified with a Qiagen gel extraction kit (Qiaquick columns; Qiagen, Chatsworth, CA) and stored at -20^0 C until for sequencing.

2.5.Immunohistochemical analysis

Immunohistochemical staining for tumour-associated gene products (Tp53, MDM2) was performed using a streptavidin peroxidase procedure after an antigen retrieval process using microwave. Tp53 and MDM2 anti-antibodies were selected among various commercially available antibodies, after test procedures using human control slides for immunohistochemistry (Abchem, Japan, DO-1, ab8645). Cut-off values for Tp53 and MDM2 antibodies were based on previously established cut-off values (Miyazaki et al., 2000). For Tp53 and MDM2, the immunolabelling pattern of each case was scored as positive (strong labelling), weakly positive (faint staining), or negative (absence of staining), and the extent of immunolabelling was also categorized as diffuse (if the entire neoplasm was labelled) or focal (if focal

loss of expression was noted) (Iacobuzio-Donahue et al., 2002). For statistical analysis of this large amount of data (Bouras et al., 2002), the results of immunostaining were considered to be positive if 10% the cancer cells showed staining

Exon	Primer (5' -3')	Product size (bp)	Ta (^{0}C)
1	F: CACAGCTCTGGCTTGCAGA R: AGCGATTTTCCCGAGCTGA	442	53
2	F: AGCTGTCTCAGACACTGGCA R: GAGCAGAAAGTCAGTCCCATG	317	54
3-4	F: AGACCTATGGAAACTGTGAGTGGA R: GAAGCCTAAGGGTGAAGAGGA	631	59
5-6	F: CGCTAGTGGGTTGCAGGA R: CACTGACAACCACCCTTAAC	550	58
7	F: CTGCTTGCCACAGGTCTC R: TGGATGGGTAGTAGTATGGAAG	283	59
8-9	F: GTTGGGAGTAGATGGAGCCT R: GGCATTTTGAGTGTTAGACTG	455	63
10	F: CTCAGGTACTGTGTATATACTTAC R: ATACTACGTGGAGGCAAGAAT	351	60
11	F: TCCCGTTGTCCCAGCCTT R: TAACCCTTAACTGCAAGAACAT	476	58

Table 5.1. Primers designed in the present study for amplification of exons and splice sites in TP53 gene.

in nucleus. MDM2 protein expression status was considered to be positive when more than 10% of the cancer cells showed cytoplasmic staining.

2.6.SSCP analysis

The TP53 all exon region was screened for mutations with a rapid, sensitive method of SSCP (Gayther et al., 1998). An aliquot of 0.75 µl of each PCR product was diluted with an equal volume of water and mixed with 1.5 µ1 of 95% formamide. This mixture was denatured at 95°C for 5 min, cooled on ice and 2 µl was used for loading on SSCP gel (8% non-denaturing polyacrylamide gels). SSCP Gels were pre-run at 400 V, 20 mA, 2 W, for 10 or 50 volt-hours (Vh). Electrophoresis was performed at 400 V, 20 mA, 2 W, for 200-300 Vh. Electrophoresis was carried out at either 4, 10, 15 or 20°C depending on the optimal temperature for a given PCR fragment. The gels were ethidium bromide stained, and gel documented by Syngen-G-BOX(USA) machine.

2.7.Sequence analysis

The samples exhibiting mutation or polymorphism and instability, after SSCP analysis, was taken for further sequencing and mutation analysis. All products were sequenced from opposite directions to ensure reading accuracy. Sequences and chromatograms obtained were examined by chromas software version 2.13, DNA baser and align by BLAST [www.ncbi.nlm.nih.gov/blast]. All the sequences containing the mutation ware evaluated for their potential pathogenicity using the following algorithms: DNA baser version 3.5.4.2, Codon Code aligner version V.4.2.2, The Tp53 all exon will be checked from

the Tp53 database [Tp53 data retrieving Excel file of International agency for Research on Cancer, World Health Organization] and from Gene card data base [HGNC: 11998; Entrez Gene: 7157; Ensembl: ENSG00000141510; OMIM: 191170; UniProtKB: P04637].Mutation taster [www.mutationtaster.org/], PolyPhen-2 [http://genetics.bwh.harvard.edu/pph], SIFT [http://sift.jcvi.org], Mutation Assessor [http://mutationassessor.org/]. The MEGA Align algorithm was used at two depths of alignment [Cancer to Normal and Normal to database sequences]. The results of PolyPhen-2 was retrieved from the original webpage [version 2.2.2] but also from version 2.0.22 run by PON-P and version 1 run by Condel, which use them for weighted average scores.

2.8.PCR-Based Pathogen Detection

The presence of *H. pylori* infection was determined in patients with GC by multiplex PCR amplification of 16S rRNA and CagA genes using Hp1-Hp2 and CagAF-CagAR primers, respectively (Table 5.1) (Ghatak et al., 2016). The PCR products were electrophoresed with 1.5% agarose gel. *H. pylori* infection was defined by the presence of an intact band of 109 bp (16S rRNA) and 400 bp (CagA gene). The presence of EBV type1/type 2 infections was carried out by a standard PCR assay crossways type-specific regions of EBNA3C gene using previously described primer sets (Table 5.1) (Kingma et al., 2997). The PCR amplification was carried out for 35 cycles at 95 °C for 45 seconds, 56 °C for 45 seconds, 72 °C for 1 minutes followed by a final extension at 72 °C for 10 minutes. The PCR products were analyzed by 12% polyacrylamide gel, and the presence of 153 and 246 bp was considered as type 1 and type 2 EBV, respectively.

2.9.Reverse Transcriptase PCR (RT-PCR) for TP53 Expression

First-strand synthesis of cDNA was performed using Superscript II reverse-transcriptase (Life Technologies Inc., USA). Five µg aliquot of total cellular RNA was used for each reverse transcription reaction, and one-tenth of this reaction was used for PCR. Primers targeting TP53 exons 7 and 8 region (F: TAACAGTTCCTGCATGGGCGGC and R: AGGACAGGCACAAACACGCACC) and exons 1 and 3 of β-actin (F: ACCATGGATGATGATATCGC and R: ACATGGCTGGGGTGTTGAAG) were used (DeBruin et al., 2000). The PCR was run for 35 cycles consisting of 94° C for 30 seconds, 1 minutes at gradually decreasing temperatures using a touchdown protocol (4 cycles each at 62° C and 60° C; and 30 cycles at 58° C), and 74 °C for 5 minutes. RT-PCR products were electrophoresed in 10% polyacrylamide gel and visualized by ethidium bromide staining.

2.10. Densitometric Analysis

Semiquantitative RT-PCR gels and Western blot were analyzed to compare the banding pattern and their molecular mass using gel image analysis software Syngen G-Box (Sacramento, CA, USA) and ImageJ.

2.11. Cell cycle estimation

0.1 g of grossly gastric tumor and adjacent normal gastric mucosa tissue were used for cell cycle analysis. Cells were harvested by mechanical dis-aggregation and fine-needle aspiration. Two separate aliquots of 6 X 10^6 tumor cells were prepared for each sample. Pellets were incubated with 250 mL of 0.1 % RNAse (Sigma, St Louis, MO, USA) and 50 mg/mL Propidium iodide (presence of Sodium citrate and TritonX-100) for 30 min at 37°C and flow cytometric analysis was performed by Facs Canto and DIVA software (BD, Germany). Four distinct phases could be recognized in a proliferating cell population: the G1, S-(DNA synthesis phase), G2- and M-phase (mitosis). G2- and M-phase could not be discriminated because of the presence of identical DNA content (Blanco et al., 2013). The data obtained was analyzed

using the ModFit LT software (DNA Modeling System) version 2.0 (Verity Software House, Inc.) and single parameter histograms were obtained.

2.12. Statistical Analysis

Tp53 gene mutations among case–control subjects was tested for Hardy–Weinberg equilibrium by a chi-square (v2) test with one degree of freedom (df). The mutations in each group were estimated using odds ratios (ORs), and 95% confidence intervals (CIs). Additionally, logistic regression analyses were conducted to compute the potential confounder's influence of both genetic and environmental factors for gastric cancer. Then, the independent effect of risk factors was investigated in a multivariate model (introducing all variables and terms of interactions) keeping only those statistically significant or showing a confounding effect on the studied factors. The likelihood ratio test was used to decide whether to retain each covariate in the model. Sex, cigarette smoking, family history of gastric cancer, tuibur consumption, alcohol intake, smoked meat/vegetable intake, high consumption of salt and pickle, high consumption of fermented fat, *H. pylori* status and genotyping, EBV status, and genotyping were all considered in the regression model as potential confounders to evaluate the association of risk factors and susceptibility to gastric cancer. For all tests, a two-sided p-value <0.05 was considered statistically significant. All statistical analyses were performed using SPSS 20.0 program (SPSS, Madrid, Spain) and conformed by R-console statistical package ver3.3.0 (The R Foundation for Statistical Computing).

3. Results

3.1. Demography and Clinicopathological characteristics of gastric cancer patients

Gastric cancer was more prevalent in males (70%) among all the collected samples from Mizo population. Bivariate analysis of all the demographic factors showed that*H. pylori* infection (OR: 1.168; 95% CI: 0.365 – 3.733; P=0.044) with cagA genotyping (OR: 8.298; 95% CI: 0.365 – 6.578; P=0.005), Smoked Meat/Vegetable Intake (OR: 6.181; 95% CI: 2.528 – 12.346; P=0.001) and High fermented fat intake (OR: 3.387; 95%CI: 1.902 – 6.031; P=0.0003) were significant risk factors for gastric cancer in Mizo population (Table 4.3). High salt intake also exhibited potential factors for gastric cancer for this population (OR: 2.687; 95%CI: 0.244 – 3.936; P=0.078). No significance was found with consumption of alcohol, and cigarette smoking (Table 4.3). The median age in the young group was 36 years (range16–45) with a low proportion (35%) of patients than the older aged group (65%). The older patient group had a higher proportion of men (40%) than women (25%) (P <0.326) (Table 4.4). Tumour size and subtotal gastrostomy were more in case of older age than in younger patients. Most of the gastric cancer patients were operated with stage II tumor. The symptoms at presentation in both groups are shown in Table 4.4. The most common presenting symptom was abdominal pain (62.5%)followed by weight loss (42.5%; p = 0.05). Hemorrhage was more in young patients group and Vomiting was more in older age group.

3.2. Tp53 Gene mutations and cell cycle analysis

Screening of all the exons of TP53 gene showed that 60% patients with gastricadenocarcinoma had mutations and majority of them were somatic in nature. A total of 17 mutations were found in Exons 3, 5, 6, 8 and 11 (Table 5.2). Exon 3 exhibited 5 non-synonymous mutations including two novel mutations (11292A>T, 11156T>A), and all the variations were predicted polymorphism, except 11156T>A which was disease causing with high polyphen2 score.

Exon 5 exhibited four novel mutations including a frameshift mutation (12616_12617 insG; E204G) due to guanine insertion at g.12616 position. Due to this mutation, early stop codon was generated at the 209 codon position. Another adenine deletion was also found in codon 204 in an intestinal type of gastric cancer patient at the position of g.12619 which led to frame shift mutation, generating a stop codon at 247 codon position. Hence, the codon 204 is a mutational hot spot for TP53 gene in Mizo gastric cancer patients.

Exon 6 exhibited three mutations including 2 novel frameshift. All the mutations were predicted as disease causing by mutations taster with high polyphen2 score. Single frameshift mutation (12653_12655delTGT, S215R) at the position of g.12653 led to generating a premature 342 position stop codon. This three nucleotide deletion is highly pathogenic as predicted by SIFT, polyphen2 and was found in an intestinal type of gastric cancer patient.

Exons 8 and 11 exhibited 2 and 3 non-synonymous mutations, respectively. All the mutations in exons 8 and 11 are disease causing as predicted by mutation taster, except 13811G>A and 17873G>C. All the mutations reported in exon 11 are novel mutations for gastric cancer.

3.3.Deep intronic mutations and exon skipping

We have found a novel deep intronic heterozygous mutation (c.693-2A>G) at exon 6 splice acceptor position in 2 gastric tumor samples. The c.693-2A>G mutation in the mutant induces a fully penetrant splicing defect. Since this variant lies within the highly conserved splice donor consensus sequence at position-2 from the 5'-end of exon 6, a defect in splicing is very likelyin TP53 gene. We performed semi-quantitative RT-PCR experiment toevaluate the effect of the variant c.693-2A>G on the splicing of the

TP53 messenger RNA and found thatthe variant leads to skipping of a 24 nucleotide in exon 6 (Figure 5.1). This is a very important position for the TP53 gene because it has been assigned as DNA binding domain in uniport database. We have observed significant cell cycle alteration associated with this mutation and these samples exhibited aneuploidy with high S phase (Figure 5.2).This mutation was also significantly associated with TP53 protein and gene expression. The samples with this mutation exhibited negative expression for p53 protein and gene (Figure 5.2 and 3). This variant affects a highly conserved amino acid

Exon Number	Mutation	Frequency (%)	Amino acid change	Mutation Taster/Polyphen2 score/ SIFT	Effect of the Mutation
Exon 3	11292 A>T*	5	41D>V	Polymorphisms/ 0.027 (Benign)	• The mutation introduces more hydrophobic residues at this position. This can result in loss of hydrogen bonds and disturb correct folding. • This position regulates Sequence-Specific DNA Binding Transcription Factor Activity. So the Transcription activation (acidic) might be lost.
Exon 3	11385C>G	28.75	72P>R	Polymorphisms/ 0.083 (Benign)	• Hydrophobic interactions, either in the core of the protein or in the surface will be lost. • Interaction with histone acetyltransferases EP300 and methyltransferases (HRMT1L2) will be lost.
Exon 3	11152G>T	2.5	31V>F	Polymorphisms/ 0.273 (Benign)	• Mutant residues in bigger in size, this might lid to bumps.
Exon 3	11156T>A*	8.75	32L>Q	Disease causing/ 0.376 (Pathogenic)	• Interaction with Cell Cycle and Apoptosis Regulator 2(CCAR2) gene might be lost.
Exon 3	11298T>C	20	43L>S	Polymorphisms/ 0.209 (Benign)	• The wild type residue is located in a region annotated in uniport to form an alpha helix, but the mutated residue does not prefer alpha helix at secondary structure.
Exon 5	12616T>C*	3.75	203V>A	Polymorphisms/ 0.000(Benign)	• The mutant residue is smaller than the wild-type residue. • The mutation will cause an empty space in the core of the protein. • Mutant residue was observed more often at this position in other homologous sequences. This means that more proteins exist with that mutant residue than with the wild type residue. It is probably unlikely that your mutation of interest will be damaging for the protein.
Exon 5	12616_12617insG* E204Gfs*209	2.5	204E>G	Disease causing /0.502 (probably damaging)	• Framshift mutation, generating stop codon at 209 • This will cause a possible loss of external interactions. • The hydrophobicity of the wild- type and mutant residue

					differs. • Mutant residue is located near a highly conserved position.
Exon 5	12618G>A*[a]	13.75	204E>K	Polymorphisms/0.004 (Benign)	• The mutant residue is bigger than the wild-type residue. • The residue is located on the surface of the protein, mutation of this residue can disturb interactions with other molecules or other parts of the protein. • Mutant residue is located near a highly conserved position.
Exon 5	12619_12619 delA*	2.5	204E>G	Disease causing/ 0.538(probably damaging)	• Frameshift mutation, generating stop codon at 247 • This will cause a possible loss of external interactions. • The hydrophobicity of the wild- type and mutant residue differs. • Your mutant residue is located near a highly conserved position.
Exon 6	12659_12660 insG* V218Gfs*222	1.25	218V>G	Disease causing/ 1.000 (probably damaging)	• Frameshift mutation, generating stop codon at 222 • The hydrophobicity of the wild- type and mutant residue differs. • The mutation will cause loss of hydrophobic interactions in the core of the protein • Mutant residue is located near a highly conserved position.
Exon 6	12654G>T[a]	10	216V>L	Disease causing/0.997 (probably damaging)	• The mutant residue is bigger than the wild-type residue. • The wild-type residue was buried in the core of the protein. The mutant residue is bigger • Mutant residue is located near a highly conserved position.
Exon 6	12653_12655d elTGT S215Rfs*342	2.5	215S>R	Disease causing/1.000 (probably damaging)	• Frameshift mutation, generating stop codon at 342 • The hydrophobicity of the wild- type and mutant residue differs. • The mutation will cause loss of hydrophobic interactions in the core of the protein and mutant residue is located near a highly conserved position.

Exon 8	13811G>A	26.25	298E>K	Polymorphism/0.002 (Benign)	• The residue is located on the surface of the protein, mutation of this residue can disturb interactions with other molecules or other parts of the protein. • Mutant residue is located near a highly conserved position.
Exon 8	13834G>T*[a]	2.5	305K>N	Disease causing/0.999 (probably damaging)	• Mutant residues in bigger in size, this might lid to bumps.
Exon 11	17919G>A*[a]	2.5	391D>N	Disease causing/0.580 (probably damaging)	• The wild-type residue was buried in the core of the protein. The mutant residue is bigger
Exon 11	17925G>C*[a]	1.75	393D>H	Disease causing 1.000 (probably damaging)	• Mutant residue was observed more often at this position in other homologous sequences. This means that more proteins exist with that mutant residue than with the wild type residue.
Exon 11	17873G>C*[a]	10	375Q>H	Polymorphism/0.892 (probably damaging)	• Mutant residue is located near a highly conserved position.

Table 5.2. Mutations of TP53 geneand their pathogenicity in gastric tumor samples.

'*' - representing novel mutations; 'a' - representing heterozygous mutations.

Bibliography

Aberle H, Schwartz H, Kemler R (1996) Cadherin-catenin complex: protein interactions and their implications for cadherin function. J. Cell. Biochem. 61: 514–523.

Agudo A, Sala N, Pera G, Capellá G, Berenguer A, García N,et al. (2006) Polymorphisms in metabolic genes related to tobacco smoke and the risk of gastric cancer in the European prospective investigation into cancer and nutrition. Cancer Epidemiol. Biomarkers Prev. 15: 2427-2434.

Akhmedov AT, Marín-García J (2015) Mitochondrial DNA maintenance: an appraisal. Mol. Cell Biochem. 409: 283–305.

Alberts SR, Cervantes A, van-de CJ (2003) Gastric cancer: Epidemiology, pathology and treatment. Ann. Oncol. 14: 31–36.

Ali-Osman F, Akande O, Antoun G, Mao JX, Buolamwini J (1997) Molecular cloning, characterization, and expression in *Escherichia coli* of full-length cDNAs of three human glutathione S-transferase Pi gene variants. Evidence for differential catalytic activity of the encoded proteins. J. Biol. Chem. 272: 10004-10012.

Alonso A, Martin P, Albarran C, Aquilera B, Garcia O, Guzman A, et al. (1997) Detection of somatic mutations in the mitochondrial DNA control region of colorectal and gastric tumors by heteroduplex and single-strand conformation analysis. Electrophoresis 18: 682–685.

Anon (2011) Chapter 2: Size, Growth Rate and Rural-Urban Distribution of Populations in Mizoram, Provisional Populations Totals – Mizoram, Census of India, India. pp: 7-13.

Arand M, Mühlbauer R, Hengstler J, Jäger E, Fuchs J, Winkler L, Oesch F (1996) A multiplex polymerase chain reaction protocol for the simultaneous analysis of the glutathione S-transferase GSTM1 and GSTT1 polymorphisms. Anal. Biochem. 236: 184-186.

Arciuch AGV, Elguero EM, Poderoso JJ, Carreras CM (2012) Mitochondrial Regulation of Cell Cycle and Proliferation. Antioxid. Redox Signal 16(10): 1150–1180.

Arnold K, Bordoli L, Kopp J, Schwede T (2006) The SWISS- MODEL workspace: a web-based environment for protein structure homology modelling. Bioinformatics 26: 195-201.

Bagchi D, McGinn TR, Ye X, Bagchi M, Krohn RL, Chatterjee A, et al. (2002) *Helicobacter pylori*-induced oxidative stress and DNA damage in a primary culture of human gastric mucosal cells. Dig. Dis. Sci. 47: 1405–1412.

Ballinger SW, Schurr TG, Torroni A, Gan YY, Hodge JA, Hassan K, et al. (1992) Southeast Asian mitochondrial DNA analysis reveals genetic continuity of ancient mongoloid migrations. Genetics 130: 139–152.

Bandy B, Davison AJ (1990) Mitochondrial mutations may increase oxidative stress: Implications for carcinogenesis and aging? Free Radic. Biol. Med. 8: 523–539.

Barrientos A, Barros MH, Valnot I, Rotig A, Rustin P, Tzagoloff A (2002) Cytochrome oxidase in health and disease. Gene 286: 53–63.

Becker KF, Atkinson MJ, Reich U, Becker I, Nekarda H, Siewert JR, Hofler H (1994) E-cadherin gene mutations provide clues to diffuse type gastric carcinomas. Cancer Res. 54: 3845.

Berloco P, Russo F, Cariola F, Gentile M, Giorgio P, Caruso ML, et al. (2003) Low presence of p53 abnormalities in H. Pilori-infected gastric mucosa and in gastric adenocarcinoma. J. Gastroenterol. 38(1): 28–36.

Bianchi NO, Bianchi MS, Richard SM (2001) Mitochondrial genome instability in human cancers. Mutat. Res. 488: 9-23.

Blanco R, Rengifo CE, Cedeño M, Frómeta M, Rengifo E (2013) Flow Cytometric Measurement of Aneuploid DNA Content Correlates with High S-Phase Fraction and Poor Prognosis in Patients with Non-Small-Cell Lung Cancer. ISRN Biomarkers 2013: 1-8.

Boccia S, La-Torre G, Gianfagna F, Mannocci A, Ricciardi G (2006) Glutathione S-transferase T1 status and gastric cancer risk: a meta-analysis of the literature. Mutagenesis 21: 115–123.

Boccia S, Sayed-Tabatabaei FA, Persiani R, Gianfagna F, Rausei S, Arzani D, et al. (2007) Polymorphisms in metabolic genes, their combination and interaction with tobacco smoke and alcohol consumption and risk of gastric cancer: a case-control study in an Italian population. BMC. Cancer 7: 206.

Borrmann R (1926) Geschwelste Des Magens und Duodenums. In: F. Henke and O. Lubarsch, Verdauungsschlauch, pp. 865-1054. Berlin: Springer.

Bouras T, Southey MC, Chang AC, Reddel RR, Willhite D, Glynne R, et al. (2002) Stanniocalcin 2 is an estrogen-responsive gene coexpressed with the estrogen receptor in human breast cancer. Cancer Res. 62: 1289–1295.

Brooks-Wilson AR, Kaurah P, Suriano G, Senz J, Grehan N, Butterfield YS,et al. (2004) Germline E-cadherin mutations in hereditary diffuse gastric cancer: assessment of 42 new families and review of genetic screening criteria. J. Med. Genet. 41: 508–517.

Brunk UT, Terman A. Minireview (2002) The mitochondrial-lysosomal axis theory of aging. Accumulation of damaged mitochondria as a result of imperfect autophagocytosis. Eur. J. Biochem. 269: 1996 – 2002.

Burgart LJ, Zheng J, Shu Q, Strickler JG, Shibata D (1995) Somatic mitochondrial mutation in gastric cancer. Am. J. Pathol. 147: 1105–1111.

Cai L, Yu SZ, Zhang ZF (2001) Glutathione S-transferases M1, T1 genotypes and the risk of gastric cancer: a case-control study. World J. Gastroenterol 7: 506-509.

Camargo MC, Mera R, Correa P, Peek RM, Fontham ETH, Goodman KZ,et al. (2006) Interleukin-1β and interleukin-1 receptor antagonist gene polymorphisms and gastric cancer: a meta-analysis. Cancer Epidemiol. Biomarkers Prev. 15(9): 1674–1687.

Campos F, Carrasquilla G, Koriyama C, Serra M, Carrascal E, Itoh T, et al. (2006) Risk factors of gastric cancer specific for tumor location and histology in Cali, Colombia. World J. Gastroenterol. 12: 5772-5779.

Carew JS, Huang P (2002) Mitochondrial defects in cancer. Mol. Cancer 1: 9.

Carneiro F, Machado JC, Seruca R, Sobrinho-Simoes M (1999) E-cadherin changes in gastric carcinoma. Histopathology 35: 477-478.

Cavallaro U, Christofori G (2004) Cell adhesion and signaling by cadherins and Ig-CAMs in cancer. Nat. Rev. Cancer 4: 118–132.

Cavelier L, Jazin EE, Eriksson I, Prince J, Bave U, Oreland L, et al. (1995) Decreased cytochrome-c oxidase activity and lack of age-related accumulation of mitochondrial DNA deletions in the brains of schizophrenics. Genomics 29: 217–224.

Chatterjee A, Mambo E, Sidransky D (2006) Mitochondrial DNA mutations in human cancer. Oncogene25(34): 4663-4674.

Chaturvedi HK, Phukan RK, Zoramthanga K (1998) Tobacco use in Mizoram, India: sociodemographic differences in pattern. Southeast Asian. J. Trop. Med. Public Health 29: 66–70.

Chengbo H, Jietao M, Huawei Z (2009) Expression of Mitochondrial Transcripts in Gastric MGC803 Cell Line Subjected by Hypoxia. Clin. Oncol. Cancer Res. 6: 90-94.

Chinnery PF, Elliott HR, Patel S, Lambert C, Keers SM, Durham SE, et al. Walker M: Role of the mitochondrial DNA 16184–16193 poly-C tract in type 2 diabetes. Lancet 366(9497): 1650–1651.

Christofori G, Semb H (1999) The role of the cell-adhesion molecule E-cadherin as a tumor suppressor gene. Trends Biochemistry Sci. 24(2): 73 -76.

Claire M, Payne, Holubec H, Bernstein C, Bernstein H, Dvorak K, et al. (2005) Green Crypt-Restricted Loss and Decreased Protein Expression of Cytochrome c Oxidase Subunit I as Potential Hypothesis-Driven Biomarkers of Colon Cancer Risk. Cancer Epidemiology, Biomarkers & Prevention 14: 2066.

Correa P (2002) Human gastric carcinogenesis: a multistep and multifactorial process-first American cancer society award lecture on cancer epidemiology and prevention. Cancer Res. 52: 6735-6740.

Correa P, Fontham E, Pickle LW, Chen V, Lin YP, Haenszel W (1985) Dietary determinants of gastric cancer in south Louisiana inhabitants. J. Natl. Cancer Inst. 75: 645-654.

Correa P, Fontham ET, Bravo JC, Bravo LE, Ruiz B, Zarama G, et al. (2000) Chemoprevention of gastric dysplasia: randomized trial of antioxidant supplements and anti-*Helicobacter pylori* therapy. J. Natl. Cancer Inst. 92: 1881–1888.

Corso G, Marrelli D, Roviello F (2013) E-cadherin germ line missense mutations in diffuse gastric cancer. OA Cancer 20(1): 1-6.

Cuezva JM, Krajewska M, de Heredia ML, Krajewski S, Santamaria G, Kim H, et al. (2002) The bioenergetic signature of cancer: a marker of tumor progression. Cancer Res. 62: 6674–6681.

Darazy M, Balbaa M,Mugharbil A, Saeed H,Sidani H, Abdel-Razzak Z (2011) CYP1A1, CYP2E1, and GSTM1 gene polymorphisms and susceptibility to colorectal and gastric cancer among Lebanese. Genet. Test Mol. Biomarkers 15(6): 423-429.

DeBruin W, Wagenmans M, Peters W (2000) Expression of glutathione S-transferase alpha, P1-1 and T1-1 in the human gastrointestinal tract. Jpn. J. Cancer Res. 91: 310-316.

Del-Buono R, Pignatelli M (1999) The role of the E-cadherin complex in gastrointestinal cell differentiation. Cell Prolif. 32: 79-84.

Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. 7: 214.

Drummond AJ, Rambaut A, Shapiro B, Pybus OG (2005) Bayesian coalescent inference of past population dynamics from molecular sequences. Mol. Biol. Evol. 22: 1185–1192.

Duell EJ, Sala N, Travier N, Munoz X, Boutron-Ruault MC, Clavel-ChapelonF, et al. (2012) Genetic variation in alcohol dehydrogenase (ADH1A, ADH1B, ADH1C, ADH7) and aldehyde dehydrogenase (ALDH2), alcohol consumption and gastric cancer risk in the European Prospective Investigation into Cancer and Nutrition (EPIC) cohort. Carcinogenesis 33(2): 361-367.

Eaton DL, Bammler TK (1999) Concise review of the glutathione S-transferases and their significance to toxicology. Toxicol. Sci. 49: 156-164.

El-Rifai W, Powell SM (2002) Molecular biology of gastric cancer. Semin. Radiat. Oncol. 12(2): 128–140.

Fearon ER, Vogelstein B (1990) A genetic model for colorectal tumorigenesis. Cell61: 759-767

Fenoglio-Preiser CM, Wang J, Stemmermann GN, Noffsinger A (2003) TP53 and gastric carcinoma: A review. Hum. Mutat. 21(3): 258–270.

Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM (2010) Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. Int. J. Cancer 127: 2893-2917.

Ferlay J, Soerjomataram I, Ervik M, Dikshit R, Eser S, Mathers C, et al. (2014) GLOBOCAN 2012 v1.1, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 11 [Internet]. Lyon, France: International Agency for Research on Cancer. Available from:http://globocan.iarc.fr, accessed on 16/01/2015.

Finkel T, Holbrook NJ (2000) Oxidants, oxidative stress and the biology of aging. Nature 408: 239–247.

Fitzgerald RC, Hardwick R, Huntsman D, Carneiro F, Guilford P, Blair V, et al. (2010) Hereditary diffuse gastric cancer: updated consensus guidelines for clinical management and directions for future research. J. Med. Genet 47: 436-444.

Fliss MS, Usadel H, Caballero OL, Wu L, Buta MR, Eleff SM, et al. (2000) Facile detection of mitochondrial DNA mutations in tumors and bodily fluids. Science 287: 2017–2019.

Fox TD (2012) Mitochondrial Protein Synthesis, Import, and Assembly. Genetics 192(4): 1203–1234.

Fricke E, Keller G, Becker I, Rosivatz E, Schott C, Plaschke S, et al. (2003) Relationship between E-cadherin gene mutation and p53 gene mutation, p53 accumulation, Bcl-2 expression and Ki-67 staining in diffuse-type gastric carcinoma. Int. J. Cancer 104(1): 60–65.

Gabbert HE, Meier S, Gerharz CD, Hommel G (1991) Incidence and prognostic significance of vascular invasion in 529 gastric-cancer patients. Int. J. Cancer 49: 203-207.

García-G MA, Quintero E, Bujanda L, Nicolás D, Benito R, Strunk M,et al. (2012)Relevance of GSTM1, GSTT1, and GSTP1 gene polymorphisms to gastric cancer susceptibility and phenotype. Mutagenesis**27**: 771-777.

Garte S, Gaspari L, Alexandrie AK, Ambrosone C, Autrup H, Autrup JL, et al.(2001)Metabolic gene polymorphism frequencies in control populations. Cancer Epidemiol Biomarkers Prev 10: 1239–1248.

Ghatak S, Doris L, Mawia L, Sapkota R, Zothanpuia, Pautu JL, et al. (2014) Mitochondrial D-Loop and Cytochrome Oxidase Subunit I Polymorphisms among the Breast Cancer Patients of Mizoram, Northeast India. Curr. Genet 60: 201-212.

Ghatak S, Muthukumaran RB, Nachimuthu SK (2013) A simple method of genomic DNA extraction from human samples for PCR-RFLP analysis. J. Biomol. Tech. 24(4): 224–231.

Ghatak S, Yadav RP, Lalrohlui F, Chakraborty P, Ghosh S, Ghosh S, et al. (2016) Xenobiotic Pathway Gene Polymorphisms Associated with Gastric Cancer in High Risk Mizo-Mongoloid Population, Northeast India. Helicobacter DOI: 10.1111/hel.12308.

Ghatak S, Zothansanga, Pautu JL, Senthil-Kumar N(2015) Co-Extraction and PCR based analysis of nucleic acids from Formalin-Fixed Paraffin-Embedded Specimens. J. Clin. Lab. Anal. 29(6): 485-492.

Giardine B, Riemer C, Hardison RC, Burhans R, Elnitski L, Shah P, et al. (2005) Galaxy: a platform for interactive large-scale genome analysis. Genome Res. 15: 1451–1455.

González CA, López-Carrillo L (2010) *Helicobacter pylori*, nutrition and smoking interactions: their impact in gastric carcinogenesis. Scand. J. Gastroenterol. 45: 6-14.

Gonzalez CA, Pera G, Agudo A, Palli D, Krogh V, Vineis P, et al.(2003) Smoking and the risk of gastric cancer in the European Prospective Investigation Into Cancer and Nutrition (EPIC). Int. J. Cancer 107: 629-634.

Gonzalez MAG, Quintero E, Bujanda L, Nicolas D, Benito R, Strunk M, et al. (2012) Relevance of GSTM1, GSTT1, and GSTP1 gene polymorphisms to gastric cancer susceptibility and phenotype. Mutagenesis 27(6): 771–777.

Gopal G, Shirley S, Raja UM,Rajkumar T (2012) Endo-sulfatase Sulf-1 Protein Expression is Down-regulated in Gastric Cancer.Asian Pac. J. Cancer Prev.13(2): 641646.

Grace A, Butler D, Gallagher M, Al-Agha R, Xin Y, Leader M,et al. (2002) APC gene expression in gastric carcinoma: an immunohistochemical study. Appl. Immunohistochem. Mol. Morphol. 10(3): 221-224.

Guilford P, Hopkins J, Harraway J, McLeod M, McLeod N, Harawira N, et al. (1998) E-cadherin germline mutations in familial gastric cancer. Nature 392(6674): 402–405.

Guo W, Zhang Y, Wu B, Zhang L, Zhou D (2000) Study on the p53 gene mutation and microsatellite instability of gastric carcinoma. Zhonghua Yi Xue Yi ChuanXueZaZhi(Chinese J Med. Genet.) 17(2): 101–104.

Lu HZ, Wu YP, LuoW,Seo KW, Yoon KY, Jang KN, et al. (2009) Correlation between aneuploidy of chromosome 17, over-expression ofTP53 and TOPIIalpha, and the clinicopathological features and diagnosis of gastric adenocarcinoma. ZhonghuaZhong Liu ZaZhi(Chinese J Oncol.) 31(10): 754–758.

Habano W, Nakamura S, Sugai T (1998) Microsatellite instability in the mitochondrial DNA of colorectal carcinomas: evidence for mismatch repair systems in mitochondrial genome. Oncogene 17: 1931-1937.

Habano W, Sugai T, Nakamura SI, Uesugi N, Yoshida T, Sasou S (2000) Microsatellite instability and mutation of mitochondrial and nuclear DNA in gastric carcinoma. Gastroenterology 118: 835–841.

Hahn LW, Ritchie MD, Moore JH (2003) Multifactor dimensionality reduc-tion software for detecting gene-gene and gene-environment interac-tions. Bioinformatics 19(3): 376–382.

Han CB, Li F, Zhao YJ, Ma JM, Wu DY, Zhang YK, Xin Y (2003) Variations of mitochondrial D-loop region plus downstream gene 12S rRNA–tRNAPhe and gastric carcinomas. World J. Gastroenterol. 9: 1925–1929.

Hanby AM, Chinery R, Poulsom R, Playford RJ, Pignatelli M (1996) Downregulation of E-cadherin in the reparative epithelium of the human gastrointestinal tract. Am. J. Pathol. 148: 723-729.

Hastings WK (1970) Monte Carlo sampling methods using Markov chains and their applications. Biometrika. 57: 97–109.

Hayes JD, Pulford DJ (1995) The glutathione S-transferase supergene family: regulation of GST and the contribution of the isoenzymes to cancer chemoprotection and drug resistance. Crit. Rev.Biochem. Mol. Biol. 30: 445-600.

Hayes JD, Strange RC (2000) Glutathione S-transferase polymorphisms and their biological consequences. Pharmacology 61: 154-166.

Heerdt BG, Halsey HK, Lipkin M, Augenlicht LH (1990) Expression of mitochondrial cytochrome c oxidase in human colonic cell differentiation, transformation and risk for colonic cancer. Cancer Res. 50: 1596–1600.

Henson DE, Dittus C, Younes M, Nguyen H, Albores SJ (2004) Differential trends in the intestinal and diffuse types of gastric carcinoma in the United States, 1973-2000: increase in the signet ring cell type. Arch. Pathol. Lab. Med. 128: 765–770.

Hermanek P, Sobin LH (1987) International Union against Cancer. TNM Classification of Malignant Tumors, Ed. 4. New York: Springer.

Hibi K, Nakayama H, Yamazaki T, Takase T, Taguchi M, Kasai Y, et al. (2001) Mitochondrial DNA alteration in esophageal cancer. Int. J. Cancer 92: 319–321.

Hill C, Soares P, Mormina M, Macaulay V, Meehan W, Blackburn J, et al. (2006) Phylogeography and Ethnogenesis of Aboriginal Southeast Asians. Mol. Biol. Evol. 23 (12): 2480–2491.

Hiyama T, Tanaka S, Kitadai Y, Ito M, Sumii M, Yoshihara M, et al. (2002) p53 Codon 72 polymorphism in gastric cancer susceptibility in patients with *Helicobacter pylori*-associated chronic gastritis. Int. J. Cancer 100: 304–308.

Ho SA, Hoyle JA, Lewis FA, Secker AD, Cross D, Mapstone NP, et al. (1991) Direct polymerase chain reaction test for detection of *Helicobacter pylori* in humans and animals. J. Clin. Microbiol. 29: 2543-2549.

Hochhauser D (2000) Relevance of mitochondrial DNA in cancer. Lancet356(9225): 181182.

Horton TM, Petros JA, Heddi A, Shoffner J, Kaufman AE, Graham SD, et al. (1996) Novel mitochondrial DNA deletion found in a renal cell carcinoma. Genes. Chrom. Cancer 15: 95–101.

Hu X, Xia H, Srivastava SK, Herzog C, Awasthi YC, Ji X, Zimniak P, Singh SV (1997) Activity of four allelic forms of glutathione S-transferase hGSTP1-1 for diol epoxides of polycyclic aromatic hydrocarbons. Biochem. Biophys. Res. Commun. 238: 397-402.

Humar B, Fukuzawa R, Blair V, Dunbier A, More H, Charlton A, et al. (2007) Destabilized adhesion in the gastric proliferative zone and csrc kinase activation mark the development of early diffuse gastric cancer. Cancer Res. 67(6): 2480–2489.

Hung WY, Lin JC, Lee LM, Wu CW, Tseng LM, Yin PH, et al. (2008) Tandem duplication/triplication correlated with poly-cytosine stretch variation in human mitochondrial DNA D-loop region. Mutagenesis 23(2): 137-142.

Hung WY, Wu CW, Yin PH, Chang CJ, Li AF, Chi CW, et al. (2010) Somatic mutations in mitochondrial genome and their potential roles in the progression of humangastric cancer. Biochim. Biophys. Acta. 1800(3): 264-270.

Iacobuzio-Donahue C, Maitra A, Olsen M, Lowe A, van Heek NT, Rosty C, et al. (2003) Exploration of global expression patterns in pancreatic adenocarcinomas using cDNA microarrays. Am. J. Pathol. 162: 151–162.

Ihsan R, Devi TR, Yadav DS, Mishra AK, Sharma J, Zomawia E, et al. (2011) Investigation on the role of p53 codon 72 polymorphism and interactions with tobacco, betel quid, and alcohol in susceptibility to cancers in a high-risk population from North East India. DNA Cell Biol. 30(3):163-171.

Isidoro A, Martinez M, Fernandez PL, Ortega AD, Santamaria G, Chamorro M, et al. (2004) Alteration of the bioenergetic phenotype of mitochondria is a hallmark of breast, gastric, lung and oesophageal cancer. Biochem. J. 378: 17–20.

Jakupciak JP, Wang W, Markowitz ME, Ally D, Coble M, Srivastava S, et al. (2005) Mitochondrial DNA as a cancer biomarker. J. Mol. Diagn.7(2): 258-267.

Jemt E, Persson O, Shi S, Mehmedovic M, Uhler JP, López MD, Freyer C, et al. (2015) Regulation of DNA replication at the end of the mitochondrial D-loop involves the helicase TWINKLE and a conserved sequence element. Nucleic Acids Res. 43(19): 9262–9275.

Jessie BC, Sun CQ, Irons HR (2001) Accumulation of mitochondrial DNA deletions in the malignant prostate of patients of different ages. Experimental gerontology 37: 169-174.

Jie-Ta M, Cheng-Bo H, Yang Z, Jian-Zhu Z, Wei J, Hua-Wei Z (2012) Altered expression of mitochondrial cytochrome c oxidase I and NADH dehydrogenase 4 transcripts associated with gastric tumorigenesis and tumor dedifferentiation. Molecular medicine reports 5: 1526-1530.

Jing C, Huang ZJ, Duan YQ, Wang PH, Zhang R, Luo KS, Xiao XR (2012) Glulathione-S-transferases gene polymorphism in prediction of gastric cancer risk by smoking and *Helicobacter pylori* infection status. Asian. Pac. J. Cancer Prev. 13(7): 3325-3328.

Joseph A-M, Rungi AA, Robinson BH, Hood DA(2004) Compensatory responses of protein import and transcription factor expression in mitochondrial DNA defects. Am. J. Physiol. Cell Physiol. 286: C867 – 875.

Kagan J and Srivastava S (2005) Mitochondria as a target for early detection and diagnosis of cancer. Crit. Rev. Clin. Lab. Sci. 42(5-6): 453-472.

Kang MG, Kim YN, Lee JH, Szardenings M, Baek HJ, Kook H, et al. (2016) Clinicopathological Implications of Mitochondrial Genome Alterations in Pediatric Acute Myeloid Leukemia. Ann. Lab. Med. 36: 101-110.

Kato S, Tsukamoto T, Mizoshita T, Tanaka H, Kumagai T, et al. (2006) High salt diets dose-dependently promote gastric chemical carcinogenesis in *Helicobacter pylori*-infected Mongolian gerbils associated with a shift in mucin production from glandular to surface mucous cells. Int. J. Cancer 119: 1558-1566.

Katoh T, Nagata N, Kuroda Y, Itoh H, Kawahara A, Kuroki N, et al. (1996) Glutathione S-transferase M1 (GSTM1) and T1 (GSTT1) genetic polymorphism and susceptibility to gastric and colorectal adenocarcinoma. Carcinogenesis 17: 1855-1859.

Kaurah P, MacMillan A, Boyd N, Senz J, De Luca A, Chun N, et al. (2007) Founder and recurrent CDH1 mutations in families with hereditary diffuse gastric cancer. J. Am. Med. Assoc. 297(21): 2360–2372.

Kaurah P, MacMillan A, Boyd N, Senz J, De Luca A, Chun N, et al. (2007) Founder and recurrent CDH1 mutations in families with hereditary diffuse gastric cancer. J. Am. Med. Assoc. 297(21): 2360–2372.

Kaushal M, Ashwani K, Mishraa BS, Rajua (2010) Betel quid chewing as an environmental risk factor for breast cancer. Mutation Research703: 143–148.

Keller G, Vogelsang H, Becker I, Hutter J, Ott K, Candidus S, et al. (1999) Diffuse type gastric and lobular breast carcinoma in a familial gastric cancer patient with an E-cadherin germline mutation. Am. J. Pathol. 155(2): 337–342.

Kelley LA, Sternberg MJ(2009) Protein structure prediction on the web: a case study using the Phyre server. Nat. Protoc. 4: 363-371.

Kiefer F, Arnold K, Ku¨nzli M, Bordoli L, Schwede T (2009) The SWISS-MODEL repository and associated resources. Nucleic Acids Res. 37: 387-392.

Kim HH, Hyung WJ, Cho GS, Kim MC, Han SU, et al. (2010) Morbidity and mortality of laparoscopic gastrectomy versus open gastrectomy for gastric cancer: an interim report--a phase III multicenter, prospective, randomized Trial (KLASS Trial). Ann. Surg. 251: 417-420.

Kim YJ, Lee HS (2009) Genetic epidemiological study on single nucleotide polymorphisms associated with hepatocellular carcinoma in patients with chronic HBV infection. Korean J. Hepatol. 15: 7–14.

Kingma DW, Weiss WB, Jaffe ES, Chadburn A, Knowles DM (1997) Epstein-Barr virus latent membrane protein-1 oncogene deletions: Correlations with malignancy in Epstein-Barr virus-associated lymphoproliferative disorders and malignant lymphomas. Am. J. Pathol. 151: 805-812.

Kirches E, Krause G, Warich-Kirches M, Weis S, Schneider T, Meyer-Puttlitz B, et al. (2001) High frequency of mitochondrial DNA mutations in glioblastomamultiforme identified by direct sequence comparison to blood samples. Int. J. Cancer 93: 534–538.

Kirkland RA, Adibhatla RM, Hatcher JF, Franklin JL (2002) Loss of cardiolipin and mitochondria during programmed neuronal death: evidence of a role for lipid peroxidation and autophagy. Neuroscience 115: 587–602.

Kivisild T, Papiha SS, Rootsi S, Parik J, Kaldma K, Reidla M, et al. (2007) An Indian Ancestry: a Key for Understanding Human Diversity in Europe and Beyond. Archaeogenetics: DNA and the population prehistory of Europe, McDonaldInstitute for Archaeological Research University of Cambridge, UK, pp 261-275.

Kneller RW, Guo WD, Hsing AW, Chen JS, Blot WJ, Li JY, et al. (1992) Risk factors for stomach cancer in sixty-five Chinese counties. Cancer Epidemiol. Biomarkers Prev. 1: 113-118.

Kneller RW, McLaughlin JK, Bjelke E, Schuman LM, Blot WJ, Wacholder S, et al. (1991) A cohort study of stomach cancer in a high-risk American population. Cancer 68: 672-678.

Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D (2009) Circos: an information aesthetic for comparative genomics. Genome Res. 19: 1639–1645.

Kubicka S, Claas C, Staab S, Kuhnel F, Zender L, Trautwein C, et al. (2002) p53 mutation pattern and expression of c-erbB2 and c-met in gastric cancer: Relation to histological subtypes, *Helicobacter pylori* infection, and prognosis. Dig. Dis. Sci. 47(1): 114–121.

Ladeiras-Lopes R, Pereira AK, Nogueira A, Pinheiro-Torres T, Pinto I, Santos-Pereira R, Lunet N (2008) Smoking and gastric cancer: systematic review and meta-analysis of cohort studies. Cancer Cause Control 19: 689-701.

Lalmuanpuii R, Ghatak S, Pautu JL, Lallawmzuali D, Muthukumaran RB, et al. (2015) Mutation Profiling in Mitochondrial D-Loop Associated with Stomach Cancer and Tobacco Consumers. J. Clin. Med. Genom. 3:122.

Lan Q, Chow WH, Lissowska J, Hein DW, Buetow K, Engel LS, et al. (2001) Glutathione S-transferase genotypes and stomach cancer in a population-based case-control study in Warsaw, Poland. Pharmacogenetics11: 655-661.

Lao X, Peng Q, Lu Y, Li S, Qin X, Zhiping Chen Z, Chen J (2014) Glutathione S-transferase gene GSTM1, gene-gene interaction, and gastric cancer susceptibility: evidence from an updated meta-analysis. Cancer Cell Int. 14: 127.

Larman TC, DePalma SR, Hadjipanayis AG, Protopopov A, Zhang J, Gabriel SB, et al. (2012) Cancer Genome Atlas Research Network Spectrum of somatic mitochondrial mutations in five cancers. Proc. Natl. Acad. Sci. USA 109(35): 14087–14091.

La-Torre G, Chiaradia G, Gianfagna F, De-Lauretis A, Boccia S, Mannocci A, Ricciardi W (2009) Smoking status and gastric cancer risk: an updated meta-analysis of case-control studies published in the past ten years. Tumori 95: 13-22.

Lauren P (1965) The two histologicul main types of gastric carcinoma. Diffuse and so-called intestinal type carcinoma. An attempt al histoclinical classification. Acta. Pathol. Microbio. Scand. 64: 31-49.

Lee HC, Li SH, Lin JC, Wu CC, Yeh DC, Wei YH (2004) Somatic mutations in the D-loop and decrease in the copy number of mitochondrial DNA in human hepatocellular carcinoma. Mutat. Res. 547: 71–87.

Lee HC, Yin PH, Lin JC, Wu CC, Chen CY, Wu CW, et al. (2005) Mitochondrial genome instability and mtDNA depletion in human cancers. Ann. NucleicAcad. Sci. 1042: 109–122.

Lee J-M, Lee Y-C, Yang S-Y, Shi W-L, Lee C-J, Luh S-P, et al. (2000) Genetic polymorphisms of p53 and GSTP1, but not NAT2, are associated with susceptibility to squamous-cell carcinoma of the esophagus. Int. J. Cancer 89: 458–464.

Lewis PD, Mehrotra D, Mahdi AA, Sarin R, Kowtal P, et al. (2002) Mitochondrial DNA mutations in the parotid gland of cigarette smokers and non-smokers. Mutat. Res. 18: 47–54.

Lievre A, Chapusot C, Bouvier AM, Zinzindohoue F, Piard F, Roignot P, et al. (2005) Clinical value of mitochondrial mutations in colorectal cancer. J. Clin. Oncol. 23: 3517–3525.

Lim BH, Soong R, Grieu F, Robbins PD, House AK, Iacopetta BJ (1996) p53 accumulation and mutation are prognostic indicators of poor survival in human gastric carcinoma. Int. J. Cancer 69(3): 200–204.

Ling XL, Fang DC, Wang RQ, Yang SM, Fang L(2004) Mitochondrial microsatellite instability in gastric cancer and its precancerous lesions. World J. Gastroenterol. 10: 800-803.

Liu VW, Shi HH, Cheung AN, Chiu PM, Leung TW, Nagley P, et al. (2001) High incidence of somatic mitochondrial DNA mutations in human ovarian carcinomas. Cancer Res. 61: 5998–6001.

Loh M, Koh KX, Yeo BH, Song CM, Chia KS, Zhu F, et al. (2009) Meta-analysis of genetic polymorphisms and gastric cancer risk: variability in associations according to race. Eur. J. Cancer. 45(14): 2562–2568.

Lovat PE, Ranalli M, Corazzari M, RaffaghelloL, Pearson AD, Ponzoni M, et al. (2003) Mechanisms of free-radical induction in relation to fenretinide-induced apoptosis of neuroblastoma. J. Cell Biochem. 89: 698–708.

Lutz W, Nowakowska-Swirta E (2002) Gene p53 mutations, protein p53, and anti-p53 antibodies as biomarkers of cancer process. Int. J. Occup. Med. Environ. Health 15(3): 209–218.

Macdonald JS, Smalley SR, Benedetti J, Hundahl SA, Estes NC, Stemmermann GN, et al. (2001) Chemoradiotherapy after surgery compared with surgery alone for adenocarcinoma of the stomach or gastroesophageal junction. N. Engl. J. Med. 345: 725–730.

Machado AM, Figueiredo C, Touati E, Máximo V, Sousa S, Michel V, et al. (2009) *Helicobacter pylori* infection induces genetic instability of nuclear and mitochondrial DNA in gastric cells. Clin. Cancer Res. 15: 2995–3002.

Mahanta J, Chetia M, Hazarika NC, Narain K, Sharma SK (1998) Toxicity of tuibur, a unique form of tobacco smoke extract used in Mizoram India. Curr. Sci. 75: 381-384.

Malakar M, Devi KR, Phukan RK, Kaur T, Deka M, Puia L, et al. (2012) Genetic polymorphism of glutathione S-transferases M1 and T1, tobacco habits and risk of stomach cancer in Mizoram, India. Asian Pac. J. Cancer Prev. 13(9): 4725-4732.

Malik MA, Upadhyay R, Mittal RD, Showket AZ, Dinesh RM, Balraj M (2009) Role of xenobiotic-metabolizing enzyme gene polymorphisms and interactions with environmental factors in susceptibility to gastric cancer in Kashmir Valley. J. Gastrointest Cancer 40: 26-32.

Manuguerra M, Matullo G, Veglia F, Autrup H, Dunning AM, Garte S,et al. (2007) Multi-factor dimensionality reduction applied to a large prospective investigation on gene-gene and gene-environment interactions. Carcinogenesis 28(2): 414–422.

Mathew A, Gangadharan P, Varghese C, Nair MK (2000) Diet and stomach cancer: A case-control study in South India. Eur. J. Cancer Prev. 9: 89–97.

Matlashewski G, Tuck S, Pim D, Lamb P, Schneider J, Crawford L (1987) Primary structure polymorphism at amino acid residue 72 of human p53. Mol. Cell Biol. 7: 961–963.

Matturri L, Biondo B, Cazzullo A, Colombo B, Giordano F, Guarino M, et al. (1998) Prognostic significance of different biological markers (DNA index, PCNA index, apoptosis, p53, karyotype) in 126 adenocarcinoma gastric biopsies. Anticancer Res. 18(4B): 2819–2825.

Maximo V, Soares P, Machado JC, Seruca R, Sobrinho-Simoes M (2000) Mitochondrial DNA alteration in gastric cancer. Gastroenterology 119: 1808–1809.

Maximo V, Soares P, Seruca R, Rocha AS, Castro P, Sobrinho-Simoes M (2001) Microsatellite instability, mitochondrial DNA large deletions, and mitochondrial DNA mutations in gastric carcinoma. Genes Chromosomes Cancer 32: 136–143.

Mir MR,Shabir N,Wani KA,Shaff S, Hussain I,Banday MA, et al. (2012) Association between p16, hMLH1 and E-cadherin promoter hypermethylation and intake of local hot salted tea and sun-dried foods in Kashmiris with gastric tumors. Asian Pac. J. Cancer Prev. 13(1): 181-186.

Mocellin S, Verdi D, Pooley KA, Nitti D (2015) Genetic variation and gastric cancer risk: a field synopsis and meta-analysis. Gut 64: 1209-1219.

Moriyama N, Ishihara S, Hirose M, Watanabe S, Sato N, Kinoshita Y (2001) E-cadherin is essential for gastric epithelial restitution in vitro: A study using the normal rat gastric mucosal cell line RGM1. J. Lab. Clin. Med. 138(4): 236-242.

Murakami K, Fujioka T, Okimoto T, Mitsuishi Y, Oda T, Nishizono A, Nasu M (1999) Analysis of p53 gene mutations in *Helicobacter pylori*-associated gastritis mucosa in endoscopic biopsy specimens. Scand. J. Gastroenterol. 34: 474–477.

Murray CJ, Lopez AD (1997) Mortality by cause for eight regions of the world: Global Burden of Disease Study. Lancet 349:1269–1276.

Muta H, Noguchi M, Kanai Y, Ochiai A, Nawata H, Hirohashi S (1996) E-cadherin gene mutations in signet ring cell carcinoma of the stomach. Jpn. J. Cancer Res. 87: 843.

National Cancer Registry Programme (NCRP), (2010) Three year report of the population based cancer registries 2006-2008. (First report of 20 PBCRs in India), Bangalore, Indian Council Med. Res. 11: 71-78.

National Cancer Registry Programme (NCRP-ICMR), (2007) The North East population based cancer registries 2005-2006. Bangalore, 2007.

Nekhaeva E, Bodyak ND, Kraytsberg Y, McGrath SB, Van-Orsouw NJ, Pluzhnikov A,et al. (2002) Clonally expanded mtDNA point mutations are abundant in individual cells of human tissues. Proc. Natl. Acad. Sci. USA 99: 5521 – 5526.

Nguyen T, Brunson D, Crespi CL, Penman BW, Wishnok JS, Tannebaum SR (1992) DNA damage and mutation in human cells exposed to nitric oxide in vitro. Proc. Natl. Acad. Sci. USA 11: 3030–3034.

Noach LA, Rolf TM, Tygnat GN (1994) Electron microscopic study of association between *Helicobacter pylori* and gastric and duodenal mucosa. J. Clin. Pathol. 47: 699-704.

Nomoto S, Yamashita K, Koshikawa K, Nakao A, Sidransky D (2002) Mitochondrial D-loop mutations as clonal markers in multicentric hepatocellular carcinoma and plasma. Clin. Cancer Res. 8: 481–487.

Nomura A, Grove JS, Stemmermann GN, Severson RK (1990) A prospective study of stomach cancer and its relation to diet, cigarettes, and alcohol consumption. Cancer Res. 50: 627-631.

Nomura A, Stemmerman G N, Chyou P H, Kato I, Perez-Perez GI, Blaser MJ (1991) *Helicobacter pylori*infection and gastric carcinoma among Japanese Americans in Hawaii. N. Eng. J. Med. 325: 1132–1136.

Nose A, Tsuji K, Takeichi M (1990) Localization of specificity determining sites in cadherin cell adhesion molecules. Cell 61: 147.

Ohmiya N, Taguchi A, Mabuchi N, Itoh A, Hirooka Y, Niwa Y, Goto H (2006) MDM2 promoter polymorphism is associated with both an increased susceptibility to gastric carcinoma and poor prognosis. J. Clin. Oncol. 24: 4434–4440.

Oliveira C, Bordin MC, Grehan N, Huntsman D, Suriano G, Machado JC, et al. (2002) Screening E-cadherin in gastric cancer families reveals germline mutations only in hereditary diffuse gastric cancer kindred. Hum. Mutat. 19(5): 510–517.

Oota K, Sobin LH (1977) Histological Typing of Gastric and OesophagealTumours. International HistolÃ³gica! Classification of Tumours, No 18. Geneva, Switzerland: World Health Organization.

Ott K,Rachakonda PS,Panzram B, Keller G,Lordick F, Becker K, et al. (2011) DNA repair gene and MTHFR gene polymorphisms as prognostic markers in locally advanced adenocarcinoma of the esophagus or stomach treated with cisplatin and 5-fluorouracil-based neoadjuvant chemotherapy. Ann. Surg. Oncol. 18(9): 2688-2698.

Ozawa M, Kemler R (1990) Correctproteolytic cleavage is required for the cell adhesive function of uvomorulin. J. Cell Biol. 111: 1645.

Ozawa T (1995) Mechanism of somatic mitochondrial DNA mutations associated with age and diseases. Biochim. Biophys. Acta. 1271: 177–189.

Palli D, Caporaso NE, Shiao YH, Saieva C, Amorosi A, Masala G, et al. (1997) Diet, *Helicobacter pylori*, and p53 mutations in gastric cancer: A molecular epidemiology study in Italy. Cancer Epidemiol. Biomarkers Prev. 6(12): 1065–1069.

Palli D, Masala G, Del-Giudice G, Plebani M, Basso D, Berti D, et al (2007) CagA+ *Helicobacter pylori* infection and gastric cancer risk in the EPIC-EURGAST study. Int. J. Cancer 120: 859-867.

Palli D, Saieva C, Gemma S, Masala G, Gomez-Miguel MJ, Luzzi I, et al. (2005) GSTT1 and GSTM1 gene polymorphisms and gastric cancer in a high-risk italian population. Int. J. Cancer 115: 284-289.

Park JS, Sharma LK, Li H, Xiang R, Holstein D, Wu J, et al. (2009) Aheteroplasmic, not homoplasmic, mitochondrial DNA mutation promotes tumorigenesis via alteration in reactive oxygen species generation and apoptosis. Hum. Mol. Genet 18(9): 1578–1589.

Parkin DM, Ferlay J, Hamdi-Cherif M, Sitas F, Thomas JO, Wabinga H, Whelan SL (2003) Cancer in Africa—Epidemiology and Prevention. IARC Scientific Publications Lyons 153.

Parkin DM, Muir CS, Whelan SL, Ferlay J, Teppo L, Thomas DB (2002) Cancer incidence in five continentsVol VIII. IARC, Sci.Pub 7: 1555.

Parsonnet J, Friedman GD, Vandersteen DP, Chang Y, Vogelman JH, Orentreich N, Sibley RK (1991) *Helicobacter pylori* infection and the risk of gastric carcinoma. N. Engl. J. Med. 325: 1127-1131.

Pavithran K, Doval DC, Pandey KK (2002) Gastric cancer in India. Gastric Cancer 5: 240–243.

Pemble S, Schroeder KR, Spencer SR, Meyer DJ, Hallier E, Bolt HM, et al. (1994) Human glutathione S-transferase theta (GSTT1): cDNA cloning and the characterization of a genetic polymorphism. Biochem. J. 300: 271-276.

Penta JS, Johnson FM, Wachsman JT (2001) Mitochondrial DNA in human malignancy. Mutat. Res. 488: 119-133.

Pham XH, Farge G, Shi Y, Gaspari M, Gustafsson CM, Falkenberg M (2006) Conserved sequence box II directs transcription termination and primer formation in mitochondria. J. Biol. Chem. 281: 24647-24652.

Phukan RK, Zomawia E, Narain K, Hazarika NC, Mahanta J (2005) Tobacco use and stomach cancer in Mizoram, India. Cancer Epidemiol. Biomarkers Prev. 14: 1892– 1896.

Phukan RK, Narain K, Zomawia E, Hazarika NC, Mahanta J (2006) Dietary habits and stomach cancer in Mizoram, India. J. Gastroenterol. 41(5):418-424.

Piao JM, Shin MH, Kweon SS, Kim HN, Choi JS, Bae WK, et al (2009) Glutathione-S-transferase (GSTM1, GSTT1) and the risk of gastrointestinal cancer in a Korean population. World J. Gastroenterol 15: 5716–5721.

Pinkham JL, Dudley AM, Mason TL (1994) T7 RNA polymerasedependent expression of COXII in yeast mitochondria. Mol. Cell. Biol. 14: 4643–4652.

Polyak K, Li Y, Zhu H, Lengauer C, Willson JK, Markowitz SD, et al. (1998) Somatic mutations of the mitochondrial genome in human colorectal tumours. Nat. Genet 20: 291–293.

Rastogi T, Devesa S, Mangtani P, Mathew A, Cooper N, Kao R (2008) Cancer incidence rates among South Asians in four geographic regions: India, Singapore, UK and US. Int. J. Epidemiol. 37: 147–160.

Richard SM, Bailliet G, Paez GL, Bianchi MS, Peltomaki P, Bianchi NO (2000) Nuclear and mitochondrial genome instability in human breast cancer. Cancer Res. 60: 4231–4237.

Robinson JT, Thorvaldsdottir H, Winckler R, Guttman M, Lander ES, Getz-Mesirov JP (2011) Integrative genomics viewer. Nat. Biotechnol. 29: 24–26.

Rodriguez-Hernandez A, Cordero MD, Salviati L, Artuch R, Pineda M, Briones P, et al. (2009) Coenzyme Q deficiency triggers mitochondria degradation by mitophagy. Autophagy 5: 19–32.

Roth MJ, Abnet CC, Johnson LL, Mark SD, Dong ZW, Taylor PR, et al. (2004) Polymorphic variation of CYP1A1 is associated with the risk of gastric cardia cancer: a prospective

case–cohort study of cytochrome P-450 1A1 and GST enzymes. Cancer Cause Control 15: 1077-1083.

Ruiz-Pesini E, Lott MT, Procaccio V, Poole JC, Brandon MC, Mishmar D, Yi C, et al. (2007) An enhanced MITOMAP with a global mtDNA mutational phylogeny. Nucleic Acids Res. 35: 823–828.

Russo A, Maconi G, Spinelli P, Felice GD, Eboli M, Andreola S, Ravagnani F, Settesoldi D, Ferrari D, Lombardo C, Bertario L (2001) Effect of lifestyle, smoking, and diet on development of intestinal metaplasia in *H. pylori*-positive subjects. Am. J. Gastroenterol 96: 402-408.

Saadat M (2006) Genetic polymorphisms of glutathione S-transferase T1 (GSTT1) and susceptibility to gastric cancer: a meta-analysis. Cancer Sci. 97: 505-509.

Sakurai S, Sano T, Nakajima T (1995) Clinico-pathological and molecular biological studies of gastric adenomas with special reference to p53 abnormality. Pathol. Int. 45(1): 51–57.

Sanchez-Cespedes M, Parrella P, Nomoto S, Cohen D, Xiao Y, Esteller M, et al. (2001) Identification of a mononucleotide repeat as a major target for mitochondrial DNA alterations in human tumors. Cancer Res. 61: 7015–7019.

Santos AM, Sousa H, Pinto D, Portela C, Pereira D, Catarino R, et al. (2006) Linking TP53 codon 72 and P21 nt590 genotypes to the development of cervical and ovarian cancer. Eur. J. Cancer 42 (7): 958–963.

Satoh K, Kihira K, Kawata H, Tokumaru K, Kumakura Y, Ishino Y, et al. (2001) p53 expression in the gastric mucosa before and after eradication of *Helicobacter pylori*. Helicobacter 6: 31–36.

Satyanarayana L, Asthana S (2008) Life time risk for development of ten major cancers in India and its trends over the years 1982 to 2000. Indian J. Med. Sci.62: 35–44.

Schipper JH, Frixen UH. Behrens J, Unger A, Jahnke K, Birchmeier W (1991) E-cadherin expression in squamous cell carcinomas of head and neck: inverse correlation with tumor deditferenlialion and lymph node metastasis. Cancer Res. 51: 6328-6337.

Shadel GS, Clayton DA (1997) Mitochondrial DNA maintenance in vertebrates. Annu. Rev. Biochem. 66: 409–435.

Shay JW, Werbin H (1987) Are mitochondrial DNA mutations involved in the carcinogenic process? Mutation Res. 186: 149-160.

Shiao YH, Palli D, Caporaso NE, Alvord WG, Amorosi A, Nesi G, et al. (2000) Genetic and immunohistochemical analyses of p53 independently predict regional metastasis of gastric cancers. Cancer Epidemiol. Biomarkers Prev. 9(6): 631–633.

Shimoyama T, Fukuda S, Liu Q, Nakaji S, Fukuda Y, Sugawara K (2002) Production of chemokines and reactive oxygen species by human neutrophils stimulated by *Helicobacter pylori*. Helicobacter 7: 170–174.

Shimoyama Y, Hirohashi S (1991) Cadherin intercellular adhesion molecule in hepatocellular carcinomas: loss of E-cadherin expression in an undifferentiated carcinoma. Cancer Lett. 57: 131-135.

Shiozaki H, Tuhala H, Oku H. Miyata M. Kobuyashi K, Tamura S, et al. (1991) Expression of immunoreactive E-cadherin adhesion molecules in human cancers. Am. J. Palhol. 139: 17-23.

Shoubridge EA, Karpati G, Hastings KEM (1990) Deletion mutants are functionally dominant over wild-type mitochondrial genomes in skeletal muscle fiber segments in mitochondrial disease. Cell 62: 43 – 49.

Singer TP (1994) Determination of the activity of succinate, NADH, choline and aglycerophosphate dehydrogenase. Methods Biochem. Anal. 22: 123–175.

Singha K (2012) Identity, contestation and development in north east India: a study of Manipur, Mizoram and Nagaland. JCPP 3: 403-424.

Smith PR, Cooper JM, Govan GG, Harding AE, Schapira AH (1993) Smoking and mitochondrial function: a model for environmental toxins. Q. J. Med. 86: 657-660.

Soares P, Ermini L, Thomson N, Mormina M, Rito T, Röhl A, et al. (2009) Correcting for Purifying Selection: An Improved Human Mitochondrial Molecular Clock. Am. J. Hum. Genet 84(6): 740–759.

Solcia E, Fiocca R, Luinetti O, Villani L, Padovan L, Calistri D, et al. (1996) Intestinal and diffuse gastric cancers arise in a different background of *Helicobacter pylori* gastritis through different gene involvement. Am. J. Surg. Pathol. 20 Suppl 1: S8-22.

Soya SS, Vinod T, Reddy KS, Gopalakrishnan S, Adithan C (2007) Genetic polymorphisms of glutathione-S-transferase genes (GSTM1, GSTT1 and GSTP1) and upper aerodigestive tract cancer risk among smokers, tobacco chewers and alcoholics in a an Indian population. Eur. J. Cancer 43: 2698-2706.

Stappert J, Kemler R (1994) A short core region of E-cadherin is essential for catenin binding and is highly phosphorylated. Cell Adhes. Commun. 2: 319.

Sugimura T, Kawaki T (1973) Experimental stomach cancer. Methods Cancer Res. 7: 245–308.

Suliman HB, Welty-Wolf KE,Carraway M,Tatro L, Piantadosi CA (2004) Lipopolysaccharide induces oxidative cardiac mitochondrial damage and biogenesis. Cardiovasc. Res. 64: 279 – 288.

Sullivan BL, Chandel SN (2014) Mitochondrial reactive oxygen species and cancer. Cancer Metab. 2: 17.

Sun AS, Cederbaum AI (1980) Oxidoreductase activities in normal rat liver, tumor-bearing rat liver and hepatoma HC-252. Cancer Res. 40: 4677–4681.

Sun AS, Sepkowitz K, Geller SA (1981) A study of some mitochondrial and peroxisomal enzymes in human colonic adenocarcinoma. Lab. Invest. 44: 13–17.

Szczepanowska K, Trifunovic A (1847) Different faces of mitochondrial DNA mutators. Biochim. Biophys. Acta. 2015: 1362–1372.

Tahara T, Shibata T, Nakamura M, Okubo M, Yamashita H, Yoshioka D, et al. (2011) Polymorphisms of DNA Repair and Xenobiotic Genes Predispose to CpG Island Methylation in Non-Neoplastic Gastric Mucosa. Helicobacter 16: 99-106.

Takahashi N, Joh T, Yokoyama Y, Seno K, Nomura T, Ohara H, et al. (2000) Importance of gap junction in gastric mucosal restitution from acid-induced injury. J. Lab. Clin. Med. 136: 93-99.

Takeichi M (1993) Cadherins in cancer: im- plications for invasion and metasta- sis. Curr. Opin. Cell Biol.5(5): 806–811.

Tamura G (2004) Promoter methylation status of tumor suppressor and tumor-related genes in neoplastic and non-neoplastic gastric epithelia. Histol. Histopathol. 19: 221-228.

Tanaka M, Cabrera VM, González AM, Larruga JM, Takeyasu T, Fuku N, et al. (2004) Mitochondrial genome variation in eastern Asia and the peopling of Japan. Genome Res. 14: 1832–1850.

Taylor RW, Barron MJ, Borthwick GM, Gospel A, Chinnery PF, Samuels DC, et al. (2003) Mitochondrial DNA mutations in human colonic crypt stem cells. J. Clin. Invest. 112: 1351–1360.

Thiery JP (1996) The saga of adhesion molecules. J. Cell Biol. 61: 489–492.

Tolbert DC, Fenoglio-Preiser A, Noffsingert G, De-Voe J, MacDonald (1999) The relation of p53 gene mutations to gastric cancer subsite and phenotype. Cancer Causes and Control 10: 227-231.

Tolk HV, Barac L, Pericic M, Klaric IM, Janicijevic B, Campbell H, et al. (2001) The evidence of mtDNA haplogroup F in a European population and its ethnohistoric implications. Eur. J. Human. Genet 9: 717–723.

Torroni A, Miller JA, Moore LG, Zamudio S, Zhuang J, Droma T, et at. (1994) Mitochondrial DNA analysis in Tibet: Implications for the origin of the Tibetan population and its adaptation to high altitude. Am. J. Phys. Anthropol. 93: 189–199.

Tripathi S, Ghoshal U, Mittal B, Chourasia D, Kumar S,Ghoshal UC (2011) Association between gastric mucosal glutathione-S-transferase activity, glutathione-Stransferase gene polymorphisms and *Helicobacter pylori* infection in gastric cancer. 30(6):257-263.

Trounce IA, Kim YL, Jun AS, Wallace DC (1996) Assessment of mitochondrial oxidative phosphorylation in patient muscle biopsies, lymphoblasts and transmitochondrial cell lines. Methods Enzymol. 264: 484–509.

Tsugane S, Sasazuki S (2007) Diet and the risk of gastric cancer: review of epidemiological evidence. Gastric Cancer 10: 75-83.

Tsukihara T, Aoyama H, Yamashita E, Tomizaki T, Yamaguchi H, Shinzawa-Itoh K, (1996) The whole structure of the 13-subunit oxidized cytochrome *c* oxidase at 2.8 A. Science 272: 1136–1144

Uemura N, Okamoto S, Yamamoto S, Matsumura N, Yamaguchi S, Yamakido M, et al. (2001) *Helicobacter pylori* infection and the development of gastric cancer. N. Engl. J. Med. 345: 784-789.

VanOven M, Kayser M (2009) Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. Hum. Mutat. 30(2): E386-E394.

Verma M, Kumar D (2007) Application of mitochondrial genome information in cancer epidemiology. Clin. Chim. Acta. 383: 41–50.

Vleminckx K, Vakaet L, Marcel M, Fiers W, Van RF (1991) Genetic manipulation of E-cadherin expression by epithelial tumor cells reveals an invasive suppressor role. Cell 66: 107-119.

Vumson, Thawagn NT, Venghlui (1986) Zo History. Vumson Publisher, Aizwal, Mizoram pp: 26–39.

Wade M, Li YC, Wahl GM (2013) MDM2, MDMX and p53 in oncogenesis and cancer therapy. Nat. Rev. Cancer 13(2): 83–96.

Wallace DC, Brown MD, Lott MT (1999) Mitochondrial DNA variation in human evolution and disease. Gene 238(1): 211-230.

Ward MH, López-Carrillo L (1999) Dietary factors and the risk of gastric cancer in Mexico City. Am. J. Epidemiol. 149: 925-932.

Wink DA, Kasprzak KS, Maragos CM, Elespuru RK, Misra M, Dunams TM, et al. (1992) DNA deaminating ability and genotoxicity of nitric oxide and its progenitors. Science 11: 1001–1003.

Xu Y, Yao L, Ouyang T, Li J, Wang T, Fan Z, et al. (2005) p53 Codon 72 polymorphism predicts the pathologic response to neoadjuvant chemotherapy in patients with breast cancer. Clin. Cancer Res. 11: 7328–7333.

Ye C, Shu XO, Pierce L (2010) Mutations in the mitochondrial DNA D-loop region and breast cancer risk. Breast Cancer Res. Treat. 119: 431–436.

Ye Y, Li X, Yang J, Miao S, Wang S, Chen Y, etal. (2013) MDM2 is a useful prognostic biomarker for resectable gastric cancer. Cancer Sci.104(5): 590–598.

Yeh JJ, Lunetta KL, van Orsouw NJ, Moore FD, Jr, Mutter GL, Vijg J, et al. (2000) Somatic mitochondrial DNA (mtDNA) mutations in papillary thyroid carcinomas and differential mtDNA sequence variants in cases with thyroid tumours. Oncogene 19:2060–2066.

Yeole BB (2008) Trends in Cancer incidence in Oesophagus, Stomach, Colon, Rectum and Liver Cancer in Males in India. Asian Pacific J. Cancer Prev. 9: 97-100.

Yin PH, Lee HC, Chau GY, Wu YT, Li SH, Lui WY, et al. (2004) Alteration of the copy number and deletion of mitochondrial DNA in human hepatocellular carcinoma. Br. J. Cancer 90: 2390–2396.

Yoneda M, Chomyn A, Martinuzzi A, Hurko O, Attardi G (1992) Marked replicative advantage of human mtDNA carrying a point mutation that causes the MELAS encephalomyopathy. Proc. Natl. Acad. Sci. USA 89: 11164 –11168.

Yu JY, Li L, Ma H, Liu K, Cheng X, Li YL, Song XL (2013) Tumor necrosis factor-α 238 G/A polymorphism and gastric cancer risk: a meta-analysis. Tumor Biology 10: 1007-1012.

Zendehdel K, Bahmanyar S, McCarthy S, Nyren O, Andersson B, Ye W (2009) Genetic polymorphisms of glutathione S-transferase genes GSTP1, GSTM1, and GSTT1 and risk of esophageal and gastric cardia cancers. Cancer Cause Control 20: 2031-2038.

Zhang Y, Sun LP, Xing CZ, Xu Q, He CY, Li P, et al. (2012) Interaction between GSTP1 Val allele and *H. pylori* infection, smoking and alcohol consumption and risk of gastric cancer among the Chinese population. PLoS One 7(10): e47178.

Zhang Z, Fu G, Wang M, Tong N, Wang S, et al. (2008) P53 codon 72 polymorphism and ovarian cancer risk: a meta-analysis. JNMU 22 (5): 279–285.

Zhao Y, Deng X, Song G, Qin S, Liu Z (2013) The GSTM1 null genotype increased risk of gastric cancer: a meta-analysis based on 46 studies. PLoS One 8(11): e81403.

Zhao Y, Yang H, Zhang X, Chen G (2005) Mutation in D-loop region of mitochondrial DNA in gastric cancer and its significance. World J. Gastroenterol 11(21): 3304-3306.

Zhou S, Kachhap S, Singh KK (2003) Mitochondrial impairment in p53-deficient human cancer cells. Mutagenesis 18(3): 287–292.

Zhou S, Kachhap S, Sun W, Wu G, Chuang A, Poeta L, et al. (2007) Frequency and phenotypic implications of mitochondrial DNA mutations in human squamous cell carcinoma of the head and neck. PNAS. 104: 7540–7545.

Zimniak P, Nanduri B, Pikuła S, Bandorowicz-Pikuła J, Singhal SS, Srivastava SK, et al. (1994) Naturally occurring human glutathione S-transferase GSTP1-1 isoforms with isoleucine and valine in position 104 differ in enzymic properties. Eur. J. Biochem. 224: 893-899.

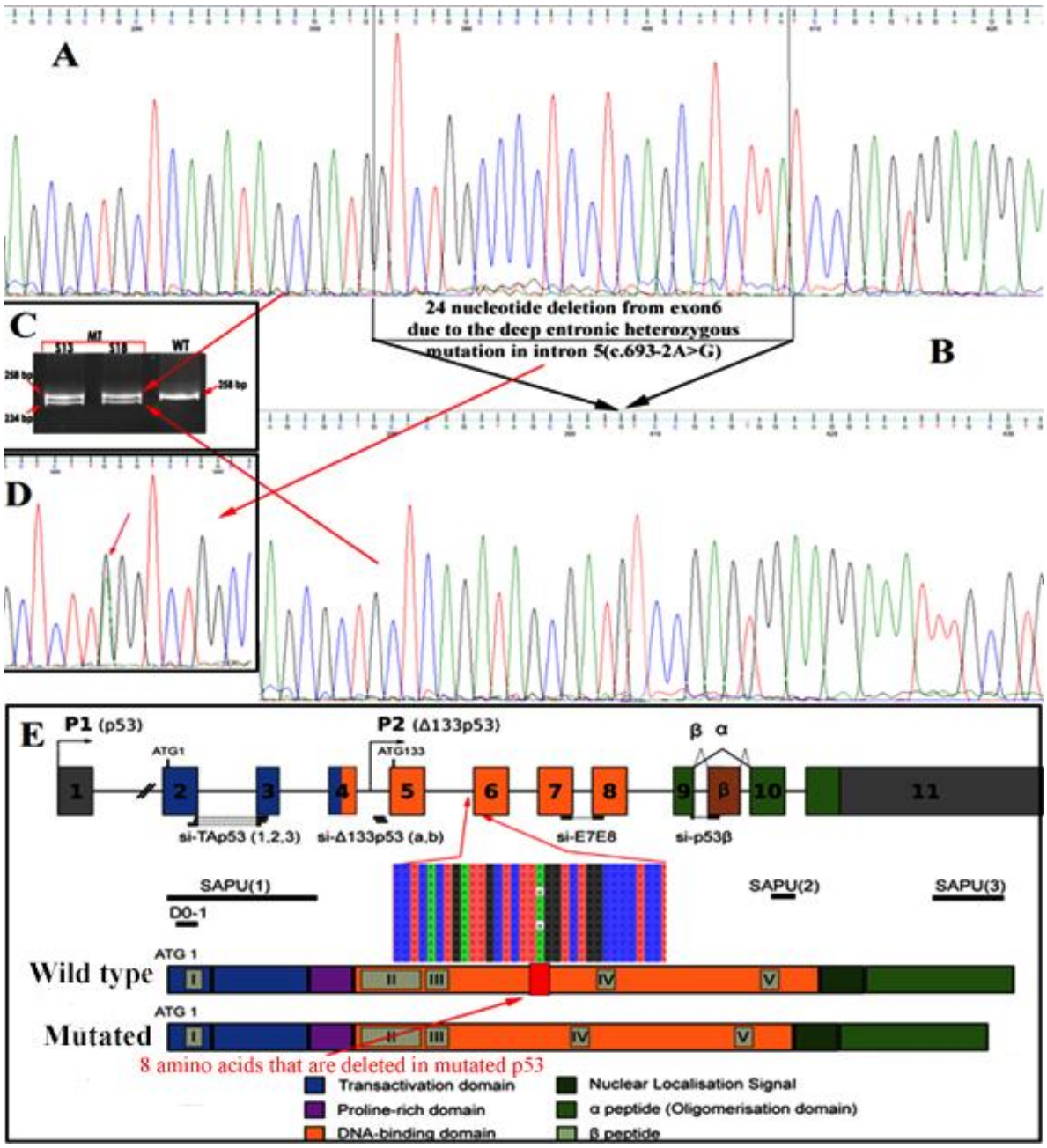


Figure 5.1. The c.693-2A>G mutation in the mutant induces a fully penetrant splicing defect. (A) cDNA sequence of exon 5 and 6; (B) cDNA sequence of the exon 5 and 6 after 24 nucleotide deletion; (C) amplified product for tumor and normal samples (WT- Wild type having normal 258 bp allele, MT- Mutant type having heterozygyous 234 bp allele); (D) genomic DNA sequence and the potential splice site deep intronic mutation (c.693-2A>G) (E) Overview about the localization of the mutation and its consequence upon transcript and protein domain processing.

position and was classified to be disease-causing by different prediction tools (MutationTaster, Human splicing finder). The splice-site variant lies within the consensus splice acceptor site or deep intronic position and therefore is predicted to affect the normal splicing of exon 6 (Figure 5.1).

3.4.Immunohistochemical analysis of TP53 and MDM2 genes

In 80 unrelated primary gastric carcinomas, we compared common prognostic markers such as age, sex, tumor location, Lauren type, tumor grading, TNM staging,*H. pylori* status and expression of p53 and MDM2 proteins.A total of 68 tumors stained positive for p53 (Figures 5.3 and 5.5) and 12 tumors with a expression of less than 20% were treated as negative expression.There was no obvious relation between p53 staining and the age or sex of the patients. No significant correlations were found between p53 tissue status and Lauren type, tumor grading, and depth of invasion (T status). In contrast, there was a significant association between p53 tissue status and the metastatic spread to lymph nodes. TP53-positive tumors were associated with a higher incidence of metastasis to lymph nodes (91.4%) than were p53-negative tumors (71.4% ; P= 0.021) . p53 protein overexpression was observed in 27.5% (22/80; 95% CI, 4.7%–7.4%) of gastric tumor. Forty-six tumor samples exhibited moderate or low expression, whereas 12 tumor samples exhibited negative expression of p53 proteins. But all the adjacent normal cells (of negatively expressing p53 tumor samples) exhibited moderate level of p53 protein expression. High expression of MDM2 proteins was observed in ten tumor samples which had negative p53 expression. But most of the tumor samples which had high positive p53 expression, exhibited moderate level of MDM2 proteins.

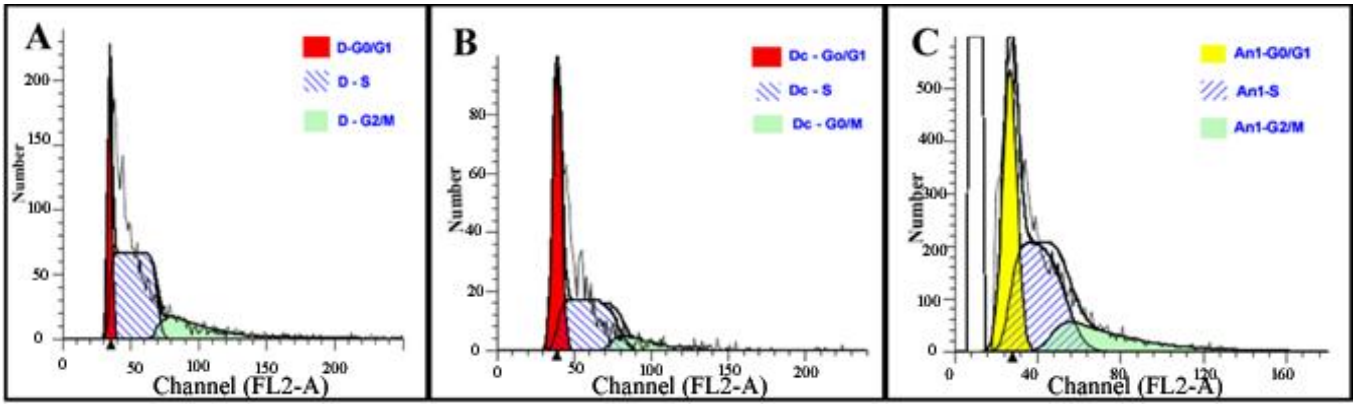


Figure 5.2.DNA content analysis in gastric cancer samples measured by the flowcytometry. Histogram of A) Diploid Normal sample with G0/G1 peak (red), S-phase (shaded peak) and G2/M peak (green) B) Diploid gastric cancer tumour sample C) Aneuploid gastric cancer tumour sample with aneuploidy peak (yellow).

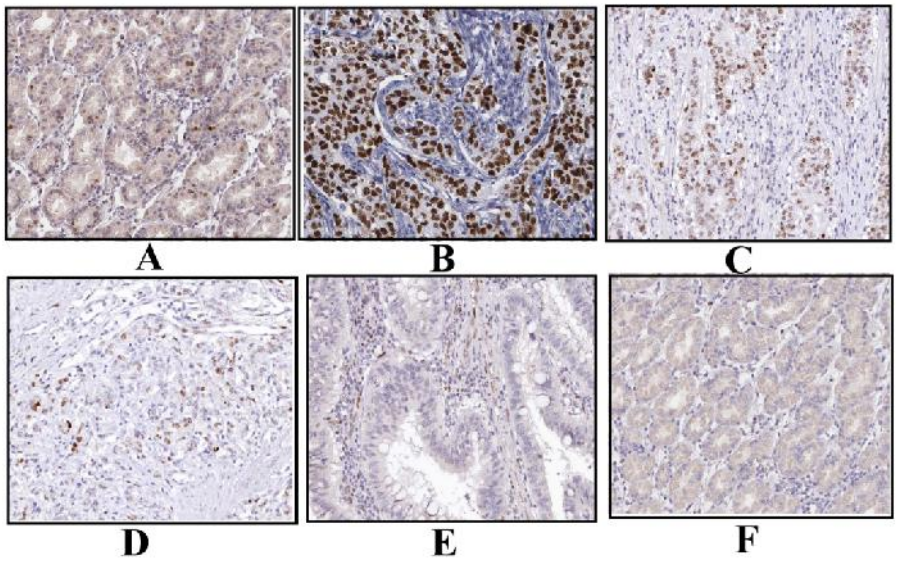


Figure 5.3. (A) Positive high immunoexpression of anti-TP53 in Normal cell (B) Positive high immunoexpression of anti-TP53 antibody in tumor cell (C) Positive moderate immunoexpression of anti-TP53 antibody in tumor cell (D) Positive low immunoexpression of anti-TP53 antibody in tumor cell (E) Negative immunoexpression of anti–TP53 antibody in tumor cell, represented by the brownish colour in

the nucleus. (F) Positive moderate immunoexpression of anti-TP53 antibody in adjacent normal cell (from negative immunoexpression cancer cell).

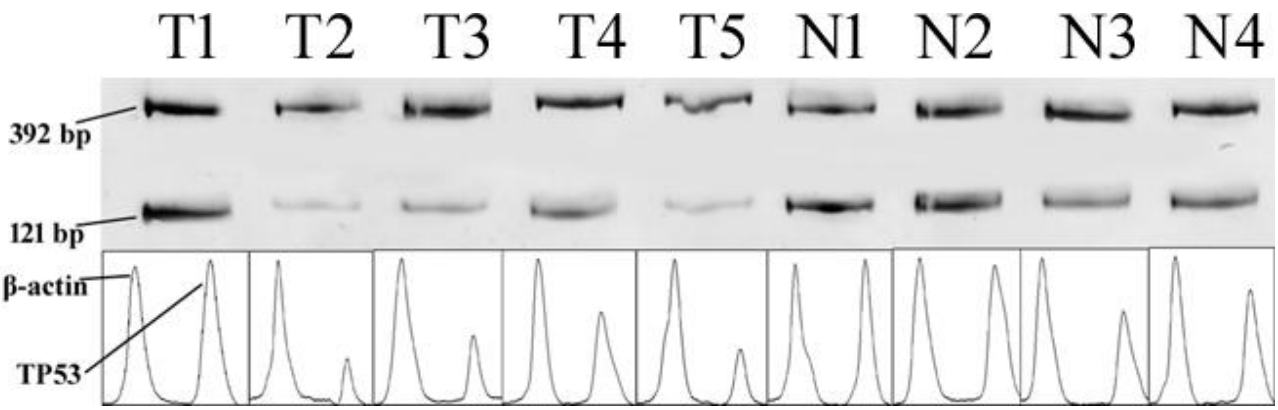


Figure 5.4. Representative tumor samples of TP53(121 bp) expression in 10% polyacrylamide gel. Densitometric quantitative band intensity for TP53 and β-actin (392 bp) amplified products. T1-T5: Tumor samples; N1-N4: Adjacent normal samples.

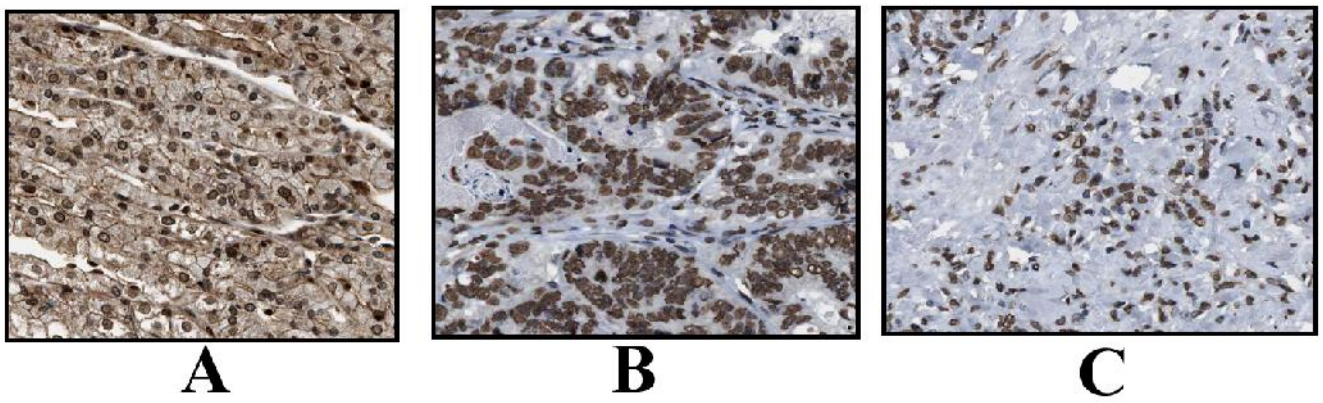


Figure 5.5 (A) Positive high immunoexpression of anti-MDM2 in Normal cell (B) Positive high immunoexpression of anti-TP53 antibody in tumor cell (C) Positive moderate immunoexpression of anti-TP53 antibody in tumor cell

3.5.Clinicopathological influence in TP53 gene

Immunohistochemical analysis of p53 expression in *H. pylori*-positive GC resected tissue sections showed that the average positive expression in tumor and adjacent normal tissues were 44.3% and 6.9%, respectively, while in *H. pylori*-negative tissues they were 16.4% and 4.4% respectively. The average p53 expression was significantly higher in the tumor tissue in the presence of *H. pylori* than in the absence of *H. pylori*. In the logistic regression analysis, *H. pylori* negative subjects, and those infected with cagA negative *H. pylori* were combined and compared with cagA positive subjects to find the exact influence of presence of cagA genotyping in TP53 mutation and expression. Most of the samples containing cagA positive exhibited TP53 mutation with high p53 protein expression.

4. Discussion

In the current study, we attempted to identify novel Missense and splice site mutations in TP53 gene that are associated with demographic factors and gastric cancer risk. Gastric cancer in Mizo population has been shown to have a positive association with consumption of smoked and salted meat and vegetables. Smoke-drying and preservation leads to formation of carcinogenic N-nitroso compounds by the reaction of Nitrites with amines and amides in meats and other proteins (Correa, 1985; Ghatak et al., 2016). Because of the presence of nitrite in processed smoked meats, its role in the development of stomach cancer cannot be ignored. This indirectly contributes to the development of chronic atrophic gastritis, leading to the development of stomach cancer (Nomura, 1990). Studies in the past have also shown

positive association of high intake of smoked meats as potential confounder for gastric cancer (Ward et al., 1999; Kneller et al., 1991). High salt intake showed significantly positive association with diffuse type gastric cancer. Intragastric high salt concentration destroys the mucosal barrier through the increase of surface mucous cell mucin and decrease of gland mucous cell mucin (Tamura et al., 1994), leading to inflammation and damage such as diffuse erosion, atrophic gastritis and decreased acidity of the stomach (Tsugane et al., 2007). It creates a condition favoring *H. pylori* infection (Kato et al., 2006). Salt may also directly damage gastric mucus, improve inflammatory responses of the gastric epithelium (Campos et al., 2006) leading to increased epithelial cell proliferation as part of the repair process. This potentiates the action of carcinogens and increases the probability of endogenous mutations (Kim et al., 2010). Infection with *H. pylori* and *CagA+* strains were more frequent in gastric cancer patients than in controls (OR: 1.168; 95% CI: 0.365 – 3.733; P = 0.094 and OR: 8.298; 95% CI: 10.365 – 6.578; P<0.005, respectively). However, very less differences in the prevalence of *Epstein–Barr virus* (*EBV*) strains were observed between patients and controls (22.5 vs. 7.5%, respectively). Frequent consumption of fermented fat (sa-um) was found to be associated with the risk of developing stomach cancer. This is a food cuisine uniquely consumed in Mizoram (Phukan et al., 2006). Dietary intakes of total or saturated fat have been shown to be associated with stomach cancer (González and Lopez-Carrillo, 2010; Kneller et al., 1992). Boiled pork fat, in addition to being a rich source of saturated fat, may form carcinogenic compounds during long storage, as in other stored meats.

A recent meta-analysis including 203 relevant studies (assessing 225 polymorphisms across 95 genes) found a total of 37 polymorphisms across 27 genes to be significantly associated with gastric cancer in Asians, and 12 polymorphisms across 11 genes in Caucasians (Camargo et al., 2006). In terms of salt consumption, studies *in vitro* have shown that several *H. pylori* genes associated with virulence

(including *cagA*) were up-regulated when the bacterium was cultured in a medium with high salt concentration. As a result, increased expression of CagA was observed leading to alteration of gastric epithelial cell morphology and function (Loh et al., 2007).

TP53 abnormalities, including protein over-expression, missense mutations, frameshifts and Clinico-pathological factors, have been assessed in studies of gastric cancer usually involving populations at high risk for cancers arising in the setting of multifocal gastritis (e.g., Japan, china, northeast India). The majority of gastric cancers in these high-risk populations are intestinal type tumors that arise in the pyloric antrum, with a significant minority of diffuse tumors arising in oxyntic mucosa showing severe superficial gastritis. The antral and oxyntic mucosa of these patients is usually quite normal in appearance, and their risk factors (obesity, abdomen pain and weight loss) (MacDonald et al., 2001) differ from those of more distal gastric cancers (*H. pylori* infection, high salt and nitrate intake, inadequate intake of antioxidant micronutrients) (Nomura, 1991).Mutations in p53 represent one of the critical events in the development of many human neoplasias (Lutz and Nowakowska-Swirta, 2002) and are often associated with the acquisition of greater tumoral invasiveness (Sakurai et al., 1995). GCs show a variable involvement of p53 mutations. Reports published in the last few yearsshows variable p53 mutation rate ranging from 8 to 55% of the cases (Berloco et al., 2003; Frickeet al., 2003). In our own study, overall 60% presented p53 mutations, mostly within the conserved areas (59%) or in domains, which were important for the functionality of the protein (41% in L3-LSH). Although several authors report that p53 mutations tended to be located in the upper portion of stomach associated with advanced age (Palli et al., 1997) or with advanced stage (Guo et al., 2000); in our present study, we did not find any significant association between p53 mutations and the traditional clinicopathological variables (age, sex, site and size of the tumor, stage, grading, invasiveness). Moreover, we found significant association

between Tp53 gene mutations with presence of *H. pylori* with or without cagA genotyping. Contradictory findings have been reported (El-Rifai and Powell, 2002) with regard to the presence of p53 alterations and the prognosis of GC patients. A great many studies, however, have used immunohistochemical analysis to detect overexpression of p53 as an indirect means of evaluating mutations of this gene and this assay does not appear to have consistent prognostic significance in GC patients (Fenoglio-Preiser et al., 2003). Few studies so far conducted to evaluate the prognostic impact of p53 mutations in GCs do not appear to provide any further information regarding this aspect (Lim et al., 1996; Matturri et al., 1998; Kubicka et al., 2002).The more frequent association of p53 mutations with gland forming, or intestinal type, cancers than in diffuse tumors reproduces the findings of Solcia et al. (1996), who suggest that *H. pylori*-induced gastric cancers follow two induction pathways: glandular and mixed glandular/ diffuse tumors arise in intestinalized mucosa, whereas diffuse tumors arise directly from the neck cells of non-metaplastic glands through primary involvement ofgenes affecting cell±cell and cell±matrix junctional proteins.

We report in this study, new mutations located deep into TP53 intron 5, c.693-2A>G, through the analysis of transcripts of an index case belonging to high-risk gastric cancer population. This mutation activates a cryptic 34 splice site and leads to the exclusion of a 24-nt from exon 6 that disrupts the TP53 ORF and introduces several premature termination codon (PTC) in the coding sequence. According to the present accepted rules, most of the mutant transcripts are degraded by the NMD pathway. c.693-2A>G was subsequently identified only for this region. Some splicing defects might not affect all the produced transcripts but merely enhance the occurrence of an alternative or a cryptic splicing event. The causality of variants leading to such partial effects is difficult to assess in molecular diagnostic as it is usually not known how detrimental they are. We can deduce from these observations that the cryptic exon is nearly

systematically included in transcripts deriving from the mutant TP53 allele in carriers of the c.693-2A>G mutation, which can therefore be considered as a true deleterious mutation.

Several studies have reported an association of codon 72 variants with tumor susceptibility. The Pro/Arg genotype are more likely to develop gastric cancer than those with other genotypes (Kim et al., 2009; Hiyama et al., 2002), especially in patients who are *H. pylori* infected. An increased frequency of the pro allele (pro/pro or arg/pro genotypes) is also found in patients with gastric cancer. In the present study we have found 28.75 % patient exhibited 72 codon variation, most of the variation are somatic in nature except for 5 patients exhibited germline in nature. TP53 codon 72 mutation was higher in tumors with the Arg/Arg variant (homozygous). Earlier studies suggested that the resistance to chemotherapy in the homozygous Pro/Pro tumors was largely due to the Pro form of wild-type TP53 rather than the heterozygous Arg/Arg mutation (Santos et al., 2006), which was found in Mizo population. However, in a study of gastric adenocarcinoma, Zhang et al. (2004) reported that the frequency of the Arg homozygous allele was positively correlated to the patient's age at the baseline. In our study most of the patients with 72 codon variation with Arg/Arg were in young age. This polymorphism is balanced, although the selective pressure maintaining this is not known. The difference in the allele causes changes in the primary structure of the protein (Matlashewski et al., 1987). Some researchers have shown that the codon 72 polymorphism of the p53 contributes to susceptibility to various cancers, including the lung, esophagus, breast, uterus, and vulva (Lee et al., 2000). All the patients containing Pro/Arg or Arg/Arg allele in 72 codon, were infected with *H. pylori* with or without cagA genotyping.

C to T or T to C and A to G or G to A mutations are induced by nitric oxide (Nguyen et al., 1992; Wink et al., 1992), a substance known to be produced during *H. pylori* infections. G:C-A:T transitions are also specifically induced by N-methyl-N_2-nitro-N-nitrosoquanidine and N-nitroso compounds found in foods,

substances considered to be carcinogens involved in gastric carcinogenesis (Sugimura and Kawaki, 1973) which was found a potential confounder for our study. Most of the mutated patients used to consume high amount of smoked meat, processed meat and fermented pork fat. These foods are commonlyconsumed in Mizo populations. In our study, the smoked meat and fermented fat consumption also significantly associated with gastric cancer which can induced N-nitroso compounds as a carcinogen. In gastric benign lesions and gastric cancer, our research group assessed the p53 overexpression and occurrence of aneuploidy for TP53 gene mutations. In gastric tumor, immunohistochemistry revealed p53 overexpression in 12% of the analyzed cases, as well as TP53 gene mutations in 56% of the cases, but normal adjacent mucosa was exhibited moderate level of p53 expression with diploid cell cycle. Lu et al. (2009) showed that in tumor tissues aneuploidy occurred in GC samples overexpressed p53 protein, significantly higher than those in non-tumor strict mucosa. All the overexpression of p53 protein was associated with the size of tumors that may help in diagnosis and prognostic prediction of GC.The role of *H. pylori* in the expression of p53 was studied earlier by Satoh et al. (2001), who found that p53 positive cells in *H. pylori*-infected gastric mucosa before treatment decreased significantly one month after *H. pylori* eradication. All the samples containing *H. pylori* cagA-positive strains contribute significantly to p53 alteration in GC. Murakami et al. (1999) reported that *H. pylori* infection can induce p53 point mutations in gastric mucosa of gastritis patients that in turn leads to dysplasia or carcinoma. Most of the patients have intestinal type of tumor which was infected with *H. pylori* with or without cagA genotyping. The p53 mutations mostly exhibited in intestinal type of gastric cancer rather than diffuse type gastric cancer for Mizo population. The degree of p53 expression correlates with the proliferative rate of the tumors, perhaps explaining the higher incidence of p53 positivity in intestinal versus diffuse GC (Fenoglio-Preiser et al., 2003).

In this study, we focused on the participation of genetic alterations of TP53 gene, such as mutational inactivation, SNPs, and expression of mutant form of p53 protein in the esophageal and gastric carcinogenesis. The studies emphasize the fundamental role of molecular alterations of "the guardian of the genome" in these neoplasms, with serious consequences for the deregulation of the cell cycle, loss of pro-apoptotic function. Considering the involvement of TP53 alterations both in early stages as in tumor progression, it is an important biomarker for the diagnosis, tumor progression, and poor prognosis associated with lymph nodes metastasis. Gastric carcinogenesis maybe related to the existenceof chronic *H. pylori* infection, which leads to imbalance of proliferation and apoptosis in the early stage, and furthermore mutationof p53-tumor-suppressor system and finally gastric cancer tumorigenesis. Hence, this molecular pathology mechanism can be applied in routine diagnostic procedures, classification systems, disease monitoring.

Chapter VI

Novel mutations and inactivation of the E-Cadherin Gene in diffuse Type Gastric Cancer

1. Introduction

Development of malignant tumours, in particular the transition from benign lesions to invasive, metastatic cancer is characterized by a tumour cell's ability to overcome cell–cell adhesion and invade the surrounding tissue. Many different cell-adhesion molecules are implicated in human carcinogenesis. During the transition from normal cells to highly malignant tumour cells the expression of some of these adhesion molecules is switched off, whereas that of others is induced (Thiery, 1996).The E-cadherin is a member of the transmembrane glycoprotein family responsible for calcium-dependent, cell-to-cell adhesion and plays a fundamental role in the maintenance of cell differentiation and the normal architecture of epithelial tissues (Moriyama et al., 2001). The protein is encoded by the CDH1 gene which is located on chromosome 16q22.1 and consists of 16 exons. Mutations in CDH1 geneare known to be associated with Hereditary Diffuse Gastric Cancer syndrome (HDGC) (Kaurah et al., 2007; Humar et al., 2007; Brooks-Wilson et al., 2004).

Loss of intercellular adhesion is a hallmark of migratory cells. Given the central role of E-cadherin in this process, the close association between E-cadherin loss and the acquisition of an infiltrating, invasive phenotype of a tumor is not surprising (Cavallaro et al., 2004). However, the effect of E-cadherin loss on epithelial tissue lacking a preexisting neoplastic phenotype is less well understood. Gastric cancer is usually classified into intestinal and diffuse types. Diffuse gastric cancer can be further subdivided into poorly differentiated carcinoma and signet-ring cell carcinoma (SRCC). Recent evidence suggests early SRCC is an initial, differentiated form of diffuse gastric cancer that may evolve into poorly differentiated

carcinoma (Humar et al., 2007). Although the worldwide incidence of gastric cancer is declining, the incidence of diffuse gastric cancer, particularly SRCC is rising in the United States (Henson et al.,2004).

The majority of human cancers (~80– 90%) originate from epithelial cells. Normally, epithelial cells are tightly interconnected through several junctional structures, including tight junctions, adherens-type junctions and desmosomes which are initimately associated with the actin and intermediate cytoskeleton. Crucial for the establishment and maintenance of these junctional complexes are Ca^{2++}-dependent homophilic interactions mediated by the cell-adhesion molecule E-cadherin in the zonulaadherens (Aberle et al., 1996).The proliferation and migration of epithelial cells are crucial for gastric mucosalrecompense. In the migration of epithelial cells, cell-to-cell adhesion seems to be essential. Among the molecules known to be involved in adhesion between epithelial cells, E-cadherin in cooperation with cytoskeletal structures plays an important role in cell differentiation, morphogenesis (Del Buono et al., 1999) and oncogenesis (Carneiro et al., 1999). Several molecules that make up tight junctions (Noach et al., 1994) and gap junctions (Takahashi et al., 2000) have been shown to be involved in mucosal recompense. However, the role of E-cadherin in the recompense of gastric epithelial cells after gastric mucosal damage is unknown. E-cadherin expression is reported to be weak at the edge of intractable intestinal ulcers in patients with Crohn's disease. This observation suggests an important role of E-cadherin in ulcer healing (Hanby et al., 1996).

Patients with germline mutations in the CDH1 have a high risk of developing diffuse gastric cancer (Keller et al., 1999). The estimated penetrance of CDH1 mutations is 70–80% for stomach cancer (Kaurah et al., 2007; Oliveira et al., 2009). Because of high penetrance of CDH1 mutations and almost

100% mortality of patients with symptomatic DGC prophylactic total gastrectomy during the second decade of life is recommended for CDH1 mutation carriers (Guilford et al., 2007).

Here, we describe the identification of inactivating germline and somatic CDH1 mutations in diffuse, poorly differentiated gastric cancer for Mizo population. We also examined the expression of E-cadherin on a panel of primary gastric carcinomas encompassing a wide range of tumor stages and on autologous normal mucosae. We also compared the pattern of E-cadherin expression on patients with signet-ring cell gastric carcinoma.

2. Materials and methods

2.1.Subjects and sample collection

Tumor tissue, adjacent normal and blood specimens from untreated patients with primary gastric cancer (n = 80) obtained up to 30 min after surgical removal were immediately collected from the Civil Hospital, Aizawl, Mizoram and stored at -80°C until sectioning and nucleic acid extraction. In tumors of advanced stages, tissue samples of regional lymph node metastases (n = 43) were available. In addition, 80 unrelated healthy controls (aged from 31 to 73 years) blood samples belonging to the same ethnic group were also collected. The tumor samples were all confirmed histologically and all subjects provided informed consent for obtaining the study specimens. Clinical staging for each gastric carcinoma was evaluated according to the TNM staging system indicating the extent of tumor spread (Hermanek and Sobin, 1987). Gross appearance of the tumors was described according to the Borrmann classification (Borrmann, 1926). Histomorphological tissue architecture of the tumor samples expressed according to the Lauren (Lauren, 1965) and the WHO classifications (Oota and Sobin, 1977) was evaluated on both paraffin sections and hematoxylin-eosin stained sections. Since histology often varied within the same

tumor, the diagnosis was based on the dominant pattern. Alterations of tumor cell shape were described by grading (Hermanek, 1986). Furthermore, tumor diameter and lymphatic vessel invasion (Gabbert et al., 1991) were determined.Medical charts were reviewed using a standard protocol to obtain information on cancer treatment, clinical stage, dietary habits, previous disease history, physical activity, tobacco, and alcohol use by an in-person interview using a structured questionnaire (Ghatak et al., 2014). The study protocol was approved by the Institutional Review Board of all institutes involved in the study.

2.2. DNA extraction from the blood sample

The lymphocytes from whole blood were separated by lysing the RBCs using a hypotonic buffer (ammonium bicarbonate and ammonium chloride, Hi-media) with minimal lysing effect on lymphocytes. Three volumes of RBC lysis buffer were added to the blood sample, mixed by vortexing and inverting thoroughly for 5 min and centrifuged (Eppendorf 5415R, Germany) at 2,000 X g for 10 min. The lymphocytes was used for DNA extraction by modified protocol of Ghatak et al. (2013).

2.3.DNA and RNA extraction from the tissue samples

Deparaffinization was carried out by adding 1 ml of xylene to the tissue section in each microfuge tube followed by vigorous vortexing for 10 minutes. The mixture was centrifuged at 12000 rpm for 10 minutes. The supernatant was discarded and the deparaffinization steps were repeated once again followed by rehydration through subsequent washings with 100%, 90% and 70% absolute ethanol diluted in RNase free DEPC treated water, respectively. The deparaffinised tissue and surgical resection specimen was used for the DNA and RNA co-extraction by the modified protocol of Ghatak et al. (2015).

2.4.PCR amplification of the CDH1 exon region with splicing site

The CDH1 different exon region with splicing site was amplified by polymerase chain reaction (PCR) using eight different primers pair (Table 6.1). PCR was carried out in 25 µl total reaction volumes each containing 100 ng of template DNA, 0.2 pM of each primer, 1X PCR buffer, 1.5 mM $MgCl_2$, 200 mMdNTPs, and 1 unit of Taq DNA polymerase (Fermentas, Germany). The reaction mixture was heated to 94^0 C for 5 min, followed by 30 cycles each consisting of 1 min denaturation at 94^0 C, 1 min annealing at respective temperature (Table 6.1), 1 min of extension at 72^0 C and a final 5 min extension at 72^0 C. The PCR amplification products (3 µl) were subjected to electrophoresis in a 1.2 % agarose gel in 1X TBE buffer at 80 Vh for 30 min, stained with Ethidium Bromide and images were obtained in GBOX gel documentation systems (UK). PCR products were purified with a Qiagen gel extraction kit (Qiaquick columns; Qiagen, Chatsworth, CA) and stored at -20^0 C until for sequencing.

2.5.Immunohistochemical analysis

Immunohistochemical staining for tumour-associated CDH1 gene products (E-cadherin) was performed using a streptavidin peroxidase procedure after an antigen retrieval process using microwave. E-cadherin anti-antibodies were selected among various commercially available antibodies after test procedures using human control slides for immunohistochemistry (Abchem, Japan, ab1416). Cut-off values for E-cadherin antibodies were based on previously reported (Miyazaki et al., 2000). For E-cadherin, the immunolabelling pattern of each case was scored as positive (strong labelling), weakly positive (faint staining), or negative (absence of staining), and the extent of immunolabelling was also categorized as diffuse (if the entire neoplasm was labelled) or focal (if focal loss of expression was noted) (Iacobuzio-Donahue et al., 2000). For statistical analysis of this large amount of data (Bouras et al., 2002), the results

of immunostaining were considered to be positive if 10% the cancer cells showed staining in cytoplasm and membranes.

Exon	Primer (5' -3')	Product size (bp)	Ta (^{0}C)
1	F: TGCCTCCGGGGCTCACCT R: AGAGGAAGGGGCGCAGGG	291	65
2	F: GTTTCGGTGAGCAGGAGGG R: CCTACTCCGCCCAGGGAC	216	63
3	F: CTCTGTGATTTCTGCCCTGC R: CGCACTAAAACAACAGCGAAC	290	60
4-5	F: CTTGTTCCTCATCTTCTTTCC R: GACCAAAAAATCCTGGGTGG	511	57
6	F: CCTTCTCCCATGTTTTCTTCC R: GTATTCAGATTCTCATTAGTGG	232	54
7	F: CAGCTTGTCTAAACCTTCATC R: CCTCCACACCCTCTGGATC	248	56
8	F: GCTAGTGTTCCTGGTCCTG R: AGTGGTGACACTTAGTTCATG	245	58
9	F: GACACATCTCTTTGCTCTGC R: GGGACAAGGGTATGAACAGC	269	59
10	F: AGTCATGGCAGAAACCACAG R: CAGTTGCTGCAAGTCAGTTG	418	60
11	F: CCAGAGCTTGTCCCCGTTC R: GTCGAGGCAGCAAAGGCTC	203	62
12	F: GTCTGGTGGAAGGCAATGG R: GAAGCATGGCAGTTGGAGC	345	61
13	F: TTCCTCCCCTGGTCTCATC R: TCAAAGGCTGAGTCACTTGC	304	60
14	F: TTGCTCTGTCTCCCCCACC R: TCCGAATAAAGAGATCACCAC	205	57
15	F: AGTGAAGGCATCATCCAACC R: AGCTTAGAGATGAGCCATGC	294	59
16	F: AGATGCTTTTGTCCCTTCTTC R: AGCACATGGGTCTGGGGC	272	59

Table 6.1. Primers designed in the present study for amplification of exons and splice sites in CDH1 gene.

2.6.SSCP analysis

The CDH1 all exon region was screened for mutations with a rapid, sensitive method of SSCP (Gayther et al., 1998). An aliquot of 0.75 µl of each PCR product was diluted with an equal volume of water and mixed with 1.5 µ1 of 95% formamide. This mixture was denatured at 95°C for 5 min, cooled on ice and 2 µl was used for loading on SSCP gel (8% non-denaturing polyacrylamide gels). SSCP Gels were pre-run at 400 V, 20 mA, 2 W, for 10 or 50 volt-hours (Vh). Electrophoresis was performed at 400 V, 20 mA, 2 W, for 200-300 Vh. Electrophoresis was carried out at either 4, 10, 15 or 20°C depending on the optimal temperature for a given PCR fragment. The gels were ethidium bromide stained, and gel documented by Syngen-G-BOX(USA) machine.

2.7.Sequence analysis

The samples exhibiting mutation or polymorphism and instability, after SSCP analysis, was taken for further sequencing and mutation analysis. All products were sequenced from opposite directions to ensure reading accuracy. Sequences and chromatograms obtained were examined by chromas software version 2.13, DNA baser and align by BLAST [www.ncbi.nlm.nih.gov/blast]. All the sequences containing the mutation ware evaluated for their potential pathogenicity using the following algorithms: DNA baser version 3.5.4.2, Codon Code aligner version V.4.2.2, The CDH1 all exon was checked from the E cadherin Gene card data base [HGNC:17481, Entrez Gene:9992, Ensembl: ENSG000000390687, OMIM: 1920905, UniProtKB: P128303].Mutation taster [www.mutationtaster.org/], PolyPhen-2 [http://genetics.bwh.harvard.edu/pph], SIFT [http://sift.jcvi.org], Mutation Assessor [http://mutationassessor.org/]. The MEGA Align algorithm was used at two depths of alignment [Cancer

to Normal and Normal to database sequences]. The results of PolyPhen-2 was retrieved from the original webpage [version 2.2.2] but also from version 2.0.22 run by PON-P and version 1 run by Condel, which use them for weighted average scores.

2.8.Reverse Transcriptase PCR (RT-PCR) for CDH1 Expression

First-strand synthesis of cDNA was performed using Superscript II reverse-transcriptase (Life Technologies Inc., USA). Five μg aliquot of total cellular RNA was used for each reverse transcription reaction, and one-tenth of this reaction was used for PCR. Primers targeting TP53 exons 4 and 5 region (F: CTCACATTTCCCAACTCCTCTC and R: TGGCAATGCGTTCTCTATCC) and exons 1 and 3 of β-actin (F: ACCATGGATGATGATATCGC and R: ACATGGCTGGGGTGTTGAAG) were used (DeBruin et al., 2000). The PCR was run for 35 cycles consisting of 94° C for 30 seconds, 1 minutes at gradually decreasing temperatures using a touchdown protocol (4 cycles each at 62° C and 60° C; and 30 cycles at 58° C), and 74 °C for 5 minutes. RT-PCR products were electrophoresed in 10% polyacrylamide gel and visualized by ethidium bromide staining.

2.9.Densitometric Analysis

Semiquantitative RT-PCR gels and Western blot were analyzed to compare the banding pattern and their molecular mass using gel image analysis software Syngen G-Box (Sacramento, CA, USA) and ImageJ.

2.10. Statistical Analysis

CDH1 gene mutations among case–control subjects was tested for Hardy–Weinberg equilibrium by a chisquare (v2) test with one degree of freedom (df). The mutations in each group were estimated using odds ratios (ORs), and 95% confidence intervals (CIs). Variant distributions in CDH1 gene and their

association with gastric cancer stages, depth, grading, CDH1 gene and protein expression and demographic factors was assessed by fisher exact test. Additionally, logistic regression analyses were conducted to compute the potential confounder's influence of both genetic and environmental factors for gastric cancer. Then, the independent effect of risk factors was investigated in a multivariate model (introducing all variables and terms of interactions) keeping only those statistically significant or showing a confounding effect on the studied factors. For all tests, a two-sided p-value <0.05 was considered statistically significant. All statistical analyses were performed using SPSS 20.0 program (SPSS, Madrid, Spain) and conformed by R-console statistical package ver3.3.0 (The R Foundation for Statistical Computing).

3. Results

3.1. E-Cadherin Expression in Primary Gastric Cancer samples

Most of the investigated primary diffuse type gastric carcinomas exhibited a low level or negative E-cadherin gene and protein expression (92%), and this was heterogeneous both in terms of the fraction of positive cells and the staining intensity (Figures 6.1 and 6.2). In contrast, most of the intestinal type of gastric carcinoma exhibited a strong and moderate level E-cadherin expression (72%) (Table 6.2). Comparison with various clinicopathological features revealed no significant relationship between E-cadherin-staining pattern and tumor diameter (<5 cm versus >5 cm, P = 0.264) or lymphatic vessel invasion (+ versus -, P = 0.182), whereas the Borrmann classification, reflecting the gross appearance of the carcinomas, showed a slight, significant, correlation between the presence of E-cadherin and Borrmann types I and II (type I/II versus type III/IV, P = 0.0038). However, comparison of E-cadherin staining with the TNM staging system showed a significant correlation with the T stage, which refers to the infiltration depth of the primary tumor. Carcinomas extending through the serosa (stage T3) and

involving contiguous structures (stage T4) exhibited <50% E-cadherin-positive cells more frequently than did stage T1/T2 tumors (P = 0.0134). A trend was also observed between the presence of lymph node metastasis and the loss of E-cadherin expression (N0 versus N1/2, P = 0.0479). While no correlation could be demonstrated with the M stage (M0 versus Ml. P = 0.328), which was evaluated at the time of diagnosis when metastatic disease was obviously evident. The most significant correlation was found with differentiation parameters (Table 6.2). According to the Lauren classification, gastric carcinomas were classified as intestinal, intermediate, and diffuse tumor types. Intestinal tumors which are characterized by a distinct glandular formation generally demonstrated positive or weakly positive E-cadherin staining, although the fraction of E-cadherin-positive cells varied (Table 6.2). Intestinal carcinomas could be further subdivided into well-differentiated (n = 15), moderately differentiated (n = 18), and poorly differentiated (n = 5) tumors (Table 6.2). Well-differentiated tumors composed of tight monolayer glands all showed a strong and homogeneous E-cadherin expression similar to the autologous normal mucosa. Most of the moderately differentiated tumors also showed moderate level of E-cadherin staining, although the fraction of positive cells was always less than in normal mucosa. The poorly differentiated carcinomas all demonstrated weak E-cadherin expression in only a small fraction of tumor cells.

In contrast to the glandular formations of diffuse type tumors, the polymorphic cells of diffuse carcinomas form spots and clumps in the surrounding tissue. None of these tumors (n = 28) showed any E-cadherin staining (n = 21) or low level (n = 7) (Table 6.2). Intermediate type tumors contain both intestinal and diffuse areas. In these tumors, glandular structures were always E-cadherin positive, while specimens showing the typical diffuse morphology were E-cadherin negative. Intensely stained, well-differentiated and weakly stained, poorly differentiated tumor areas could be seen adjacent to each other on the same tissue section. The correlation between E-cadherin expression and glandular differentiation

in gastric cancer was also apparent when the tumors were classified according to the WHO system (Table 6.2). This histological classification distinguishes various types of adenocarcinomas such as papillary, tubular, and mucinous, as

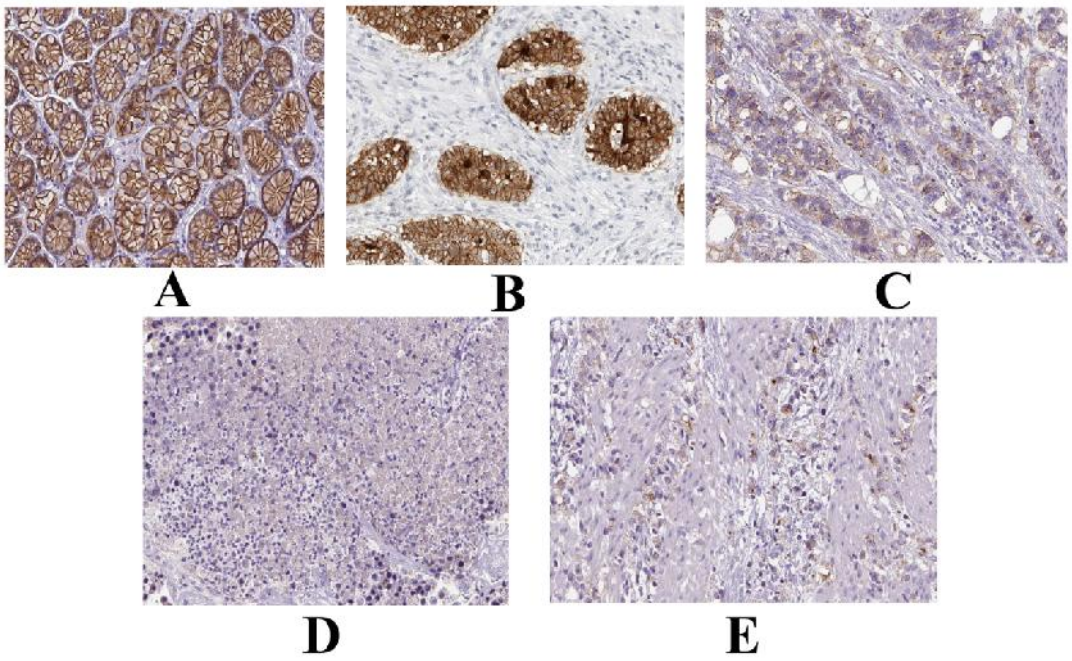


Figure 6.1. (A) Positive high immunoexpression of anti-CDH1in Normal cell; (B) Positive high immunoexpression of anti-CDH1 antibody in tumor cell (Signet ring cell); (C) Positive moderate immunoexpression of anti-CDH1 antibody in tumor cell;(D) Negative immunoexpression of anti–CDH1 antibody in tumor cell represented by the brownish colour in the nucleus (Signet ring cell); (E) Positive lowimmunoexpression of anti-CDH1 antibody in adjacent normal cell (from negative immunoexpression cancer cell).

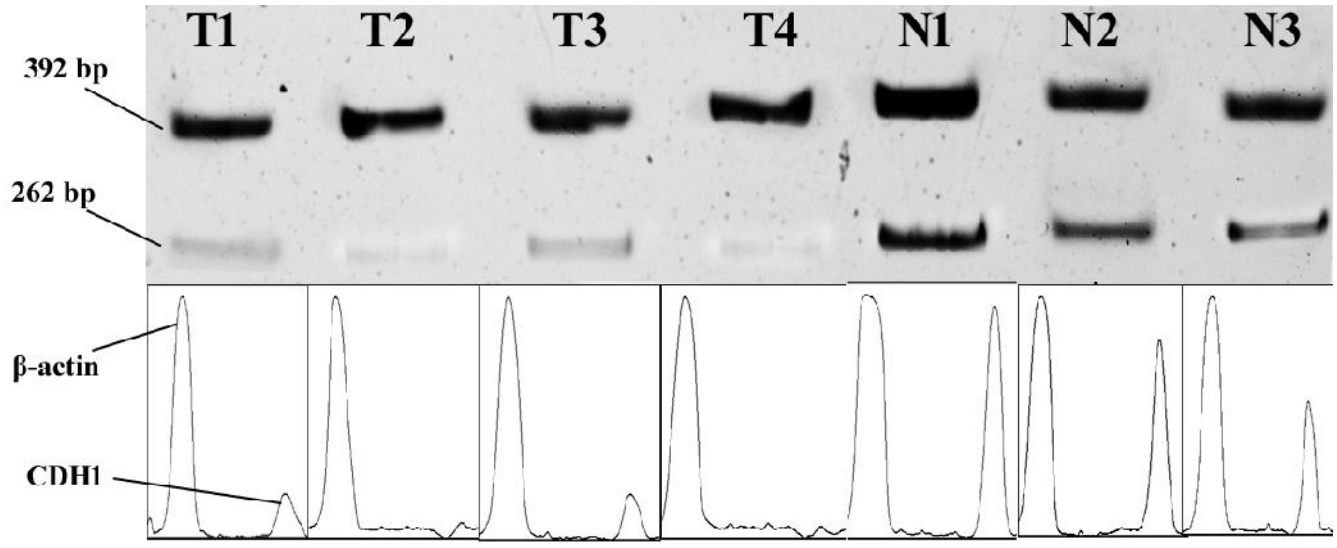


Figure 6.2. Representative tumor samples of CDH1(262bp) expression in 10% polyacrylamide gel. Densitometric quantitative band intensity for CDH1 and β-actin (392 bp) amplified products. T1-T4: Representative tumor samples; N1 - N3: Representative adjacent normal samples.

Conference/Seminars/Workshops Attended

- Training course of **"Bioinformatics-Protein and their Structure Prediction"** Sponsored by Department of Biotechnology, Ministry of Science and Technology, Government of India, Mizoram University, 23-24th November, 2011.
- Participated as a member of organizing committee in "**5th Interactive Meeting of North East Bioinformatics Center Network (NEBInet)"** from 11-12th October, 2012.
- **"Advanced Molecular Biology Techniques**" organized by Institutional Biotech Hub, Departmentof Biotechnology, Gauhati University, Guwahati 781014, Assam, India. From 19 to 25th November, 2012.
- Department of Biotechnology, Government of India Sponsored National workshop of hands on "**Training Course on Tools in Molecular Biology and Bioinformatics**" November 26 – 30, 2012 Department of Biotechnology, Assam University, Silchar- 788 011.
- Participated as a Delegate in the **"1st Indian Cancer Genetics Conference & Training Workshop (Sequencing Technology)"** from January 22-31, 2013 at ACTREC & Tata Memorial Hospital, Mumbai.
- Participated as a participant in the **"Workshop on the Barcoding and PCR Techniques"** from February 22-36, 2013 at State Biotech Hub, Tripura University, Agartala, Tripura.
- Participated as a participant in the **"Structure Determination of Macromolecules"** from March 26-28, 2013 at DBT-BIF Centre, Department of Biotechnology, Mizoram University.

- Participated as a Resource Person in the **"A Primer in Bioinformatics and Biotechnology to School Students"** from June 08, 2013 at DBT-State Biotech HUB, Department of Biotechnology, Mizoram University.
- Oral presentation in the **"BIOINFOGEN - 2013"** from September 6 - 8, 2013 at Genetic – Metabolic Unit, Department of Pediatrics PGIMER, Chandigarh, India.
- Participated as a participant in the **"India-UK Scientific Seminar"** from January 14-15, 2014 at Department of Biotechnology, Mizoram University.
- Poster presentation in the **" Advances in Cancer Genomics"** from May 30 – 31, 2014 Organized jointly by Mizoram State Cancer Institute and Department of Biotechnology, Mizoram University, Aizawl, Mizoram.
- Participated as a participant in the **"Workshop on Capacity Building in Effective Management of Intellectual Property Right (IPRs)"** from August 27 - 28, 2014 at Mizoram University, Aizawl, Mizoram.
- Participated as a Delegate in the **"2nd Indian Cancer Genetics Conference & Workshop"** from 27th October to 2nd November, 2014 at ACTREC & Tata Memorial Hospital, Mumbai.
- Oral presentation in the **"2nd Indian Cancer Genetics Conference & Workshop"** from 27th October to 2nd November, 2014 at ACTREC & Tata Memorial Hospital, Mumbai.
- Participated as a participant in the **"Basic and Clinical Flow Cytometry course"** jointly organized by the Cytometric Society of India, Indian Institute of Technology, Guwahati and Cachar Cancer Hospital and Research Center from November 10 - 11, 2014 at Cachar Cancer Hospital and Research Centre, Silchar, Assam.

- Participated as a Resource Person in the **"Cancer Mutations - Detection and Analysis"** from April 17 – 18, 2015 at DBT-State Biotech HUB, Department of Biotechnology, Mizoram University.
- Participated as a Resource Person in the **"Browsing genome with Ensemble and UCSC genome browser"** from 30th May, 2015 at DBT-State Biotech HUB, Department of Biotechnology, Mizoram University.
- Participated as a participant in the **"Advanced Research Training Workshop for North East Region Scientists on Understanding Human Disease and Improving Human Health Using Genomics-Driven Approaches"** organized by the **National Institute of Biomedical Genomics (NIBMG)**, Kalyani from October 05 - 113, 2015 at NIBMG, Kalyani, West Bengal.

		Laurein-classification			WHO-classification		Grading (G)		
CDH1 expression (%)	**Adjacent Normal Mucosa (N = 80)**	**Intestinal (N = 38)**	**Intermediate (N = 14)**	**Diffuse (N = 28)**	**Adenocarcinoma (N = 46)**	**Signet ring cell carcinoma (N = 34)**	**G1/G2 (N = 23)**	**G3 (N = 51)**	**G4/GX (N = 6)**
>70	57	7	-	-	7	-	6	1	-
>50-70	11	10	3	-	13	-	8	5	-
>30-50	12	13	2	-	9	6	5	10	-
≤30	-	4	1	7	5	7	3	8	1
0	-	4	8	21	12	21	1	27	5

Table 6.2. E-cadherin expression in primary gastric cancer correlated with various histopathological classifications. Staining intensity E-cadherin-positive cells was evaluated in comparison to the adjacent normal, N= number of tumors of each type investigated. According to the WHO system: papillary, tubular, and mucinous tumor types were combined under the term "Adenocarcinoma". High E-cadherin expression was associated with intestinal tumor type (intestinal versus diffuse P = 0.0039), adenocarcinoma tumor type (adenocarcinoma versus signet ring cell carcinoma, P = 0.0021), and a high degree of cellular differentiation (G1/G2 versus G3/G4/GX, P = 0.0001).Histo-morphological parameters reflecting the degree of glandular and cellular differentiation showed a significant correlation with the E-cadherin expression.

well as signet ring cell types. All but one of the signet ring cell carcinomas (n = 34) were E-cadherin negative (n = 21 and low level (n = 7) (Figure 6.1, Table 6.2), while most of the other adenocarcinoma types (34 of 46) were E-cadherin positive. As in the intestinal type, the staining reaction depended on the degree of differentiation of tumor glands and thus, the E-cadherin-negative adenocarcinomas lacked any glandular formations. Another differentiation parameter which correlated with E-cadherin expression was the grading, which describes the morphological deviations of tumor cells from the typical epithelial phenotype of normal mucosa cells. Stomach carcinomas mainly composed of cells characterized by a columnar and polarized cell shape (Gl, G2) showed intensive E-cadherin staining enriched at the lateral cell borders. G3 carcinomas, consisting of cells with abnormal mitosis and nuclear pleomorphism or large prominent signet ring cells generally exhibited a reduced circumferential staining or a complete loss of the cell adhesion molecule. Small atypical malignant cells (G4, GX) were uniformly E-cadherin negative (Table 6.2). E-cadherin expression was also correlated with tumor recurrence and mortality.

3.2. CDH1 mutation pattern in lobular and ductal tumors

Screening of all the exons of CDH1 gene showed that 23.75 % patients with gastricadenocarcinoma had mutations and majority of them were germline in nature (due to presence in adjacent normal and blood sample). A total of 24 mutations were found in Exons 3, 4, 5, 7, 8, 9, 10, 11, 12, 14 and 15 (Table 6.3). Out of all 24 mutations, 15 were non-synonymous, 6 were frameshift and 3 were synonymous in nature. Exon 3 exhibited 2 non-synonymous mutations including one novel frameshift mutation (64502_64503insG), and it was predicted disease causing with high polyphen2 score. Due to this mutation, early stop codon was generated at the 94 codon position (Table 6.3).

Exon 4 exhibited four novel mutations including a frameshift mutation (71239_71240insC) due to cytosine insertion at g.71239 position. Due to this mutation, early stop codon was generated at the 168 codon position. All the mutations in exon 4 were predicted as disease causing with high polyphen2 score. Exon 5 exhibited four novel mutations, all are predicted as polymorphisms by mutations taster, except 71311G>A non-synonymous mutation which was disease causing with high polyphen2 score. Due to these mutations, the function of cadherin 1 binding Domain might be lost (Table 6.3).

Exons 7 and 8 exhibited 3 and 2 mutations which was predicted as disease causing with high polyphen2 and SNP-GO score, respectively. Two frameshift mutations (74485_74486insA) at the position of g.74485 in exon 7 and (74922_74923insG) at the position of g.74922 in exon 8 led to generating premature stop codon at 293 and 350 positions, respectively.

Exons 9 exhibited 3 novel mutations including two frameshift mutation (76145_76146insGA and 76175_76176insT) which was predicted as disease causing with high polyphen2 and SNP-GO score. Two frameshift mutations (76145_76146insGA) at the position of g.76145 and (76175_76176insT) at the position of g.76175 led to generating premature stop codon at 418 and 419 position, respectively.

Exons 11, 12, 14 and 15 exhibited single novel mutation each, as predicted polymorphism, disease causing, polymorphism and disease causing with high polyphen2 and SNP-GO score, respectively. Due the mutations of exon 11, 12, 14 and 15, function of cytoplasmic binding TOPO domain might be lost for E-cadherin protein.

All the mutations are homozygous in nature except 74922_74923insG in exon 8 is heterozygous in nature. All the 6 frameshift mutations exhibited germline in nature and negative gene and protein expression by a

Exon Number	Mutation	Frequency	Amino acids change	Mutation Taster/Polyphen2/ SNP&GO	Effect of the Mutations
Exon 3	64491C>A	18%	70S>Y	Polymorphisms/0.976(Probably damaging) / Disease(RI-1)	• The new residues that is introduced, located in the signal peptide differ in the properties and it is possible this mutation disturb recognition of signal peptide. • Sequence of this peptide is important because it is recognized by other proteins and often cleaved off to generate the mature protein. • This domain annotated in Uniport as calcium ion binding domain.
Exon 3	64502_64503insG* F75Ifs*19	2.5%	75F>I	Disease causing/ 0.050(Benign)/Disease (RI-5)	• Stop codon is generated at the 94 position. • Sequence of this peptide is important because it is recognized by other proteins and often cleaved off to generate the mature protein. • This mutation might disturb recognition of signal peptide.
Exon 4	71240C>A*	20	143P>H	Disease causing / 0.999(Probably damaging)/ Neutral(RI-5)	• Hydrophobic interactions, either in the core or on the surface, will be lost. • Prolines are known to have a very rigid structure, sometimes forcing the backbone in a special conformation. Possibly this mutation changes a proline with such a function into another residue, thereby disturbing the local structure.
Exon 4	71239_71240insC * N144Qfs*24	8.75	144N>Q	Disease causing/ 0.000(Benign)/Neutral(RI-5)	• Stop codon is generated at the 168 position. • The new residue that is introduced in the signal peptide differ in the properties and it is possible this mutation disturbs recognition of signal peptide. • The wild-type and mutant amino acids differ in size. • The mutant residue is bigger, this might lead to bumps. • Calcium ion binding site may be disturb for this mutation
Exon 4	71261G>A*	13.75	150R>K	Disease causing / 0.015(Benign)/ Neutral(RI-4)	• The new residue that is introduced in the signal peptide differ in the properties and it is possible this mutation disturbs recognition of signal peptide.

					• The mutant residue is smaller, this might lead to loss of interactions.
Exon 4	71268G>T*	1.25	152Q>H	Disease causing / 0.027(Benign)/ Disease (RI-1)	• The mutation is located within the single peptide. The sequence of this peptide is important because it is recognized by other proteins and often cleaved off to generate the mature protein. • The mutant residue is smaller, this might lead to loss of interactions.
Exon 5	71311G>A*	3.75	167E>K	Disease causing/ 0.859(Probably damaging)/ Neutral(RI-4)	• Function of cadherin 1 binding Domain might be lost.
Exon 5	71356G>A*	2.5	Synonymous	Polymorphisms	• Function of cadherin 1 binding Domain might be lost.
Exon 5	71378A>G	2.5	Synonymous	Polymorphisms	• Splicing acceptor increased at g.71374 wt: 0.33 / mu: 0.51 (marginal change - not scored) • Extracellular TOPO domain function might be lost due to splicing error.
Exon 5	71493G>T*	2.5	186G>C	Polymorphisms / 0.412 (Benign)/ Disease (RI-2)	• Function of cadherin 1 binding Domain might be lost.
Exon 7	74485_74486insA*[a] T287Nfs*6	10	287T>N	Disease causing / 0.994 (probably damaging)/ Disease (RI-7)	• Frameshift mutation generates stop codon at 293. • Function of Extracellular binding TOPO_DOM might be lost. • Function of cadherin 2 binding Domain might be lost.
Exon 7	74486A>C*	2.5	287T>P	Disease causing / 1.000 (probably damaging)/ Disease (RI-9)	• Function of cadherin 2 binding Domain might be lost. • Function of Extracellular binding TOPO_DOM might be lost.
Exon 7	74487C>A	2.5	287T>K	Disease causing / 1.000 (probably damaging)/ Disease (RI-8)	• Function of cadherin 1 binding Domain might be lost.
Exon 8	74922_74923insG*[a] Y341Vfs*9	1.25	341Y>V	Disease causing / 1.000 (probably damaging)/ Disease (RI-8)	• Stop codon is generated at the 350 position. • Frameshift mutation generates stop codon at 350. • Function of cadherin 2, cadherin 3 and cadherin 4 binding Domain might be lost.
Exon 8	75039G>A	10	Synonymous	Disease causing	• Splicing Donor lost at g.75038 positionwhich was leading to sequence motif lost.

					• Function of Extracellular binding TOPO domain might be lost.
Exon 9	76144T>G*	26.25	398V>G	Disease causing / 0.999 (probably damaging)/ Disease (RI-7)	• Function of cadherin 3, cadherin 4 and cadherin 5 binding Domain might be lost.
Exon 9	76145_76146insGA* T399Efs*19	2.5	399T>E	Disease causing /0.999 (probably damaging)/ Disease (RI-8)	• Frameshift mutation generates stop codon at 418. • Function of Extracellular binding TOPO domain might be lost. • Function of cadherin 3, cadherin 4 and cadherin 5 binding Domain might be lost.
Exon 9	76175_76176insT* W409Lfs*10	2.5	409W>L	Disease causing / 1.000 (probably damaging)/ Disease (RI-8)	• Frameshift mutation generatesg stop codon at 419. • Function of Extracellular binding TOPO domain might be lost.
Exon10	78348G>C*	1.75	460V>L	Polymorphisms / 0.005 (Benign)/ Neutral (RI-4)	• Function of cadherin 3 binding Domain might be lost.
Exon10	78364C>G*	2.5	465S>C	Polymorphisms / 0.421 (Benign)/ Neutral (RI-1)	• Function of Extracellular binding TOPO domain might be lost. • Function of cadherin 3 and cadherin 4 binding Domain might be lost.
Exon11	82149T>G*	2.5	553F>L	Polymorphisms / 0.000 (Benign)/ Neutral (RI-4)	• Function of Extracellular binding TOPO domain might be lost. • Function of cadherin 4 binding Domain might be lost
Exon12	84874T>G*	1.25	603C>W	Disease causing / 1.000 (probably damaging)/ Disease (RI-7)	• Function of Extracellular binding TOPO domain might be lost. • Function of cadherin 5 binding Domain might be lost.
Exon14	90976C>T*	2.5	731L>F	Polymorphisms / 0.760 (probably damaging)/ Neutral (RI-6)	• Function of cytoplasmic binding TOPO domain might be lost.
Exon15	92458G>C*	2.5	775G>A	Disease causing / 1.000 (probably damaging)/ Neutral (RI-1)	• Function of cytoplasmic binding TOPO domain might be lost.

Table 6.3. Mutations of CDH1 geneand their pathogenicity in gastric tumor samples.

'*' - representing novel mutations; 'a' - representing heterozygous mutations.

frameshift mediated RNA decay and/or result in truncated gene products (Figure 6.1, Figure 6.2). One missense mutation in exon 7 occurred in the region encoding 1 of the 5 extracellular domains, containing Ca2-binding motifs. Multiple alignment of E-cadherin from 4 species (human, mouse, rat and xenopus) showed that Threonine at codon 287 is invariant in these species. This change would be predicted to influence the sensitivity of E-cadherin-mediated cell adhesion inresponse to changes in extracellular Ca^{2++} concentration. All mutations were absent in the blood samples of the patients (carrying tumors with the mutation), indicating their somatic origin. Among 19 cases with primary gastric tumors, 14 exhibited germline CDH1 alterations (all the 14 mutations was found in tumor, adjacent normal and blood samples).

2. Discussion

Diffuse gastric carcinoma is featured by a low CDH1 expression, dependent on genetic and epigenetic mechanisms (Muta et al., 1996). In the case of genetic lesions, around 35% of patients display mutations in the central protein region coding for the extracellular domain of CDH1 and in particular for the Ca^{++} binding site (Becker et al., 1995). It was previously demonstrated that such mutations are sufficient to disrupt CDH1 function by hampering the correct protein folding and orientation (Ozawa et al., 1990). With respect to the DNA mutations reported in our study, mutations in exons 8 and 9 were found in the extracellular domain of CDH1. Both mutations in exon 8 (74922_74923insG and 75039G>A) are sited in the Ca^{++} binding site of EC-3 domain, whereas mutations at exon 9 (76145_76146insGA and 76175_76176insT) are always in the EC-3 domain, but away from the Ca^{++} binding site. To our knowledge,these mutations have not been previously reported and possible functional consequences resulting from such mutations are unknown. None of the mutation was found for exons 1, 2, 6, 13 and 16 in Mizo population. Exon 11, 12, 14 and 15 exhibited single disease causing missense mutations, each for

2.5% diffuse type gastric tumor samples (Table 6.3). This sequence encodes for the cytoplasmatic tail of the protein where interaction CDH1/-catenin occurs (Nose et al., 1990). At this level, there is a cluster of 8 serine whose function is modulated by phosphorylation with the consequent regulation of CDH1 adhesivity(Stappert et al., 1994) and mutations in this region is likely to affect the CDH-1-dependent signal transduction pathway. Two missense mutation found at exon 10 (78348G>C and 78364C>G) maps in the same CP-2 domain, interacting with catenins, but also such a mutation was never previously described. However, mutations found in tumor specimens were absent in germ-line cells from the 21.25% patients, thereby indicating that CDH-1 mutations occurred during the process of tumorigenesis.

The cell adhesion molecule E-cadherin plays a key role in the establishment and maintenance of epithelial tissue structure and its down-regulation is potentially important in the formation of metastases from carcinomas. To examine how the expression of E-cadherin changes during the development and progression of gastric carcinomas, 80 primary tumors were immunohistochemically analyzed. The E-cadherin expression patterns were correlated with a number of histopathological parameters and compared with that observed in adjacent normal mucosae and solid metastasesof different localizations. Comparison with adjacent normal mucosae, the tumor revealed a reduction of E-cadherin expression in 92% of the primary gastric carcinomas. In both primary and metastatic tumors, low expression of E-cadherin (<30% positive cells) and dispersion of the molecule over the entire cell surface significantly correlated with cellular de-differentiation and glandular disintegration, terminating in the complete loss of E-cadherin in undifferentiated carcinomas. These findings are consistent with the observations of Shiozaki et al. (1991). Although Shimoyama and Hirohashi (1991) could not find any correlation between E-cadherin expression and differentiation of primary gastric carcinomas, such a correlation has now been reported for a wide variety of human carcinomas (Schipper et al., 1991; Gabbertet al., 1991).

Since these events are fundamental to the establishment of epithelial structures it would appear that E-cadherin itself induces cellular differentiation, and therefore, its down-regulation leads to the disintegration of glandular structures. The reduced expression of E-cadherin and its dispersion over the entire cell surface, possibly associated with the disarrangement of the cytoskeletal network are suggested to free the cells from a variety of contact-mediated controls and to encourage the successful escape of tumor cells from the primary tumor. This speculation is supported by in vitro experiments showing that the loss of E-cadherin is correlated with dedifferentiation and invasive behavior of epithelial cell lines (Vleminckx et al., 1991). As well as by the correlation between reduced E-cadherin expression and tumor recurrence and mortality demonstrated in the present study. Most of the gastric carcinoma metastatic lesions were found to express E-cadherin. This is not necessarily in opposition to a role for downregulation of this molecule in metastasis formation since minor changes in E-cadherin expression and/or distribution have been shown to be sufficient to cause a loss of cell adhesion. We have found most of the patients with large size tumor and total or subtotal gastrectomy were significantly associated with CDH1 mutations with low or negative expression in tumor samples.

Taking into consideration our findings, it appears that germ-line CDH1 cannot independently explain diffuse type gastric cancer and that other genetic alterations are involved in this disease. In fact, somatic mutations, enhanced by gene silencing, underlie the CDH-1 down-regulation in eightof seventeen patients. Remarkably, in four cases CDH1 was down-regulated in cancer specimens, without any evidence of either CDH1 mutations. Therefore, other genetic unknown alterations able to induce CDH1 down-regulation could be present in the gastric cancer patients. This finding is well in keeping with recent data indicating a mixed prevalence of hereditary and sporadic CDH1 genetic alterations in gastric

cancer and prompts further research in this field to isolate unknown genes involved in CDH1 down-regulation.

Chapter VII

Xenobiotic pathway gene polymorphisms associated with gastric cancer in high risk Mizo-Mongoloid population, Northeast India

1. Introduction

Stomach cancer is the second leading cause of death and the fourth-most common malignancy in the modern world and highly associated with dietary and lifestyle habits (Murray and Lopez, 1997; Ferlay et al.,2010). Considerable differences in the incidence of stomach cancer were found worldwide and the highest incidence among men has been reported from China (AAR=145.0) and among women from Japan (AAR=38.9) per 10^5 population (Ferlay et al., 2010). It is generally accepted that cancer risk is influenced by the interaction between genetic components and environmental factors, such as diet and lifestyle. Previous studies on risk factors for stomach cancer have shown a positive correlation between *GST* polymorphism and consumption of alcohol, tobacco smoking, unique food habits and pathogen infection (Parkin et al., 2002; Correa, 2002]. Thus, the studies of genes involved in the detoxification of environmental carcinogens are important for the determination of inter-individual susceptibility to cancer (Phukan et al., 2006).

Stomach cancer is a multifactorial disease and arises from the interactive combinations of allelic variants (in low-penetrance genes) and relevant environmental risk factors. Genetic polymorphism in xenobiotic metabolizing genes may act as susceptibility factors for stomach cancer risk. Glutathione-S-transferases (*GST*s) constitute a superfamily of phase-II metabolizing enzymes that play a key function in detoxification of a wide variety of potentially cytotoxic/ genotoxic compounds (Hayes and Strange, 2000). The *GST* isoenzymes (*GSTM-1*, *GSTT-1*, and *GSTP-1*) are normally expressed along the human gastrointestinal tract (Hayes and Pulford, 1995). The phenotypic absence of *GSTM1* and *GSTT1* activity

is due to homozygosityfor an inherited deletion of these genes, termed as null genotype (Lan et al., 2001). Individuals with the homozygous *GSTM1* null genotype show no protein expression and are expected to have reduced detoxification ability for hazardous compounds (Pemble et al., 1994). Furthermore, single-nucleotide polymorphisms (SNPs) in the *GSTP1* gene at codons 105 (Ile→Val) and 114 (Val→Ala) were found to be associated with reduced GST enzyme activity for several classes of substrates [Cai et al., 2001; Zimniak et al., 1994; Tahara et al., 2011].

However, the expression of *GSTP1* influenced by polymorphism and gene - gene and gene - environment interactions among tobacco-associated gastric cancer patients is not well understood. An individual's difference in susceptibility to chemically induced carcinomas may possibly be attributed to the variations in the detoxification pathways due to polymorphic variants of GST genes (DeBruin et al., 2000). In European population, the *GSTM1* deletion frequency was found to be higher than *GSTT1*(DeBruin et al., 2000), and a significant association was found between the *GSTT1* null genotype and GC risk with all other *GST* isoenzymes polymorphism (Garte et al., 2001). According to previous studies, when stratified by ethnicity a significant association between *GSTT1* null genotype and GC riskwas also found in Chinese and other Asian populations, but not in Caucasians (Garte et al., 2001).

Mizoram, Northeast India records the highest incidence of gastric cancer in India (Mahanta et al., 1998). Consumption of tobacco in various traditional forms such as "tuibur" (tobacco smoke–infused aqueous solution) and Chaini (traditional name 'Sada') are common habits which are important risk factors for stomach cancer (Mahanta et al., 1998; Russo et al., 2001). Therefore, correlation of consumption of tobacco with the incidence of stomach cancer in Mizoram cannot be ruled out. We performed a case–control study to evaluate the relevance of *GSTM1*, *GSTT1* and *GSTP1* gene

polymorphisms to Gastric cancer (GC) susceptibility and to assess their interaction with other environmental and lifestyle factors, namely *H. pylori*, *EBV* infection and smoking habits for this high GC risk state of Mizoram, Northeast India.

2. Materials and Methods

2.1. Subjects

High incidence of stomach cancer is found in the Mizo-Mongloid population, Northeast India with age-adjusted rates (AAR) of 42.9 and 20.5 per 10^5 population in male and female, respectively (NCRP, 2010). The present study consisted of 80 random gastric cancer patients (aged from 37 to 79 years) with or without a family history of any cancer and 80 unrelated healthy controls (aged from 31 to 73 years) belonging to the same ethnic group. The samples were collected from patients who received treatment from Mizoram State Cancer Institute, Zemabawk, Mizoram, India between September 2013 and February 2015. The study design and data collection methods have been described previously. Patients with gastric neoplasms including adenocarcinoma (MALT lymphoma, stromal or carcinoid tumors). Patients with secondary or recurrent GC, previous history of other malignancies or refusal to participate were excluded from the study (Ghatak et al., 2014). Medical charts were reviewed using a standard protocol in order to obtain information on cancer treatment, clinical stage, dietary habits, previous disease history, physical activity, tobacco and alcohol use by an in-person interview using a structured questionnaire (Ghatak et al., 2014). The study protocol was approved by the Institutional Review Board of all institutes involved in the study.

2.2. Nucleic Acid extraction

Blood samples (50 µl) were processed freshly and served as subjects for DNA isolation (Ghatak et al., 2013). Haematoxylin and eosin-stained slides were prepared after microdissection from paraffin

embedded tumor tissue for conforming the adenocarcinoma, whereas the blood samples were used directly. Co-extraction of RNA and DNA was performed for the paraffin embedded tumor tissue (Ghatak et al., 2015).

2.3. Genotyping of GST M1/ T1/ P1 gene polymorphisms

GSTM1 and *GSTT1* null polymorphism were detected by multiplex polymerase chain reaction (PCR) (Table 7.1) (Arand et al., 1996). The reaction mixture (25 µl) contained 50-100 ng of genomic DNA in 1 x Taq buffer, 200 µM of each dNTP, 0.15 µM of each primer and 1 U of Taq DNA polymerase. Amplified products were analyzed by electrophoresis on 8% polyacrylamide gels resulting in a 219-bp fragment for *GSTM1*, 459-bp fragment for *GSTT1* and a 349-bp fragment of the albumin gene (as an internal control). The absence of the specific *GSTM1* and/or *GSTT1* fragments indicated the corresponding null genotype, whereas the presence of the 349 bp albumin band ensured that the null genotype was not documented due to failure of PCR. In addition, two SNPs in the *GSTP1* gene for amino acid substitution at codons 105 (Ile→Val) and 114 (Val→Ala) were genotyped by PCR-RFLP method (Table 7.1) (Zimniak et al., 1994; Ali-Osman et al., 1997; García-G et al., 2012). The fragment containing the *GSTP1* Ile105Val (rs1695) polymorphic site was amplified according to the following parameters: 94°C for 5 min., followed by 35 cycles at 94°C for 30 sec., 58°C for 30 sec., 72°C for 1 min. and a final elongation at 72°C for 7 min.. Ten µl of PCR product was digested with 1 unit of *BsmAI* (New England Biolabs, Barcelona, Spain) for 6 h at 55°C. Digests were electrophoresed on 8% polyacrylamide gels resulting in three fragments of 305, 135 and 128 bp (allele A) or in four fragments of 222, 135, 128 and 83 bp (allele G). PCR cycling conditions were 5 min. at 94°C followed by 35 cycles at 94°C for 30 sec., 52°C for 30 sec. and 72°C for 30 sec. with a final elongation step at 72°C for 7 min. for *GSTP1* Val114Ala (rs1138272) polymorphism. Ten µl of PCR products were digested with 1 unit of *AciI* (New England Biolabs, Barcelona, Spain) for 6 h at 37°C and electrophoresed on 8% polyacrylamide gel. The

T allele was defined by the presence of an intact fragment of 170 bp and the C allele by the presence of two fragments of 143 and 27 bp.

Gene	**Primer (5' to 3')**	**Annealing temperature**	**Type of polymorphism**	**Amplified products/Alleles**
GSTM1	F: GAACTCCCTGAAAAGCTAAAGC R: GTTGGGCTCAAATATACGGTGG	60°C – 1 min	Gene deletion	− Null genotype + 219 bp
GSTT1	F: TTCCTTACTGGTCCTCACATCTC R: TCACCGGATCATGGCCAGCA		Gene deletion	− Null genotype + 459 bp
Albumin	F: GCCCTCTGCTAACAAGTCCTAC R: CCCTAAAAAGAAAATCGCCAATC		-	Internal control 349 bp
GSTP1 I105V (rs1695)	F: AATACCATCCTGCGTCACCT R: TGAGGGCACAAGAAGCCCCTT	59°C -40 sec	BsmAI RFLP	A - Digested fragment (308, 258 bp) G - Digested fragment (258, 219 and 89 bp)
GSTP1V114A (rs1138272)	F: ACAGGATTTGGTACTAGCCT R: AGTGCCTTCACATAGTCATCCTTG	50°C -40 sec	AciI RFLP	C - Digested fragment (143, 27 bp) T - Intact fragment (170 bp)
H. pylori (16srRNA)	F: CTGGAGAGACTAAGCCCTCC R: ATTACTGACGCTGATTGTGC	60°C-45 sec	Multiplex PCR	109 bp
H. pylori (CagA)	F: AATACACCAACGCCTCCAAG R: TTGTTGCCGCTTTTGCTCTC			400 bp
EBNA3C	F: AGAAGGGGAGCGTGTGTTGT R: GGCTCGTTTTTGACGTCGGC	56°C -45 sec		153 bp for type 1 EBV 246 bp for type 2 EBV
GSTP1 For expression	F: ATGACTATGTGAAGGCACTG R: AGGTTCACGTACTCAGGGGA	55°C-45 sec	Multiplex PCR	433bp band (Intensity less – less expiration; Intensity more – optimum expiration)
β-Actin Control	F: ACCATGGATGATGATATCGC R: ACATGGCTGGGGTGTTGAAG			392 bp band Band Intensity should be equal

Table 7.1. Primers tested for polymorphism pattern in the present study

2.4. PCR based Pathogen detection

The presence of *H. pylori* infection was determined in GC patients by multiplex PCR amplification of *16S rRNA* and *CagA* genes by using Hp1-Hp2 and CagAF-CagAR primers, respectively (Table 7.1) (Ghatak et al., 2015; Ho et al., 1991). The PCR products were electrophoresed with 1.5 % agarose gel. *H. pylori* infection was defined by the presence of an intact band of 109 bp (16S rRNA) and 400 bp (*CagA* gene). The presence of *EBV* type1/ type 2 infections was carried out by a standard PCR assay crossways type-specific regions of EBNA3C gene using previously described primer sets (Table 7.1) (Kingma et al., 1997). The PCR amplification was done for 35 cycles at 95°C for 45 sec., 56°C for 45 sec., 72°C for 1 min. followed by a final extension at 72°C for 10 min.. The PCR products were analysed by 12% polyacrylamide gel and the presence of 153 and 246 bp was considered as type 1 and type 2 EBV, respectively.

2.5. Reverse transcriptase PCR (RT-PCR) for *GSTP1* expression

First-strand synthesis of cDNA was performed using Superscript II reverse-transcriptase (Life Technologies Inc., USA) (Table 7.1). Five µg aliquot of total cellular RNA was used for each reverse transcription reaction and one-tenth of this reaction was used for PCR. Primers targeting exons 6 and 7 of *GSTP1* and exons 1 and 3 of β-actin were used (DeBruin et al., 2000). The PCR for *GSTP1* was run for 38 cycles consisting of 94°C for 30 sec., 1 min at gradually decreasing temperatures using a touch-down

protocol (4 cycles each at 62°C and 60°C; and 30 cycles at 58°C), and 74°C for 5 min.. RT-PCR products were electrophoresed in 10% polyacrylamide gel and visualized by ethidium bromide staining.

2.6. Densitometric analysis

Semiquantitative RT-PCR gels were analyzed to compare the banding pattern and their molecular mass using gel image analysis software-Syngen G-Box (Sacramento, CA, USA) and ImageJ. The *GSTP1*expression was compared with *GSTP1* genotyping by HEMI 1.0.1-Heatmap Illustrator software.

2.7. Statistical analysis

GST gene polymorphism among case-control subjects were tested for Hardy–Weinberg equilibrium by a chi-square ($\chi 2$) test with one degree of freedom (df). The polymorphisms in each group were estimated using odds ratios (ORs) and 95% confidence intervals (CIs) for each genotyping. *GSTT1* and *GSTM1* null genotyping were checked by the presence and absence of the bands (*GSTT1*-null and *GSTM1*-null genotype, respectively), whereas *GSTP1* rs1695 and rs1138272 polymorphisms were assessed using co-dominant model. Additionally, logistic regression analyses were conducted to compute the potential confounder's influence of both genetic and environmental factors for gastric cancer. Then, the independent effect of risk factors was investigated in a multivariate model (introducing all variables and terms of interactions) keeping only those statistically significant or showing a confounding effect on the studied factors. The likelihood ratio test was used to decide whether to retain each covariate in the model. Sex, gender, Cigarette smoking, family history of gastric cancer, tuibur consumption, alcohol intake, smoked meat/vegetable intake, high consumption of salt and pickle, high consumption of fermented fat,*H. pylori* status and genotyping, *EBV* status and genotyping were all considered in the regression model as potential confounders to evaluate the association of risk factors and susceptibility to gastric cancer.For all tests, a two-sided P-value <0.05 was considered statistically significant. All statistical

analyses were performed using SPSS 20.0 program (SPSS Ibérica, Madrid, Spain) and SYSTAT 13.0. (Systat Software Inc., USA). Heat map analysis was done by Hemi IBP software.

2.8. Multifactor dimensionality reduction (MDR) analysis

MDR analysis (MDR 3.0.2) was performed to assess the risk of gastric cancer by studying the gene-gene and gene-environment effect on the basis of GST genotyping and *GSTP1* gene expression. It is a non-parametric, model-free statistical approach which is used to generate one-dimensional model to predict gastric cancer susceptibility. The test is used to overcome the limitations encountered by parametric methods like logistic regression (LR) due to small sample size. Data was generated using 10-fold cross-validation procedure and 10 times random seed number to reduce the chance of false positives. The best model was selected on the basis of maximum cross validation consistency (CVC) and testing balance accuracy (TBA). All MDR results with $p \leq 0.05$ were considered to be significant (Hahn et al., 2006).

2.9. Interaction entropy graph

Interaction entropy graph were constructed using MDR software package (MDR 3.0.2). The graphs comprised of a node for each variable with pairwise linkage between them. The percentage of entropy removed by each variable is visualized for each node. The entropy estimates calculated were used to interpret the independent effect of each variable in comparison to the interaction effect of the variables and to determine if the interactions were synergistic (Manuguerra et al., 2007). The potential confounding factors, used in the multivariate model, were considered for interaction entropy estimation.

2.10. Homology modelling of GSTP1 Ile105Val (rs1695) and GSTP1 Val114Ala (rs1138272)

Healthy control and mutant GSTP1 protein sequences were submitted to the SWISS MODEL (http://swissmodel. expasy.org/) and Phyre2 server for the prediction of three-dimensional structure (Arnold et al., 2006; Kiefer et al., 2009; Kelley and Sternberg, 2009). Phyre2 uses the hidden Markov method to generate alignments of a submitted protein sequence against proteins with published structures. The resulting alignments were used to predict homology-based three-dimensional structure. The model was judged to be accurate when over 90 % of the submitted residues were modelled at greater than 90 % confidence.

3. Results

3.1. Clinical and demographic characteristics

Gastric cancer was more prevalent in males (70%) among the Mizo-Mongloid population (Table 4.3). Infection with *H. pylori* and *CagA*+ strains were more frequent in GC patients than in controls (OR: 1.168; 95% CI: 0.365 – 3.733; P = 0.094 and OR: 8.298; 95% CI: 10.365 – 6.578; P<0.005, respectively). However, no differences in the prevalence of *Epstein–Barr virus* (*EBV*) strains were observed between patients and controls (22.5 vs. 7.5%, respectively). Among the various demographic factors in the bivariate analysis, extra intake of fermented fat (OR: 3.387; 95% CI: 1.902 – 6.031; P<0.0003) and smoked meat consumption (OR: 6.181; 95% CI: 2.528 – 12.346; P<0.001) were also found to be risk factors for GC (Table 4.3). No significance was found with consumption of alcohol and familial history with gastric cancer risk.

3.2. Genotyping

Significant differences in genotype distribution were observed between patients and control samples (Table 7.2). The frequencies of *GSTM1* (67.5 vs. 35%) and *GSTT1* (47.5 vs. 12.5%) null genotypes in

GC patients was significantly high from those observed in controls (Figure7.1A). Moreover, the simultaneous presence of both the *GSTM1* and *GSTT1* null genotypes was significantly different between patients and control samples (30% in GC vs. 7.5% in HC; OR: 2.29, 95% CI: 1.36 –9.53). Significant differences in genotype or allele frequencies of the *GSTP1* IIe105Val and *GSTP1* Ala114Val polymorphisms were also observed. Presence of heterozygous allele A/G in rs1695 was 32.5 % in GC and 20 % in HC (OR: 1.93; 95% CI: 0.70 – 5.33) and homozygous allele G/G in rs1995 was 27.5 % in GC and 5 % in HC (OR: 7.21; 95% CI: 1.48 – 13.07). In case of rs1138272, C/T heterozygous was 25 % in GC and 15% in HC (OR: 1.88; 95% CI: 0.62 – 5.81) and T/T homozygous allele was 5 % in GC and absent in HC (Table 7.2). We also examined

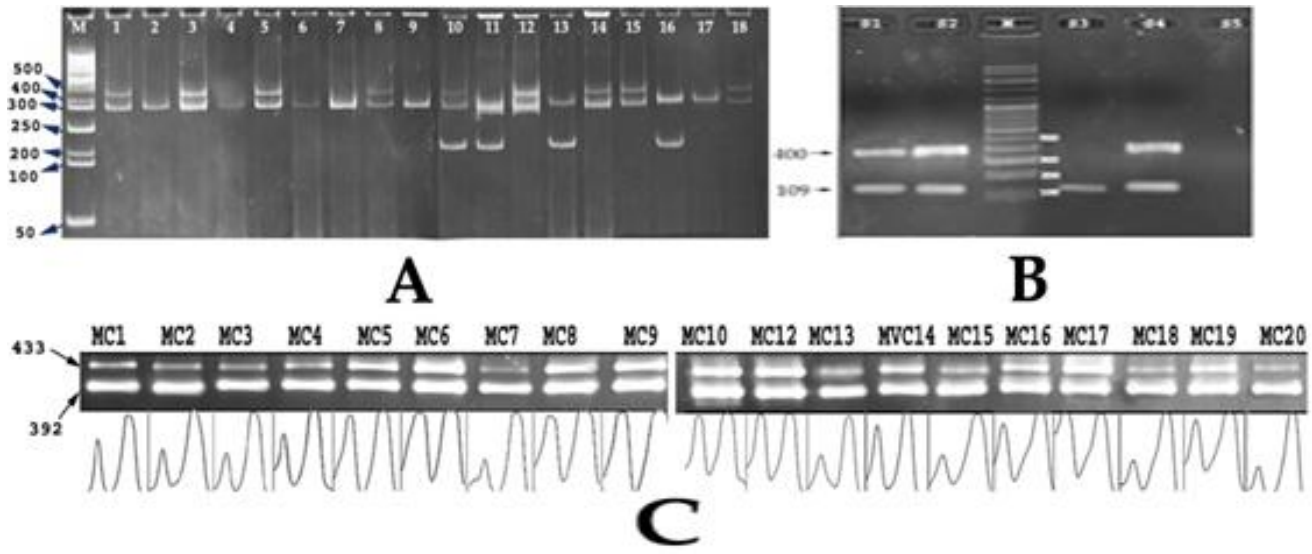


Figure7.1.(A)Multiplex PCR amplified products of *GSTT1* (459 bp), *GSTM1* (219 bp) and internal control gene albumin (349 bp); (B) *16SrRNA* (109 bp) and*CagA*(400 bp) region of *H. pylori* amplified by Multiplex PCR; (C) *GSTP1* (433 bp) expression in 10% polyacrylamide gel, β-Actin (392 bp) as control. M – Low range ruler plus (100 – 3000 bp); (A) 1 – 9 Gastric cancer Sample; 10 – 18 Healthy control Samples; (C - Digested fragment); S4- *AciI* RFLP (T - Intact fragment); (B) S2 to S4 - *H. pylori* positive

samples, S1 and S5 - positive and negative control, respectively; S3 – *H. pylori* positive but CagA negative.; (C) MC1 to MC20 – gastric cancer sample.

Genotype	**Total HC[a] (n=80)**	**Total GC[b] patients (n=80)**	**[c]OR**	**[d]95% CI**
GSTM1				
(+)	52 (65%)	26 (32.5%)		
(-)	28 (35%)	54 (67.5%)	0.208	1.53 – 9.75
P value = 0.973				
GSTT1				
(+)	70 (87.5%)	42 (52.5%)		
(-)	10 (12.5%)	38 (47.5%)	0.389	0.052– 2.924
P value = 1.05				
GSTM1/T1				
(+/+)	44 (55%)	12 (15 %)	0.043	0.008– 0.249
(−/+)	6 (7.5 %)	14 (17.5 %)	0.389	0.052– 2.924
(+/−)	24 (30 %)	30 (37.5%)	0.208	0.039– 1.116
(−/−)	6 (7.5 %)	24 (30 %)	2.29	1.36 – 9.53
P value = 0.049				
GSTP1 105 (rs1695)				
A/A	60 (75%)	28 (35%)	0.22	0.09 – 0.58
A/G	16 (20%)	32 (32.5%)	1.93	0.70 – 5.33
G/G	4 (5%)	20 (25%)	7.21	1.48 - 13.07
P value = 0.057				
GSTP1 114 (rs1138272)				
C/C	68 (85%)	44 (55%)	0.22	0.74 – 0.63
C/T	12 (15%)	20 (25%)	1.88	0.62 – 5.81

T/T	0 (0%)	16 (5%)	NA	NA
		P value = 0.261		

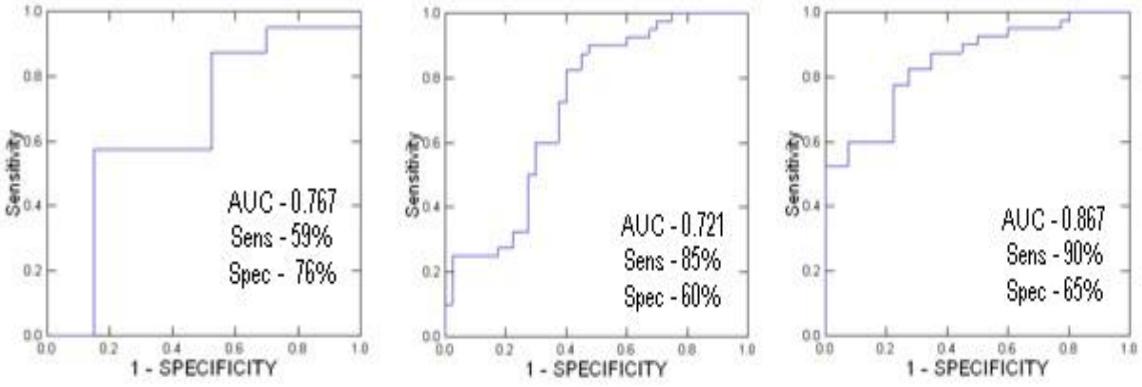
[a]HC – Healthy Control
[b]GC – Gastric Cancer
[c]OR – Odd ratio
[d]95% CI – 95 % Confidence Interval
P value (< 0.05) indicates significant association of the genotype with gastric cancer

Table 7.2. Genotype pattern for GST gene polymorphism in the samples

whether plausible contribution of GST polymorphisms to GC risk could be modified by interaction with other risk factors (smoking habit, *H. Pylori* and *EBV* infection, and family history of GC) identified in our study population (Tables 7.2 and 7.3). The exposure variables were: smoking habit (current smokers vs. non-smokers), *H. pylori* infection (positive vs. negative), *CagA*amplification status (positive vs. negative) and family history of GC (positive vs. negative; positive history was defined as any reported GC in first-degree relatives or two or more GC cases in second-degree relatives). Smoking status or family history of GC did not modify the association between GST gene polymorphisms and GC risk (Table 4.3). *H. pylori* and *CagA*positive status was significantly associated with the GST polymorphism. Moreover, tests for interaction under a multiplicative model showed strong statistically significant interaction between GST genotypes and the risk factors.

In summary, logistic regression analysis identified *H. pylori* infection (OR: 1.168; 95% CI: 0.365 – 3.733; P = 0.094) with *CagA*+ strains (OR: 8.298; 95% CI: 0.365 –36.578; P < 0.005), smoking habit (OR: 0.595; 95% CI: 0.373 – 1.947; P =0.029) as independent risk factors for the development of GC (Figure7.1B, Table 4.3). Concerning the genetic factors, *GSTT1* and *GSTM1/T1* Null genotype is strongly related with the GC phenotype with OR: 0.389; 95% CI: 0.052 – 2.924, and OR: 2.29; 95% CI: 1.36 –

9.53, respectively. *GSTP1*105 was significantly related to GC phenotype (Table 7.2). Significant association of GST polymorphisms and *H. pylori* infection with *CagA*was observed. Presence of *GSTT1* null alleles in GC patients was significantly associated with the presence of *H. pylori* and *CagA*(OR: 3.00; 95% CI: 0.85 – 10.54 and OR: 6.65; 99% CI: 1.98 – 2.27). Presence of *GSTM1* null alleles in GC patients was also significantly associated with presence of *CagA* genotype (OR: 5.93 95% CI: 1.53 – 2.90). *GSTM1/ T1* null genotyping was significantly associated with the presence of *H. pylori* (OR: 4.67 95% CI: 1.19 – 8.33) and *CagA* (OR: 3.08 95% CI: 0.75 – 2.61) (Table 7.2). Multiplex PCR analysis of *GSTM1/T1* and *GSTP1*



(A) (B) (C)

Figure7.2. Recevier operating Characteristic curve based on **(A)** multiplex analyses of GSTM1, GSTT1 and GSTP1 polymorphisms; **(B)** GST polymorphism associated with food habits; **(C)** GST polymorphism associated with *H. pylori.*

The predicted probability from the final logistic regression model are calculated to differentiate between healthy control and Mizoram Gastric cancer. AUC - area under the curve; Sens – Sensitivity; Spec – Specificity.

Polymorphism	Genotype	*H. pylori* Negative		*H. pylori* Positive		*CagA* status Negative		*CagA* status Positive	
		HC/GC	**OR (95% CI)**	**HC/GC**	**OR (95% CI)**	**HC/GC**	**OR (95% CI)**	**HC/GC**	**OR (95% CI)**
GSTM1	+	38/8	1 (Reference)	14/18	1.37 (0.45 - 4.12)	50/16	1 (Reference)	2/10	5.57 (0.62 – 0.03)
	−	10/16	0.179 (0.02 – 1.61)	18/38	3.11 (1.8 - 8.19)	22/28	0.33 (0.09 – 1.17)	6/26	5.93 (1.53 – 2.90)
		P value for interaction =0.125				P value for interaction = 0.087			
GSTT1	+	46/6	1 (Reference)	24/36	1.90 (0.76 – 4.78)	70/40	1 (Reference)	0/2	NA
	−	2/18	0.871 (0.31 – 2.44)	8/20	3.00 (0.85 –10.54)	4/2	0.894 (0.35 – 2.26)	8/34	6.65 (1.98 – 2.27)
		P value for interaction =0.793				P value for interaction =0.813			
GSTM1/T1	+/+	28/8	1 (Reference)	16/4	0.21 (0.04 – 1.16)	44/8	1 (Reference)	0/4	NA
	−/+	2/6	3.16 (0.31 – 3.77)	4/8	2.11 (0.36 - 2.45)	6/4	0.64 (0.10 – 4.11)	0/10	NA
	+/−	18/6	0.27 (0.06 – 1.12)	6/24	5.28 (1.36 – 0.53)	22/24	1.12 (0.42 – 2.97)	2/6	3.16 (0.31 – 1.77)
	−/−	0/2	NA	6/22	4.67 (1.19 – 8.33)	0/8	NA	6/16	3.08 (0.75 – 2.61)
		P value for interaction =0.046				P value for interaction =0.053			
GSTP1 105	A/A	46/8	1 (Reference)	14/24	2.02 (0.70 – 5.82)	58/24	1 (Reference)	2/8	4.33 (0.46 – 0.60)
rs1695	Carrier G	2/16	9.75 (1.15 – 8.10)	18/32	2.29 (0.87 – 6.08)	14/20	1.57 (0.53 – 4.65)	6/28	6.64 (1.73 – 5.46)
		P value for interaction =0.042				P value for interaction =0.035			
GSTP1 114	C/C	48/10	1 (Reference)	20/34	2.21 (0.85 – 5.74)	64/18	1 (Reference)	4/26	9.14 (1.90 – 3.89)
rs1138272	Carrier T	0/14	NA	12/22	2.14 (0.70 – 6.53)	8/26	4.33 (1.27 – 4.77)	4/10	2.71 (0.49 – 4.90)
		P value for interaction =0.485				P value for interaction =0.582			

Table 3. Mediation effect of *H. pylori* infection and their association with GST gene polymorphism

polymorphisms gave 0.767 area under the curve value (AUC) with a sensitivity of 59 % and specificity of 76 %, respectively (Figure 7.2A). In addition, validation of *GST* polymorphisms associated with food habits (AUC = 0.721, sensitivity- 85%, specificity - 60%) and *H. pylori* (AUC = 0.867, sensitivity - 90%, specificity - 65%) was found to be significant (Figure 7.2B, 7.2C).

3.3. MDR analysis

MDR analysis was used to further validate our logistic regression based study on gene-gene and gene-environment interaction by using a model free approach. A four order interaction model was chosen along with their cross validation consistency (CVC) and testing balance accuracy (TBA). The best model was selected based on the highest CVC and TBA value among all the models. Among the entire dataset, the presence of *H. pylori* with smoked meat consumption is the best one factor model that was found to be statistically significant ($p < 0.001$) with a CVC of 10/10 and testing accuracy of 0.54. The combination of *GSTT1* null genotype and *H. pylori* presence was found to be the best two factor model with CVC of 5/10 and TBA of 0.33 ($p<0.001$). The combination of *GSTT1* null genotype, *H. pylori* with *CagA* genotype and smoked meat consumption and smoking was found to the best three-factor model, which was also the best overall model with a CVC of 10/10 and TBA of 0.54 ($p<0.001$). Presence of G/G genotype (-0.141) of *GSTP1* rs1695 and presence of T/T genotype (-0.183) of *GSTP1* rs1138272 was significantly associated with less expression of *GSTP1* gene (Figure 7.3B).

3.4. Interaction entropy graph

The interaction entropy graph shows the presence of *CagA* genotyping of *H. pylori* (14.2 %) with smoked meat or vegetable and oral snuff as the large independent effectors in gastric cancer among the various

environmental factors (Figure 7.3A). The *GSTT1* null genotype (7.7 %) was found to contribute the highest

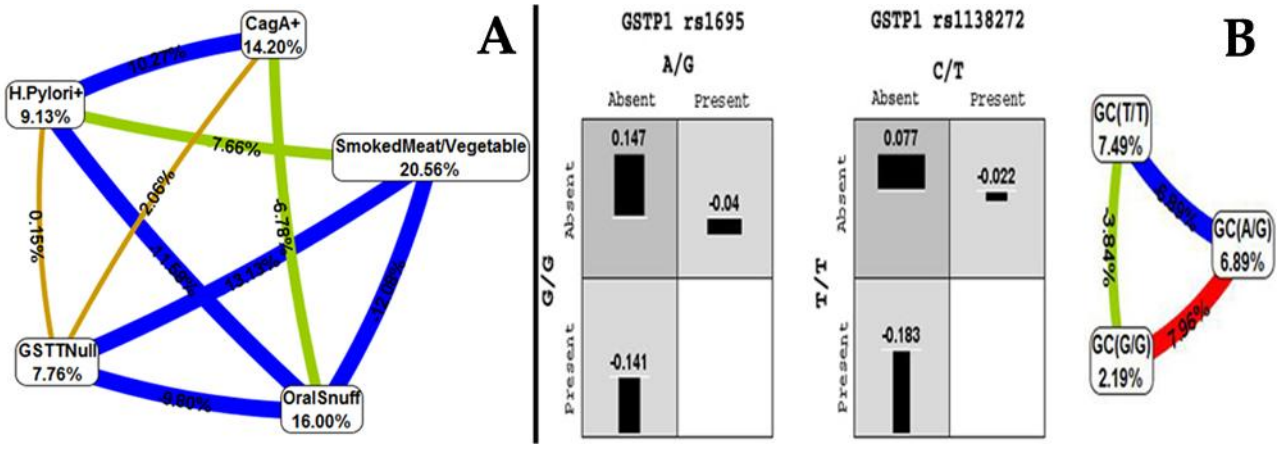


Figure 7.3.(A) Interaction entropy graph to find the gene-environment interaction with Gastric cancer risk. The percent of the entropy for independent factors as well as their interactions are represented in the graph. (B) Summary of the two SNPs GSTP1 rs1695: A>G (A/G and G/G) and GSTP1 rs1138272: C>T (C/T and T/T) relation with GSTP1 gene expression in MDR analysis. Black bar height shows GSTP1 expression.

Bar wide shows frequencies of sample. Positive percentage of entropy denotes synergistic interaction while negative percentage denotes redundancy. Here, the red color denotes the high degree of synergistic interaction, golden color denotes the moderate degree of synergistic interaction, green color represents moderate redundancy while blue color denotes the highest. GC – Gastric cancer.

independent effect among all the genetic factors, while a synergistic effect was found in the case of *GSTT1* gene with *H. pylori* and *CagA* genotyping by removing 0.15 and 2.06 % of entropy. Consumption of smoked meat was also found to contribute small independent entropy of 1.66 %. G/G and A/G genotype of *GSTP1* rs1695 showed high degree of synergistic interaction with the *GSTP1* gene expression (Figure 7.3B).

3.5. Comparative modelling of GSTP1 polymorphisms

The natural variants were substituted into the wild-type sequence for comparative modelling. The variance *GSTP1* Ile105Val (rs1695) and Val114Ala (rs1138272) indicates interruption of H-site substrate binding pocket for chemical binding and were found to affect the metabolising activity of the protein. Estimated protein folding by phyre2 and Swiss model servers also differed from normal and mutated *GSTP1*.

3.6. Densitometric analysis

Densitometric analysis of *GSTP1* expression in relation to *GSTP1* genotyping for healthy control and gastric cancer individuals was done by Image J and represented in Heat map by HEMI 1.0.1 software. Majority of the gastric cancer sample showed less *GSTP1* expression than healthy controls (Figure 7.1C, 7.4A, 7.4B). Samples (25%) exhibiting low expression of *GSTP1* showed homozygous mutant type G/G and T/T for rs1695 and rs1138272, respectively. Heterozygous A/G and C/T for *GSTP1* rs1695 and rs1138272, respectively had moderate expression. The A/A for *GSTP1* 105 (rs1695) and C/C for *GSTP1*

114 (rs1138272) wild type genotyping showed approximately similar expression like healthy controls (Figure 7.4A, 7.4B).

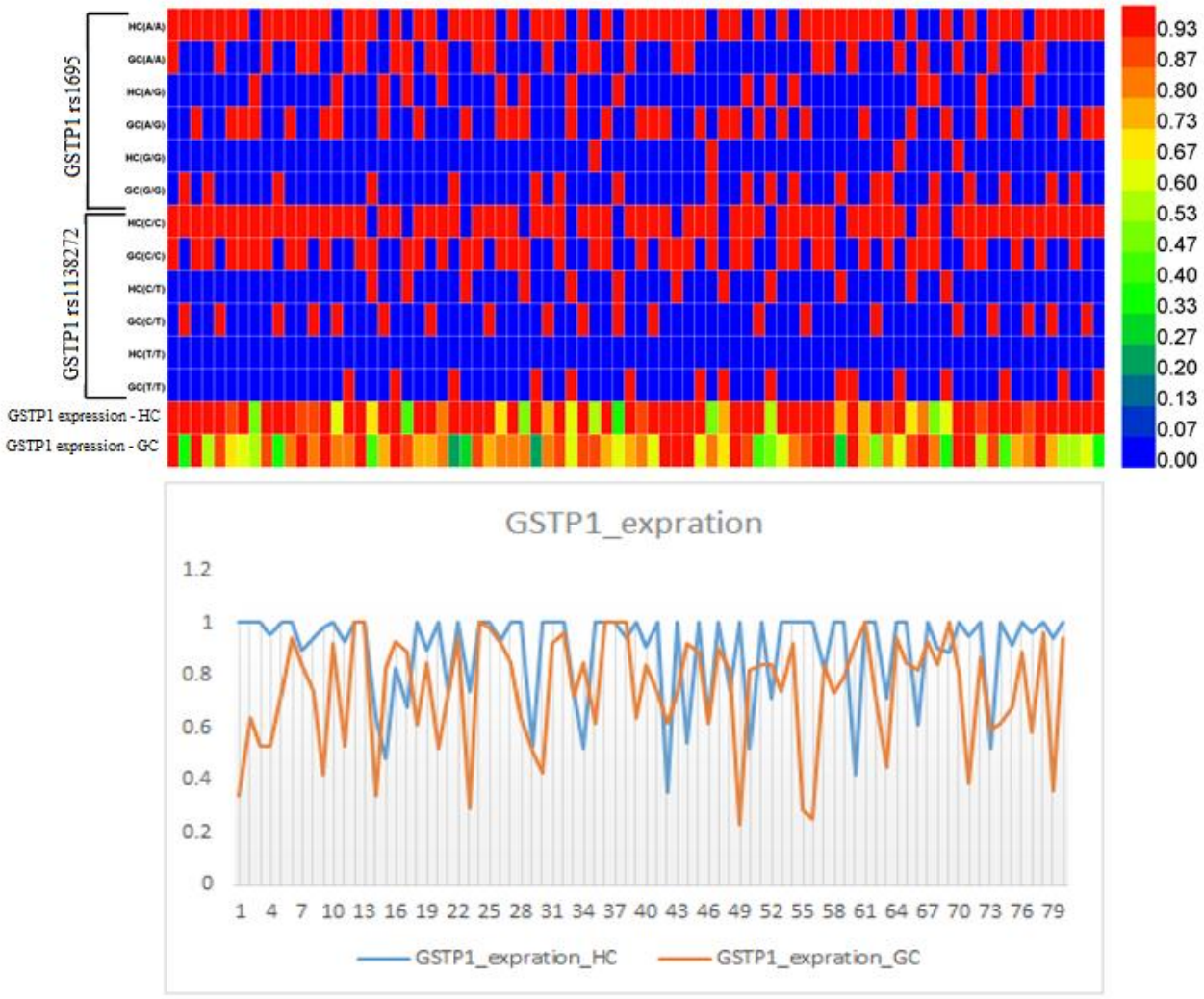


Figure 7.4. Heat map and graphical representation of GSTP1 gene expression relation with two different SNPs and their genotyping.

HC – Helathy Control, GC – Gactric Cancer.

4. Discussion

In our study, *GSTM1/T1* gene deletions were found to be associated with a higher risk of GC. Moreover, the simultaneous presence of both the *GSTM1* and *GSTT1* null genotypes was not identical in GC and control groups (30% in GC vs. 7.5% in HC; OR: 2.29, 95% CI: 1.36 – 9.53). A significant increase in risk of GC for individuals with deletions in both *GSTM1* and *GSTT1* genes were observed by collecting data from seven different planned studies (319 cases and 656 controls) that were *GSTT1* and *GSTM1* genotype combinations (Saadat, 2006; Boccia et al., 2006; Jing et al., 2012). However, simultaneous absence of *GSTM1* and *GSTT1* genes were not significantly associated to GC risk in some other studies carried out in different geographic populations (Piao et al., 2009).

In the *GSTpi-1* gene, *GSTP1* 105 (rs1695) polymorphism evaluated was related to GC susceptibility or phenotype. The homozygous status of *GSTP1* 105 (rs1695) (A>105G) was most susceptible to GC risk in the Mizo population (OR: 7.21; 95% CI: 1.48 – 35.07). The allele frequencies of the *GSTP1* Ile105Val (rs1695) and Val114Ala (rs1138272) polymorphisms were not similar to European or Western populations (Zendehdel et al., 2009), but similar in other Asian population (Mocellin et al., 2015). The variant *GSTP1* Val105 allele was found to be linked to the risk of esophageal squamous cell carcinoma with a predisposition associated with tobacco smoking in Swedish patients (Gonzalez et al., 2003). A probabilistic description for this inconsistency is the important ethnic and geographic variation observed in the GST profile. Such variation is of particular interest in the case of *GSTM1* and *GSTT1* null frequencies, which differ considerably between Asians and Caucasians. These geographical variations in

allele frequencies of GST gene along with reasonable connections with other gene polymorphisms could be the probable explanation for these observations.

A large number of epidemiological studies have reported the association between food habit, tobacco smoking and GC risk (La-Torre et al., 2009; Ladeiras-Lopes et al., 2008; Agudo et al., 2006; Soya et al., 2007; Malakar et al., 2012). However, very few studies have published the potential interaction between tobacco smoking and GST gene polymorphisms with respect to GC risk (Malik et al., 2009; Katoh et al., 1996). In the present study, we aimed to investigate whether the reasonable GST gene variants associated with GC risk could be triggered by other environmental risk factors like *H. pylori* infection, consumption of smoked food and tobacco related products. In our study, smoking habit was not significantly associated between *GST* variants and GC risk when compared with other factors. This finding is in concurrence to a recent meta-analysis where no association between the *GSTT1*and *GSTM1*null genotype with increased risk of GC in either smokers or non-smokers were observed (Boccia et al., 2007; Zhao et al., 2013).

In accordance to previous reports, *H. pylori* infection was identified as a potential risk factor associated with GST polymorphisms and GC risk (Palli et al., 2007; Parsonnet et al., 1991; Uemura et al., 2001). A significant effect of *H. pylori* status was observed in modifying the contribution of GST polymorphisms to GC risk. Palli*et al*. reported a significant increase of GC among *H. pylori* infected individuals with deletions in both the *GSTM1* and the *GSTT1* genes in a high-GC risk Italian population (Palli et al., 2005). In the present study, *CagA+* strain was highly associated with *GSTM1/T1* null allele and *GSTP1* 105 (rs1695) with OR: 3.08; 95% CI 0.75 – 2.61 and OR: 6.64 95% CI: 1.73 – 5.46, respectively (Table 7.2). In addition, studies on *GSTP1* allelotypes found that the isoenzymes with Valine in position 105

(*GSTP1*) were able to metabolize the diol epoxides of polycyclic aromatic hydrocarbons more efficiently and those with a A114V transition in the presence of V105 are able to more effectively metabolize benzo(a)pyrene(Hu et al., 1997; Eaton and Bammler, 1991). The MDR analyses also explained the gene–environment interaction (Figure 7.3A); and identified the combination of presence of *H. pylori* with *CagA* genotype and smoked meat consumption as the best model for gastric cancer risk associated with *GSTT1* null genotype in Mizo population (Figure7.3A, Table 4.3). Interaction entropy graph was drawn for visualization and interpretation of MDR interactions. Presence of *H. pylori* with *CagA* genotype showed significant individual effect with *GSTT1* gene polymorphisms as well as strong synergistic effect among each other in gastric cancer risk and validated the results of gene–environment interaction. The relatively small sample size in our study might be a shortcoming for predicting high-order interactions and multiple comparison analysis; however, MDR approach improves statistical power to overcome small sample size limitation by using cross-validation and permutation testing strategy. *GSTP1*, on the other hand, is expressed at high levels in extra hepatic tissues, including the kidney, liver and the gastrointestinal tract suggesting an important role in the protection against carcinogens and other xenobiotics in these tissues (DeBruin et al., 2000). Abnormal or low expression might not protect the cells from carcinogens or other xenobiotics(DeBruin et al., 2000). In this study, majority of the samples showed significant negative *GSTP1*expression for gastric cancer samples in the presence of G/G, A/G (rs1695) and T/T (rs1138272) genotypes. The mutant homozygous condition for these two rs-IDs might be potential risk factor of low *GSTP1* expression followed by gastric cancer development in Mizo population. Low level *GSTP1* expression had significant relation with homozygous mutant type G/G for GSTP1 105 (rs1695) and T/T for GSTP1 114 (rs1138272) (Figure 7.4A, 4B). These results indicate that an individual's detoxification capacity and exonic mutation in the xenobiotic gene also plays a crucial role in the initiation of gastric cancer (Roth et al., 2004). However, no differences in the occurrence of

Epstein–Barr virus (*EBV*) strains were observed between patients and controls associated with the *GSTP1* gene polymorphisms. After adjustment for potential confounding factors, a statistically significant difference was observed in the frequency of GSTM1/ T1 null genotyping with the presence of *H. pylori* and CagA like other study (Zhang et al., 2012). Extra intake of fermented fat (P<0.0003) and smoked meat consumption (P<0.001) were also found to be potential confounders for gastric cancer risk in relation to GSTM1/T1 null genotyping and GSTT1 null genotyping. *GSTP1*105 (rs1695) for homozygous G/G was also significantly related to GC phenotype as a potential confounder like other population (Lao et al., 2014). Further adjustment for other potential confounders did not substantially change the estimated ORs; therefore, we present only the estimates adjusted for significant matching variables.

In this study, the limitations concerning sample size can be ignored because of the fact that the Mizo ethnic group has a very small population size of about 10, 91,014 with 5,52,339 males and 5,38,675 females (Anon, 2011). The fact that our study is one of the first performed in the unique and high gastric cancer risk Mizo-Mongoloid population from North-East India will provide insights for early diagnosis of this critical disease as well help to evaluate the genetic epidemiology of the disease for other populations as well.

In conclusion, our data showed that *GSTM1*, *GSTT1* and *GSTP1* polymorphism analyzed in this study were found to be significantly associated with the demographic factors and GC risk in Mizo population. Similar to several other complex diseases, it is very devious to consider the weight of each demographic factor concerned in its pathogenesis separately, particularly the contribution of genetic factors. Presence of *H. pylori* with *CagA* genotype showed significant individual effect with *GSTT1* polymorphisms as well as strong synergistic effect among each other in gastric cancer risk and validated the results of gene–

environment interaction. Significant negative expression of *GSTP1* for gastric cancer samples with presence of G/G, A/G (rs1695) and T/T genotype (rs1138272) might not protect the cell from carcinogens or other xenobiotics. Further, prospective studies with larger sample size analysing gene-environment relations in different geographic areas and ethnic groups are needed to assess the significance of each factor for gastric cancer.

Chapter VIII

Summary

1. This study reports the first complete mitochondrial genome sequence data for a novel F1d2 haplogroup. Overall, these new data enriches the human maternal ancestry phylogeny and discloses novel mitochondrial genome signatures in this Indian Asian population.
2. 6395C >G heterozygous COXI mutation is present in 60% of the gastric cancer samples. Among the 27 COXImutations, 22 (14 non-synonymous and 8 synonymous) were novel mutations. All the 14non-synonymous mutations were not previously reported in the literature or the public mtDNA mutation databases. 6 missense mutations are damaging mutation with 1.00 polyphen2 and SIFT score and two domain (Heme a3/CuB – 6395,6962, 7196; K-pathways – 6621,6665, 7175) functions may be altered.
3. Our studies showed that the residual activity of the H240Y and S255P mutation was observed in the gastric cancer tumor sample. This observation indicates that impaired proton uptake in the K pathway cannot be overcome by protons supplied from the exterior, as seen in the D132 mutant.
4. Moreover, the variation in the amino acid residues of COXI in Gastric carcinogenesis manifests in the alteration of the molecular oxygen reduction site (Heme a3/CuB).
5. Total 49 mutations are found in hyper variable region I (HVI), 6 are in non-coding region and 17 are inhyper variable region II (HVII). In three of the gastric cancer patients, we found three distinct point mutations (228G>A, 235A>G and 248Adel) in the D-loop CSB1 and transcription factor X binding site region of the mtDNA
6. Most of the gastric cancer samples showed diploidy, except in six samples where aneuploidy (containing 6395C>G change) resulted in less DNA content in G2/M phase, whereas high DNA content in S phase. For most of the gastric tumor and adjacent normal tissues, high level of ROS generation (MFI) was observedthan endoscopic control group.

7. This study suggests that mitochondrial mutation in gastric tumor cell inactivation leads to reduced complex I, II and IV activity. The complex II and IV enzymes activity were significantly reduced for tumor, D-loop, COXI mutated and MSI cells than the adjacent normal cells.
8. The cells harboring the 6395C>G, 7103C>G and 7194C>G mutations causes more than50% decreasein the activity of complex IV. This COX activity in the adjacent normal control cells are normal indicating that these mutationsin gastric cancer influences the activity of a different complexin the respiratory chain.
9. The 303 and 16185 poly (C)n tract exhibited C_7TC_6 and C_7TC_4 for Mizo population, respectively. Seventeen cancer patients (Male -13; Female -5) and three healthy control(Male – 2; Female -1) exhibited abnormal bands with different mobilities for the 303 poly (C)n tract (Figure 4.4; Table 4.8). Five cancer patients (Male -3; Female -2) and two healthy control (Male – 1; Female -1) exhibited abnormal bands with different mobilities for the 16185 poly (C)n tract.
10. The coding regions are NDI$(C)_6$, NDII$(A)_7$, COI$(A)_7$, COIII$(T)_7$, NDV$(C)_6$ and $(A)_8$, NDV(CCT)3 and (AGC)3 and $tRNA^{ser}(A)_6$ and $(C)_6$ (Table 4.9). As shown in Table 4.9, four cases (male -3, female -1) revealed microsatellite instability in NDI region. Four samples (male -3, female -1) exhibited a cytidine deletion in the $(C)_6$ tract within the NADH dehydrogenase 5 (NDV) gene (Figure 4.4).
11. Two samples (male -2) and four samples (male -3, female -1) exhibited a adenine and cytidine insertion in the $(A)_8$and $(C)_6$ for NDV and NDI coding regions, respectively. These samples also exhibited frameshift mutations inthe $(A)_8$and $(C)_6$ tracts and created mitochondria specific stop codons AGA and AGG, respectively. The samples exhibited NDI and NDV coding region MSI and also D-loop 303 (C)n and 16185 (C)n tract instability.

12. The results show significant differences in the levels of mRNA in COXI mutated tumor and adjacent normal cells. The results show that the protein level was significantly reduced in COXI mutated tumor cell (Figure 4.6B an 4.6C). Taken together, the dis-coordinate expression of COXI mRNA and protein suggest post-transcriptional regulation by COXI gene mutations.
13. A significant relationship was observed with presence of *H. pylori* infection and MSI along with cagA genotypein gastric tumor tissue. High consumption of smoked meat and smoked vegetables also significantly affected the mitochondrial MSI and moderate association was found with high fermented fat intake.
14. Most of the negative COX I immunostaining was observed in cases of tumor having 6395C>G, 7103C>G mutations together and at stage-III only. Eight samples (10%) were negative, 46 (57.5%) were with moderate expression and 26 (32.5%) were positive for COXI protein expression in gastric adenocarcinoma tissues. After performing the Fisher exact test, the MT-COXI expression was significantly correlated with the Stages III of gastric cancer tissues ($p = 0.038$).
15. Screening of all the exons of TP53 gene showed that 60% patients with gastricadenocarcinoma had mutations and majority of them were somatic in nature. A total of 17 mutations were found in Exons 3, 5, 6, 8 and 11 (Table 5.2).
16. Exon 5 exhibited four novel mutations including a frameshift mutation (12616_12617 insG; E204G) due to guanine insertion at g.12616 position. Due to this mutation, early stop codon was generated at the 209 codon position. Another adenine deletion was also found in codon 204 in an intestinal type of gastric cancer patient at the position of g.12619 which led to frame shift mutation, generating a stop codon at 247 codon position. Hence, the codon 204 is a mutational hot spot for TP53 gene in Mizo gastric cancer patients.

17. Exons 8 and 11 exhibited 2 and 3 non-synonymous mutations, respectively. All the mutations in exons 8 and 11 are disease causing (as predicted by mutation taster), except 13811G>A and 17873G>C. All the mutations reported in exon 11 are novel mutations for gastric cancer.
18. We have found a novel deep intronic heterozygous mutation (c.693-2A>G) at exon 6 splice acceptor position in 2 gastric tumor samples. The c.693-2A>G mutation in the mutant induces a fully penetrant splicing defect leading to 24 bp deletion in exon 6. Since this variant lies within the highly conserved splice donor consensus sequence at position-2 from the 5'-end of exon 6, a defect in splicing is very likelyin TP53 gene.
19. This mutation was also significantly associated with TP53 protein and gene expression. The samples with this mutation exhibited negative expression for p53 protein and gene (Figure 5.2 and 5.3). This variant affects a highly conserved amino acid position and was classified to be disease-causing by different prediction tools.
20. No significant correlations were found between p53 tissue status and Lauren type, tumor grading, and depth of invasion (T status). In contrast, there was a significant association between p53 tissue status and the metastatic spread to lymph nodes.
21. Immunohistochemical analysis of p53 expression in *H. pylori*-positive GC resected tissue sections showed that the average positive expression in tumor and adjacent normal tissues were 44.3% and 6.9%, respectively, while in *H. pylori*-negative tissues they were 16.4% and 4.4% respectively. The average p53 expression was significantly higher in the tumor tissue in the presence of *H. pylori* than in the absence of *H. pylori*.
22. Most of the investigated primary diffuse type gastric carcinomas exhibited a low level or negative E-cadherin gene and protein expression (92%), and this was heterogeneous both in terms of the fraction of positive cells and the staining intensity (Figures6.1 and 6.2).

23. Comparison with various clinicopathological features revealed no significant relationship between E-cadherin-staining pattern and tumor diameter (<5 cm versus >5 cm, P = 0.264) or lymphatic vessel invasion (+ versus -, P = 0.182), whereas the Borrmann classification (reflecting the gross appearance of the carcinomas) showed low significant correlation between the presence of E-cadherin and Borrmann types I and II (type I/II versus type III/IV, P = 0.0038).
24. A trend was also observed between the presence of lymph node metastasis and the loss of E-cadherin expression (N0 versus N1/2, P = 0.0479). While no correlation could be demonstrated with the M stage(M0 versus Ml. P = 0.328), which was evaluated at the time of diagnosis when metastatic disease was obviously evident.
25. Most of the moderately differentiated tumors also showed average level of E-cadherin staining, although the fraction of positive cells was always less than in normal mucosa. The poorly differentiated carcinomas demonstrated weak E-cadherin expression in only a small fraction of tumor cells.
26. As in the intestinal type, the staining reaction depended on the degree of differentiation of tumor glands and thus, the E-cadherin-negative adenocarcinomas lacked any glandular formations. Another differentiation parameter which correlated with E-cadherin expression was the grading, which describes the morphological deviations of tumor cells from the typical epithelial phenotype of normal mucosa cells.
27. Screening of all the exons of CDH1 gene showed that 23.75 % patients with gastricadenocarcinoma had mutations and majority of them were germline in nature (due to presence in adjacent normal and blood sample). A total of 24 mutations were found in Exons 3, 4, 5, 7, 8, 9, 10, 11, 12, 14 and 15 (Table 6.3). Out of all 24 mutations, 15 were non-synonymous, 6 were frameshift and 3 were synonymous in nature.

28. Exons 7 and 8 exhibited 3 and 2 mutations, respectively which were predicted as disease causing (with high polyphen2 and SNP-GO score). Two frameshift mutations (74485_74486insA) at the position of g.74485 in exon 7 and (74922_74923insG) at the position of g.74922 in exon 8 led to generating premature stop codon at 293 and 350 positions, respectively.
29. Exon 9 exhibited 3 novel mutations including two frameshift mutation (76145_76146insGA and 76175_76176insT) which was predicted as disease causing with high polyphen2 and SNP-GO score. Two frameshift mutations (76145_76146insGA) at the position of g.76145 and (76175_76176insT) at the position of g.76175 led to generating premature stop codon at 418 and 419 position, respectively.
30. Exons 11, 12, 14 and 15 exhibited single novel mutation each, as predicted polymorphism, disease causing, polymorphism and disease causing with high polyphen2 and SNP-GO score, respectively. Due the mutations of exon 11, 12, 14 and 15, function of cytoplasmic binding TOPO domain might be lost for E-cadherin protein.
31. Our data showed that Presence of *H. pylori* with CagA genotype showed significant individual effect with GSTT1 polymorphisms as well as strong synergistic effect among each other in gastric cancer risk and validated the results of gene–environment interaction.
32. Significant negative expression of GSTP1 for gastric cancer samples with presence of G/G, A/G (rs1695) and T/T genotype (rs1138272) might not protect the cell from carcinogens or other xenobiotics.
33. Extra intake of fermented fat ($P<0.0003$) and smoked meat consumption ($P<0.001$) were also found to be potential confounders for gastric cancer risk in relation to GSTM1/T1 null genotyping and GSTT1 null genotyping.

Conclusion

This study has identified six novel mtDNA variations: two in the protein coding, two in control region and one variation in noncoding region (Table 3.1 and Figure 3.5). These variants returned no hits in Mitomap or mtDB database, thus these positions are deemed to be novel (Ruiz-Pesini et al., 2007). The variants were noted in all the individuals and are thus probably ancestral variations which led to novel F1d2 haplogroup.

To our knowledge, the present study is a novel finding in terms of the possible role of mtDNA COXI and D-loop mutations in gastric carcinogenesis. In summary, we have identified novel mutations in mitochondrial control region and cytochrome c oxidase subunit I genes (which is frequently less expressed) in gastric tumor cells. Mitochondrial microsatellite instability is also a major factor for progression of gastric tumor and it was significantly associated with *H. pylori* infection with cagA positive or negative and smoked meat consumption in Mizo population. All the tumor samples containing COXI 6395C>G heterozygous mutation are producing high level of ROS. The COXI mutations significantly altered the conserved amino acids, and they included both new heteroplasmic somatic and recurrent homoplasmic germ-line mutations. This might be major cause of low activity of complex IV enzymes. Due to the COI $(A)_7$ tract alteration, the complex III and IV enzyme activities might be altered and reflects the mitochondrial pathogenicity in gastric cancer. Mitochondrial gene alterations may attribute for gastric cancer risk in Mizoram, Northeast Indian population. The analysis of mtMSI and D-loop alteration might help to identify patients at high risk for gastric cancer diagnosis. The mutations of D-loop and COXI takes part in the carcinogenesis and progression of gastric cancer through the effect of increased ROS and altering the mitochondrial ETC complex enzyme activity and cell cycle. The mutations in COXI markedly decreased overall immunostaining of cytochrome c oxidase subunit

I may prove to be reliable biomarkers of gastric cancer risk, because they occur at elevated frequency in the stomach mucosa of patients with, and at high risk for cancer. We have also demonstrated that mtMSI is an early and important event in the progression of gastric carcinogenesis, especially in intestinal-type gastric cancer. *H. pylori* infection contributes to mtMSI in gastric cancer development. Our results support a role for mtMSI in different mechanisms of gastric carcinogenesis. Because the majority of patients with *H. pylori* infection will not progress to cancer and only a subset of these patients harbor mtMSI, it is conceivable that patients with *H. pylori* infection displaying mtMSI are at greater risk of developing gastric cancer than those without instability. Further, prospective studies with larger sample size analyzing gene–environment relations, mitochondrial mutations and microsatellites instability in different geographic areas, and ethnic groups are needed to assess the significance of each factor for gastric cancer.

In this study, we focused on the participation of genetic alterations of TP53 gene, such as mutational inactivation, SNPs, and expression of mutant form of p53 protein in the esophageal and gastric carcinogenesis. This study emphasizes the fundamental role of molecular alterations of "the guardian of the genome" in these neoplasms, with serious consequences for the deregulation of the cell cycle, loss of pro-apoptotic function. Considering the involvement of TP53 alterations both in early stages as in tumor progression, it is an important biomarker for the diagnosis, tumor progression, and poor prognosis associated with lymph nodes metastasis. Gastric carcinogenesis may be related to the existence of chronic *H. pylori* infection, which leads to imbalance of proliferation and apoptosis in the early stage, and furthermore mutation of p53-tumor-suppressor system and finally gastric cancer tumorigenesis. Hence, this molecular

pathology mechanism can be applied in routine diagnostic procedures, classification systems, disease monitoring.

Taking into consideration our findings, it appears that germ-line CDH1 cannot independently explain diffuse type gastric cancer and that other genetic alterations are involved in this disease. In fact, somatic mutations, enhanced by gene silencing, underlie the CDH-1 down-regulation in eight of seventeen patients. Remarkably, in four cases CDH1 was down-regulated in cancer specimens, without any evidence of either CDH1 mutations. Therefore, other genetic unknown alterations able to induce CDH1 down-regulation could be present in the gastric cancer patients. This finding is well in accordance with recent data indicating a mixed prevalence of hereditary and sporadic CDH1 genetic alterations in gastric cancer and prompts further research in this field to isolate unknown genes involved in CDH1 down-regulation.

In conclusion, our data showed that *GSTM1*, *GSTT1* and *GSTP1* polymorphism were found to be significantly associated with the demographic factors and GC risk in Mizo population. It is very devious to consider the weight of each demographic factor concerned in its pathogenesis separately, particularly the contribution of genetic factors. Presence of *H. pylori* with *CagA* genotype showed significant individual effect with *GSTT1* polymorphisms as well as strong synergistic effect among each other in gastric cancer risk and validated the results of gene–environment interaction. Significant negative expression of *GSTP1* for gastric cancer samples with presence of G/G, A/G (rs1695) and T/T genotype (rs1138272) might not protect the cell from carcinogens or other xenobiotics. Further, prospective studies with larger sample size analyzing gene-environment relations in different geographic areas and ethnic groups are needed to assess the significance of each factor for gastric cancer.

List of publications

1. **Ghatak S**, Muthukumaran RB, Senthil-Kumar N. (2013). A Simple Method of Genomic DNA Extraction from Human Samples for PCR-RFLP Analysis. Journal of Biomolecular Techniques. 24:224–231. (**Impact factor: 1.92).**
2. **Ghatak S**, Zothansanga, Pautu JL, Senthil-Kumar N. (2015). CoExtraction and PCR based analysis of nucleic acids from Formalin-Fixed Paraffin-Embedded Specimens. Journal of Clinical Laboratory Analysis. 29(6):485-92. (**Impact factor: 1.56).**
3. **Ghatak S**, Yadav RP, Lalrohlui F, Chakraborty P, Ghosh S, Ghosh S, Pautu JL, Zohmingthanga J, Senthil Kumar N. (2016). Xenobiotic Pathway Gene Polymorphisms Associated with Gastric Cancer in High Risk Mizo-Mongoloid Population, Northeast India. Helicobacter. DOI: 10.1111/hel.12308.(**Impact factor: 3.96).**
4. Ghosh S, Ghosh S, Bankura B, Saha ML, Maji S, **Ghatak S**, Pattanayak AK, Senthil-Kumar N, Das M. (2016). Association of DNA repair and xenobiotic pathway gene polymorphisms with genetic susceptibility to gastric cancer patients in West Bengal, India. Tumor Biology. 37(7):9139-9149.**(Impact factor: 2.98).**
5. Lalmuanpuii R, **Ghatak S**, Pautu JL, Lallawmzuali D, Muthukumaran RB, Senthil-Kumar N. (2015). Mutation Profiling in Mitochondrial D-Loop Associated with Stomach Cancer and Tobacco Consumers. Journal of Clinical & Medical Genomics. 3(1):1-5.**(Impact factor: 0.56).**
6. N. Senthil-Kumar, **Ghatak S**, Pautu JL. Mitochondrial and nuclear polymorphisms associated with gastric cancer in F1d2 Mizo haplogroup, Northeast India. Annals of Oncology**[Abstract only]**. 26(s9): ix4-ix4. (**Impact factor: 7.06).**

7. **Ghatak S**, Lallawmzuali D, Mawia L, N. Senthil-Kumar. (2014). Mitochondrial DLoop and Cytochrome Oxidase C Subunit I Polymorphisms Among The Patients Breast Cancer Of Mizoram, Northeast India. Current Genetics. 60: 201–212. (**Impact factor: 3.38).**
8. **Ghatak S**, Lallawmzuali D, Mukherjee S, Lalmawia, Pautau JL,Senthil-Kumar N. (2014). Polymorphism in mtDNA control region of Mizo-mongloid Breast Cancer samples as revealed by PCR-RFLP analysis. Mitochondrial DNA. 27(3): 2205–2208.(**Impact factor: 1.76).**
9. Thapa S, Lalrohlui F, **Ghatak S**, Zohmingthanga J, Lallawmzuali D, Pautu JL, Senthil-Kumar N. (2015). Mitochondrial complex I and V gene polymorphisms associated with breast cancer in mizo-mongloid population. Breast Cancer. 23(4):607–616.(**Impact factor: 1.59).**
10. Lallawmawma H, Sathishkumar G, Sarathbabu S, **Ghatak S**, Sivaramakrishnan S, Gurusubramanian G, Senthil-Kumar N. (2015). Synthesis of silver and gold nanoparticles using Jasminum nervosum leaf extract and its larvicidal activity against filarial and arboviral vector Culex quinquefasciatus Say (Diptera: Culicidae). Environmental Science and Pollution Research. 22(22):17753-17768.(**Impact factor: 2.83).**
11. Kimi L, **Ghatak S**, Yadav RP, Chhuani L, Lallawmzuali D, Pautu JL, Senthil-Kumar N. (2015). Relevance of GSTM1, GSTT1 and GSTP1 Gene Polymorphism to Breast Cancer Susceptibility in Mizoram Population, Northeast India. Biochemical Genetics. 54(1):41-49.(**Impact factor: 0.95).**
12. **Ghatak S,** Yadav RP, Rathore HS, Thou K, Jakha F, Sumi KT, Sanga Z, Senthil-Kumar N. (2016). Mitochondrial control region and GSTP1 polymorphism associated with familial urinary bladder cancer in Karbi-Anglong tribe of Assam, Northeast India.

Egyptian Journal of Medical Human Genetics. DOI.10.1016/j.ejmhg.2016.02.002.**(Impact factor: 0.40).**

13. Lalrohlui F, Thapa S, **Ghatak S**, Zohmingthanga J, Senthil-Kumar N. (2016). Mitochondrial complex I and V gene polymorphisms in type II diabetes mellitus among high risk Mizo-Mongoloid population, Northeast India. Genes and Environment 38(1):1–10.
14. Mandal SD, Passari AK, **Ghatak S**, Mishra VK, Senthil-Kumar N, Singh BP. (2015). Total Phenol Content, Antioxidant and Antimicrobial Capability of Traditional Medicinal Plants of Mizoram, Eastern Himalayas, Northeast India. EC Agriculture. 2.3: 350-357.
15. Zodinpuii D, **Ghatak S**, Mukherjee S, Senthil-Kumar N. (2013). Genetic relatedness of genus Oryza from Eastern Himalayan region as revealed by chloroplast matK gene. Asian Journal of Conservation Biology 2 (2):144-151.
16. Lalhmangaihi R, **Ghatak S**, Laha R, Gurusubramanian G, Senthil-Kumar N. (2014). Protocol for Optimal Quality and Quantity Pollen DNA Isolation from Honey Samples. Journal of Biomolecular Techniques. 25(4): 92–95(**Impact factor: 1.92).**

Communicated Papers

17. **Ghatak S**, Lalnunhlimi S, Lalrohlui F, Yadav RY, Chakraborty P, Pautu JL, Zohmingthanga J, Roy NK, Kunnumakkara AB, Lalhruaitluanga H, Senthil-Kumar N. (2016). Novel AKT1 mutations associated with cell cycle abnormalities in Gastric carcinoma. Genes and Genomics.

18. **Ghatak S**, Pautu JL, Zohmingthanga J, Rochamliana, Sharma S, Senthil-Kumar N. (2016). Novel F1d2 maternal lineage from Mizo tribal population of Northeast India. Plos One.
19. Yadav R, **Ghatak S**, Lalrohlui F, Kannan R, Kumar R, Sanga Z, Pautu JL, Zomingthanga J, Muthukumaran R, Senthil-Kumar N. (2016). Association of SNPs in cell cycle regulatory and MMR genes with susceptibility to gastric cancer. PeerJ.
20. **Ghatak S**, Chakraborty P, Yadav RP, Lalrohlui F, Mukherjee S, Pautu JL, Zohmingthanga J, Senthil-Kumar N. (2016). Pathogenicity of novel mitochondrial DNA mutations and mitochondrial dysfunction in gastric tumorigenesis. GUT.
21. **Ghatak S**, Chakraborty P, Yadav RP, Lalrohlui F, Sarathbabu S, Pautu JL, Zohmingthanga J, Senthil-Kumar N. (2016). Characterization of novel Missense and splice site mutations inTP53 gene and their genetic and functional analysis in gastric tumorigenesis. Gastroenterology.
22. **Ghatak S**, Chakraborty P, Yadav RP, Lalrohlui F, Sarathbabu S, Pautu JL, Zohmingthanga J, Senthil-Kumar N. Novel mutations and inactivation of the E-Cadherin Gene in diffuse Type Gastric Cancer. (2016). Human Mutation.