Genetic Analysis and Gene Expression Profile in Gastric Cancer in a High-Risk Northeast Region of India

Synopsis of thesis

Submitted in partial fulfillment

of the requirements for the degree of

DOCTOR OF PHILOSOPHY

by

THOUDAM REGINA DEVI

Under the Supervision of

Dr. Sunita Saxena

BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE

PILANI (RAJASTHAN) INDIA

2012

Synopsis

Background

(Chapter 1 and 2)

Gastric cancer (GC) is a leading cause of death worldwide, and patients have an overall 5-year survival rate of less than 10% [1]. Mortality of GC is a significant burden not only on patients but also on the whole health system[2]. Gastric carcinogenesis is a multistep process progressing from chronic gastritis, through glandular atrophy, intestinal metaplasia and dysplasia [3]. There are geographic and ethnic differences in the incidence of GC around the world as well as with its trends for each population over time. The incidence patterns observed among immigrants change according to where they live. All of these factors serve to indicate the close association of GC with modifiable factors such as diet [4]. Epidemiological and experimental studies have revealed vegetables and fruits decrease the risk of GC and that a high intake of smoke, salted and nitrated foods is a risk factor for the disease [5].

Frequency of GC, however, is very low in India, Bangladesh, Pakistan and Thailand compared to that in Japan and China [6]. However, in Mizoram, Northeast (NE) India, medical practitioners observed very high prevalence of GC [7]. The age adjusted incidence rates (AARs) from cancer data from both population-based and hospital-based cancer registries in India, showed the highest incidence of GC in the NE region of the country. Among males in Aizawl district (AAR 57.3), Mizoram state (AAR 50.6), Mizoram state excluding Aizawl district (AAR 46.9) and Sikkim state (AAR 14.2) had higher AARs among all the PBCRs. Among females all the three identified areas in Mizoram state i.e.

Aizawl district (AAR 33.6), Mizoram state (AAR 23.3) and Mizoram state excluding Aizawl district (AAR 17.2) had higher AARs followed by Chennai (AAR 5.6) [8].

The etiology of GC is multifactorial. Risk factor such as smoking has been reported to impairs healing and promote recurrence of precursor lesions for the development of GC like dysplasia, chronic atrophic gastritis and intestinal metaplasia [9]. Smokers had been shown to have twofold increase risk of GC compared with nonsmokers [10]. The mechanism by which tobacco smoke causes GC is not well known. Tobacco smoke contains several well-known chemical carcinogens that could act through direct contact with the gastric mucosa or indirectly through the blood flow [9]. The GST enzymes are involved in the detoxification of many xenobiotics, including several carcinogens from tobacco smoke. Sequence variations in genes coding for these enzymes may potentially alter individual susceptibility to cancer. Certain genes within the GSTM and GSTT (GSTM1 and GSTT1) subfamilies exhibit homozygous deletion (null genotype) polymorphisms that are considered important modifiers of individual risk for environmentally induced cancers. Individuals who have the homozygous deletion in one of these genes have no GSTM1 and GSTT1 enzyme activity, and thus are more susceptible to carcinogens. The prevalence of *GSTM1* and *GSTT1* null genotypes was found to vary among ethnic groups. Since Deakin et al. first investigated the relationship between *GSTT1* deficiency and GC in 1996, several studies have appeared in the literature, and most of them have refuted an association between *GSTT1* deficiency and GC risk [11, 12]. Meta-analysis have also reported *GSTM1* null genotype to increase GC risk [11]. One of the major problems with the published studies is that most of them were based on small numbers of cases and controls. Furthermore, because the *GSTs*

genotype is presumed to affect GC risk by influencing detoxification of activated environmental carcinogens and by interaction with other unfavourable GST polymorphisms, the potential modifying effect of *GSTs* status on the relationship between tobacco smoking and GC is of particular interest.

Beside the genes coding the xenobiotic detoxifying enzyme another important gene which has been reported to be frequently mutated in human cancer is p53 tumor suppressor gene. It has been found to be mutated in more than 50% of human cancers, it has attracted the interest of numerous researchers. The capacity of p53 for multiple biological functions can be attributed to its ability to act as a sequence-specific transcription factor to regulate expression of over one hundred different targets, and thus to modulate various cellular processes including apoptosis, cell cycle arrest and DNA repair [13]. The p53 belongs to an unique protein family which includes three members: p53, p63 [14] and p73 [15]. Although these proteins are structurally and functionally related to each other, p53 seems to have evolved in higher organisms to prevent tumor development, whereas p63 and p73 have clear roles in normal developmental biology [16]. Because p53 plays a key role in regulation of the cell cycle and induction of apoptosis, there has been enthusiasm about its potential for therapeutic application. The fact that p53 was originally described as an oncogene has recently come full circle with mutant p53 having been shown to exhibit gain–of–function properties that actually drive tumour progression and metastasis [17]. The p53 codon 72 Arg right curved arrow Pro polymorphism has been suggested to be associated with risk for different kind of cancers [18-22], but the data on GC is very limited [23]. Codon 72 polymorphism is a single base substitution of cytosine for guanine, leading to

arginine (A72) being replaced by proline (P72) [24]. The results are conflicting with Pro/Pro genotype showing association with lung cancer [25-28] breast cancer [29, 30] and GC [31] whereas Arg/Arg genotype being more prevalent in cervical cancer [32, 33]. Literature available from India is limited and inconsistent. Two different studies have reported both Arg/Arg and Pro/Pro genotypes to be associated with risk of lung cancer [28, 34] whereas no association was reported with oral cancer [18, 35]. There are no reports on association of *p53* codon 72 polymorphism with GC from India. Studies on codon 72 polymorphism have revealed striking ethnic differences [36, 37] and have demonstrated that frequency of *p53* variant allele varies with latitude, increasing in a linear trend as populations near the equator. Thus ethnicity might be related to allelic distribution of the gene and its determinacy in disease involvement [38]. Studies are needed to substantiate and to explore the contribution of p53 codon 72 polymorphism in the etiology of GC.

*Helicobacter pylori (*H*. pylori)* is considered the most prevalent infectious agent among humans, and it causes gastric inflammation, gastroduodenal ulcers, and a risk of GC [39]. *H. pylori* are a gram-negative spiral bacterium that colonizes the human stomach. The bacterium is discovered in 1983 by Robin Warren and Barry Marshall. Once acquire *H. pylori* usually persists for life, unless eradicated by antimicrobial therapy. Even though the prevalence of infection may be very high (70-90% in developing countries, 25-50% in developed countries), most humans infected with *H.pylori* are asymptomatic and only a few patients develop peptic ulcer or GC [40]. In Japan, there is a strong correlation of *H. pylori* infection and GC. On the contrary, in Africa (“the African enigma”), *H. pylori* infection does not always correlate with the risk

for peptic ulceration and GC [41]. Host genetic predisposition and local environmental factors, together with bacterial genotypes, may play an important role in the development of disease [40]. When infected with *H. pylori*, the relative risk for the development of GC increases to 2.1. On assuming that approximately 50% of the world is infected with this organism, this bacteria becomes responsible for 42% of GCs worldwide [42]. Obviously, more research on this bacteria and its effect on cancer needs to be done.

High throughput methods, such as microarrays and Next Generation Sequencing (NGS) are increasingly being used to systematically compare molecular features of individual cancers to key clinical parameters. Previous studies have documented the importance of genetic alteration affecting known oncogene, tomor suppresser genes, and mismatch repair genes in the development of GC [43, 44]. More and more studies have shown that these technologies are a powerful and revolutionary tool for biological and medical researches. The technology makes it possible to understand collective gene functions rather than just those of individual genes, and significantly contributes to advances in fundamental questions in biology as well as in clinical medicine. As the dysregulated expression and variations of genes lies at the origin of tumors, its measurements via the microarray and NGS technology can be very helpful to predict the clinical behavior of malignancies. Many works have shown that cancer diagnosis based carcinogenesis into the clinical decision-making process [45-47].

<u>Gap in existing research</u>

Very little is known about the underlying mechanisms involved in the development of gastric carcinoma. Nonetheless, the knowledge of the molecular events

involved in the development of gastric carcinoma is far from complete. Technologies such as microarray analysis and NGS may be useful in identifying new molecular genetic markers, and further work may determine whether these markers can be employed to help stratify patients into different multimodal treatment regimens. The technique we are using provides a rapid, robust, and sensitive platform to elucidate the molecular mechanisms underlying GC metastasis but also may identify candidate diagnostic markers and therapeutic targets. However, the validation and translation of these genomic classifiers as biomarkers into a completed 'bench-to-bedside' cycle for tailoring treatment to individuals is a major challenge and limits inflated expectations. Inspite of the high incident rate of GC in Northeastern regions like Mizoram no work at the molecular level has been investigated so far. Genetic study will provide an insight into the mechanism involved in the prevalence and development of GC in this region where a vast difference in food habits and culture from the rest of India were observed. The current study will be implicated to correlate genetic alteration and gene expression profile in GC with prevalent risk factor in high risk Northeastern region of India.

Aim and Objectives

(Chapter 3)

The aim of the study is to understand the underlying mechanisms involved in the carcinogenesis of GC in NE India where a very high incidence of GC is reported and search for possible markers to assist in both diagnosis and therapeutic approaches. The specific objectives of the present study are defined as under.

1. ***To determine the association of T1, M1 and P1 polymorphism in glutathione S – transferase genes and gastric cancer risk in northeast population of India***

Genotypes responsible for interindividual differences in ability to activate or detoxify genotoxic agents are recognized as biomarkers of susceptibility. Among the most studied genotypes are human glutathione transferases. The relationship of genetic susceptibility was studied especially in relation to the genetic polymorphism of glutathione S-transferase genes by using PCR-RFLP method and confirmation was done by sequencing for each genotype.

2. ***To determine the association of tumor suppressor p53 Arg72Pro polymorphism and risk of gastric cancer in the northeast general population of India***

   p53 is an important tumor suppressor, normally preventing cancer development via apoptosis. A genomic Arg72Pro substitution in the *p53* protein has important influence on cell death via apoptosis, which could be beneficial. We therefore tested the hypotheses that this polymorphism influences the risk of GC in NE population of India. PCR-RFLP method was used and confirmation was carried out by sequencing.

3. ***To evaluate PCR assays for detection of the presence of Helicobacter pylori in gastric cancer patients of northeast India***

   Several techniques such as culture, histology, rapid urease test etc have varying sensitivity and specificity for *H . pylori* detection. PCR being a highly efficient and reliable molecular technique for detection of various microorganism, different PCR using *H. pylori* specific genes Viz *UreA, GlmM* and *16SrRNA* were compared for their specificity and sensitivity for detection of *H. pylori*.

4. ***To study the gene expression profile of gastric cancer tissues in association with environmental risk factors***

Microarray technology is capable of determining the expression levels of thousands of genes in a biological sample simultaneously. This makes it widely used in cancer research. Characterization of these genes will help to elucidate the pathways and processes of carcinogenesis. It can be used to help clinical decision making, such as predicting therapy response, etc.

'OciChip Human A' chip (Ocimum Biosolution, Hyderabad, India) which contained 20160 genes has been used for the study.

5. ***To study genomic alteration involved in the process of carcinogenesis using next generation sequencing technology in matched normal and gastric tumor tissue***

Mutations are hallmark of cancers and identification of the mutations is imperative in our understanding of the disease. The advance in next generation sequencing (NGS) has transformed the way to identify mutations. It enables identification of somatic mutations, including base substitutions and indels. The rapid increase in NGS publications recently illustrated the potential of the technology, reporting rare mutations in various cancers, many previously undetected. Solexa plateform was used and specific regions of one hundred and sixty nine genes were analyzed by sequencing based on NGS technology.

Chapter 4: Determination of association of T1, M1 and P1 Polymorphism in Glutathione S –Transferase Genes and Gastric Cancer Risk in Northeast Population of India

In the present study, the association of polymorphism of *GSTT1*, *GSTM1*, and *GSTP1* genes with GC risk was evaluated using PCR-RFLP to find out if this could explain the unusually high prevalence of GC in the NE region of India. Exposure to the type and amount of environmental toxins is variable not only in different geographic regions, but also in different ethnic groups within the same geographic region. We have also analyzed the data of different ethnic groups separately as well as a combined group.

Results: The frequency of *GSTT1* and *GSTM1* null genotype was 38% and 37% in samples obtained from patients with GC and 32% and 45% in controls, respectively. Variant genotypes (Ile/Val and Val/Val) of *GSTP1* were found more frequently in GC cases (44%) when compared with controls (36%), but the difference was not statistically significant (OR=1.29, 95% CI: 0.80–2.07, p=0.29). No significant association of *GSTT1*, *GSTM1* and *GSTP1* polymorphism either singly or in combination were observed with GC risk. However, *GSTM1* was seen to have 26% lesser chance of developing GC in Mizoram population (OR = 0.65, 95%CI = 0.29-1.12, p = 0.10) and 35% lesser chance when analysis was carried out with NE population as one group (OR = 0.74, 95%CI = 0.47-1.16, p = 0.19). When data were analyzed for each geographical region, the prevalence of *GSTT1* null genotype in Assam was found to be significantly higher (OR=3.07, 95% CI: 1.33–7.09, p=0.009) in GC cases (27%) when compared with controls (12%).

Betel quid chewing habits were higher in GC cases when compared with a control population, but this difference was statistically insignificant (OR=1.45, 95% CI: 0.90-2.35, p=0.12). Region wise analysis reveals that betel quid chewing increase threefold risk of developing GC in Assam population when compared with controls and it was statistically significant (OR=3.61, 95%CI=1.06-12.21, p=0.04).

Chapter 5: Determination of association of p53 gene, its interaction with tobacco, Betel quid and alcohol consumption and risk of gastric cancer: a case-control study in Northeast population to understand the etiology

The fact that p53 was originally described as an oncogene has recently come full circle with mutant p53 having been shown to exhibit gain–of–function properties that actually drive tumour progression and metastasis. Lack of data on *p53* codon 72 polymorphism and high incidence of GC in NE region of India incited us to explore and evaluate any relevance of this polymorphism in this ethnic population. We carried out a case control study and the role of *p53* codon 72 polymorphism and its interaction with tobacco, betel quid and alcohol was analyzed.

Results: No significant increase in risk of GC was observed in a univariate or in a multivariable analysis for dominant and recessive models of inheritance. Frequency of Arg/Arg, Arg/Pro, and Pro/Pro genotypes was 20.1%, 61.9%, and 17.9% in the cases and 22.7%, 51.4%, and 25.9% in controls. Conditional logistic regression analysis showed higher risk associated with Arg/Pro genotype, whereas the Pro/Pro appeared to be a protective genotype. These observations were represented by both OR1 and OR2 but lacked statistical power. Analysis for interactions of risk habits with p53 genotypes did not reveal any combination to be

significant toward GC. However, interaction of betel quid chewing with Arg/Pro genotype reached a near significance level, with a twofold risk of GC (OR2=2.40, 95% CI=0.91–6.26, p=0.07).

Chapter 6: Evaluation of PCR assays for detection of the presence of Helicobacter pylori in gastric cancer patients of northeast India

The primer choosen for the study were for *GlmM*, *UreA* and *16S rRNA*. The aims of the present study were to compare the accuracy of the reported PCR primer pairs using GC biopsy specimens known to either contain *H. pylori* or to be *H. pylori* negative by serological tests in high risk region of NE India.

Results: The PCR methods using *GlmM, UreA* and *16SRNA* genes have shown a sensitivity of 85%, 65% and 53% and specificity of 63%, 63%, 66% respectively. The positive likelihood ratio of *GlmM, UreA* and *16SRNA* are 2.29, 1.75 and 1.53 and the negative likelihood ratio are 0.24, 0.56 and 0.72 respectively. The sensitivities of different combination of primers i,e of *GlmM+ve/16SrRNA+ve GlmM+ve/UreA+ve, UreA+ve/16SrRNA+ve* and are 50%, 49%, 48% and whereas their specificities are 74%, 69% and 73% respectively. The positive likelihood ratio of *Glm+ve/16SrRNA+ve*, *GlmM+ve/UreA+ve* and *UreA+ve/16SrRNA+ve* are 1.94, 1.59 and 1.85 respectively. The negative likelihood ratios are 0.67, 0.73 and 0.71 respectively. Among the three primers choosen *GlmM* seems to be a promising one for detection of *H. pylori*, however the various combinations of primers offer no increase in sensitivity or specificity.

Chapter 7: Gene Expression Profiling of gastric cancer in North East India: Role of Helicobacter pylori and tobacco

In this chapter we have evaluated the gene expression profiles of GC patients and correlation with *H.pylori* infection were examined in this region. Validation of microarray results was done on eight selected genes by quantitative real-time RT-PCR analysis.

Results: Of the 108 significantly up-regulated genes, genes involved in apoptosis (*RYK, SH2D2A, CSNK1E, LRP12, HSPE1, FGB, PXN, CHP*), proinflammatory response (*IL9*), angiogenesis (*SH2D2A, ANG*), cell adhesion (*CD151, PXN, WASF1)*, cell proliferation *(CDC2L1, JAG2, IL9),* regulators of I-kappa B kinase / NF-kappa B cascade (*CXXC5*), Wnt signaling (*CHP, CSNK1E, RYK*), Notch signaling (*JAG2*), Hedgehog signaling (*CSNK1E, BMP7*), metobolic pathway (*LDHB*), thyroid cancer (*RET)* and epithelial cell signaling in *H. pylori* infection (*ATP6V1G2*) were found to be biologically relevant in tumorigenesis. Of the 62 significantly down-regulated genes, genes involved in base-excision repair (*RAD51L3*), inhibition of cellular proliferation (*INSM1),* immune response (*RFX1, CXCL10, C3AR1, MALT1*), viral response (*ATP6V1G2, HBXIP, ACE2*), epithelial cell signaling in *H. pylori* infection (*JAM3*) were found to be relevant in tumorigenesis.

For studying the role of *H.pylori*-environment interaction that might modify susceptibility of GC, potential interactions of *H. pylori* with known risk factors were analysed. Although, betel quid chewing do not emerged as a risk factor, interestingly on interaction with *H. pylori* it causes a significant threefold risk of GC (OR2=3.52, 95% CI=1.16-10.68, *p*=0.026). When relationship between tobacco

chewers were analysed with *H. pylori*, significant interaction were observed with all the three combination but the significant level was enhanced in those individuals who were *H. pylori* positive as well as chewers (OR2=7.12, 95% CI=5.02-36.31, $p \leq 0.0001$).

Chapter 8: Assessing matched normal and tumor pools in next-generation sequencing studies to indentify genomic alteration in gastric adenocarcinoma in high risk northeastern region of India

In the current study Human, Illumina-Solexa platform has been used for NGS study to reveal the genomic alteration involved in the carcinogenesis of GC with enhanced pace.

Results: The total number of SNPs observed were 178 in cases whereas those of indels were 23 (12 insertion and 11 deletion). The *ACTL6A*, *PARK2*, *PTPN1*, *TP63*, *TSC2*, *VHL* and *NGFR* were the exclusive genes with SNPs only in cases. No SNPs were observed in these genes in the normal tissue samples. Among *FLNC, FAT1, ACTL6A, CUL2, MSH6, DLEC1* and *MET* with SNP targeted at the coding region *FAT1* was of high importance leading to non synonymous changes of amino acid at codon 482 position. Of the novel SNPs, four were located in the coding region of *TSG101, APC, FLNC and CASP10* genes. The novel indels observed in cases alone was at chromosomes 11 with ref Seq ID NM_003682 involving *MADD* gene.

Conclusion

Although our results do not show consistent significant association of the candidate gene polymorphism in the risk of GC, we cannot totally rule out the possibility of GSTs gene effect on the carcinogenesis of GC as they are important with the

detoxification of tobacco related carcinogens. Moreover a significant association of *GSTT1* null genotype was observed with GC in Assam population. Notably one of the major advances in our understanding of the risk factors and *H. pylori* interaction is that the bacterial factors are influence by the risk factors such as tobacco, betelquid and alcohol in determining disease outcome. The frequency of *H. pylori* positive cases did not show significant difference among cases and controls. Though *H.pylori* together with tobacco and betel quid chewing was found to confer a significant risk for GC. The environmental and genetic system of the host influence *H. pylori* infection that might have elicited an overly aggressive response resulting in a series of morphological changes that lead to cancer. Pathways identified in our study i,e wnt, notch, hedgehog and epithelial signaling in *H. pylori* infection have specific role during the process constituting the genetic decomposition of the host. The mechanism of GC carcinogenesis in this region involve other genetic alteration such as those in *ACTL6A, PARK2, PTPN1, TP63, TSC2, VHL*, *NGFR*, *FAT1*, *APC*, *TSG101, FLNC, CASP10* and *MADD* as evident in our NGS study. These supported the multiplicative risk factor hypothesis of gastric carcinogenesis. It has been speculated that many risk factors such as *H. pylori* and environmental factor already in existence may work in part.

Future Scope

We have identified genes that are implicated in GC in NE region of India. The development of bioinformatics in biological research is developing rapidly; and the amplification procedures could be pushed further. The identification of the linked between abnormal gene expression and mutation in diseased tissues will lead to

the discovery of new drug targets as well as advances in diagnosis. If in the future these elements can be successfully integrated, it's a realistic possibility that we will be able create a window through which to view gene expression and candidate gene at the cellular level in normal and diseased tissues. Intervention to eradicate *H. pylori* and control over betelquid and tobacco consumption early in this pathological process might prevent progressively more severe disease.

References

1. Chen, Y.R., et al., *Quantitative proteomic and genomic profiling reveals metastasis-related protein expression patterns in gastric cancer cells.* J Proteome Res, 2006. 5(10): p. 2727-42.
2. Vieth, M. and M. Stolte, *Elevated risk for gastric adenocarcinoma can be predicted from histomorphology.* World J Gastroenterol, 2006. 12(38): p. 6109-14.
3. Steenport, M., et al., *Association of polymorphisms in myeloperoxidase and catalase genes with precancerous changes in the gastric mucosa of patients at inner-city hospitals in New York.* Oncol Rep, 2007. 18(1): p. 235-40.
4. Tsugane, S. and S. Sasazuki, *Diet and the risk of gastric cancer: review of epidemiological evidence.* Gastric Cancer, 2007. 10(2): p. 75-83.
5. Ming, S.C., *Cellular and molecular pathology of gastric carcinoma and precursor lesions: A critical review.* Gastric Cancer, 1998. 1(1): p. 31-50.
6. Singh, K. and U.C. Ghoshal, *Causal role of Helicobacter pylori infection in gastric cancer: an Asian enigma.* World J Gastroenterol, 2006. 12(9): p. 1346-51.
7. Phukan, R.K., *High prevalence of stomach cancer among the people of Mizoram, India.* Current Science, 2004. 87(3): p. 10.
8. ICMR, *Consolidated report of population based cancer registries 2001-2004 , Incidence and distribution of cancer. National cancer registry programme. .* Indian council of medical research (ICMR), Bangalore, 2006.
9. Gonzalez, C.A., et al., *Smoking and the risk of gastric cancer in the European Prospective Investigation Into Cancer and Nutrition (EPIC).* Int J Cancer, 2003. 107(4): p. 629-34.
10. Pavithran, K., D.C. Doval, and K.K. Pandey, *Gastric cancer in India.* Gastric Cancer, 2002. 5(4): p. 240-3.
11. Boccia, S., et al., *Glutathione S-transferase T1 status and gastric cancer risk: a meta-analysis of the literature.* Mutagenesis, 2006. 21(2): p. 115-23.
12. Hatagima, A., et al., *Glutathione S-transferase polymorphisms and oral cancer: a case-control study in Rio de Janeiro, Brazil.* Oral Oncol, 2008. 44(2): p. 200-7.
13. Zhu, L.B.a.W.-G., *P53: Structure, Function and Therapeutic Applications. .* Journal of Cancer Molecules 2006. 2(4): p. 141-153.
14. Schmale, H. and C. Bamberger, *A novel protein with strong homology to the tumor suppressor p53.* Oncogene, 1997. 15(11): p. 1363-7.
15. Kaghad, M., et al., *Monoallelically expressed gene related to p53 at 1p36, a region frequently deleted in neuroblastoma and other human cancers.* Cell, 1997. 90(4): p. 809-19.
16. Irwin, M.S. and W.G. Kaelin, *p53 family update: p73 and p63 develop their own identities.* Cell Growth Differ, 2001. 12(7): p. 337-49.
17. Whibley, C., P.D. Pharoah, and M. Hollstein, *p53 polymorphisms: cancer implications.* Nat Rev Cancer, 2009. 9(2): p. 95-107.
18. Tandle, A.T., V. Sanghvi, and D. Saranath, *Determination of p53 genotypes in oral cancer patients from India.* Br J Cancer, 2001. 84(6): p. 739-42.

19. Rogounovitch, T.I., et al., *TP53 codon 72 polymorphism in radiation-associated human papillary thyroid cancer.* Oncol Rep, 2006. 15(4): p. 949-56.
20. Mitra, S., et al., *Association of specific genotype and haplotype of p53 gene with cervical cancer in India.* J Clin Pathol, 2005. 58(1): p. 26-31.
21. Papadakis, E.D., N. Soulitzis, and D.A. Spandidos, *Association of p53 codon 72 polymorphism with advanced lung cancer: the Arg allele is preferentially retained in tumours arising in Arg/Pro germline heterozygotes.* Br J Cancer, 2002. 87(9): p. 1013-8.
22. Wu, M.T., M.C. Chen, and D.C. Wu, *Influences of lifestyle habits and p53 codon 72 and p21 codon 31 polymorphisms on gastric cancer risk in Taiwan.* Cancer Lett, 2004. 205(1): p. 61-8.
23. Shen, H., et al., *P53 codon 72 polymorphism and risk of gastric cancer in a Chinese population.* Oncol Rep, 2004. 11(5): p. 1115-20.
24. Pietsch, E.C., O. Humbey, and M.E. Murphy, *Polymorphisms in the p53 pathway.* Oncogene, 2006. 25(11): p. 1602-11.
25. Wang, Y.C., et al., *Prognostic significance of p53 codon 72 polymorphism in lung carcinomas.* Eur J Cancer, 1999. 35(2): p. 226-30.
26. Birgander, R., et al., *p53 polymorphisms and haplotypes in nasopharyngeal cancer.* Hum Hered, 1996. 46(1): p. 49-54.
27. Kawajiri, K., et al., *Germ line polymorphisms of p53 and CYP1A1 genes involved in human lung cancer.* Carcinogenesis, 1993. 14(6): p. 1085-9.
28. Jin, X., et al., *Higher lung cancer risk for younger African-Americans with the Pro/Pro p53 genotype.* Carcinogenesis, 1995. 16(9): p. 2205-8.
29. Papadakis, E.N., D.N. Dokianakis, and D.A. Spandidos, *p53 codon 72 polymorphism as a risk factor in the development of breast cancer.* Mol Cell Biol Res Commun, 2000. 3(6): p. 389-92.
30. Sjalander, A., et al., *p53 polymorphisms and haplotypes in breast cancer.* Carcinogenesis, 1996. 17(6): p. 1313-6.
31. Hiyama, T., et al., *p53 Codon 72 polymorphism in gastric cancer susceptibility in patients with Helicobacter pylori-associated chronic gastritis.* Int J Cancer, 2002. 100(3): p. 304-8.
32. Dokianakis, D.N. and D.A. Spandidos, *P53 codon 72 polymorphism as a risk factor in the development of HPV-associated cervical cancer.* Mol Cell Biol Res Commun, 2000. 3(2): p. 111-4.
33. Storey, A., et al., *Role of a p53 polymorphism in the development of human papillomavirus-associated cancer.* Nature, 1998. 393(6682): p. 229-34.
34. Sreeja, L., et al., *p53 Arg72Pro polymorphism predicts survival outcome in lung cancer patients in Indian population.* Cancer Invest, 2008. 26(1): p. 41-6.
35. Nagpal, J.K., S. Patnaik, and B.R. Das, *Prevalence of high-risk human papilloma virus types and its association with P53 codon 72 polymorphism in tobacco addicted oral squamous cell carcinoma (OSCC) patients of Eastern India.* Int J Cancer, 2002. 97(5): p. 649-53.
36. Sjalander, A., et al., *p53 polymorphisms and haplotypes in different ethnic groups.* Hum Hered, 1995. 45(3): p. 144-9.

37. **Beckman, G., et al., *Is p53 polymorphism maintained by natural selection?* Hum Hered, 1994. 44(5): p. 266-70.**
38. **Fan, R., et al., *The p53 codon 72 polymorphism and lung cancer risk.* Cancer Epidemiol Biomarkers Prev, 2000. 9(10): p. 1037-42.**
39. **Thye, T., et al., *Genomewide linkage analysis identifies polymorphism in the human interferon-gamma receptor affecting Helicobacter pylori infection.* Am J Hum Genet, 2003. 72(2): p. 448-53.**
40. **Maggi Solca, N., et al., *Population genetics of Helicobacter pylori in the southern part of Switzerland analysed by sequencing of four housekeeping genes (atpD, glnA, scoB and recA), and by vacA, cagA, iceA and IS605 genotyping.* Microbiology, 2001. 147(Pt 6): p. 1693-707.**
41. **Yamazaki, K., et al., *Tumor differentiation phenotype in gastric differentiated-type tumors and its relation to tumor invasion and genetic alterations.* World J Gastroenterol, 2006. 12(24): p. 3803-9.**
42. **Kim, K.E., *Gastric Cancer in Korean Americans: Risks and Reductions.* Korean Korean Am Stud Bull, 2003. 13(1/2): p. 84-90.**
43. **Lin, W., et al., *Tyrosine kinases and gastric cancer.* Oncogene, 2000. 19(49): p. 5680-9.**
44. **Werner, M., et al., *Gastric adenocarcinoma: pathomorphology and molecular pathology.* J Cancer Res Clin Oncol, 2001. 127(4): p. 207-16.**
45. **Goh, L., et al., *Assessing matched normal and tumor pairs in next-generation sequencing studies.* PLoS One, 2011. 6(3): p. e17810.**
46. **Schuster, S.C., *Next-generation sequencing transforms today's biology.* Nat Methods, 2008. 5(1): p. 16-8.**
47. **Hippo, Y., et al., *Global gene expression analysis of gastric cancer by oligonucleotide microarrays.* Cancer Res, 2002. 62(1): p. 233-40.**
48. ***cagA-Positive Helicobacter pylori Populations in China and The Netherlands Are Distinct.* INFECTION AND IMMUNITY, 1998. 66(5): p. 1822–1826.**
49. **Clayton, C., K. Kleanthous, and S. Tabaqchali, *Detection and identification of Helicobacter pylori by the polymerase chain reaction.* J Clin Pathol, 1991. 44(6): p. 515-6.**
50. **Smith, S.I., et al., *Comparison of three PCR methods for detection of Helicobacter pylori DNA and detection of cagA gene in gastric biopsy specimens.* World J Gastroenterol, 2004. 10(13): p. 1958-60.**
51. **Ho, S.A., et al., *Direct polymerase chain reaction test for detection of Helicobacter pylori in humans and animals.* J Clin Microbiol, 1991. 29(11): p. 2543-9.**

Genetic Analysis and Gene Expression Profile in Gastric Cancer in a High-Risk Northeast Region of India

THESIS

Submitted in partial fulfillment

of the requirements for the degree of

DOCTOR OF PHILOSOPHY

By

Thoudam Regina Devi

Under the Supervision of

Dr. Sunita Saxena

BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE

PILANI (RAJASTHAN) INDIA

2012

BIRLA INSTITUTE OF TECHNOLOGY & SCIENCE

PILANI RAJASTHAN

CERTIFICATE

This is to certify that the thesis entitled "**Genetic Analysis and Gene Expression Profile in Gastric Cancer in a High-Risk Northeast Region of India**" submitted by **Thoudam Regina Devi,** ID No. 2007PHXF029P for award of Ph.D. Degree of the Institute embodies original work done by her under my supervision.

Signature in full of the Supervisor

Name in capital block letters DR. SUNITA SAXENA

Designation SCIENTIST 'G' (DIRECTOR)
NATIONAL INSTITUTE OF PATHOLOGY

Date: **17-09-2012**

Acknowledgement

I take this opportunity with much pleasure to thank all the people who have helped me through the course of my journey towards producing this thesis. I wish to express my deepest gratitude to my supervisor, DR SUNITA SAXENA, Director, National Institute of Pathology, ICMR. Ma'am, Thank you very much for giving me the opportunity to work in the intriguing field of cancer genetics. Your positive attitude, dedication to research and belief in me has been a never ending source of inspiration. Your generous support, great patience in reading and revising manuscripts and invaluable criticism has been very helpful.

This thesis would not have been possible without the help, support and patience of Dr. Sujala Kapur, Deputy Director, National Institute of Pathology, ICMR.Ma'am, Thank you for all the guidance and support I have received from you and for the constant support, encouragement and for enlightening me the first glance of research.

I am immensely thankful to Prof. B.N. Jain, Vice-Chancellor, B.I.T.S, Pilani for providing me this opportunity to pursue the off-campus PhD of the Institute.

I express my gratitude to Dr. S.K. Verma (Dean), Prof. Ashish K. Das (Ex Dean), Prof. Ravi Prakash (Ex- Dean), Research and Consultancy Division (RCD), BITS, Pilani for his constant official support, encouragement and making the organization of my research work through the past few years easy. I also thank my Doctoral Advisory Committee (DAC) members, Dr. Vishal Saxena and Dr. Uma Dubey. I thank Dr. Sharad Shrivastava Mr. Dinesh Kumar, Dr. Hemanth Jadav, Ms. Monica Sharma, Mr. Gunjan Soni, Mr. Amit Kumar and Ms. Sunita Bansal, nucleus members of RCD and all office staff of RCD B.I.T.S, Pilani, for their cooperation and guidance in all official procedures associated with the doctoral program.

I would not have succeeded in this effort, had it not been for the support in sample collection from Dr. Eric Zomawai (Head of Department Civil Hospital, Pathology Department, Aizawl), Dr. Jagannath Sharma (Chief consultant Pathologist BB Barooah Cancer Institute, Guwahati), Dr. Yogesh Verma (Consultant Pathologist & Head Sir Thutab

Namgyar Memorial Hospital, Gangtok, Sikkim). Special thanks to for your support in recruitment and diagnosis of the patients and for providing clinical inputs.

I would also like to thank Mrs R Sharda (Administrative Officer) for her help in various matters.

I am thankful to Prof. Indira Nath, for her being excellent source of inspiration. Ma'am, thank you for your kind suggestions and advices. I sincerely thank to Dr Usha Agarwal, Dr Avninder, Dr Aruna Mittal, Dr Laxman and Dr Poonum Salholtra for their kind support.

I am very much thankful to Dr L C Singh, for his keen support and immense help. I am also thankful to Dr Raju, for giving valuable comments and support. I am also very much thankful to Dr A K Mishra, for always been there for all questions, big and small. Without your statistical expertise I would have been at a loss. I am also very much thankful to Dr Anju Bansal for the important comments and support she has given me during my thesis writing. I sincerely thank to lab members Mrs Valsamma Mathew and Mr Jagdish Pant for their caring and supportive attitude. I also thank to Mr Mohan Singh, Mr Jagat and Mr. Sajid Hussain for their assistance and help during the research work.

Special heartfelt thanks to my dear friends Abha Soni, Pragya Singh, Apruv Rashmi and Dhirendra Singh Yadav. Thank you for friendship, support and for being there with me in every circumstances. You all have been the driving force behind my efforts. I sincerely thank my Uncle and Aunty (Mr.and Mrs. Virender Prakash) for their kind support and encouragement.

This thesis work was not possible without financial support from Indian council of medical research (ICMR, JRF and SRF), and National Institute of Pathology (ICMR), India for proving me research fellowships and funding for research work. Thus their immense support is gratefully acknowledged.

I would like to thank my senior Dr.Anurupa Chakraborty, Dr. Indranil Chattopadhya and Dr. Anand Verma for their valuable suggestions and inputs. I warmly thank all my colleagues, Pradeep Singh Chauhan, Rakhshan Ihsan, Mishi Kaushal,

Shreshtha Malviya, Meena Lakhanpal, P Shantilatha, Ashima Khana, Dhananjay Pathak and Neetu Mishra for their help and many good advices during these years. I would also like to thank Dr Bharat Bhshan,Dr Hemchandra Jha, Dr Harsh Vardhan, Dr Rajneesh Jha, Dr Rishen Gupta, Dr Dhiraj Gupta, Dr Paresh Sharma and Rashmi Tomar for their help and support.

Finally, **my parents;** thank you for all your love and amazing support you have shown me throughout my life. During these long years of academic degree and with things that were not necessarily meant for my disposal; your patience and tolerance was really incredible. You let my imagination and curiosity for life to run free. You had faith in me when I had none and you set me free to follow my dreams. Last but not least, God Almighty, for the wonderful life that you has given me and the great set of human beings that you have made me friends with, I thank you.

Place: New Delhi

Date: 17-09-2012 Thoudam Regina Devi

Abbreviations

°C	Degree Celsius
AA	Aromatic Amines
AAR	Age Adjusted Incidence Rate
AHH	Aryl Hydrocarbon Hydroxylase
APC	Adenomatous Polyposis Coli
AR	Amphiregulin
Arna	Amplified RNA
CDH	Competitive DNA Hybridization
cDNA	Complementary DNA
CYPs	CytochromeP-450
DNA/RNA	Deoxyribose/Ribose Nucleic Acid
dNTP	Deoyribose Nucleotide Triphosphate
EBV	Epstein-Barr Virus
EDTA	Ethylene Diamine Tetra-Acetic Acid
ELISA	Enzyme Linked Immunosorbent Assay
FAP	Familial Adenomatous Polyposis
GC	Gastric Cancer
GSTs	Glutathione S-transferases
GWAS	Genome-Wide Association Study
H. pylori	*Helicobacter pylori*
HDGC	Hereditary Diffuse Gastric Cancer
HGC	Hereditary Gastric Carcinoma
HGF	Hepatocyte growth factor

IgG	Immunoglobulin G
LOH	Loss of Heterozygosity
MAD	Median Absolute Deviation
MALT	Mucosa Associated Lymphoid Tissue
Mg	Milligram
Ml	Millilitre
mM	Millimolar
MMRs	Mismatch Repair Genes
MSI	Microsatellite Instability
NE	North East
Ng	Nanogram
OD	Optical Density
OR	Odd Ratio
PAHs	Polycyclic Aromatic Hydrocarbons
PBS	Phosphate Buffer Saline
PCR	Polymerase Chain Reaction
Prot.K	Proteinase K
PTP	Picotiter Plat
RFLP	Restriction Fragmnt Length Polymorphism
RIN	RNA Integrity Number
Rpm	Revolution per minute
RT	Room Temperature
RT PCR	Real Time PCR
SD	Standard Deviation
SNPs	Single Nucleotide Polymorphisms

SNVs	Single nucleotide variants
TAE	Tris Acetate Ethylene diamine tetra acetate
TBE	Tris Boric Ethylene diamine tetra acetate
UV	Ultra Violet
XMEs	Xenobiotic Metabolizing Enzymes
Α	Alpha
Γ	Gamma
Θ	Theta
Μ	Mu
Π	Pi

List of Table

2.1.	Cumulative risk for tobacco related cancer (TRC) in north eastern states of India for men and women upto 64 yr (2003-2004)	13
2.2.	Possible Dietary and Other Factors Associated With Cancer in India	22
4.1A	Association of tobacco, betel quid and alcohol consumption with risk of gastric cancer in Northeast population	86
4.1B.	Association of *GSTT1*, *GSTM1* and *GSTP1* genotypes with gastric cancer in Northeast population	86
4.2A.	Region specific association of tobacco, betel quid chewing and alcohol consumption with risk of gastric cancer	87
4.2B.	Region specific distribution of *GSTT1*, *GSTM1* and *GSTP1* genotypes among gastric cancer cases and controls	88
4.3.	Combined effects of *GSTT1*, *GSTM1* and *GSTP1* genotypes in the study subjects	89
5.1.	Oligonucleotide primer sequences used for amplification of *p53* gene	100
5.2.	Distribution of demographic variables and genotypes between gastric cancer cases and controls	108
5.3.	Risk estimates for *p53* codon 72 polymorphism and co-variates in cancer patients	109
5.4.	Estimate of the effect of the *p53* codon 72 polymorphism on cancer risk modeled with logistic regression	110
5.5.	Gene-Environment interactions odds ratio for *p53* codon 72 genotypes and risk habits for cancer patients	111-112
5.6.	Frequency of *p53* Polymorphism in Gastric Cancer: Worldwide Scenario	113
6.1.	Oligonucleotide primer sequences for amplification of for *UreA, GlmM* and *16SRNA*	123
6.2.	Results of three PCR methods and ELISA for the detection of *H. pylori* from GC cases	134
6.3.	Performance of PCR methods for the detection of *H.pylori*	135

6.4.	Performance of combination of PCR methods for the detection of *H. pylori*	135
6.5.	Probability of positivity and negativity for *GlmM*, *UreA* and *16SRNA*	136
6.6.	Probability of positivity and negativity for combination of GlmM+ve- UreA+ve, UreA+ve-16S rRNA+ve, and GlmM+ve-16S RNA+ve	137
7.1.	Demographic and clinical characteristics of gastric adenocarcinoma cases	161-162
7.2.	Risk estimates for *H. pylori* status, tobacco, betel quid and alcohol consumption habits in GC patients	163
7.3.	Biologically relevant and statistically significant up-regulated and down-regulated genes in gastric cancer patients	164-168
7.4A.	Locations, function, assay ID and amplicon size of eight genes that were validated by Real-Time RT-PCR	169
7.4B.	Differential expression between *H. pylori* positive (HP+) patients.and *H. pylori* negative (HP-) patients	170
7.5A.	Matrix of gene expression of *RAD51L3*, *PXN*, *ANG*, *BMP7*, *CXCL10*, *RET*, *RYK* and *LDHB* in *H. pylori* positive patients	171
7.5B.	Matrix of gene expression of *RAD51L3*, *PXN*, *ANG*, *BMP7*, *CXCL10*, *RET*, *RYK* and *LDHB* in *H. pylori* negative patients	172
7.6.	Distribution of interactions terms and odds ratio for *H. pylori* and risk habits	173
8.1.	Demographic and clinical characteristics of gastric cancer	199
8.2.	Total number of SNPs and indels in normal gastric cancer cases	200
8.3.	List of genes with SNPs in gastric cancer	201
8.4.	Known SNPs in gastric cancer cases	202
8.5.	Novel SNPs in gastric cancer cases	203
8.6.	Known and novel indels in gastric cancer cases	204
8.7.	GO analysis of genes identified in gastric cancer	205-209

List of Figures

2.1.	Trends of cancer death rates among males for selected cancers in US population	15
2.2.	Trends of cancer death rates among females for selected cancers in US population	15
2.3.	Incidence and mortality rate of gastric cancer in male and female of developed and developing countries	17
2.4.	Comparison of age adjusted incidence rates of gastric cancer in males and females across all PBCRs	20
2.5.	Complication due to *H .pylori* leading to gastric cancer	30
2.6.	Schematic representation of the factors contributing to gastric pathology and disease outcome in *H. pylori* infection	31
2.7.	Partial gastrectomy	32
2. 8.	Total gastrectomy	32
2.9.	Genetic alterations in gastric cancer	36
2.10.	Relationship between phase 1 and 2 detoxification reaction	44
2.11.	Possible pattern of divergence in the GST superfamily resulting in multiple GST classes is shown	46
2.12.	Genetic and epigenetic changes leading to tumorigenesis	54
2.13.	Experimental workflows for performing gene expression analysis using oligonucleotide and cDNA microarrays	58
2.14.	Experimental workflow for performing mutation analysis using oligonucleotide microarrays	60
2.15.	Illumina workflow	65
4.1.	Estimation of quality of genomic DNA as visualized on an ethidium-bromide stained 1% agarose gel	77
4.2.	Agarose gel electrophosresis showing null mutations in *GSTM1* and *GSTT1* genes in gastric cancer	79

4.3.	Representative gel picture of PCR-RFLP method carried out to map polymorphism of GSTP1. Lane 1-100bp ladder (M/s Fermentas, Lithuania). Lanes 2, 3, 4 - Ile/Val (heterozygous) allele. Lanes 5, 6, 7 - Ile/Ile (homozygous) allele	81
4.4.	Electrogram showing the sequence analysis of different genotypes of *GSTP1* polymorphism in gastric cancer samples	83
5.1.	Representing PCR amplification of *p53* gene (199bp)	103
5.2.	RFLP analysis for *p53* gene codon 72 polymorphism	103
5.3.	Electrogram showing the sequence analysis of different genotypes of p53 codon 72 polymorphism	104
6.1.	(A) Schematic representation of gene orientation in *GlmM* (B) 2 % agarose gel electrophoresis of 296 bp amplicon (C) Representative picture showing sequencing of the 296 bp *GlmM* gene	129
6.2.	(A) Schematic representation of gene orientation in *UreA* (B) 2 % agarose gel electrophoresis of 411 bp amplicon (C) Representative picture showing sequencing of the 411 bp *UreA* gene	130
6.3.	(A) Schematic representation of gene orientation in *16S rRNA* (B) 2 % agarose gel electrophoresis of 428 bp amplicon (C) Representative picture showing sequencing of the 428 bp *UreA* gene	131
7.1.	Representative picture of electropherogram showing RNA quality	147
7.2.	Experimental design	148
7.3.	Box plot for raw and log2 transformed data for each array	149
7.4.	Box plot for each sample after median centering and MAD scaling	150
7.5.	Volcano plot showing up and down regulated genes	151
7.6.	Heatmap showing the clustering of samples and probes as a measure of proximity	152

7.7.	Regression plots for fold change by microarray (Y-axis) and quantitative real-time PCR assay (X-axis) for *LDHB* (A), *PXN* (B), *RYK* (C), *RET* (D), *ANG* (E), *BMP7* (F) and *CXCL10* (G)	159
8.1.	Bioanalyzer profiles of sonicated samples	188
8.2.	Process of library preparation	189
8.3.	Data analysis flow for human targeted resequencing	192
8.4.	Chromosomal map of known SNPs in gastric cancer cases. Red colour indicate the location of the SNPs	196
8.5.	Chromosomal map of known indels in gastric cancer cases. Red colour indicate the location of the SNPs	197
8.6.	Figurative representation of the catalogue of mutations in gastric cancer samples.	198

Abstract

Gastric cancer (GC) is the fourth most common cancer worldwide and remains a major health issue and a leading cause of death. Geographic and ethnic differences in its incidence has been observed and this indicate the close association of GC with modifiable factors such as enviroment. Frequency of GC is very low in India. However, a very high prevalence of GC has been observed in Mizoram, NE India. Despite the high incidence, there is no comprehensive data on the molecular pathogenesis of the disease in this region. The aim of this study was to understand the underlying mechanism involved in the carcinogenesis of GC in NE India and search for possible markers to assist in both diagnosis and therapeutic approaches.

Tobacco, betel quid and alcohol consumption habits were higher in GC cases when compared with controls, but this difference was statistically insignificant. However centre-wise analysis revealed betel quid to impart significant risk (OR=3.61, 95% CI 1.06-12.21, p=0.04) for GC in Assam population of NE India. The form of tobacco and betel quid used within different areas of NE India varies widely and this could be one of the reason for the different in their influence with the risk of GC in these areas.

The association of polymorphisms in detoxifying GST family genes including *GSTT1, GSTM1* and *GSTP1* were analysed by PCR-RFLP. Null genotypes of *GSTT1, GSTM1* and variant *GSTP1* genotypes were not found to be associated with risk of GC. However, null genotypes of *GSTT1* were found to

impart significant risk (OR=3.07, 95% CI 1.33-7.09, P=0.009) for GC in Assam population of NE India. This inconsistency of results may be due to different ethnicity or interaction between different environmental and genetic factors. Furthermore, the role of tumor suppressor *p53* gene polymorphism using PCR-RFLP method showed no significant relationship between *p53* codon 72 polymorphisms and GC risk. Additionally, We observed that among the three primers choosen (*GlmM, UreA, 16SrRNA*) that of *GlmM* seems to be a promising one for detection of *H. pylori* with higher sensitivity of 85% as compared to that of *UreA* (65%) and *16SRNA* (53%). The frequency of *H. pylori* positive cases did not show significant difference among cases and controls. Though *H.pylori* together with tobacco and betel quid chewing was found to confer a significant risk for GC. Gene expression profiling with 'OciChip Human A' microarray chip was performed. Pathway analysis revealed differentially expressed genes to be associated with pathways viz Wnt (*CHP, CSNK1E, RYK*, Notch (*JAG2*), Hedgehog (*CSNK1E, BMP7*) and epithelial signaling pathway in *H. pylori* (*ATP6V1G2, JAM3*) and are functionally relevant to carcinogenesis. To analyze the genomic alterations involved in GC in NE India targeted re-sequencing of 169 potentially important genes was carried out. The *ACTL6A, PARK2, PTPN1, TP63, TSC2, VHL* and *NGFR* were genes with SNPs only in GC and not in the normal gastric mucosa. Among the known SNPs, the SNP in *FAT1* was of high importance leading to non-synonymous change at codon position 482. Novel SNPs detected were located in coding region of *APC*, *TSG101*, *FLNC* and *CASP10.* The only novel indel detected targeted the *MADD* gene and it is an insertion of guanine base at position 47291817.

Our study does not show consistent association of the candidate gene polymorphism with GC risk, however interaction analysis reveals the significance of ethnicity, environmental and genetic factors. Further advances shows that the bacterial factors are influence by tobacco and betelquid in determining disease outcome. Pathways identified i,e wnt, notch and hedgehog have important role in carcinogenesis constituting the genetic decomposition of the host. Genes i,e *ACTL6A, PARK2, PTPN1, TP63, TSC2, VHL*, *NGFR*, *FAT1*, *APC*, *TSG101, FLNC, CASP10, MADD* might play role in bringing about an substitutional mechanism in gastric carcinogenesis as suggested by our study. It has been speculated that *H. pylori* and environmental factors may work in part with the genetic factors. Intervention to eradicate *H. pylori* and control over betelquid and tobacco consumption early in this pathological process might prevent progressively more severe disease.

Contents **Page no**

1.	**Chapter 1.**	**Introduction**	1–8
2.	**Chapter 2.**	**Review of Literature**	9–64
	2.1.	**Etiology of gastric cancer in north east India**	9-12
	2.2.	**Incidence, geographic variability and overall survival of gastric cancer**	12-15
	2.3.	**Indian scenario of gastric cancer**	16-18
	2.4.	**Diet and gastric cancer in India**	19-20
	2.5.	**Precursors of gastric cancer**	21-22
	2.5.1.	***Chronic atrophic gastritis***	21
	2.5.2.	***Intestinal metaplasia***	21
	2.5.3.	***Dysplasia***	22
	2.6.	**Pathobiology**	23-25
	2.7.	**Etiology of gastric cancer**	25
	2.7.1.	***Environmental factors***	25
	2.7.2	***Associated pathological conditions***	26-29
	2.7.2.1.	*Helicobactor pylori*	26-29
	2.8.	**Symtoms, staging and treatment**	29-33
	2.9.	**Genetic events in gastric oncogenesis**	33
	2.9.1.	***Microsatellite and chromosomal instability***	34
	2.9.2.	***Cell-adhesion and Cell-cycle regulators***	35
	2.9.3.	***Growth factors and cytokines***	36
	2.9.4.	***Oncogenes***	37
	2.9.5.	***Tumor suppressor gene***	37-40
	2.10.	**Gene polymorphisms**	40-41
	2.11.	**Xenobiotic-metabolizing genes**	41-42
	2.11.1.	***Glutathione S-transferases (GSTs)***	43-44

	2.11.1.1.	*GSTM1 null and GSTT1 null*	45-46
	2.11.1.2.	*GSTP1*	46-47
	2.11.2.	***Cytochrome P450 (CYP)***	47-48
	2.11.3.	***NAD (P)H:quinone oxidoreductase 1 (NQO1)***	48-49
	2.11.4.	***N-acetyltransferase (NAT)***	49-51
	2.12.	**Promising highthroughput techniques**	52-53
	2.12.1.	***Microarray***	53-54
	2.12.1.1.	*cDNA microarrays*	54-55
	2.12.1.2.	*Oligonucleotide microarrays*	56-58
	2.12.2.	***Microarray-based gene expression profiling in gastric cancer***	58-61
	2.12.3.	***Next-generation sequencing (NGS)***	61-62
	2.12.3.1.	*Illumina genome analyzer*	62-64
3.	**Chapter 3.**	**Aims and Objectives**	65-67
4.	**Chapter 4.**	**Determination of association of *T1*, *M1* and *P1* Polymorphism in Glutathione S Transferase Genes and Gastric Cancer Risk in Northeast Population of India**	68-90
	Introduction		
	Experimental methods		
	Results		
	Discussion		
5.	**Chapter 5.**	**Determination of association of *p53* gene, its interaction with tobacco, Betel quid and alcohol consumption and risk of gastric cancer: a case-control study in Northeast population to understand the etiology**	91-111
	Introduction		
	Experiment al methods		
	Results		

	Discussion		
6.	**Chapter 6.**	**Evaluation of PCR assays for Detection of the presence of *Helicobacter pylori* in Gastric Cancer Patients of Northeast India**	112-134
	Introduction		
	Experiment al methods		
	Results		
	Discussion		
7.	**Chapter 7.**	**Gene Expression Profiling of gastric cancer in North East India: Role of Helicobacter pylori and tobacco**	135-173
	Introduction		
	Experiment al methods		
	Results		
	Duscussion		
8.	**Chapter 8.**	**Assessing matched normal and tumor pools in next-generation sequencing studies to indentify genomic alteration in gastric adenocarcinoma in high risk northeastern region of India**	174-209
	Introduction		
	Experiment al methods		
	Results		
	Duscussion		
9.	**Chapter 9.**	**Conclusion and Future Scope of Work**	210-211
10.		**References**	212-244
11.		**Appendix**	245-247
12.		**List of Publications**	248-251
13.		**Biography of Candidate**	252-257
14.		**Biography of Supervisor**	258-274

Chapter 1: Introduction

Gastric cancer (GC) is the fourth most common cancer worldwide with 930,000 cases diagnosed in 2002 [1] and remains a major health issue and a leading cause of death [1, 2]. GC is associated with high mortality rate of ~800,000 per year [2]. One in a million people under the age of 55 seeking medical attention for indigestion has GC and one in fifty of all ages seeking medical attention for burping and indigestion have GC [3]. Decline in the incidence of GC has been observed worldwide and has been attributed to domestic refrigeration, decrease in the intake of salted, smoked and pickled foods and the greater availability and consumption of fresh fruit and vegetables. There are other environmental factors apart from diet, that increase the risk of GC in susceptible individuals including tobacco smoking [4], alcohol consumption and *Helicobacter pylori* (*H. pylori*) infection [5].

Gastric carcinogenesis is a multistep and multifactorial process, result due to complex interaction between environmental and genetic factors [6]. In addition, there is remarkable regional and ethnic variation in the incidence of GC compared to cancers in other organs. Countries from South East Asia, including India reported lower incidence of GC [7, 8], however, high prevalence of GC has been reported in Northeast (NE) India. The people of NE India have lifestyle and dietary habits different from those in other parts of India [9]. The area report high incidence of tobacco usage in different forms [10]. NE part of India, due to its unique, strategic geographic location and the presence of linguistically,

culturally and demographically diverse populations is a hotspot for population genetics. This area shows the world's highest incidences of cancers of tongue and stomach particularly in men, associated with tobacco use [11].

Inter-individual and inter-ethnic differences play an important role in determining chemical exposure risk and detoxification. Inter-individual variation is seen for effects of xenobiotics because their detoxification and excretion from the body depends on the presence, efficiency and concentration of detoxifying enzymes, which is ultimately decided by the genetic constitution of that person. This seems a reason behind variability in the occurrence of various diseases in different ethnic groups. Humans are exposed to a number of xenobiotics. Effects of these toxic chemicals on the body are determined by quantitative description of absorption, distribution, metabolism and excretion [12, 13]. Many polymorphism on the xenobiotic detoxifying genes have been observed and its association studies in relation to different cancer is conflicting [14-17].

The glutathione S-transferase (*GST*) gene superfamily encodes enzymes that catalyze the conjugation of glutathione to electrophilic compounds. These enzymes generally detoxify endogenous and exogenous agents, but also participate in the activation and inactivation of oxidative metabolites of carcinogenic compounds [18]. Variant *GST* alleles have been identified within the general population. The most extensively studied variant *GST*s include two *GST* deletion alleles (i.e., *GSTM1**0/*0 [GenBank: BC024005.2] and *GSTT1**0/*0 [GenBank: BC007065.1]) and the *GSTP1* Val allele which is characterized by an adenine to guanine substitution at position -313 ($A^{-313}G$) in

exon 5 [GenBank: BC010915.1; dbSNP: rs1695] [19]. The functional consequences of the *GSTM1* and *GSTT1* (*0/*0) genotypes are obvious in terms of enzyme activity; gene deletion results in loss of conjugation potential. The *GSTP1* polymorphism, resulting from an isoleucine to valine substitution within the active site of the enzyme at codon 105 ($I^{105}V$), is linked to altered substrate-specific thermostability and conjugation activity [18]. For instance, the *GSTP1* 105 Val variant has been associated with lower efficiency for diol epoxides of some polycyclic aromatic hydrocarbons, therefore resulting in decreased detoxification of these compounds compared to the Ile allele [20]. Genetic variations in polymorphic *GST* genes have been implicated in the etiology of numerous cancers [21-25]. As mention earlier, GC has a sharp variation in its geographic distribution [26]. The high incidence in special areas indicates the importance of environmental factors in gastric carcinogenesis. However, only a small part of individuals in the high-risk area for GC develop into GC, although all the residents in that area share very similar environment-related risk factors and lifestyle, suggesting that host susceptibility factors, such as the polymorphisms of detoxifyimg enzymes, may play an important role in increased risk for GC.

Beside the genes coding the xenobiotic detoxifying enzyme another important gene which has been reported to be frequently mutated in human cancer is p53 tumor suppressor gene. It has been found to be mutated in more than 50% of human cancers, it has attracted the interest of numerous researchers. The capacity of p53 for multiple biological functions can be attributed to its ability to act as a sequence-specific transcription factor to

regulate expression of over one hundred different targets, and thus to modulate various cellular processes including apoptosis, cell cycle arrest and DNA repair [27]. The p53 belongs to an unique protein family which includes three members: p53, p63 [28] and p73 [29]. Although these proteins are structurally and functionally related to each other, p53 seems to have evolved in higher organisms to prevent tumor development, whereas *p63* and *p73* have clear roles in normal developmental biology [30]. Because *p53* plays a key role in regulation of the cell cycle and induction of apoptosis, there has been enthusiasm about its potential for therapeutic application. The fact that *p53* was originally described as an oncogene has recently come full circle with mutant *p53* having been shown to exhibit gain–of–function properties that actually drive tumour progression and metastasis [31]. The *p53* codon 72 Arg right curved arrow Pro polymorphism has been suggested to be associated with risk for different kind of cancers [32-36], but the data on GC is very limited [37]. Codon 72 polymorphism is a single base substitution of cytosine for guanine, leading to arginine (A72) being replaced by proline (P72) [38]. The results are conflicting with Pro/Pro genotype showing association with lung cancer [39-42] breast cancer [43, 44] and GC [45] whereas Arg/Arg genotype being more prevalent in cervical cancer [46, 47]. However, no association between either genotype and cancer risk have also been reported for head and neck [48] and cervical cancer [49]. Literature available from India is limited and inconsistent. Two different studies have reported both Arg/Arg and Pro/Pro genotypes to be associated with risk of lung cancer [42, 50] whereas no association was reported with oral cancer [32, 51]. There are no reports on

association of *p53* codon 72 polymorphism with GC from India. Studies on codon 72 polymorphism have revealed striking ethnic differences [52, 53] and have demonstrated that frequency of *p53* variant allele varies with latitude, increasing in a linear trend as populations near the equator. Thus ethnicity might be related to allelic distribution of the gene and its determinacy in disease involvement [54]. Studies investigating the association between *p53* codon 72 polymorphism and GC risk report conflicting results. Recent meta-analysis suggests that the *p53* codon 72 polymorphism may be associated with GC among Asians [55]. Studies are needed to substantiate and to explore the contribution of *p53* codon 72 polymorphism in the etiology of GC.

Helicobacter pylori (*H. pylori*) causes gastritis and peptic ulceration and it is an important risk factor for gastric adenocarcinoma [1]. *H. pylori* strains differ, and possession of specific virulence factors greatly increases the risk of disease [2]. It was observed that the early acquisition of *H. pylori* infection in childhood resulted in pangastritis in adulthood. This pattern of gastritis is usually associated with mucosal atrophy which is a precancerous condition for GC [3]. Therefore, acquiring the infection at an early age is a recognised risk factor for the development of GC [56]. The identification of *H. pylori* has not only revolutionized our understanding of peptic ulcer disease but is also changing our understanding of GC. *H. pylori* is able to create its own microenvironment, resistant to gastric acid, and is therefore difficult to eradicate. The World Health Organization in 1984 classified *H. pylori* as a class 1 carcinogen. It is found in 70 to 95% of all GCs. When infected with *H. pylori*, the relative risk for the

development of GC increases to 2.1. On assuming that approximately 50% of the world is infected with this organism, this bacteria becomes responsible for 42% of GCs worldwide [57]. Obviously, more research on this bacteria and its effect on cancer needs to be done.

Attempt has also been made to identify novel GC–related genes using microarray to understand the molecular mechanism associated with gastric carcinogenesis. Array technologies are accurate and comprehensive ways of simultaneously analyzing the expression of thousands of genes and have been rapidly applied in many research fields [58] . Several studies have been carried out to clarify gene expression changes in GC tissues [59]. To be mention study by Jeong-Min Kim et al has identified forty genes as either up-regulated or down-regulated genes in human GC cells. Among these, genes such as *SKB1*, *NT5C3*, *ZNF9*, *p30*, *CDC20*, and *FEN1*, were confirmed to be up-regulated genes and genes such as *MT2A* and *CXX1* were identified as down-regulated genes. Study by Kim, J.M. et al had identified novel genes viz *CDC20* and *MT2A* related to human GC using cDNA microarray [60]. The gene identified by microarray technique are valuable resource for understanding the molecular mechanism associated with tumorigenesis of gastric carcinogenesis and for the discovery of potential diagnostic markers of GC. Recent advances in microarray techniques have enabled to study the expression of many genes simultaneously. Another high throughput technology i,e Next generation sequencing (NGS) has moved the expression analysis to the genomic level, which is leading to new discoveries and understandings at much faster pace. Gastric carcinogenesis is a multistep process

involving genetic and epigenetic alteration [61]. Single nucleotide variations (SNPs), present as either germline or somatic point mutations, are essential drivers of tumorigenesis and cellular proliferation in many human cancers [61]. NGS provide the plateform where a large number of genes can be analysed for mutation detection in a short period of time with an effective cost [62]. High throughput and lower per base cost in the technique of NGS allows researchers to focused in designing sequencing projects to fulfill the aim of their research. NGS has revolutionized biomedical research [63]**.** The NGS study on GC will reveal the genomic landscape of the disease.

It has been demonstrated that different genetic pathways lead to GC [64], the individual genetic susceptibility to GC probably involves many genes [65], although their effects may only be small. As a reason study of genetic alteration and gene expression using microarray techniques and NGS which take in a large number of genes reveal that high throughput methods, such as microarrays and NGS are increasingly being used to systematically compare molecular features of individual cancers to key clinical parameters. More and more studies have shown that microarray and NGS technology is a powerful and revolutionary tool for biological and medical researches. The knowledge of the molecular events involved in the development of gastric carcinoma is far from complete. These technologies may be useful in identifying new molecular genetic markers, and further work may determine whether the markers identified can be employed to help patients into different treatment regimens. Little is known about the exact expression changes in tumorigenesis, which will help us identify the events that

leads to the initiation and progression of cancer development. Combined use of these advance technique may offer improved outcomes for GC patients and might provide new threads of hope for GC treatment.

In order to reduce the risk of GC, we need to understand its biology. Its pathogenesis can be modeled a stepwise progression from normal mucosa to cancer. GC is a cause of significant morbidity and cancer-related mortality worldwide. Although the understanding of the biology of this disease is increasing, the development of biologically targeted therapies for GC has been inadequate. Thus, studies involving different facet are needed to confirm the findings and to examine the interaction between genotypes and environmental factors in the causation of GC.

Chapter 2: Review of Literature

2.1. Etiology of GC in Northeast India

Among human cancer, gastric carcinogenesis appears to be a complex multistep process with a multi-factorial etiology, where environmental, geographical, and genetic factors appear to play major roles [66]. Several tobacco constituents, including nitrosamines, polycyclic aromatic hydrocarbons, aromatic amines, various aldehydes and phenols, may be causally related to GC [67]. In Mizoram, NE India, medical practitioners observed very high prevalence of GC. In Mizoram, the prevalence of GC among females (AAR = 14.4/105) exceeds that of breast (AAR = 3.4/105) and cervix (AAR = 1.8/105) cancer. Prevalence-wise, Mizoram stands at the first position among the states of India and at the fifth position in the world (both sexes) as far as GC is concerned [9]. Mizoram is situated between 92.15´– 93.29´ E long and 21.58´–24.35´N lat. The Mizo people have their ancestral origin in China [68]. The people of NE are culturally and ethnically distinct from the other tribes and communities of India. There is peculiar smoking habits and use of other tobacco products. A number of smoking and smokeless tobacco products are in use all over the world. But unlike other smokeless tobacco products, a unique tobacco smoke–infused water is used in Mizoram and is locally known as tuibur. This product is made locally by passing smoke, generated by burning tobacco, through water until the preparation turns cognac in color and has a pungent smell. In vitro studies using the allium root test show the toxic nature of tuibur [69]. Indigenous crude devices are used for the

production of tuibur on small scale. Users take about 5 to 10 mL tuibur orally and keep it in the mouth for some time and then spit it out. Most of the users take it several times a day. Meiziol, a local cigarette made from vaihlo (Nicotiana dadacum) tobacco. After plucking, the tobacco leaves are thrashed by feet until the leaves become soft and most of the juices flow out. Then they are dried in the sun or sometimes in a warm place like over the fireplace without applying direct heat. Then they are cut into small pieces and rolled directly using a thin paper. The tobacco content of each meiziol is about 0.8 to 1 g. The length of each meiziol is 6 to 7 cm [10]. There are more than 60 carcinogens in cigarette smoke and at least 16 in unburned tobacco. Among these, tobacco specific nitrosamines (such as 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone and *N′*-nitrosonornicotine), polycyclic aromatic hydrocarbons (such as benzo[*a*]pyrene) and aromatic amines (such as 4-aminobiphenyl) seem to have an important role as causes of cancer [70]. Tredaniel J et al have estimated that over 80,000 cases of GC (11% of all estimated cases) may be attributed to tobacco smoking each year [71]. Tobacco use in the NE region was observed to be very high [72]. This is likely to be the major reason for a higher incidence of GC in this region**.** Aizwal district of Mizoram and Kamrup district of Assam showed highest lifetime risks with one out of every 13 men and one out of 15 men are likely to develop tobacco related cancer (TRC) [72]. The cumulative risks were computed for various regions of NE and shown in the **Table 2.1**.

Table 2.1. Cumulative risk for tobacco related cancer (TRC) in Northeastern states of India for men and women upto 64 yr (2003-2004)

	NE region							
Male:	Silchar	Dibrugarh	Kamrup	Imphal West	Mizoram	Aizwal	Mizoram excluding Aizwal	Sikkim
TRC cum. risk %	3.9	3.3	6.6	2.4	4.4	7.9	2.4	1.6
No.of males of whom one may develop TRC	26	30	15	42	23	13	42	62
All cancers cum. risk %	7.0	5.6	10.8	5.5	11.4	15.7	8.9	4.7
No.of male of whom one may develop cancer	14	18	9	18	9	6	11	21
Female:								
TRC cum. risk %	1.5	1.3	3.1	1.5	2.1	3.2	1.5	1.2
No.of female of whom one may develop TRC	65	78	32	68	48	31	69	84
All cancers cum. risk %	4.9	5.0	10.5	6.5	10.5	14.5	8.1	5.8
No.of females of whom one may develop cancer	21	20	10	15	10	7	12	17

All cancers: cancers of all sites
TRC, Tobacco related cancer; Cum risk, Cumulative risk

(Source: Satyanarayana *et al* Indian J Med Res 128, September 2008, pp 318-319)

The cumulative risk per cent for TRC among males ranges from 1.6 (*i.e.*, one out of every 62) in Sikkim to 7.9 (one out of every 13) in Aizwal district of Mizoram. Among females the risk ranges from 1.2 (one out of every 84) in Sikkim to 3.2 (one out of every 31) in Aizwal. There were noticeable differences between the NE region and other regions of India. The risk of developing TRC was 1 out of every 26 males in Bhopal and 1 out of every 47 males in Bangalore. For females, this was 1 out of every 67 in Chennai and 1 out of every 112 in Delhi. The risk in Barshi rural registry was one out of 104 every males and one out of every 213 females [73]. Out of risk for developing all types of cancers in the NE regions of India, the proportion of TRC risk ranged from 27 to 61 per cent among males and from 19 to 31 per cent among females **(Table 2.1).** Aizwal district of Mizoram

and Kamrup district of Assam showed highest lifetime risks with one out of every 13 men and one out of 15 men are likely to develop TRC respectively. The risk among female of these districts was also to the extent of one out of 31. Health education and tobacco cessation services should be targeted at these high risk populations to prevent sufferings and deaths that could be caused due to tobacco use [72].

The large variation in the incidence of GC in different geographic regions has often been thought to be due to variation in exposure to environmental factors. The different in the etiology of GC in NE population from western population and other Indian population might be due to wide variations in dietary habits or nutritional factors, tobacco and alcohol habits. A better understanding to understand the complex mechanisms involved in the carcinogenesis of GC in NE region of India in association with the different environmental and diet factors are a need of the day.

2.2. Incidence, geographic variability and overall survival of gastric cancer

GC is one of the most common cancers in Asia. Although its incidence in other regions is lower, it is still a major health problem worldwide. [74]. Most cases were reported from developing countries. In Japan, the age-standardized incidence of GC ranged from 60 to 92 per 100,000 in men and from 24 to 39 per 100,000 in women. Among the white population in the U.S.A., the incidence was one-tenth of that observed in Japan i,e 6.6 per 100,000 in men and 2.6 per 100,000 in women. Marked geographic variability is also observed [75-79]. The

trends of death rates in US population due to different cancers showed a remarkable decline in GC **Figures 2.1** and **2.2.**

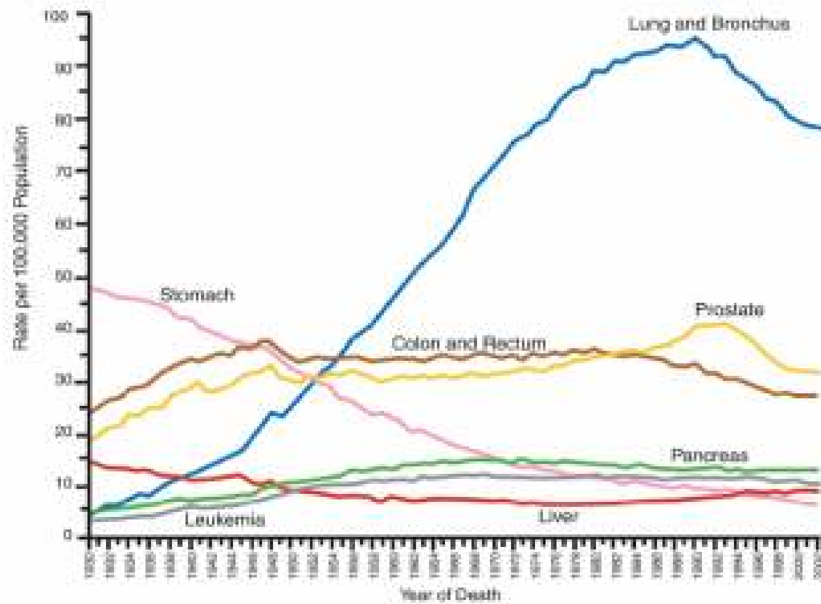


Figure 2.1. Trends of cancer death rates among males for selected cancers in US population. Source: National Center for Health Statistics, Centers for Disease Control and Prevention.©2009

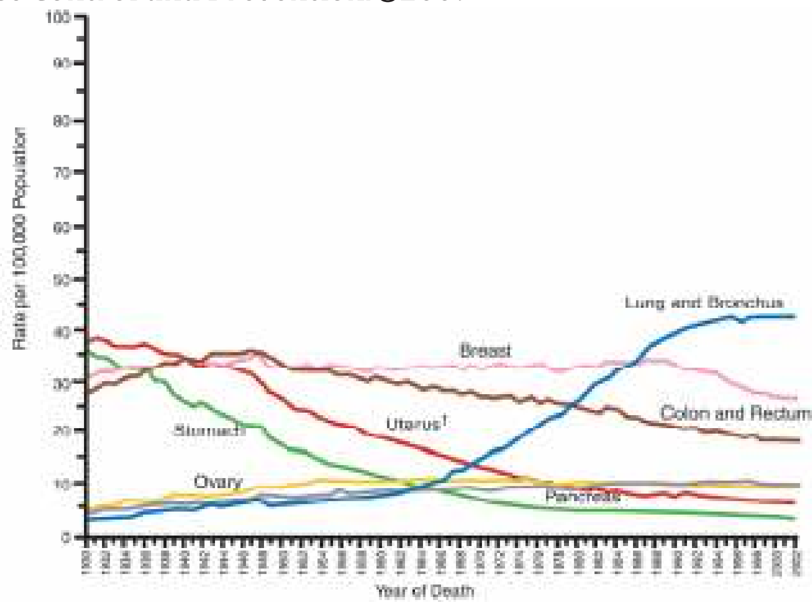


Figure 2.2. Trends of cancer death rates among females for selected cancers in US population. Source: National Center for Health Statistics, Centers for Disease Control and Prevention.©2009

The geographical areas with more cancer deaths (per 100,000) in 2007 were Japan, China, Latin America, parts of Eastern Europe, and Portugal [80]. The peak incidence of GC is estimated to occur at 50-70 years, as this tumor is rare before 30 years of age [81, 82]. Males are affected more often than females [76, 77]. The incidence and mortality rate of GC is higher in both males and females in developing countries [1] **(Figure 2.3)**. Prognosis is generally rather poor, with 5-year relative survival below 30% in most countries [83]. The 5-year relative survival rate for all cases in US from 1995 to 2001 is only 23% [80]. In Europe, the relative survival from GC in 1990-1994 was poor in both sexes: 42% at 1 year and 23% at 5 years [84]. The younger patients (under 45 years) had the higher 5-year survival (35%), while in patients over 74 years it was only 17%. In Finnish Cancer Registry the 5-year survival is closely dependent on the stage of the tumor, which means only 3% survival for lesions with distant metastasis and 61% for still localized malignancies [85, 86]. Although this disease is better understood now, low survival rates persist due to the lack of suitable and specific biomarkers for early detection, with most cases being diagnosed in the late stages.

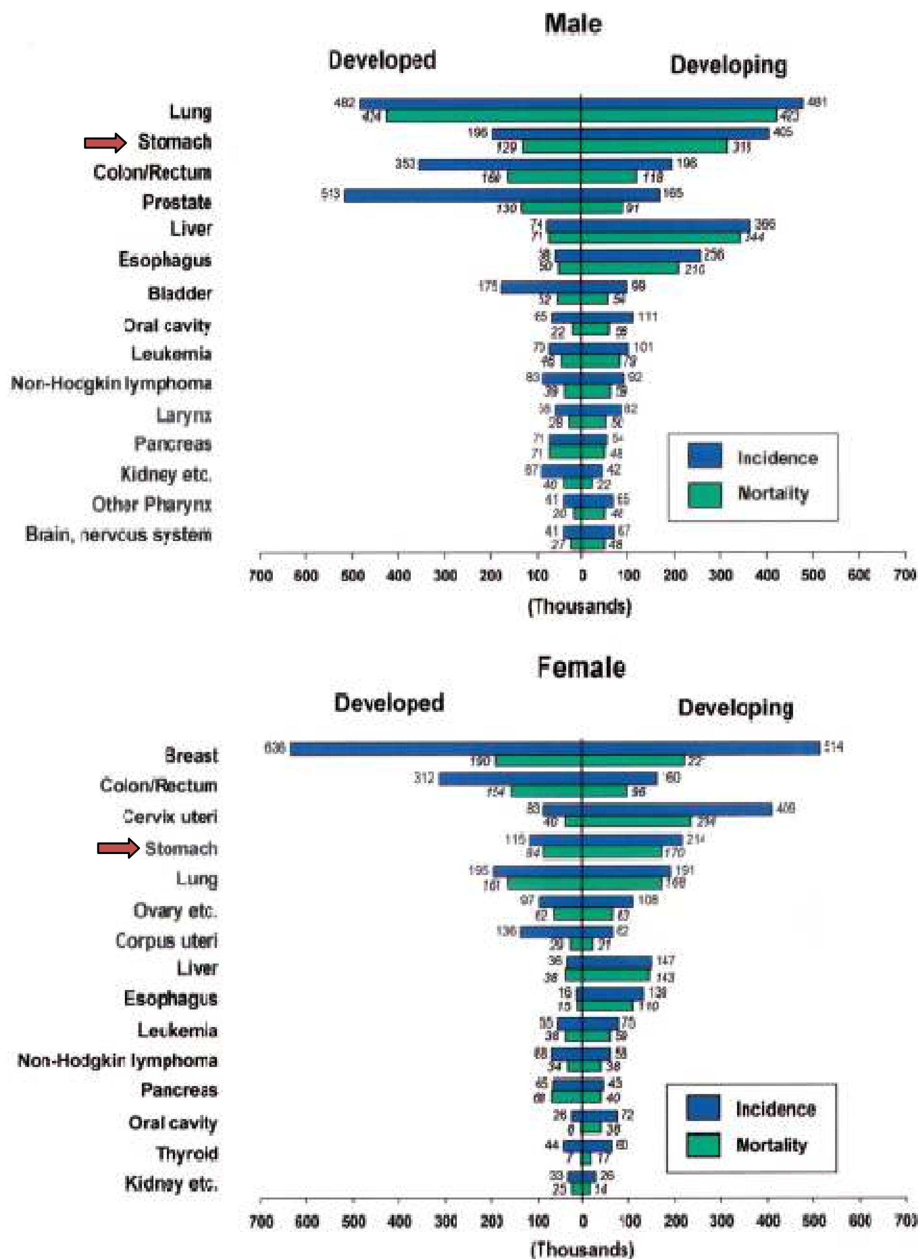


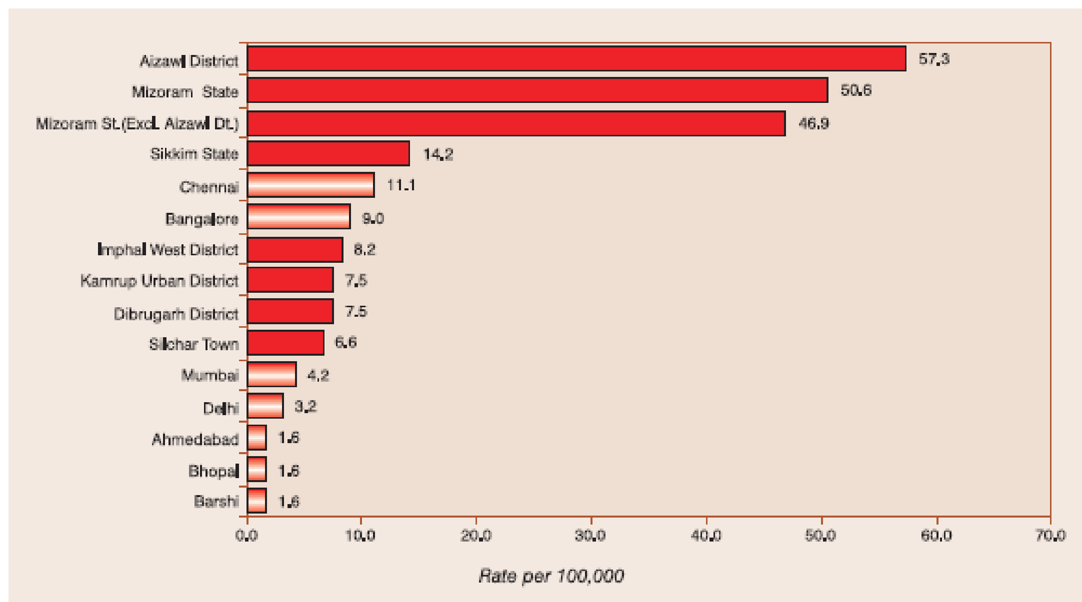
Figure 2.3: Incidence and mortality rate of gastric cancer in male and female of developed and developing countries. Source: Global Cancer Statistics, 2002: CA Cancer J Clin. 2005 Mar-Apr; 55(2):74-108

2.3. Indian scenario of Gastric Cancer

India is a developing country with one of the most diverse populations and diets in the world. Overall cancer rates in India are lower than those seen in Western countries, but are rising with increasing industrialization, urbanization, migration of rural population to the cities, increase in life expectancy and changes in diet habbits and lifestyle. Cancer rates in India are rising as development progresses, with a changing profile of burden at different cancer sites. Data from population based cancer registries (PBCRs) in India show that the most frequently reported cancer sites in males are lung, oesophagus, gastric, and larynx [87]. Cancer data from both population-based and hospital-based cancer registries in India, showed the highest incidence of GC in Mizoram in the NE of the country. Among males in Aizawl district (AAR 57.3), Mizoram state (AAR 50.6), Mizoram state excluding Aizawl district (AAR 46.9) and Sikkim state (AAR 14.2) had higher AARs among all the PBCRs. Among females all the three identified areas in Mizoram state i.e. Aizawl district (AAR 33.6), Mizoram state (AAR 23.3) and Mizoram state excluding Aizawl district (AAR 17.2) had higher AARs followed by Chennai (AAR 5.6) **(Figure 2.4)** [88]. Mizoram is situated between 92.15′ and 93.29′E longitude and 21.58′ to 24.35′N latitude; it is a virtually land-locked area situated between Myanmar in the east and Bangladesh in the west [89]. The people of this region are culturally and ethnically distinct from the other tribes and communities of NE India. The Mizo people have their ancestral origin in China [90]. Fish, pork, beef, and other meats are popular nonvegetarian foods in Mizoram. They preserve them by smoke-drying and

salting for future consumption. Sa-um (fermented pork fat) and bekang (fermented soya bean) are two other unusual foods in Mizoram [89]. With the distinct life styles and cultures, this region is the paradise for epidemiologists who seek the etiological clues from the geographic variations in the occurrence of the disease and the relationship between the genetic and environmental factors. Inspite of high incidence rate of GC in NE India, India ranked moderate to low in the incidence GC compared to other country.

MALES



FEMALES

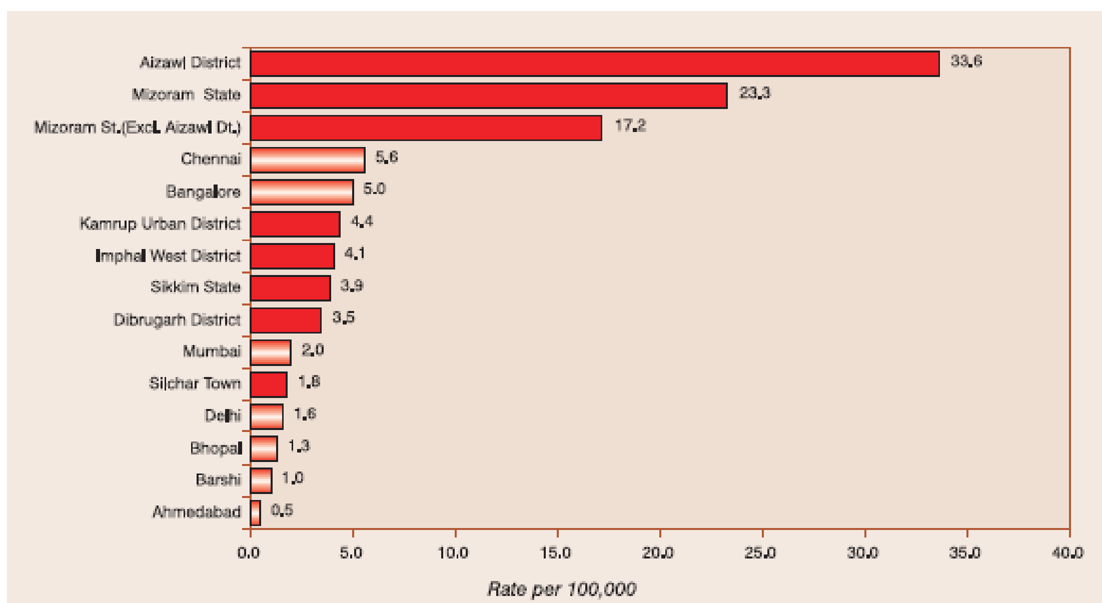


Figure 2.4: Comparison of age adjusted incidence rates of gastric cancer in males and females across all PBCRs. Source: NCRP report: 2006-2008

2.4. Diet and gastric cancer in India

Dietary changes, reductions in physical activity, and increasing obesity due to urbanization have been associated with the risk of GC occurrence. Obesity and lack of physical activity are associated with increased risk at various cancer sites. Among urban populations, energy intake has increased at the same time that energy expenditures have decreased, due in part to in industries reliant on mechanization [91]. Diets high in saturated fats have been associated with increased risk for cancer. Large epidemiological studies have identified a possible association between increased dietary fibre and a decreased risk for cancers of the colon and breast. The Indian diet, which generally includes adequate levels of vegetables, fruits, and fiber-rich grains, may provide some protection against increased risk for GC [91]. Turmeric also has been found to inhibit the growth of 19 clinical strains of *H. pylori* [92]. Amrita Bindu, a dietary supplement that is a salt-spice-herbal mixture, was found to protect rats against cancer induced by *N*-methyl-*N*-nitrosoguanidine, a potent carcinogenic nitrosamine [93].

A case-control study in Mumbai found that consumption of dried fish (OR=12.4) increased the risk while green tea consumption (OR=0.4) decreased the risk of having GC [94]. A prospective case-control study from Trivandrum evaluated dietary risk factors for GC and found that high consumption of rice (OR=3.9), spicy food (OR=2.3), chili (OR=7.4), and high-temperature food (OR=7.0) increased the risk of developing GC [95]. Studies of mice fed with deep-fried vegetables and salted or sundried Ribbonfish, which is commonplace in Indian cooking, reported a 20 percent increase in gastric carcinoma. The rise has

been attributed to the presence of polycyclic aromatic hydrocarbons produced by the cooking method [96]. As mentioned earlier, the use of the spice turmeric is associated with a reduced risk of GC, in part because of its protective effect against the carcinogenic bacterium *H. pylori*, a major risk factor for GC [97]. Several prospective epidemiological research studies exploring the relationships between diet and lifestyle and cancer have been carried out in India and it is an important factor in cancer aetiology (**Table 2.2**).

Table 2.2: Possible Dietary and Other Factors Associated With Cancer in India

	Decreased Risk	**Increased Risk**
Oral cancer	Diet high in vegetables and fruits [98], Fish [98], Eggs [98]	Betelquid chewing [99], Reverse smoking (palate)[100]
Esophageal cancer	Diet high in vegetables [101]	Betel quid chewing [101],Chillies [102],Salted tea [102], Kalakhar [103]
Endometrial cancer	Diet high in vegetables and fruits [104], Diet high in carotenoids [104]	High body mass index [104], Saturated fat intake [104], Human papillomavirus [105]
Cervical cancer	Vitamins C and E [104]	Human papillomavirus [105], Tobacco use [104]
Ovarian cancer	Diet high in fish [104]	Saturated fat intake [104], Human papillomavirus [105]
Breast cancer	Diet high in vegetables and fruits [99], High physical activity [104]	Diet high in saturated fats [104],High body mass index [99], Saturated fat [104]
Stomach cancer	Green tea [94],Turmeric [92], Cumin [92], Basil.[92], Tapioca [95]	Dried fish [106],High-temperature foods [95], Chillies [95],Spicy foods[95], High consumption of rice[95], *H. pylori* [92]

2.5. Precursors of gastric cancer

Certain clinical or histological conditions have been recognised as precursors of GC. These include chronic atrophic gastritis, adenomatous gastric polyps, previous gastric resections and possibly chronic gastric ulceration [107].

2.5.1 *Chronic atrophic gastritis*

Chronic atrophic gastritis is a very common histological finding, but there is evidence of it to be a preliminary stage in the histogenesis of GC. Imai et al [108] found chronic gastritis in the antrum of 58% japaneses and in the corpus of 41% Japanese, but in the antrum of 18% American and in the corpus of 22% American. Surgical resection specimens showed more frequent gastritis. These findings suggest a positive association between chronic gastritis and GC. Similar results were found by Cheli et al [109] comparing Italian and Hungarian populations by endoscopic biopsy of asymptomatic subjects. Siurala et al [110] found that GC developed in 9.5% of patients with atrophic gastritis, 1% of those with superficial gastritis, and none of those with normal histology. Walker et al [111] found a cancer incidence of 10% after a mean of 15 years follow up after endoscopic biopsy diagnosis of atrophic gastritis.

2.5.2. *Intestinal metaplasia*

The epidemiological evidence linking atrophic gastritis to GC also indicates an association between intestinal metaplasia and GC, since the prevalence of intestinal metaplasia correlates well with the incidence of atrophic gastritis especially in the high risk areas [112]. Certain GC, particularly of the intestinal type, are believed to arise from areas of intestinal metaplasia. Lauren

[113] found atrophic gastritis and intestinal metaplasia bordering almost all of his cases of intestinal type GC. Rubio et al [114] found a significantly higher intestinal metaplasia index in patients with intra-mucosal adenocarcinoma, especially those of intestinal type, than in patients with gastric ulcer. Matsukura et al [115] examined the distribution of intestinal metaplasia in stomach specimens and found that 95% of intestinal type carcinomas were surrounded by intestinal metaplasia. Morson [116] showed that cancer bearing tissue of stomach contain more intestinal metaplasia than age matched benign tissue.

2.5.3. *Dysplasia*

The main features of dysplasia is characterized by abnormalities of cytology, differentiation and glandular architecture. Dysplasia is a more selective marker of increased cancer risk unlike that of atrophic gastritis and intestinal metaplasia which are found in high frequencies and good evidences are available of atrophic gastrities and intestinal metaplasia linking with the development of GC [117, 118]. Cytological abnormalities of dysplasia include nuclear pleomorphism, prominent nucleoli, loss of nuclear polarity within the cell, as well as an increased nuclear/cytoplasmic ratio, pseudostratification of the nuclei, and cytoplasmic basophilia. Dysplasia range from mild, moderate to severe dysplasia falling just short of the diagnosis of carcinoma [119]. Cuello et al [120] found a positive correlation between dysplasia, and the gastric juice nitrite content, which may be relevant to the in vivo production of carcinogenic nitroso compounds. Jass [121] recognised heterogeneity of dysplasia with special reference to mucin histochemistry. Dysplasia is of two type, type I and type II.

Type I dysplasia was the classical form resembling a colonic adenoma. Type II dysplasia was found in association with incomplete intestinal metaplasia and appeared to have two distinct cell populations: goblet cells, and an intermediate cell. This form was found in association with poorly differentiated intestinal type carcinomas. Jass speculated that type I dysplasia could proceed to malignant transformation by progresssive deterioration, whilst type II might be more unstable and capable of direct malignant transformation even when of an apparently mild grade. Dysplasia especially if severe, is generally accepted as a precancerous lesion with a high likelihood of a malignant outcome and requiring repeated biopsy and endoscopic surveillance [121].

2.6. Pathobiology

GCs are classified according to their gross and histologic morphology and additionally according to their location in the stomach. Overwhelmingly, the most common type of GC is gastric adenocarcinomas and its frequent site is gastric antrum with lesser curvature involved more often than greater curvature. Gastric adenocarcinoma may be classified into intestinal or diffuse based on criteria purposed by Lauren wherein the proportion of intestinal type accounts for approximately 50%, that of the diffuse type 35% and the remainder 15% is characterized as ''unclassified'' or mixed type cancer. The intestinal type is characterized by cohesive neoplastic cells forming glandlike tubular structures, whereas in diffuse type cell cohesion is absent, so that individual cells infiltrate and thicken the stomach wall without forming a discrete mass. This difference in

microscopic growth pattern is also reflected in the different macroscopic appearance of the two histological subtypes [113, 122-124].

Both tumour types however, can have large extracellular mucin lakes that dissect tissue plane. However, the macroscopic margins in the intestinal type correspond approximately to the microscopic spread whereas the diffuse type can extend submucosally far beyond its macroscopic borders. This difference in tumor spread of the two types of Lauren-classification is of clinical importance in decision-making about appropriate treatment option. In the intestinal type, there is a progression from normal mucosa to chronic inflammation, leading to chronic gastritis. This can develop into atrophic gastritis to intestinal metaplasia to dysplasia and finally to frank carcinoma. In the diffuse type normal mucosa progresses eventually from chronic gastritis directly to cancer. This type of gastric adenocarcinoma is much more difficult to diagnose and treat as compared to intestinal type. The decline in overall incidence of gastric carcinoma during this century appears to be largely attributable to a decrease of the intestinal type lesions, while the occurrence of diffuse type is thought to have remained more stable [122, 123, 125].

Beside adenocarcinoma occurring much more rarely are gastric lymphomas, carcinoids and GIST [126]. Among all gastric maglignancy nearly 5% are lymphomas and these are often referred to as mucosa associated lymphoid tissue (MALT) or MALTomas which usually arise at sites of chronic inflammation, most common of which is chronic *H. pylori* infection. MALTomas can however transform into more aggressive tumour similar to large B-cell lymphomas.

Carcinoids arise from the endocrine cells that release peptide and nonpeptide hormones to coordinate gut function. Grossly, carcinoids form small polyploid lesions, yellow or tan in colour. Gastric carcinoid tumours may arise in association with atrophic gastritis while others without predisposing factors are more aggressive. As regard other non epithelial tumours, gastrointestinal stromal tumour is the most common mesenchymal tumour of the abdomen, and more than half of these occur in the stomach [126, 127].

2.7. Etiology of gastric cancer

Due to demographic variability and recent changes in GC incidence, much emphasis has been placed on studying etiologic and risk factors in GC. Environmental factors, occupational factors, associated pathological conditions, genetic and epigenetic factors all play role in the development of this disease.

2.7.1. *Environmental factors*

Dietary intake data support the role of certain foods as risk factors for the development of GC. Lack of, or infrequent consumption of vegetables and fruits, is a risk factor for GC [128]. Fried, very salted, cured and smoked foods are diet components contributing to the development of gastric carcinomas. Carcinogens in cooking fumes, possibly heterocyclic amines formed during high temperature cooking certain foods such as red meat and meat sauce may play role in the development of GC [129]. Several N-nitroso compounds, present in foods and beverages or formed in the stomach from their precursors, act as alkylating agents. Mean dietary nitrate intake was significantly higher in GC patients supporting that N-nitroso compounds from dietary sources may play a role in the

etiology of GC [130]. Cigarette smoking and alcohol may promote gastric carcinogenesis, and preventive measures addressing these factors could considerably reduce the incidence of GC. In a recent population-based, prospective cohort study in Norway, no statistically significant associations between various degrees of exposure to alcohol and risk of GC was revealed, but combined high use of cigarettes and alcohol increased the risk of noncardial GC nearly 5-fold, compared to nonusers [131]. It has been suggested that polymorphisms, detoxifying enzymes and smoking may alter the susceptibility to cancer development in the stomach [132]. Prospective studies on cigarette smoking and GC have suggested that GC is a tobacco-related disease [133, 134]. Some authors have described a possible effect of ethanol in promoting GC at the distal segment in patients abusing alcohol [135].

2.7.2. *Associated pathological conditions*

Particular medical conditions, such as infection by *H. pylori* and/or Epstein-Barr Virus (EBV), adenomatous gastric polyps, pernicious anemia [136-138], intestinal metaplasia [136], partial gastrectomy [136], which decreases gastric acid or achlorhydria [136] and chronic atrophic gastritis, could promote GC.

2.7.2.1. *Helicobactor pylori*

H. pylori is a spirally-shaped, microaerophilic, Gram-negative bacterium with urease, catalase and oxydase activities and with a tuft of sheated unipolar flagella [139, 140] that colonizes the stomach in about 50% of all humans. These features, together with an unusual resistance to acidic pH conditions, allow the

bacterium to survive in the stomach lumen. The gastric antrum is its most favorite site, but other parts of the stomach may be colonized. *H. pylori* is the first formally recognized bacterium as a category 1 carcinogen by the World Health Organization International Agency for Research on Cancer [141] and is also one of the most successful human pathogens, as over half of the world's population is colonized with this bacterium [139].

In countries with high socio-economic standards infection is considerably less common than in developing countries. Infection is typically contracted in early childhood, frequently and the bacteria may remain in the stomach for the rest of the person's life and this chronic infection is initiated in the lower part of the stomach [142]. As first reported by Robin Warren, the presence of *H. pylori* is always associated with an inflammation of the underlying gastric mucosa as evidenced by an infiltration of inflammatory cells [143].

In most individuals *H. pylori* infection is asymptomatic. However, about 10-15% of infected individuals will some time experience peptic ulcer disease. Such ulcers are more common in the duodenum than in the stomach itself. Severe complications include bleeding and perforation **(Figure 2.5)**. *H. pylori* represents an additional risk factor for GC. This infection may lead to inflammation of the superficial layer of the stomach, inducing changes such as atrophic and chronic gastritis, peptic ulcer etc in the cells that line the stomach which eventually results in GC **(Figure 2.6)**. It should be stressed, however, that most patients with *H. pylori* infection do not develop GC. The independent association between peptic ulcer disease and GC remains questionable. Although bacterial factors

have an important role in disease pathogenesis, most of the evidences suggest that host factors are paramount in determining progression to GC.

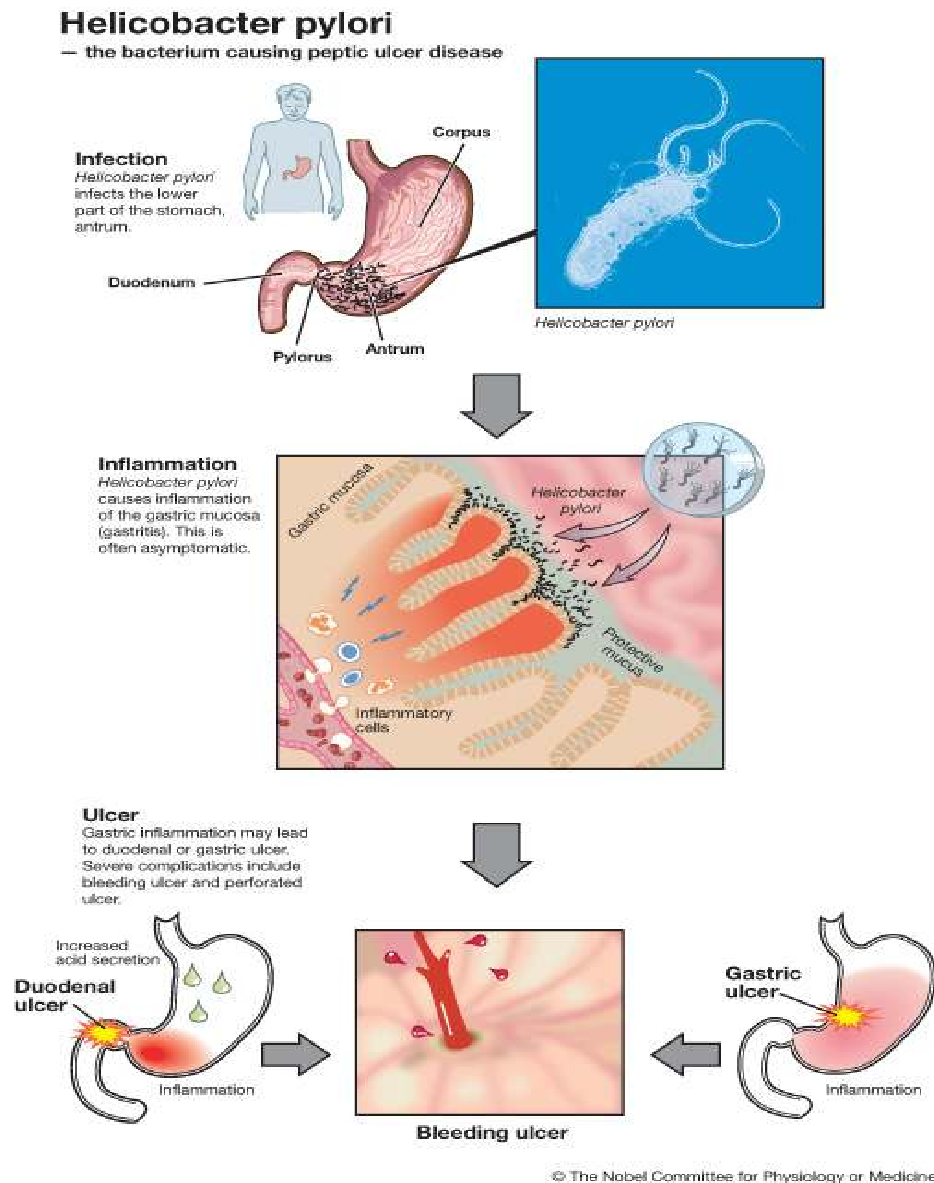


Figure 2.5: Complication due to H .pylori leading to gastric cancer. Source: ©The Nobel Committee for Physiology or medicine. Website: Nhipcauykhoa.net

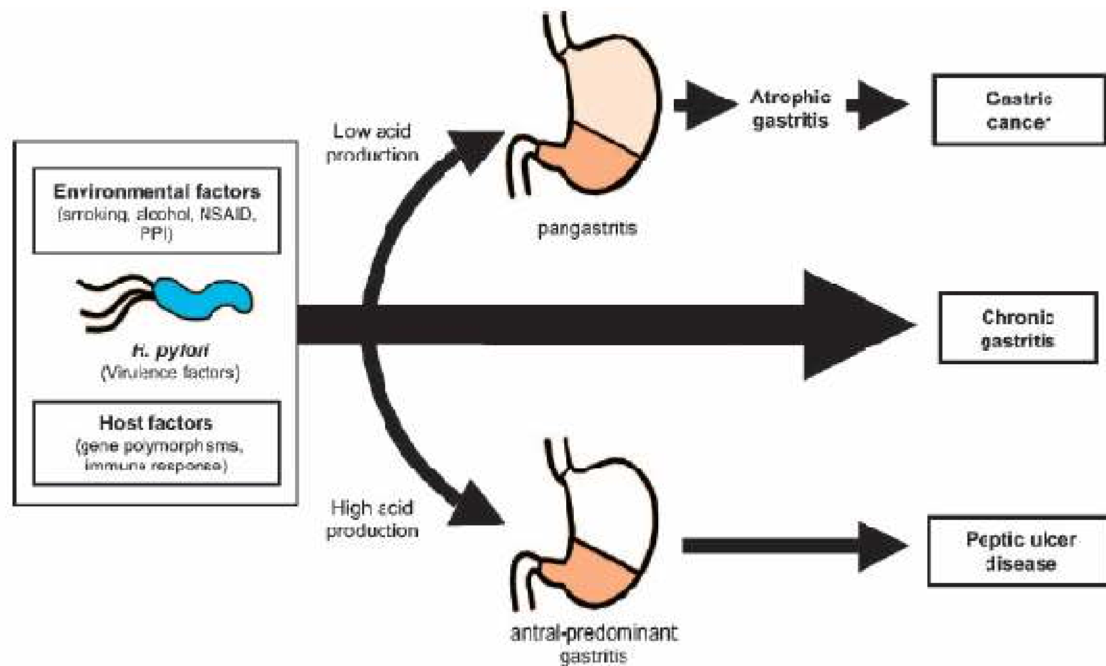


Figure 2.6. Schematic representation of the factors contributing to gastric pathology and disease outcome in H. pylori infection. Source: Kusters et al., 2006

2.8. Symtoms, staging and treatment

The symptoms of GC often mimic those of peptic ulcer disease, namely a vague upper abdominal pain aggravated by food, heartburn, and indigestion. These symptoms are often initially treated with antacids and H-2 blockers. Loss of appetite and early satiety are other common symptoms. Large GCs may result in partial blockage of the digestive tract and cause patients to vomit after meals. GC may also bleed, resulting in the vomiting of blood or the appearance of black, tarry stools [144].

GC is staged by the TNM system. **T** refers to the depth of penetration of the tumor through the wall of the stomach and whether it invades into local structures; **N** refers to the presence or absence of lymph node involvement near

the stomach; and **M** refers to the presence of absence of metastatic spread of tumor to distant sites in the body. This staging system is important in guiding treatment decisions and offering prognostic information to the patient. For example, early stage GCs has 50 percent chance to be cured, whereas cancers which have metastasized to distant organs have a very slim chance of long-term survival (< 3 percent) [145-148].

From a treatment standpoint, the first question to address is whether the cancer is operable or not. This depends on a number of factors, including its location, how bulky the tumor is, whether it invades surrounding structures, the extent of lymph node involvement, and whether the individual is strong enough to undergo a major operation [148]. Patients with GC are surgically treated via gastrectomy, which could be total or partial **[Figure 2.7 and 2.8]**.

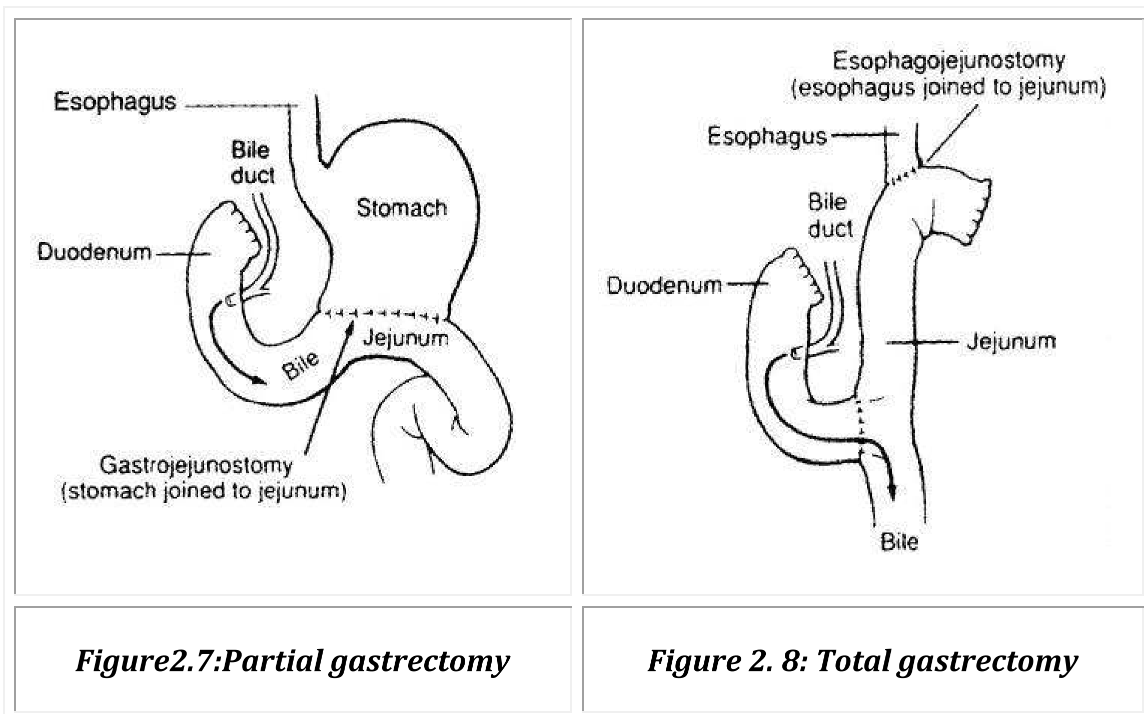


Figure2.7:Partial gastrectomy	*Figure 2. 8: Total gastrectomy*

Total gastrectomy is associated with a risk of mortality and postoperative complications in 2-4% and 10-20% of patients, respectively [149]. It has been proposed that laparoscopic subtotal gastrectomy for distal GC is a safe oncological procedure with benefits such as reduced blood loss, shorter time to resumption of oral intake and earlier discharge from hospital [150].

Another recent study showed that an alternative strategy for resectable patients is to administer combination chemotherapy both before and after surgical resection. For patients with metastatic disease at the time of diagnosis, the mainstay of treatment is chemotherapy. Following surgery, adjuvant treatment has been shown to improve outcomes. This treatment typically consists of a combination of external-beam radiation and concurrent chemotherapy [148, 151-154]. Single agents such as epirubicin, mitomycin, doxorubicin, cisplatin, etoposide, fluorouracil, irinotecan, hydroxyurea, taxanes and the nitrosoureas have low response rates (15% to 30%), brief duration of response, few complete responses and little impact on survival. Combinations of drugs are more widely used than single agents, largely because of higher response rates, more frequent complete responses and the theoretical potential of longer survival. Drug combination therapy has been shown to improve median survival by about six months in patients with metastatic disease. Chemotherapeutic drug combinations mostly used in clinical practices can be: leucovorin, etoposide and fluorouracil (*ELF*) or methotrexate, fluorouracil, leucovorin, doxorubicin (*FAMTX*) or hydoxyurea, leucovorin, fluorouracil and cisplatin. The last treatment has been shown to have slightly better results since

the response rate is 62% and the median survival time is 11 months [155]. Other modalities of combined therapies such as radiotherapy and leucovorin/fluorouracil chemotherapy have been applied depending on patient response and tolerance. Over the last few years, the use of preoperative or neo-adjuvant chemotherapy has been found to convert unresectable tumors to resectable ones. Biological therapy with trastuzumab, a recombinant humanized anti-HER-2/neu, combined with chemotherapeutic agents (i.e., doxorubicin) can be applied in GC treatment successfully [156]. Other non-chemotherapeutic drugs have been recently used in clinical practice. Of these, non-steroidal anti-inflammatory drugs (*NSAIDS*) are used for prevention or regression of cancer since they may target the cycloxygenase-2 (*COX-2*) enzyme [157]. Other studies have reported that immunological therapy using a combination of antibodies against the receptors of vascular endothelial growth factor (*VEGF*) and epidermal growth factor (*EFG*) could represent a powerful tool for the therapy of GC since either anti-VEGF or anti-EFG or their combination could effectively inhibit tumor cell growth. These findings support the hypothesis that inhibiting multiple biological pathways that mediate tumor growth is an effective therapeutic strategy for the treatment of GC [158]. Conventional adjustments in the dose of chemotherapeutic treatment could be ineffective in preventing toxicity and response variability in GC patients. New strategies for individualizing treatment for cancer patients are becoming an emerging issue in clinical practice. Pharmacogenetics could be an important source of information in this respect, by clarifying the complex correlation existing between an individual genetic profile

and the response to therapy in terms of toxicity and activity. It is currently speculated that some host gene polymorphisms involved in metabolism, cellular transport and interaction with molecular targets of the drugs used in GC therapy viz , *XRCC1*, *ERCC1*, *GSTs*, *MTHFR, UGT1A1*, *MDR1* and *MRP2* might be prognostic factors in the clinical outcome of specific chemotherapeutic treatments. Clinical applications of pharmacogenetics could represent a powerful tool in determining the appropriate drug and dose to be used in each individual patient with GC [159].

2.9. Genetic events in gastric oncogenesis

Multiple genetic and epigenetic alterations of oncogenes, tumor suppressor genes, DNA repair genes, cell cycle regulators, cell adhesion molecules, growth factor/receptor systems and microsatellite instability are involved over the course of the multi-step conversion of normal epithelial cells to GC **(Figure 2.9)**. Identification of specific genetic pathways in GC may have an impact on prognosis and selection of treatment strategies.

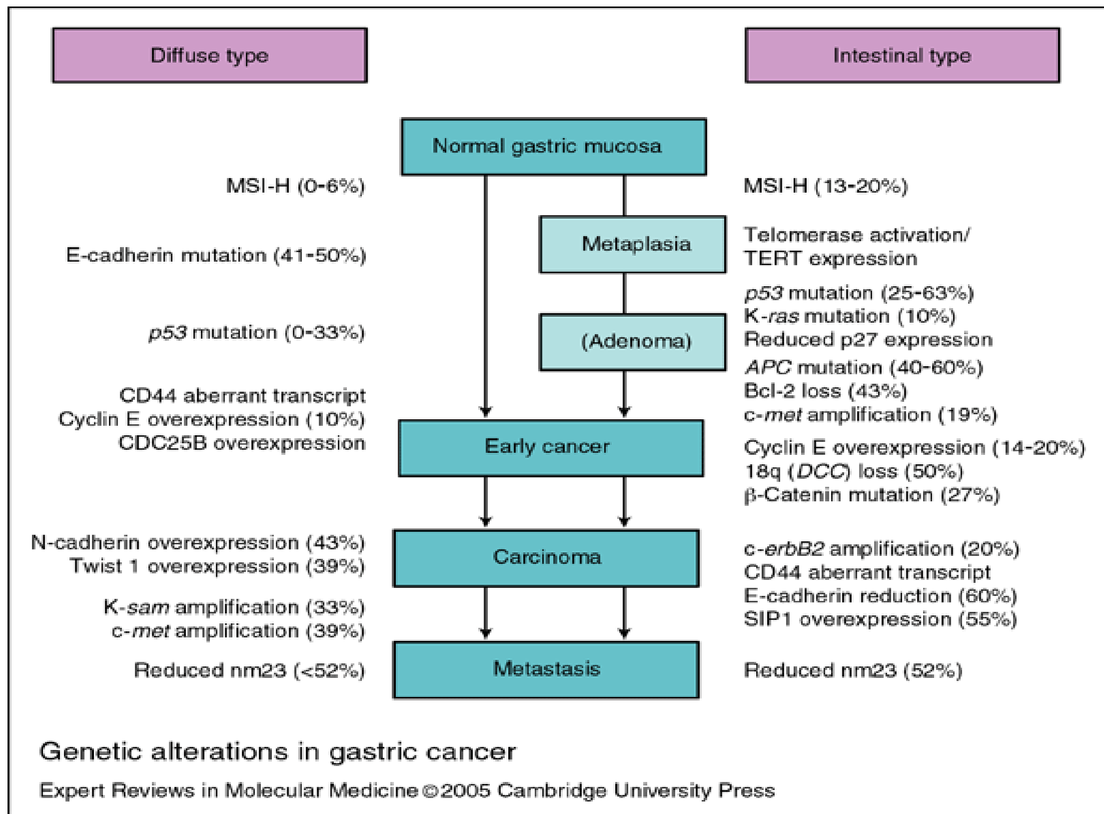


Figure 2.9. Genetic alterations in gastric cancer. Source: Expert reviews in molecular medicine 2005 cambridge university press

2.9.1. *Microsatellite and chromosomal instability*

Widespread tumor-associated microsatellite instability (MSI) is believed to be caused by altered repair of spontaneous DNA replication errors after mutational inactivation or epigenetic silencing of at least one of various mismatch repair genes (MMRs), including *hMLH1*, *hPMS1*, *hPMS2*, *hMSH2*, and *hMSH6/GTBP*. Ottini *et al.* have demonstrated that high frequency of MSI in GC is associated with female sex, antral location, intestinal-type histology, advanced tumor stage, vascular invasion, and positive family history [160]. Leung *et al.* suggested that high-frequency MSI in sporadic GC is mostly due to epigenetic inactivation of *hMLH1* gene, and the loss of HMLH1 protein is a significant event

in the development of invasive tumor [161]. In a high proportion of GC cases loss of heterozygosity (LOH) is observed at chromosomes 1p, 5q, 7q, 11p, 13q, 17p, and 18p, which are possible sites of tumor suppressor genes [162, 163]. Usually LOH is also required to inactivate a MMR in GC. Semba *et al.*, 1998 suggested high MSI in young patients with diffuse GC. Defect of DNA repair system was frequently associated with LOH [164].

2.9.2. *Cell-adhesion and Cell-cycle regulators*

Mutations in genes encoding for cell-adhesion molecules have been described in GC as well. Inactivation or down-regulation of E-cadherin protein, which belongs to the functionally related trans-membrane glycoprotein family, is found in GC and contributes to an increase in cell motility, the first step of cancer invasion and metastasis. A product of the *CDH1* gene (locus 16q22.1), is responsible for the Ca (2+)-dependent cell-cell adhesion mechanism, its inactivation has been suggested to play an important role in the growth and invasion either in hereditary gastric carcinoma (HGC) or in hereditary diffuse gastric cancer (HDGC) [165, 166]. Besides, rare genetic alterations of Motif-Containing GTPase-Activating Protein 1 gene, a negative regulator of cell-cell adhesion at adherens junctions, have been found especially in diffuse GCs [167]. Osteopontin, a protein ligand of *CD44*, is up-regulated in and together with abnormal *CD44* result in lymphatic invasion and metastasis [168, 169]. Galectin-3, a galactoside-binding protein is another molecule related to cell adhesion implicated in gastric tumor metastasis [170]. Many papers also demonstrated that *E-cadherin* mutations together with those of β-*catenin* and γ-*catenin to be*

involved in the development and progression of diffuse and schirrhous-type cancers [171-173].

2.9.3. *Growth factors and cytokines*

A broad range of growth factors and cytokines are produced in the gastric tumor environment by different cells accounting for complex cell interactions and for regulation of differentiation, activation, and survival of multiple cell types. *EGF, TGFα* members of the EGF family are overexpressed in the intestinal-type of gastric carcinomas [174]. TGFβ growth factor is more prevalent in diffuse type carcinomas with diffusely productive fibrosis [174, 175]. *IGF II* and *bFGF* growth factors are overexpressed in both histotypes of GC [174]. GC cells express neuropilin-1 (*NRP-1*), which is a membrane bound coreceptor for both *VEGF-165* and *VEGF* receptor 2 (*VEGFR-2*) in endothelial cells. It is known that *NRP1* plays versatile roles in angiogenesis, axon guidance, cell survival, migration, and invasion. In the case of human GC, regulation of *NRP-1* expression is intimately associated with the *EGF/EGF-R* system. It was shown that activation of *EGF-R* may contribute to GC angiogenesis by a mechanism that involves upregulation of VEGF and *NRP-1* expression via multiple signalling pathways [176]. Angiogenic factors, such as *VEGF*, *bFGF* and *IL-8* promote neovascularisation of GC. Moreover, VEGF promotes in particular the malignant progression of the intestinal-type and amphiregulin (*AR*), another member of EGF family, is overexpressed in both types of GC [174]. *IL-8* is a member of the CXC family of chemokines, which plays a pivotal role in gastric oncogenesis; more than 80% of stomach tumors express both this cytokine and its receptor [177, 178].

2.9.4. *Oncogenes*

Many proto-oncogenes are activated in gastric malignancy. The *c-met* gene, a transmembrane tyrosine kinase receptor of hepatocyte growth factor (HGF), is found amplified in 19% of intestinal-type and 39% of diffuse-type GCs [179]. *K-sam* was first gene found amplified in the GC cell line KATO-III. Amplification of *K-sam* gene is restricted to poorly differentiated types of GC [174, 180]. Amplification of *K-sam* was found preferentially in advanced diffuse or scirrhous-type GCs (33% of all) but not in intestinal-type carcinomas [180].

The *c-erbB-2* gene is another potential cell surface receptor of the tyrosine kinase gene family. The *c-erbB-2* gene is a proto-oncogene which encodes a protein similar to but distinct from the epidermal growth factor receptor. It is commonly amplified in the intestinal-type of gastric adenocarcinoma [181]. The c-erbB-2 protein expression is enhanced in advanced stages during the progression of gastric carcinoma and is an indicator of poor short-term prognosis [182]. Mutations of *K-ras* oncogene can be found in intestinal-type cancer and the precursor lesions, intestinal metaplasia and adenoma. However, *K-ras* point mutations are uncommon in GC and are not present in diffuse gastric tumor histology [183, 184].

2.9.5. *Tumor suppressor genes*

The *p53* gene, probably the most famous tumor suppressor gene, couldn't be absent from the list of genes involved in gastric carcinogenesis. The *p*53 tumor suppressor gene is one of the most commonly mutated genes in all types of human cancer and encodes a transcription factor involved in cell cycle regulation.

It acts as a tumor suppressor gene by inducing cell cycle arrest or apoptosis and requires loss of function mutations for cancer development. It is frequently inactiveated in gastric carcinomas by LOH, missense mutations or frameshift deletions. Taken together, these genetic alterations are present in more than 60% of gastric carcinomas and are also found in intestinal metaplasia, dysplasia and adenomas. Even if *p*53 gene is highly polymorphic, with at least 13 single nucleotide polymorphisms [185]. A polymorphism in this codon, which consists in a single base pair change of either arginine (Arg) or proline (Pro), has been suggested to modulate *p*53-dependent apoptosis and modify sensitivity to chemotherapeutic agents [186, 187]. The alterations in exon 4 of the *p53* gene in GC were firstly investigated by Shepherd *et al.* in 2000 [187], who reported the following genotype frequencies: Arg/Arg (54%) Arg/Pro (33%) Pro/Pro (14%). The most intriguing aspect of the initial study is that the genotype of the codon 72 polymorphic site varied significantly with race as follows: 64% of whites had the Arg/Arg genotype compared with 24% of blacks [187]. The association between *p53* codon 72 polymorphism and GC was firstly investigated by Hiyama *et al.* in 2002 [45]. Zhou, Li et al has published a meta-analysis on *p*53 exon4 Arg72Pro polymorphism and GC including 12 case-control studies [188]. The combined results showed no significant difference in genotype distribution between GC cases and controls. A significantly lower frequency of the Arg/Arg genotype in GC in cases compared with controls was reported among Asians, while not among Caucasians. A wide range of variability in the frequencies distribution of *p53* genotypes have been observed in GC in different population

base study. So additional primary studies performed on population and collecting data on environmental and genetic co-exposures are demanded.

The *p53* gene (locus 17p13.1) frequently shows GC-AT transitions in diffuse-type GC, due to carcinogenic N-nitrosamines produced from dietary amines and nitrates in the acid gastric environment [189-191]. LOH and abnormal expression of the *p73* gene, another *p53* family member mapping at 1p36, a minimal region frequently mutated in GC, preferentially occur in the *de novo* pathway for well-differentiated adenocarcinomas of foveolar type expressing pS2, a gastric-specific trefoil factor [174, 192]. The pS2 protein is normally expressed in gastric foveolar epithelial cells. Inactivation of the *pS2* gene is observed in dysplasia, adenoma and adenocarcinoma in mice [193], suggesting its role at early steps of gastric carcinogenesis [174].

Germline mutations in tumor suppressor gene (APC) gene cause familial adenomatous polyposis (FAP), which is an autosomal-dominant colorectal cancer syndrome [194]. LOH of the two closely spaced *APC/MCC* genes has been shown to be associated with the development of gastric carcinomas [195]. Notably, *APC* gene missense mutations are present in more than 50% of the intestinal-type GC, while they are not involved in diffuse type cancers. Somatic mutations of the APC gene are observed in precursor lesions of the stomach, such as in 20-40% of gastric adenomas and in 6% of intestinal metaplasias, demonstrating its role in early steps of gastric carcinogenesis [196, 197].

LOH on chromosome 10q23.31 of tumor suppressor gene *PTEN* appears in precancerous lesions. *PTEN* mutations are restricted to advanced GC. In fact, LOH

and mutation of *PTEN* are closely related to infiltrating and metastatic GCs [198]. In a paper based on immunohistochemical analysis in a large number of patients, it is shown that BIRC5, an inhibitor of apoptosis, is positively correlated with PTEN expression in GC and is a molecular marker of lymph node metastasis, while PTEN expression is reconfirmed as a molecular marker of advanced GC [199]. *RUNX3* gene is a relatively recently discovered tumor suppressor, also involved in the complex process of gastric oncogenesis. Loss of *RUNX3* by hypermethylation of its promoter results in many tumors, including gastric malignancy. *RUNX3* methylation is observed in chronic gastritis, intestinal metaplasia and gastric adenomas, suggesting this gene as a target for epigenetic gene silencing in GC [200]. Nuclear retinoic acid receptor β, *RARβ* is another tumor suppressor gene found hypermethylated in 64% of the intestinal-type GCs, while alterations of this gene are not observed in the diffuse-type [201].

2.10. Gene polymorphisms

Individual variations in GC risk have been associated in the last decade with specific variant alleles of different genes that are present in a significant proportion of the population. Polymorphisms may modify the effects of environmental exposures, and these gene-environment interactions could partly explain the high variation of GC incidence around the world. In the last decade the association between polymorphisms and GC has been investigated within community-based genetic association studies, aiming also to explore gene-gene and gene-environment interaction. Identifying the inherited genetic variants that modify the effect of environmental exposures in GC risk could eventually lead to

more effective primary intervention. This, however, is still an open issue [202]. According to the role of the genes whose polymorphisms have been studied in association with GC, we can grouped them into the following categories: genes involved in the protection of gastric mucosa against damaging agents, in inflammatory response, in detoxification of carcinogens, in synthesis and repair of DNA, in regulation of gene expression, in cell adhesion and in cell cycle [203, 204]. Since the effect of each individual polymorphism could be small, association studies in genetic epidemiology benefit from large sample sizes. The majority of the published studies, however, are underpowered to detect a robust association.

Several polymorphisms that involve the metabolic activation or detoxification of carcinogens derived from cigarette smoke have been found to be associated with cancer risk. Many studies have focused on the relation between the distribution of polymorphic variants of different forms of the metabolic enzymes and cancer susceptibility [17, 205-208].

2.11. Xenobiotic-Metabolising genes

In general, environmental chemical carcinogens require metabolic activation by host enzymes to be genotoxic: phase 1 enzymes, such as CYPs (cytochromeP-450) activate carcinogens metabolically to form genotoxic electrophilic intermediates. These enzymes have the ability to insert an atom of molecular oxygen into their substrates, creating an active site for further detoxification by phase 2 enzymes **(Figure 2.10)**. However, certain substrates are converted into mutagenic, chemically active intermediates. Indeed, many chemical carcinogens require some metabolic activation to achieve maximum

activity, indicating the need for co-ordinated phase 1 and 2 expression to prevent accumulation of carcinogenic intermediates.

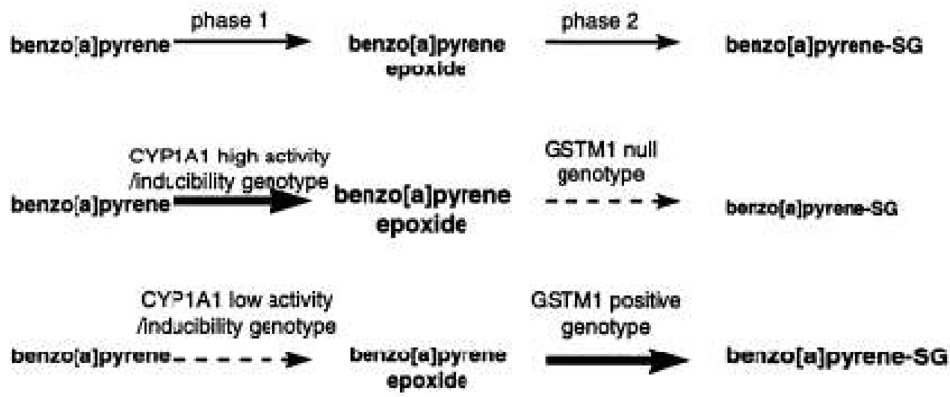


Figure 2.10. Relationship between phase 1 and 2 detoxification reaction

Activated metabolites are, in part, detoxified by phase 2 enzymes, such as GSTs, NQO1 and NAT2. Thus, the relative activity of the metabolizing enzymes, which is genetically determined to a great extent, is thought to be one of the important determinant host factors underlying cancer development [209]. Variable levels of expression of these enzymes could result in increased or decreased carcinogen activation. In fact, it is well established that genetic differences occur in expression of the xenobiotic metabolizing enzymes (XMEs) [210]. The inherited differences in the capacity of XMEs have been found to be an important factor that determines the genetic susceptibility to various malignancies [211]. Therefore, the relationship between genetic polymorphisms of these enzymes and individual susceptibility to GC is of interest. It is widely held that humans differ in their susceptibilities to cancer. Certain individuals may be more susceptible, whereas others are more resistant to cancer. This may be

due to a number of factors including health, nutritional status, and gender. From what is known about the mechanism of action of carcinogens, it is thought that genetic background could play a significant role. The obvious candidate genes are those encoding the XMEs that activate or inactivate carcinogens [212, 213].

2.11.1. *Glutathione S-transferases (GSTs)*

GSTs are a supergene family of phase 2 enzymes present in many tissues, including gastric [214]. These enzymes catalyze the detoxification (through conjugation of glutathione) of a variety of reactive electrophilic compounds, including many environmental carcinogens such as benzo [a]-pyrene and polycyclic aromatic hydrocarbons (PAHs) [215].

GSTs constitute a very ancient protein superfamily that is thought to have evolved from a thioredoxin-like ancestor in response to the development of oxidative stress [216, 217]. It is increasingly becoming clear that GSTs share sequence and structural similiarities with several stress-related proteins in a wide range of organisms [218]. It is thought that the multiple GST classes arose by a process of gene amplifcation followed by divergence, perhaps involving a mechanism similar to DNA shuffling, resulting in novel catalytic activities [219, 220] **(Figure 2.11)**. The soluble glutathione S-transferases comprise 4 main gene classes, alpha (α), mu (μ), pi (π), and theta (θ). *GST* subfamilies are widely expressed in humans: *GSTM* (μ), *GSTT* (θ) and *GSTP* (π) with overlapping substrate specificities [221]. *GSTM1* and *GSTT1 genes* exhibit homozygous deletion (null genotype) polymorphisms which lead to a lack of function and decreased ability to detoxify electrophilic carcinogens efficiently. Subjects

carrying the *GSTP1* Ile105Val Val/ Val genotype have a lower ability to detoxify electrophilic compounds than subjects carrying the wildtype genotype, Ile/Ile [222]. Variants in these genes may reduce an individual's ability to detoxify PAHs and could increase risk for various cancers, including GC [210, 223]. Individuals carrying one of these variants have no enzyme activity, and thus are more susceptible to carcinogens such as benzo[α]pyrene-7,8-diol epoxide, the activated form of benzo[α]pyrene, and smaller reactive hydrocarbons, such as ethylene oxide and diepoxybutane, which could lead to environmentally-induced cancer susceptibility [224, 225].

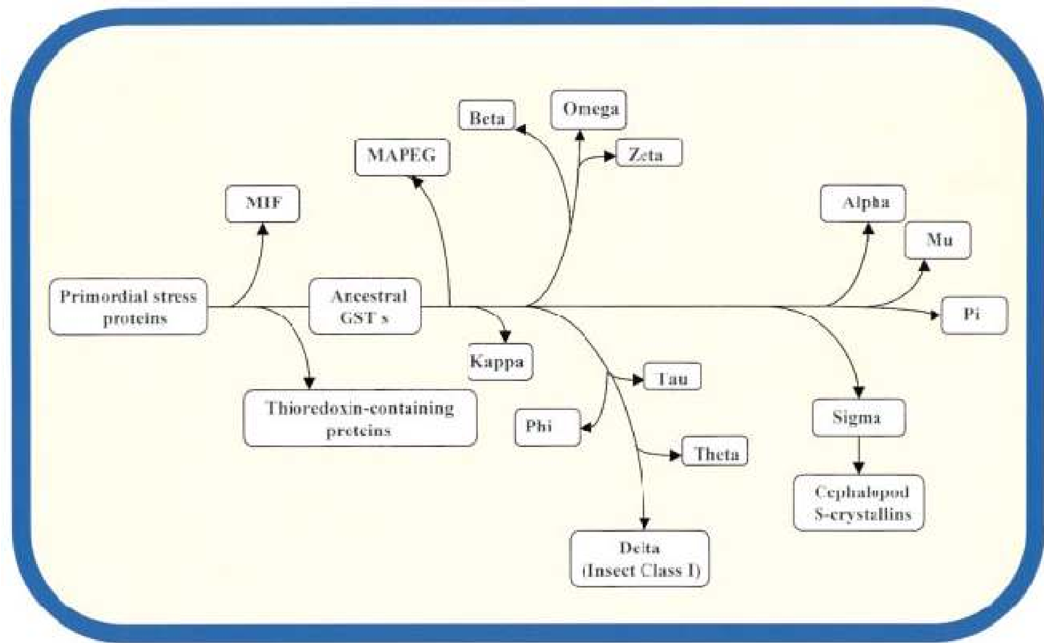


Figure 2.11. Possible pattern of divergence in the GST superfamily resulting in multiple GST classes is shown. Sourse: David Sheehan et al. Biochem. J. 2001. 360

2.11.1.1. *GSTM1 null and GSTT1 null*

GSTM1 is mainly expressed in liver, brain and stomach. The *GSTM1* null genotype is found in 10-60% of individuals ranging from 50% in Caucasians and Asians to 25% in Africans [17]. The association between *GSTM1* null genotype and GC was firstly investigated by Strange *et al.* in 1991 [226]. *GSTT1* is mainly expressed along the human gastrointestinal tract. The null genotype of *GSTT1* is present in 13-31% among caucasians and 36-55% among asians. The association between *GSTT1* null genotype and GC was firstly investigated by Deakin *et al.* in 1996 [227]. Various studies have revealed positive association between *GSTT1* and *GSTM1* null genotypes and increased risk for skin, lung, stomach, bladder, prostate and colorectal cancers [228], but there are conflicting reports also [15]. The inconsistency in results may be due to different ethnicity or interaction between different environmental and genetic factors. The first meta-analysis evaluating the association between *GSTM1* status and GC included 15 primary studies in English language and was published in 2005 [17]. The meta-OR was 1.24 for heterogeneity. The heterogeneity slightly decreased after stratifying by ethnicity, source of controls and study power. Caucasians showed a significantly increased risk (OR = 1.22) while studies with at least 80% power provided the lowest estimate (OR = 1.05).

Studies investigating the association between GST genes polymorphism and GC risk have reported conflicting results. Severals studies showed that *GSTT1* null genotype as a significance risk factor of GC [229, 230], and a significant association of *GSTM1* null genotype with GC has already been observed [231].

However Al-Moundhri MS et el reported no statistically significant associations between *GSTM1* polymorphism and GC in Arab population. Moreover, combined analysis showed that a combination of the null *GSTM1* genotype and carriers of IL-1RN*2 was associated with a statistically significant correlation with GC (odd ratio=3.6, 95% confidence interval=1.4-9.4, p=0.008 [232]. Furthermore, meta-analysis by Chen B et al suggests that *GSTT1* gene polymorphism may not be associated with increased GC risks among Europeans, Americans and East Asian [229]. Case-control studies in vietnam and Korea also provided evidence that *GSTM1* and *GSTT1* null genotype is not related related with GC risk [233, 234].

2.11.1.2. *GSTP1*

Another GST isoenzyme, GSTP1 is widely expressed in tumour cells and is responsible for the detoxification [16]. GST Polymorphisms in the *GSTP1*, located on chromosome 11q13 in humans, have been associated with a reduction in enzymatic activity toward several substrates, including both chemotherapy agents (such as cisplatin, a common agent used in lung cancer treatment) and carcinogens found in tobacco smoke [235, 236]. Two single nucleotide polymorphisms in *GSTP1* that result in a change in amino acids have been identified. A single nucleotide polymorphism in exon 5 (Ile105Val, 313A/313G), the A-to-G transition that results in an amino acid change from isoleucine to valine, results in significantly lower conjugating activity among individuals who carry one or more copies of the guanine allele (Ile/Val or Val/Val) compared with those who have the A/A (adenine/adenine; Ile/Ile) genotype [19, 237]. Having at least one copy of the guanine allele at this locus is also associated with increased

levels of hydrophobic adducts and higher levels of PAH-DNA adducts in human lymphocytes [238]. A second single nucleotide polymorphism in exon 6 (Ala114Val, 341C/341T) results in an amino acid change from alanine to valine, which also appears to confer lower activity [19].

2.11.2. *Cytochrome P450 (CYP)*

The CYP dependent mono-oxygenases play an important role in the metabolism of environmental carcinogens. The CYP represent the first line of defense against toxic lipophilic chemicals because they catalyze reactions involving incorporation of an atom of molecular oxygen into the substrate and the resulting increase in hydrophilicity facilitates further metabolic processing and excretion [239]. The main CYPs in humans that metabolize carcinogens are CYPlAl, CYP1A2, CYPIBI, CYP2A6, CYP2E1, CYP3A4, and CYP3A5 [240]. These enzymes have specificities for various classes of carcinogens and genetic polymorphism has been identified for most of them [239-241]. Among these, *CYP1A1* and *CYP2E1* have been reported to be involved in the pathogenesis of various malignancies in diverse ethnic groups [242]. The *CYP1A* gene family has two members: *CYPIAl*, which is predominantly expressed in extrahepatic tissues such as the lung, and *CYP1A2*, which is concentrated in the liver [243]. *CYP1A1* and *CYP1A2* have overlapping catalytic activity and are both thought to play an important role in carcinogen activation [244]. *CYP2E1* is responsible for the metabolic activation of procarcinogens such as N-nitrosamines and various other low-molecular-weight compounds into reactive intermediates that play an essential role in chemical carcinogenesis [245, 246]. The levels of expression of

CYPs vary depending on the P450 form [244]. The interindividual differences in expression have been observed and may be due to the genetic polymorphisms or the extent of induction. The CYPs activity is an important determinant of human susceptibility to toxicity and carcinogenicity of industrial and environmental chemicals [247]. Several base changes distinguishable by RFLP analyses have been found in these genes [248-250]. Although these polymorphisms do not appear to alter the primary sequence of the enzyme, an effect on gene transcription has been suggested [251]. Despite the limited knowledge of the enzyme's function, its apparent tumor-specific expression is intriguing. Further research is needed to determine if this enzyme may be a potential target for the prevention or treatment of cancers.

2.11.3. *NAD (P) H:quinone oxidoreductase 1 (NQO1)*

NQO1 is a cytosolic enzyme that catalyzes the two-electron reduction of numerous quinoid compounds into their less toxic form protecting cells against oxidative stress. It is an important enzyme in both activation and detoxification pathways known to protect against the carcinogenicity and mutagenicity of quinone compounds and their metabolites [252]. There have been more than 93 single nucleotide polymorphisms (SNPs) identified in the *NQO1* gene. The most widely studied SNP of *NQO1* is a C to T change at nucleotide position 609 (rs1800566), also known as *NQO1*2*. This results in a proline to serine amino acid change at codon 187 that is associated with a loss of enzyme activity due to instability of the protein product [253]. Thus, the enzyme activity of the homozygous variant genotype (*NQO1*2/*2*) is almost undetectable, and the

enzyme activity of the heterozygous genotype (*NQO1*1/*2*) is intermediate between the homozygous variant genotype and wild type (*NQO1*1/*1*) [254]. The *NQO1*2* allele frequency varies between different ethnic groups, ranging from 16% in Caucasians to 49% in Chinese populations. Prevalence of the *NQO1*2/*2* genotype is 4.4% in non-Hispanic whites, 5.2% in African Americans, 12.2% in Japanese, 15.5% in Mexican Hispanics, 17.9% in Native Americans, 18.8% in Koreans, and 22.4% in Chinese [254-257]. Another genetic polymorphism of *NQO1* is a single nucleotide change from C to T at nucleotide position 465 (rs4986998), also known as *NQO1*3*, which changes the amino acid at codon 139 from arginine to tryptophan. This SNP results in alternative messenger RNA splice sites that can lead to a deletion of exon 4 and create a protein lacking the quinone binding site for which enzyme activity differs according to the substrate [258, 259]. The frequency of the *NQO1*3* polymorphism is generally low and ranges from 0% to 5% among different ethnic populations [255]. Malik MA et al shows that the TT genotype and T allele of *NQO1* C609T polymorphism were significantly associated with increased risk for GC. Previous studies of the association between functional *NQO1* C609T polymorphism and several human cancers have had mixed findings [260].

2.11.4. *N-acetyltransferase (NAT)*

NAT polymorphism causes individual variations in biotransformation of various xenobiotics with primary aromatic amine or hydrazine structures [261, 262]. NAT is widely expressed in tissues [263] and cultured cells [264]. In humans, two genes, *NAT1* and *NAT2*, are responsible for N-acetyltransferase

activity [265]. Certain chemicals may be N-acetylated to a significant degree by *NAT1* and *NAT2*. These include the carcinogenic aromatic amines 2 aminofluorene, benzidine, 4-aminophenyl, 4, 4-dichloroaniline, and 2-naphthylamine [266, 267], and the cancer chemotherapeutic agent dinaline (4-amino-N-[2'-aminophenyl] benzamide) [268]. They are encoded at two distinct loci on chromosome 8p21.3-23.1 along with NATP, a pseudogene that does not encode a functional protein [269].

In *NAT2* gene, at least 23 different *NAT2* mutations have been found to date. Of these nine lead to amino acid change. Seven of the nine observed nucleotide transitions lead to amino acid changes, whereas the remaining two base substitutions exert no influence on the amino acid sequence [269]. Rapid acetylators have at least one wild-type allele, whereas slow acetylators have inherited two slow acetylation-associated alleles. Investigators have reported a wide range of values for acetylation activity in different groups [270]. The predominance of the *NAT2* slow acetylator genotype has been reported to be about 60% among Germans [271, 272], 53% among American Caucasians [272], 63% among Poles [273], and 50% among Finns [274]. In contrast, in the Japanese or Chinese populations, the rapid genotype is largely overrepresented (92 and 80%, respectively) [275, 276]. Previous studies have suggested a modifying role for *NAT* genotypes in all major cancer sites [277]. Molecular epidemiological studies demonstrated that individuals with *NAT1* rapid acetylator genotypes or *NAT2* slow acetylator genotypes in the presence of known carcinogen exposures, such as cigarette smoking, dietary exposure to heterocyclic amines (HCA) or

occupational exposure to aromatic amines (AA), were at increased risk for various types of human cancers [265, 278].

A number of previous phenotyping studies provided evidence that the NAT2 slow acetylator phenotype is a significant risk factor for the occurrence of cancer. Subsequent genotyping studies supported the important role of NAT2 slow acetylation status as a risk factor for cancers [279, 280]. NAT2 is more important than NATI for bioactivation of heterocyclic amines in vitro [281, 282]. Genotyping study indicated an increased risk of breast cancer for slow NAT2 acetylators who smoked 20 or more cigarettes per day [283]. Subsequent genotyping studies also did not give any conclusive evidence [284-286]. However, the potential role of NAT genotypes as modifiers of individual responses to environmental agents is supported in three recent studies that found that the *NAT2* slow acetylator genotype posed an increased risk of mesothelioma [287] and hepatocellular carcinoma [288]. Further addressing the potential importance of individual acetylation capacity the gene-gene interaction analysis by Stefania Boccia et al demonstrated that individuals with combined *GSTT1* null and *NAT2* slow acetylators had an additional increased risk of GC, with an OR of 3.00 (95%CI: 1.52–5.93). According to this study *NAT2* polymorphism appear to modulate individuals susceptibility to GC in the Italian population [231]. However study by Yan Wei Zhang et al on GC in Korean population showed *NAT2* acetylator genotype to be an important modifier of the effects of environmental factors on GC risk [289].

2.12. Promising highthroughput techniques

Cancer is a highly complex disease which can encompass multiple genomic alterations, including point mutations, translocations, gene amplifications, epigenetic modifications, deletions, aberrant splicing, and altered gene expression **(Figure 2.12)**. These changes may be inherited or somatically acquired during progression from a normal to a cancerous cell. Progress of phenotypes from normal to advanced carcinoma is controlled by coordination of action of hundreds of genes [290]. Conventional approaches investigating one or several candidate genes at a time can not show the whole story of carcinogenesis.

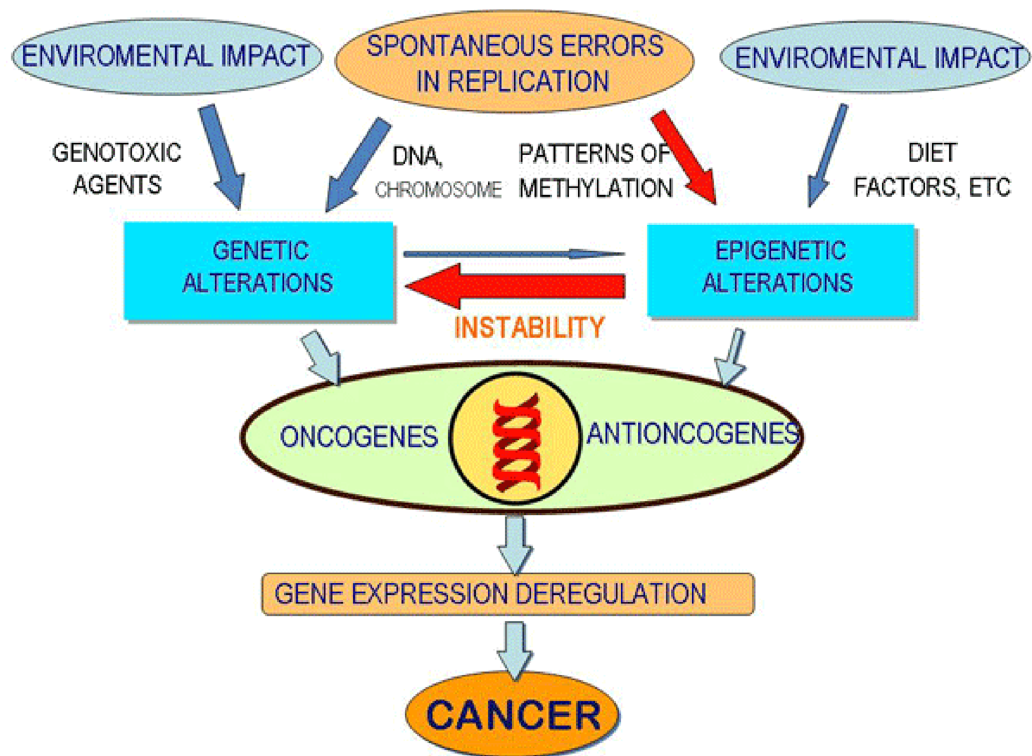


Figure 2.12. Genetic and epigenetic changes leading to tumorigenesis.

The generation of vast amounts of information, coupled with advances in technologies developed for the experimental use of such information, allows the description of biological processes with a view of global genetic perspective. However, little is known about the exact expression changes in tumorigenesis, which will help us identify the events that leads to the initiation and progression of cancer development. Analyzing alterations of gene expression profiles in neoplasia is necessary for establishing the preventive, diagnostic, therapeutic, and prognostic potential of each related gene. To illustrate the mechanisms controlling malignant changes at molecular level may provide a further understanding of tumorigenesis. One such technology, microarray which permits simultaneous monitoring of thousands of genes [291]. Global expression analysis using microarrays now allows for simultaneous interrogation of the expression of thousands of genes in a high-throughput fashion and offers unprecedented opportunities to obtain molecular signatures of the state of activity of diseased cells and patient samples [292].

2.12.1 Microarray

Microarrays have become routine methods for profiling gene expression in almost every discipline of biomedical research. Genomic scale profiling of gene expression is a potentially valuable means of evaluating changes in cancer [293, 294]. Identification of all genetic alterations is essential for a full understanding of the etiology of human cancer. Genetic analysis using a genome-wide detection tool is an essential approach to uncover all abnormalities and is also an efficient way to identify key genetic events, such as activation of oncogenes and

inactivation of tumor suppressor genes in cancer development and progression. Such an approach can lead to quick discovery of genetic markers for cancer risk assessment, diagnosis and prognosis [295]. As the cost of microarray technologies decrease, it is possible that array-based methods of genomic-scale transcript profiling may become as commonplace as PCR-based methods are currently [296]. This technology represents the most recent and exciting advance in the application of hybridization-based approaches to analysis in the biological sciences. Gene expression profiling of cancers represents the largest research category using microarrays and appears to be the most robust approach for molecular characterization of cancers. It is becoming recognized that microarray technology will be a fundamental tool for future genomic research [297]. Thus microarray technology is a powerful platform for biological exploration. Depending on the type of probes used, microarray systems are classified as either oligonucleotide or cDNA microarray.

2.12.1.1 cDNA microarrays

The cDNA microarrays are made by spotting cDNAs, usually PCR-amplified sequences from bacterial libraries, onto glass slides [298]. cDNA microarrays comprise relatively long DNA molecules immobilized on a solid surface and are mostly used for large-scale screening and expression studies. cDNA microarrays can not be used for mutation or genotyping analysis, which should be performed using oligonucleotide microarrays [299]. Spot sizes range from 80～150μm in diameter, and arrays can contain up to 80,000 spots [300]. In terms of sample preparation, RNA from cells is reverse transcribed into cDNA,

which is then fluorescently or radioactively labeled and used to probe a predetermined DNA set. Two different fluorescent dyes (usually Cy5 and Cy3) are used for cDNA microarray analysis, and a typical analysis may consist of a normal tissue sample being labeled with a green dye and a cancer tissue sample being labeled with a red dye. If both samples bind to the same target on a chip, a yellow signal is obtained, and a scanner is used to assess differing red/green/yellow emissions [301]. The greater the hybridization signal from probe-DNA binding, the higher the concentration of the RNA within the original sample [302]. In the example of target preparations, total RNA extracted from cells is fluorescently labeled by oligo dT-primed reverse transcription using nucleotides tagged with either Cy3 or Cy5 **(Figure 2.13)**. The unincorporated fluor-dUTPs are removed, the Cy3 and Cy5 probes combined, and then mixed with blockers. RNA quality and quantity can be assessed using the microcapillary-based Bionalyzer (Agilent Technologies), which can analyze as little as 5 ng of total RNA [300]. The target mixture is hybridized to the probes on the microarrays for 16～24 hours, the array is then washed and scanned [303]. Most microarray amplification methods make use of a linear-based amplification method using T7 RNA polymerase, resulting in amplified RNA (aRNA). Microarray studies that use the aRNA synthesis protocol require one amplification round, allowing as little as 10 ng or 1,000 cells to be used for the initial input [302, 304].

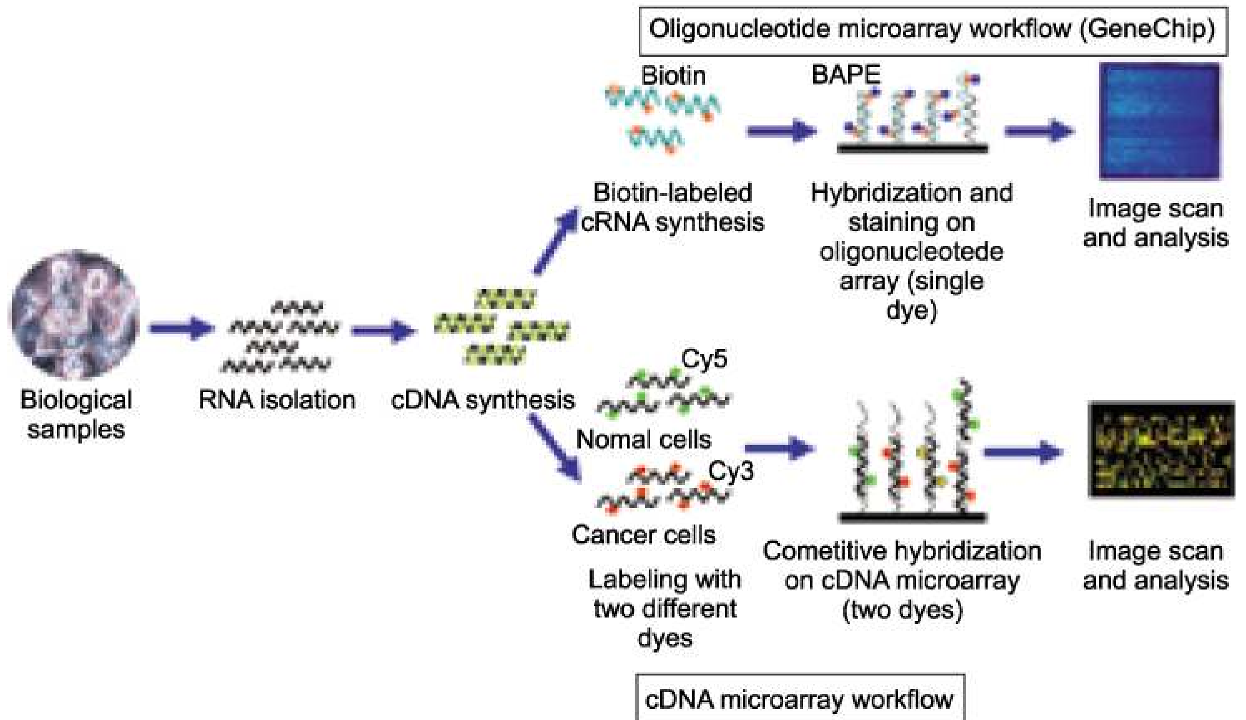


Figure 2.13. Experimental workflows for performing gene expression analysis using oligonucleotide and cDNA microarrays. Source: © Laing E, Bucca G, Smith CP, University of Surrey. 2008

2.12.1.2. Oligonucleotide microarrays

Oligonucleotide microarrays can detect mutations or SNP by discriminating between perfectly matched and mismatched signals. Few examples of oligonucleotide microarrays for mutation or genotypic analysis are RET oligonucleotide microarray, β-catenin oligonucleotide microarray, K-ras oligonucleotide microarray. The predominant *RET* mutations are missense mutations and are restricted to 10 codons in MEN2 syndromes. The *RET* oligonucleotide microarray can detect *RET* missense mutations at these 10 codons [305]. The *RET* oligonucleotide microarray can function as a fast and reliable genetic diagnostic device, which simplifies the process of detecting *RET* mutations **(Figure 2.14)**. β-catenin mutations have been identified in a variety of

human malignancies, with most of these being missense mutations restricted to hot-spot areas in exon 3. Oligonucleotide microarray have been developed for detecting β-catenin mutations at 11 codons [306]. This microarray can detect a total of 110 types of β-catenin mutation. All oligonucleotides on that array were 21 bp long and the mismatch sequence was located in the middle of the oligonucleotide [307]. Oligonucleotide microarrays provide a valid option, as they allow scientists to accurately and rapidly process large numbers of samples. Oligonucleotide microarrays also involved the Affymetrix GeneChip, which are used for both gene expression and variation detection. An Affymetrix GeneChipTM is produced by synthesizing tens of thousands of short oligonucleotides in situ on glass wafers [300]. Although originally developed for mutation detection, the same technology was adapted to measure expression levels of genes [308]. The Affymetrix P450 GeneChip is used for pharmacogenetic screening [303]. In general, 11～16 probes are selected among all possible 25-mers to represent each transcript. In addition to allowing hybridization to the most specific regions of nucleotides, the use of short oligonucleotides also allows the representation of multiple regions of a single gene in multiple spots, thus reducing the chance of false positives [309].The main advantage of the GeneChip is its ability to measure the absolute expression of genes in cells or tissues [300].

Oligonucleotide microarrays like the GeneChip allow for the differential detection of gene family members or alternative transcripts that can not be distinguished using cDNA microarrays. In the sample preparation, one or two amplification rounds are used to generate cRNA after reverse transcription. This

procedure can be carried out with significantly less starting material than is required for other methods [299].

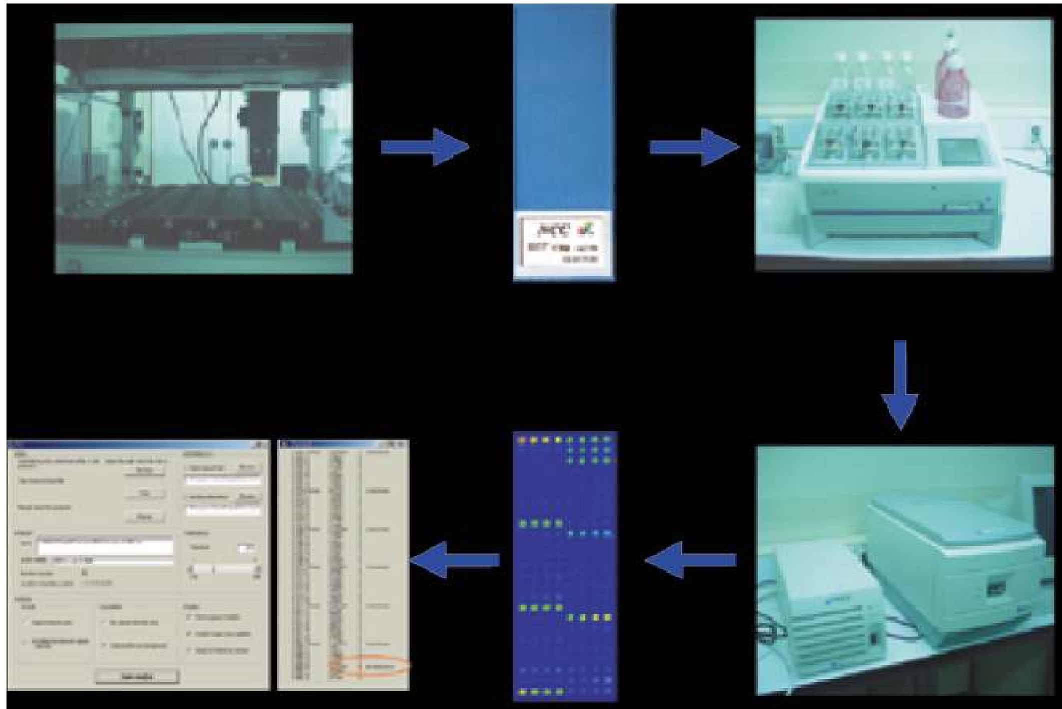


Figure 2.14. Experimental workflow for performing mutation analysis using oligonucleotide microarrays. Source: http://www.slideshare.net/helenadeus/provenance-of-microarray-experiments

2.12.2 *Microarray-based gene expression profiling in gastric cancer*

Microarrays have extended molecular research beyond the candidate gene approach and are beginning to establish a link between gene expression and functional interactions [310-312]. Understanding the differences in gene expression between normal tissue and malignant tissue, as well as the gene expression response to environmental stimuli, is central to understanding regulatory mechanisms involved in GC development and progression [311, 313, 314].

The microarray studies examining gastric adenocarcinoma have been aimed at developing exploratory gene profiles of GC cell lines to identify GC-

related genes, delineate molecular phenotypes and identify functional gene clusters as potential markers of biologic behavior [310, 315]. Recent studies have shown that microarray, in combination with statistical modeling, accurately predicted tumor behavior with respect to tumor progression, metastatic potential, tumor recurrence, and overall prognosis [316, 317]. Although in its infancy, gene expression analysis, combined holds promise in extending our understanding of gastric carcinoma. The relative paucity of data available relating GC gene profiles with prognosis and the success across various other cancers strongly reinforces the need for further exploration of this technique. With techniques capable of amplifying small quantities of tumor RNA, it is conceivable that endoscopically obtained gastric tissue samples may be used to generate preoperative predictive gene clusters. In doing so, the identification of functional gene clusters may allow improved selection of patients and identification of novel gene clusters for targeted therapy design, and improved prognostication to facilitate both clinician and patient decision-making [318].

Liu LX et al reported that the differentially expression cell cycle/growth regulator in GC showed a stronger tendency toward cell proliferation with 2.7-fold up-regulation of *CK1*. The promoter genes of apoptosis were down-regulated, including caspase-8 precursor, caspase-9 and caspase-10. Among the oncogene/tumor suppressor genes, *ABL2* was down-regulated. In addition, some genes were up-regulated, including *MMP-2, MMP-16(MT3-MMP), SKY, CD9* and semaphorin *V*. A number of genes were down-regulated, including neuroendocrine-dlg, retinoic acid receptor gamma and tumor suppressor DCC

colorectal. In general, The expression of the cancer progression genes were up-regulated, while the expression of anti-cancer progression genes were down-regulated [319]. Serial analysis of gene expression carried out by Yasui W et al shows that the commonly up-regulated genes in GC in comparison with normal gastric epithelia included *CEACAM6*, *APOC1* and *YF13H12*. By comparing gene expression profiles of GCs at early and advanced stages, several genes differentially expressed by tumor stage were also identified, including *FUS, CDH17, COL1A1* and *COL1A2,* which should be novel genetic markers for high-grade malignancy. *REGIV* is one of the most up-regulated genes in a SAGE library of a scirrhous-type GC [320]. According to the cDNA array experiments performed by El-Rifai W et al between tumor samples and normal gastric epithelial tissue reveals that the up-regulated genes had expression ratios ranging from 2.5 to 16, whereas the down-regulated genes had a range from -2.5 to -16. No variation in gene expression was detected in the analysis of the xenografted tumors versus the primary tumors, indicating that the xenografts represented the primary tumors well. Thirty-eight genes showed altered gene expression in 5 or more samples (>45%). Thirty-one genes were up-regulated and seven genes were down-regulated. The most abundantly up-regulated genes included genes such as *S100A4*, *CDK4*, *MMP14* and beta catenin [321]. Shao Y et al reported that genes related to cell cycle, growth factor, cell adhesion, and matrix remodeling were differentially expressed in gastric adenocarcinoma tissues [322]. Study by Zhang XQ et al shows that most of the overexpressed genes were those related to cell adhesion, cell motility, matrix reconstruction, cell

proliferation and/or signal transduction; while genes related to defense response, toxicoid metabolism, DNA repairing, nuclear-cytoplasmic transport and/or anti-apoptosis made up the main list of the underexpressed genes [323]. The quick and high-throughout method of profiling gene expression by cDNA or oligonucleotide array can provides an overview of key factors that may involved in GC, and may aid the study of GC carcinogenesis and provide molecular targets for diagnosis and therapy. The precise relationship between the altered genes and gastric carcinogenesis is a matter for further investigation.

2.12.3. *Next-generation sequencing (NGS)*

A new generation of non-Sanger-based sequencing technologies has delivered on its promise of sequencing DNA at unprecedented speed, thereby enabling impressive scientific achievements and novel biological applications. With the ultimate goal of deciphering the human genome, the throughput requirement of DNA sequencing grew by an unpredicted extent, driving developments such as automated capillary electrophoresis. Laboratory automation and process parallelization resulted in the establishment of factory-like enterprises called sequencing centers that house hundreds of DNA sequencing instruments operated by cohorts of personnel. However, even successful completion of the two competing human genome projects did not satisfy biologists' hunger for even greater sequencing throughput and, most importantly, a more economical sequencing technology [324].

Today three commercial next-generation DNA sequencing systems are available: namely Roche's (454) GS FLX Genome Analyzer marketed by Roche

Applied Sciences, Illumina's Solexa 1G sequencer, and most recently Applied Biosystem's SOLiD system [324].

2.12.3.1 *Illumina genome analyzer*

Introduced in 2006, the Illumina Genome Analyzer is based on the concept of 'sequencing by synthesis' to produce sequence reads of about 32–40 bp from tens of millions of surface amplified DNA fragments simultaneously **(Figure 2.15)**. Starting from a mixture of single-stranded, adaptor oligo-ligated DNA fragments, it involves using a microfluidic cluster station to add these fragments to the surface of a glass flow cell. Each flow cell is divided into eight separate lanes, and the interior surfaces have covalently attached oligos complementary to the specific adapters that are ligated onto the library fragments. Hybridization of these DNAs to the oligos on the flow cell is followed by a subsequent incubation with reactants and an isothermal polymerase that amplifies the fragments in a discrete area on the flow cell surfaces. The flow cell is placed into a fluidics cassette within the sequencer, where each cluster is supplied with polymerase and four differentially labeled fluorescent nucleotides that have their 30-OH chemically inactivated to ensure that only a single base is incorporated per cycle. Each base incorporation cycle is followed by an imaging step to identify the incorporated nucleotide at each cluster and by a chemical step that removes the fluorescent group and deblocks the 30 end for the next base incorporation cycle. At the end of the sequencing run, the sequence of each cluster is computed and subjected to quality filtering to eliminate low-quality reads [325].

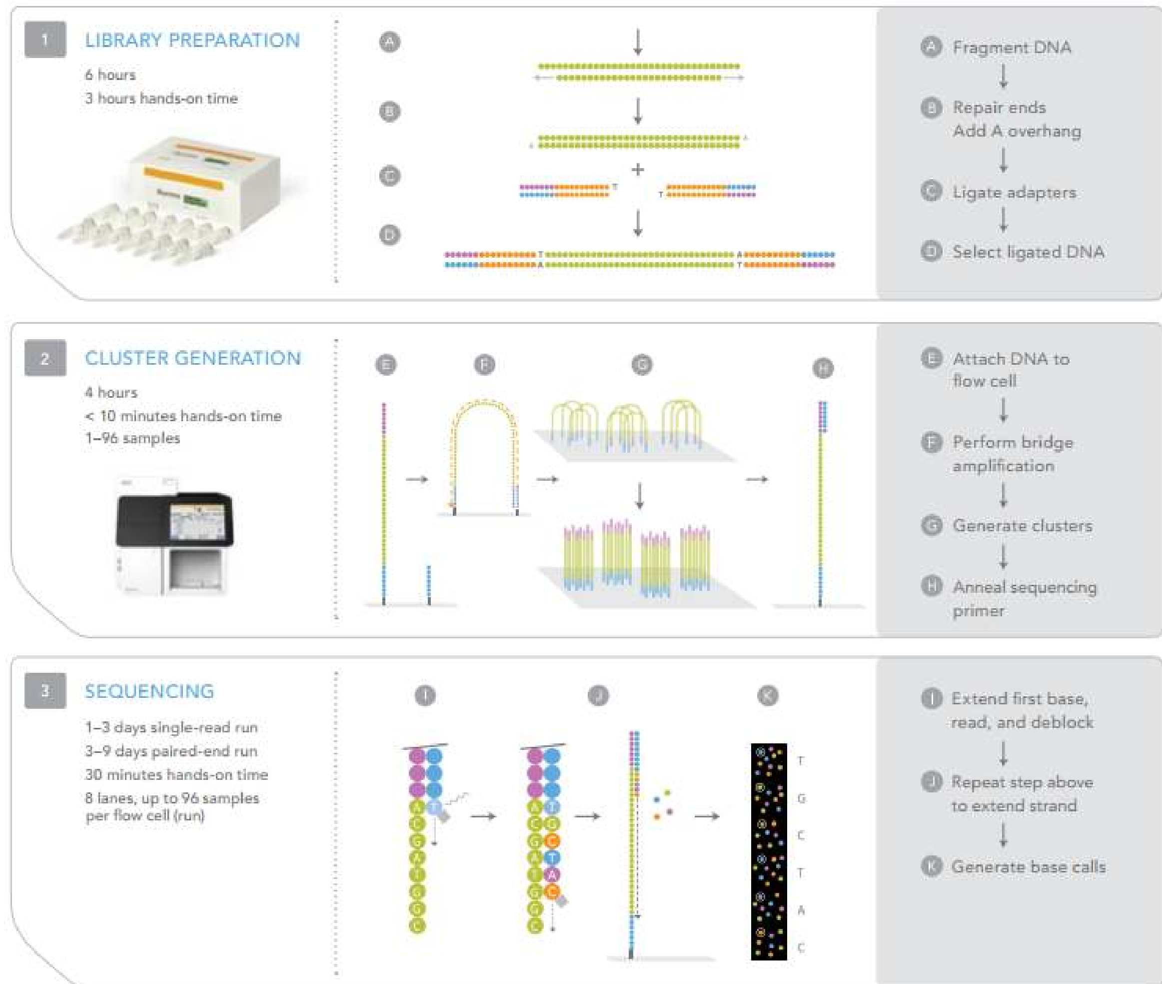


Figure 2.15. Illumina workflow. Starting from similar fragmentation and adapter ligation steps, the library is added to a flow cell for bridge amplification. The cluster fragments are denatured, annealed with a sequencing primer and subjected to sequencing. (Source: Next Generation Sequencing Brochur, wwweurofinsdna.com)

Finally, NGS has applications that are immediately relevant to the medical field. In cancer genetics, for example, specific cancer alleles can now be detected in tissues through ultra-deep sequencing of genomic DNA, in instances where previous Sanger-based trails have failed [326]. Short read length, initially deemed a major drawback of next-generation sequencing, becomes a blessing when the Sanger-based 700-bp read length is traded for a much larger number of sequence

reads [324]. NGS technology has revolutionized the study of cancers. Through matched normal-tumor pairs, it is now possible to identify genome-wide germline and somatic mutations. The generation and analysis of the data requires rigorous quality checks and filtering, and the current analytical pipeline is constantly undergoing improvements [327]. The compendium of somatic alterations in a cancer genome is shaped by multiple intrinsic and extrinsic processes, including exposure to mutagens, selective pressures active in the tissue microenvironment, genomic instability and DNA repair pathways [328]. The advent of massively parallel sequencing heralds an era in which unbiased, genomewide mutation screens allow the consequences of these processes to be discerned and decoded [327].

Chapter 3: Aims and Objectives

The aim of the study is to understand the underlying mechanisms involved in the carcinogenesis of gastric cancer (GC) in Northeast (NE) India where a very high incidence of GC is reported and search for possible markers to assist in both diagnosis and therapeutic approaches. The specific objectives of the present study are defined as under.

1. ***To determine the association of T1, M1 and P1 polymorphism in glutathione S – transferase genes and gastric cancer risk in Northeast population of India***

   Genotypes responsible for interindividual differences in ability to activate or detoxify genotoxic agents are recognized as biomarkers of susceptibility. Among the most studied genotypes are human glutathione transferases. The relationship of genetic susceptibility was studied especially in relation to the genetic polymorphism of glutathione S-transferase genes by using PCR-RFLP method and confirmation was done by sequencing 10 samples for each genotype.

2. ***To determine the association of tumor suppressor p53 Arg72Pro polymorphism and risk of gastric cancer in the Northeast general population of India***

   p53 is an important tumor suppressor, normally preventing cancer development via apoptosis. A genomic Arg72Pro substitution in the *p53* protein has important influence on cell death via apoptosis, which could be beneficial. We therefore tested the hypotheses that this polymorphism influences the risk of GC in NE population of India. PCR-RFLP method was used and confirmation was carried out by sequencing.

3. ***To evaluate PCR assays for detection of the presence of Helicobacter pylori in gastric cancer patients of Northeast India***

Several techniques such as culture, histology, rapid urease test etc have varying sensitivity and specificity for *H . pylori* detection. PCR being a highly efficient and reliable molecular technique for detection of various microorganism, different PCR using *H. pylori* specific genes Viz *UreA, GlmM* and *16SrRNA* were compared for their specificity and sensitivity for detection of *H. pylori*.

4. ***To study the gene expression profile of gastric cancer tissues in association with environmental risk factors***

Microarray technology is capable of determining the expression levels of thousands of genes in a biological sample simultaneously. This makes it widely used in cancer research. Characterization of these genes will help to elucidate the pathways and processes of carcinogenesis. It can be used to help clinical decision making, such as predicting therapy response, etc. 'OciChip Human A' chip (Ocimum Biosolution, Hyderabad, India) which contained 20160 genes has been used for the study.

5. ***To study genomic alteration involved in the process of carcinogenesis using next generation sequencing technology in matched normal and gastric tumor tissue***

Mutations are hallmark of cancers and identification of the mutations is imperative in our understanding of the disease. The advance in next generation sequencing (NGS) has transformed the way to identify mutations. It enables identification of somatic mutations, including base substitutions and indels. The

rapid increase in NGS publications recently illustrated the potential of the technology, reporting rare mutations in various cancers, many previously undetected. Solexa plateform was used and specific regions of one hundred and sixty nine genes were analyzed by sequencing based on NGS technology.

Chapter 4: Determination of Association of T1, M1 and P1 Polymorphism in Glutathione S –Transferase Genes and Gastric Cancer Risk in Various Northeastern Population of India

---

Introduction

Northeastern states in India have reported a very high prevalence of cancers when compared with other regions of India [329, 330]. Prevalence of gastric cancer (GC) (AAR: 57.3 in males and 33.6 in females) is highest highest in the Aizawl district of Mizoram [331]. The use of tobacco is very high in the Northeast (NE) India, where locally prepared tobacco products such as Tuibur, a unique tobacco smoke-infused water, and Mieziol, a local cigarette made from vaihlo, are widely used [332]. The habit of chewing betel quid, containing fresh betel nut and slaked lime wrapped in betel are also very high in Assam region of NE India. The reactive intermediates from these carcinogenic food items are neutralized and conjugated by phase II family of enzymes such as glutathione-S-transferase (GST) [333, 334]. The resultant water-soluble and less-toxic conjugated product can easily be eliminated from the cell by phase III transport mechanisms for the elimination of glutathione conjugates. The detoxification efficiency of GST enzymes is determined by the presence, amount, and nature of the isoenzymes coded by *GSTT1*, *GSTM1*, and *GSTP1* genes. The allelic polymorphism of *GSTT1* and *GSTM1* are characterized by the deletion of a part of the gene. *GSTP1* polymorphism is a single base pair substitution where adenine is replaced by guanine, resulting in an amino acid change in which isoleucine (I105) is replaced by valine (V105) [221, 335]. Electrophilic compounds are reported to

be detoxified less efficiently in individuals with null genotypes of *GSTT1* and *GSTM1* or variant genotypes of *GSTP1* (Ile/Val and Val/Val) when compared with those with wild-type genotype [222].

The presence of *GSTT1* and *GSTM1* null genotypes have been reported to be associated with increased risk for several cancers including skin, lung, bladder, prostate, colorectal, and oral cancers [228, 336]. However, several other reports have failed to confirm this association [337, 338]. In fact, *GSTT1* null genotype had been reported to be a protective factor for oral cancer in a central Indian population [339].

Polymorphic variants of *GSTP1* have also been reported to increase the risk of various cancers [210, 223]. Previous studies of gene polymorphisms and risk for tobacco-associated cancers have suggested that the polymorphisms in *GSTT1*, *GSTM1*, and *GSTP1* increase cancer risk in tobacco consumers [340, 341]. The prevalence of tobacco and betel quid chewing habits as well as the occurrence of tobacco-associated cancers is high in the NE region of India.

In a recent study, we have reported a higher prevalence of *GSTT1* and *GSTM1* null genotypes in this region when compared with other regions of India [342]. *GSTT1* null genotypes have also been reported to be associated with premalignant lesions of oral leukoplakia in the Assam region [343]. However, the prevalence of polymorphism in GST genes in GC patients from this region is not well known. The individual difference in susceptibility to chemically induced carcinomas may possibly be attributed to the genetic differences in the activation or detoxification of carcinogens due to polymorphic variants of GST genes. In the

present study, the association of polymorphism of *GSTT1*, *GSTM1*, and *GSTP1* genes with GC risk was evaluated to find out if this could explain the unusually high prevalence of GC in the NE region of India. Exposure to the type and amount of environmental toxins is variable not only in different geographic regions, but also in different ethnic groups within the same geographic region because of variations in their dietary, social, and cultural habits, although the samples included in our study belonged to a common geographical region of India, the inhabitants of this region are of different ethnic origins. As the ethnically different population inhabiting this region of India has been presumably exposed to shared environmental factors such as pesticide exposure and high level of tobacco and betel quid consumption, we also aim to analyzed the data of different racial composition separately as well as a combined group.

Materials and Methods

Selection of cases

The present study was done on samples obtained from 133 (68 from Assam, 17 from Sikkim, and 48 from Mizoram) histopathologically confirmed cases with gastric adenocarcinoma. The patients were diagnosed at three different tertiary health facilities of NE India, including Dr. Bhubneshwar Borooah Cancer Institute, Guwahati, Assam; Sir T.N.M. Hospital, Gangtok, Sikkim; and Civil Hospital, Aizawl, Mizoram, between 2006 and 2008. Questionnaires containing information on age, sex, region of origin, occupation, duration, and type of tobacco and betel quid consumption habits were recorded by interviewing all participating individuals.

Selection of controls

Samples obtained from unrelated voluntary healthy individuals who were accompanying the patients to the hospital were included as controls. The cases were matched for age, sex, and ethnicity with 267 normal healthy controls (107 from Assam, 72 from Sikkim, and 88 from Mizoram). Questionnaires containing information on age, sex, region of origin, occupation, duration, and type of tobacco and betel quid consumption habits were recorded by interviewing all participating individuals. Institutional ethical clearance was obtained as per the guidelines. An informed consent was signed and obtained from all subjects.

Collection and processing of samples

Two to 3mL of peripheral blood samples were collected in tubes containing ethylenediaminetetraacetic acid, stored in a 20°C freezer, and transported in dry ice to National Institute of Pathology.

DNA extraction from peripheral blood lymphocyte

The phenol-chloroform method has been widely used in molecular biology. Phenol-chloroform method involves the serial addition of several chemicals. Sodium dodecylsulfate (SDS) and proteinase K are added to break open the cell walls and to break down the proteins. Next a phenol/chloroform mixture is added to separate the proteins from DNA. As the DNA is more soluble in the aqueous portion of the organic-aqueous mixture, when centrifuged, the unwanted proteins and cellular debris are separated away from the aqueous phase and double stranded DNA molecules can be cleanly transferred for analysis.

Protocol

Five (5) ml of peripheral blood sample was collected from GC patients and healthy control individuals. Genomic DNA was extracted from blood using a standard phenol-chloroform extraction method. Blood was first digested with lyses buffer I (30mM Tris, 5mM EDTA, and 50mM NaCl) and lyses buffer II (20%SDS, 100μg/ml Prot.K). The mixture was incubated for 3hr at 45°C under agitation. All centrifugation steps were carried out at 10000g, 10min, 4°C. Lysate was centrifuged and the supernatant was collected and stored on ice until the phenol/chloroform/isoamyl alcohol extraction step to prevent DNA degradation that might occur during the physical cell lysis treatment. Then DNA was extracted from this supernatant by adding an equal volume of phenol/chloroform/isoamyl alcohol mixture (25:4:1 v/v/v) (Sigma, USA). The tube was briefly vortexed to obtain an emulsion then centrifuged. The aqueous layer was collected. Traces of phenol were removed by adding an equal volume of chloroform. After centrifugation, the top layer was collected and DNA was precipitated overnight at -20°C with two volumes of cold isopropanol. The pellet of crude DNA was obtained by centrifugation for 20min and was washed with 1ml of 70% cold ethanol. It was air dried and resuspended in 200μl of 1X TE buffer (10mM Tris-Cl, 1mM; Na_2EDTA, pH 8) [344]. This genomic DNA was then used for genotyping of *p53* gene polymorphism studies.

Quantitative estimation of extracted genomic DNA

The estimation of concentration of genomic DNA in solution, obtained from different sources was done using UV spectrophotometer (Shimadzu

Corporation) at 260 nm and 280 nm and the concentration of DNA was calculated according to the following formula: Concentration of genomic DNA [2 μl DNA+998 μl TE buffer] = OD at 260 nm x 50 x 500.

Quality of genomic DNA

A good quality of high molecular weight genomic DNA was obtained having a concentration of 200-500 ηg/μl. The phenol-chloroform extracted genomic DNA from the blood of normal and cancer patients were checked for their quality and quantity in an ethidium bromide stained 1% agarose gel (1g agarose in 1X TBE). High quality DNA obtained was evident from the presence of a single intact band without any smearing or degradation **(Figure 4.1)**.

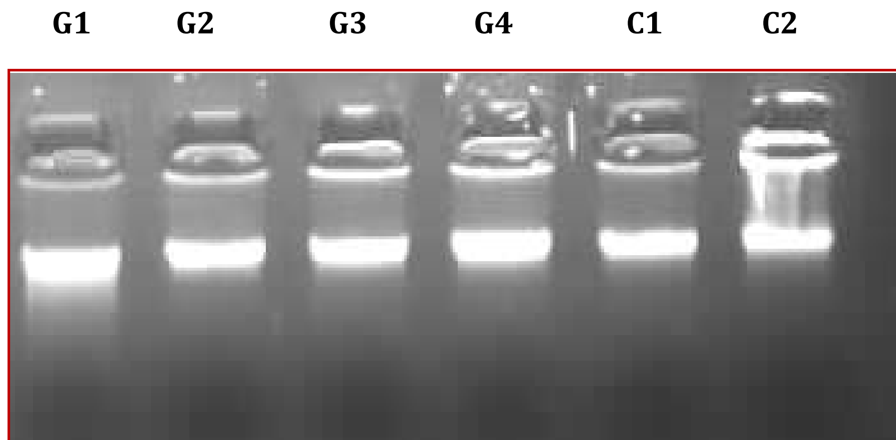


Figure 4.1: Estimation of quality of genomic DNA as visualized on an ethidium-bromide stained 1% agarose gel. Lanes G1 to G4 are showing genomic DNA extracted from GC patients while Lane C1 and C2 are showing genomic DNA from Controls

Genotyping of GSTT1 and GSTM1

A multiplex polymerase chain reaction (PCR) method was used to detect the presence or absence of the *GSTT1* and *GSTM1* genes in the genomic DNA samples of patients and controls [345]. Twenty-five microliters of PCR mixture was prepared by mixing 2.5 mL of 10X Taq buffer, 1 mL of 25mM MgCl2, 0.5 mL of 10mM dNTP mix, 0.5 mL of each forward and reverse primers (10 pM), 50–100 ng of template DNA, and 1 unit of Taq polymerase (M/s Fermentas, Lithuania). The primers were synthesized by M/s Microsynth, Germany. The primer pairs were 5’-TTCCTTACTGGTCCTCACATCTC-3’ and 5’-TCACCGGATCATGGCCAGCA-3’ for *GSTT1*, 5’-GAACTCCCTGAAAAGCTAAAGC-3’ and 5’-GTTGGGCTCAAATATACGGTGG-3’ for *GSTM1*, and 5’-CAACTTCATCCACGTTCACC-3’ and 5’-GAAGAGCCAAGGACAGGTAC-3’ for beta-globin. Beta-globin (268-bp fragment) was used as an internal control to ensure PCR amplification if the samples had null genotypes of *GSTM1* and *GSTT1*. To test for contamination, negative controls (without template) were included in every PCR run. PCR was carried out as follows: denaturation at 94°C for 4 min; followed by 20 cycles of denaturation at 93°C for 1 min, annealing at 60°C for 1 min, and extension at 72°C for 1 min; then additional 15 cycles of denaturation at 93°C for 1 min, annealing at 50°C for 1 min, extension at 72°C for 1 min; and final extension at 72°C for 10 min. The PCR products were electrophoresed in 2% agarose gels. The absence of 459 bp band indicated *GSTT1* null genotype and the absence of 219 bp indicated *GSTM1* null genotype **(Figure 4.2)**.

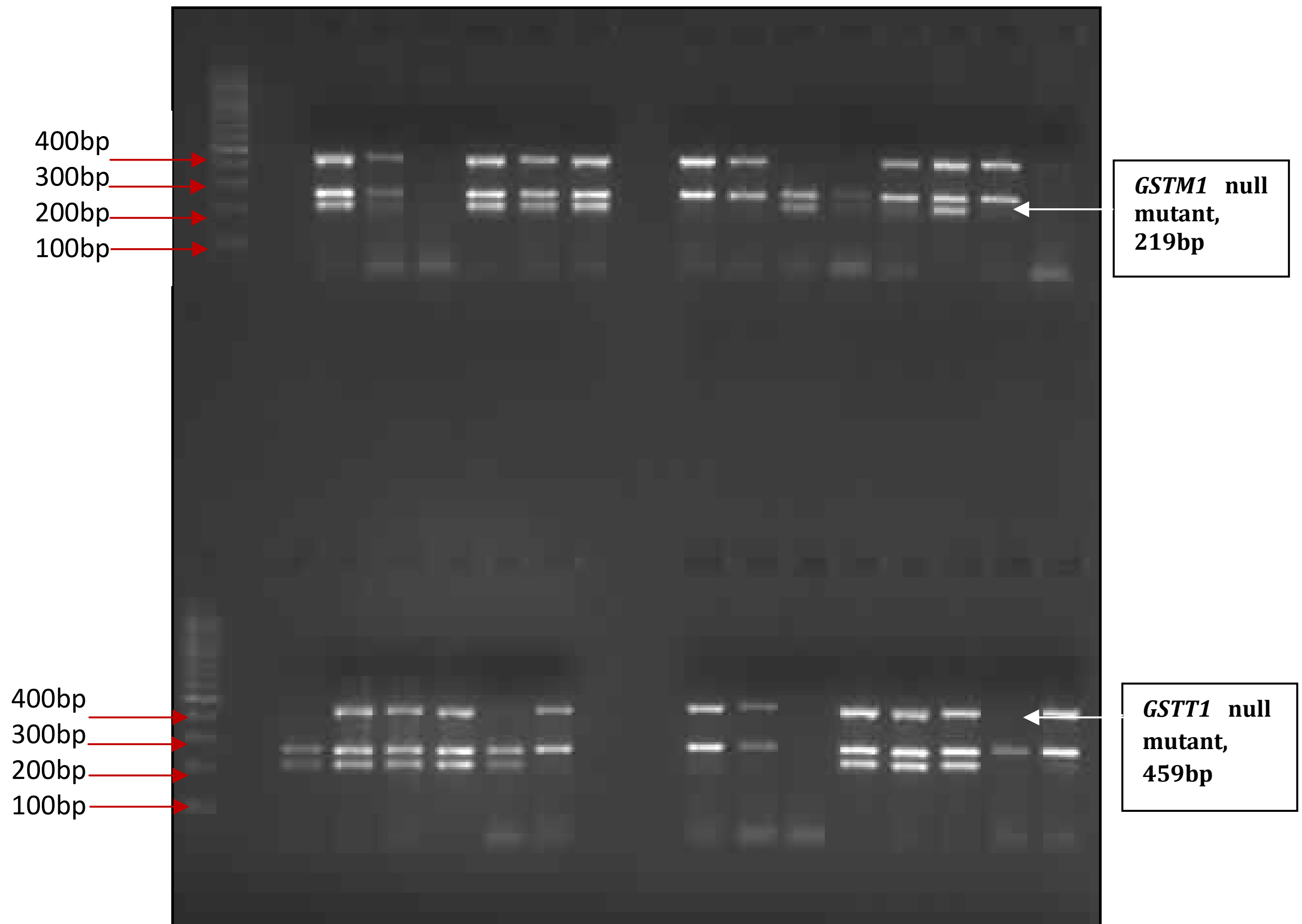


Figure 4.2 Agarose gel electrophosresis showing null mutations in GSTM1 and GSTT1 genes in gastric cancer usinpatients and control using 100bp marker (M/s Fermentas, Lithuania)

Genotyping of GSTP1

Polymorphic variants of *GSTP1* were detected by PCR–restriction fragment length polymorphism. Twenty-five microliters of PCR mixture was prepared by mixing 2.5 mL of 10X Taq buffer, 2 mL of 25mM MgCl2, 1.25mL of 10mM dNTP mix, 1.25 mL of each forward (5'-CCAGTGACTGTGTGTTGATC-3') and reverse (5'-

CAACCCTGGTGCAGATGCTC-3’) primers (10 pM) for *GSTP1*, 50–100ng of template DNA, and 1 unit of Taq polymerase. Cycling conditions were as follows: initial denaturation at 94°C for 3min; followed by 35 cycles of 94°C for 1min, 58°C for 30 s, and 72°C for 30 s; and a final extension at 72°C for 10 min. The PCR product of *GSTP1* was 189 bp in size. After testing for the amplification of PCR products in 2% agarose gel, 10 mL of PCR product was digested using BsmA1 restriction enzyme (M/s Fermentas, Lithuania) in a reaction volume of 30 mL by overnight incubation at 37°C. The products were separated by electrophoresis in 4% agarose gel in 0.6X TBE buffer. On the basis of the band patterns, three genotypic variants were identified. The wild-type genotype [*Ile/Ile* (A/A)], completely undigested, was represented by a single band at 189 bp. The genotypic variant [*Val/Val* (G/G)] was completely digested, yielding two bands of 148 and 41 bp with absence of a 189-bp fragment. The digested product that yielded all the three bands represented the heterozygous genotype [*Ile/Val* (A/G)] **(Figure 4.3)**. Genotyping procedures were validated by sequencing of representative samples **(Figure 4.4)**.

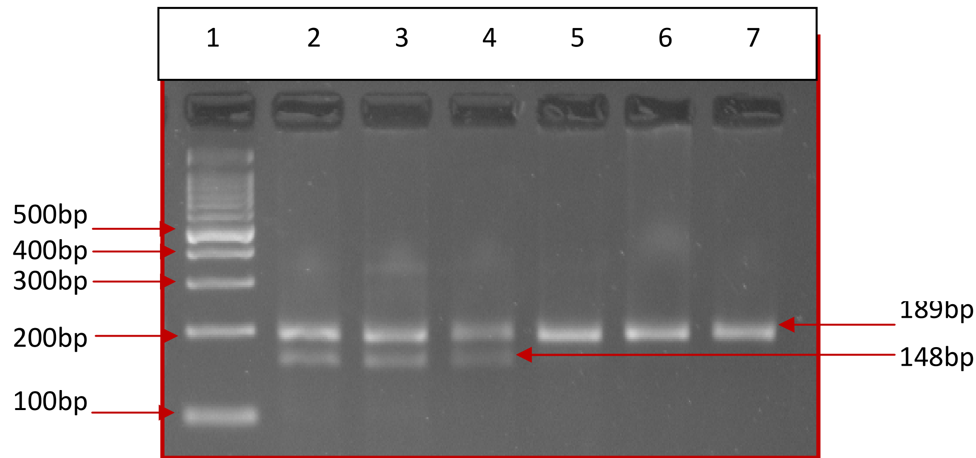


Figure 4.3: Representative gel picture of PCR-RFLP method carried out to map polymorphism of GSTP1. Lane 1-100bp ladder (M/s Fermentas, Lithuania). Lanes 2, 3, 4 - Ile/Val (heterozygous) allele. Lanes 5, 6, 7 - Ile/Ile (homozygous) allele

DNA Sequencing

Genotyping of selected cases and controls were confirmed by sequencing. No discrepancies were observed

PCR amplicon purification: The PCR product were purified using gel purification kit from Fermentas, Lithuania according to the manufacturer protocol.

Sequencing rxn: Sequencing was carried out by mixing 60-100ng (PCR amplicons), 2 mol of gene-specific primer, 1µl of BigDye Terminator III mix (Applied Biosystems, CA) and made up to a final volume of 10µl with nuclease-free water. The mixture was then cycled as follows:

Cycle conditions

94°C for 5 min

96 ^{o}C for 10 sec

50 ^{o}C for 5 se 30 Cycles

60 ^{o}C for 4 min

The sequencing products were precipitated by adding 1/50 3M Sodium Acetate pH 5.2, 1/10 125mM EDTA, 2X volume of 100% ethanol and incubated at RT for 10 min. The mixture was then centrifuged at 3000*g* for 30min and the supernatant removed. The samples were then washed with 70% ice-cold ethanol and centrifuged at 3000*g* for 5 min, followed by an invert spin and then tubes were air dry at RT. 10 µl of Hi-Di formamide was then added to each well and mixed properly. After vortexing, 10 µl of sample were aliquoted into each well of 96-Well Optical Reaction Plate (Applied Biosystems). The plate was sealed with Plate Septa 96-Well (Applied Biosystems) before placed on GeneAmp PCR System 9700 (Applied Biosystems) and heated at 95°C for 5min before the temperature was rapidly down to 4°C and held for 3min. Then, the plate was placed on plate base and covered with plate retainer.

Capillary electrophoresis

The capillary electrophoresis and sequencing was performed with Data Collection Solfware Version 2.0 on ABI PRISM 3100 *xl* Genetic Analyzer. Before operating the instrument, 1X TBE butter (Applied Biosystems) and distilled water were changed every time. Raw data was then analysed using Sequence Analysis software v5.3 *(Applied Biosystems)*.

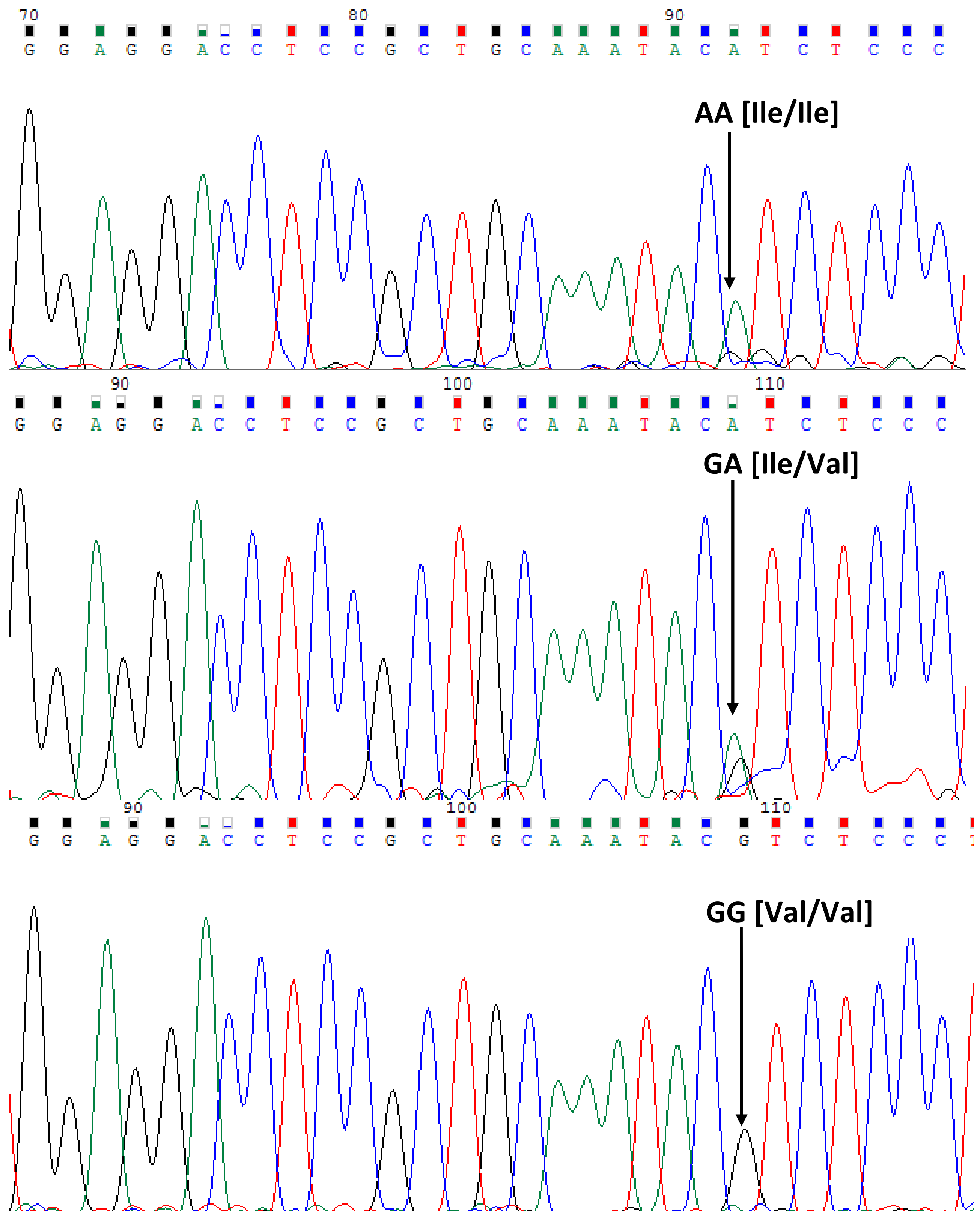


Figure 4.4. Electrogram showing the sequence analysis of different genotypes of GSTP1 polymorphism in gastric cancer samples

Statistical analysis

Stata 8.0 version software was used for statistical analysis. Hardy–Weinberg equilibrium test was done to compare the difference between the observed and expected for *GSTP1* genotype. The association for the considered covariates including tobacco use (no/yes), chewing (no/yes), smoking (no/yes), and the polymorphisms *GSTT1* (present/null), *GSTM1* (present/null), and *GSTP1* (wild type/variant) were assessed by applying the conditional logistic regression analysis (age and sex matched). The conditional logistic regression analysis was performed to get risk estimates for association of GC with these risk factors and genotypic variables and these results were interpreted in terms of adjusted odds ratios (ORs) along with their corresponding 95% confidence intervals (95% CIs). To get the estimates of regression coefficient, standard error, and statistical significance for each of the variables, enter method was used during conditional logistic regression analysis. **Table 4.1A** and **4.2B** exhibit the results of the conditional logistic regression analysis for GC with the risk factors and genotypic variables. The adjusted estimates for the specific risk factors (tobacco chewing, tobacco smoking, betel quid chewing and alcohol consumption) and genotypic variables (*GSTT1*, *GSTM1*, and *GSTP1*) were adjusted for all other risk factors and genotypes under consideration. **Table 4.2A** and **4.2B** illustrates the region-specific adjusted risk estimates for GC with the adjusted estimates in accordance to as explained for **Table 4.1A** and **4.1B** above.

Results

One hundred thirty-three patients with GC (68 from Assam, 17 from Sikkim, and 48 from Mizoram) and 267 normal healthy controls (107 from Assam, 72 from Sikkim, and 88 from Mizoram) were included in the study. The mean age of patients with GC and normal healthy controls was 52±12 and 49±10 in Assam, 60±11 and 62±10 in Sikkim, and 54±12 and 54±12 in Mizoram, respectively. Of these, 63 (47%) cases with GC and 131 (49%) controls were tobacco chewers, 79 (53%) cases and 128 (47.8%) controls were tobacco smokers, 44 (33%) cases and 83 (31%) controls were alcohol consummers and 94 (70.6%) cases and 168 (62.9%) controls were betel quid chewers. Betel quid chewing habits were higher in GC cases when compared with a control population, but this difference was statistically insignificant (OR=1.45, 95% CI: 0.90-2.35, p=0.12) **(Table 4.1A)**. Region wise analysis reveals that betel quid chewing increase threefold risk of developing GC in Assam population when compared with controls and it was statistically significant (OR=3.61, 95%CI=1.06-12.21, p=0.04). The frequency of tobacco smoking was 45% in GC cases and 50% in controls in Assam population respectively and the estimated OR showed 49% less chance of risk for GC (OR=0.51,95%CI=0.21-0.87,p=0.02) **(Table 4.2A)**.

Table 4.1 A. Association of tobacco, betel quid and alcohol consumption with risk of gastric cancer in Northeast population

Risk Factors	Cases N=133	Controls N=267	Adjusted* OR	(95% C.I.)	P-value
	n (%)	n (%)			
Tobacco Chewing	63(47)	131(49)	1.08	(0.68-1.72)	0.75
Tobacco Smoking	79(53)	128(47.8)	0.94	(0.62-1.44)	0.78
Betel quid Chewing	94(70.6)	168(62.9)	1.45	(0.90-2.35)	0.12
Alcohol Consumption	44 (33)	83(31)	0.90	(0.57-1.43)	0.68

Table 4.1B- Association of *GSTT1*, *GSTM1*and *GSTP1* genotypes with gastric cancer in Northeast population

Genotype	Cases N=133	Controls N=267	Adjusted* OR	(95% C.I.)	P-value
GSTT1					
Present	83(62)	182(68)		1.00	
Null	50(38)	85(32)	1.27	(0.78-2.06)	0.33
GSTM1					
Present	84(63)	148(55)		1.00	
Null	49(37)	119(45)	0.74	(0.47-1.16)	0.19
GSTP1					
Ile/Ile	75(56)	171(64)		1.00	
Ile/Val or Val/Val	58(44)	96(36)	1.29	(0.80-2.07)	0.29

A p-value of <0.05 is considered statistically significant.
***Adjusted with all other risk variables under consideration.**
OR, odd ratio; 95% CI, 95% Confidencial interval

Table 4.2A. Region specific association of tobacco, betel quid chewing and alcohol consumption with risk of gastric cancer

Regions	Risk Factors	Cases n/N (%)	Controls n/N (%)	Adjusted* OR (95% C.I.)	P-value
Mizoram	**Age(years±SD)**	54±12	54±12		
	Sex (M/F)	33/15	60/28		
	Tobacco Chewing	17/48(35)	30/88(36)	0.98(0.37-2.62)	0.97
	Tobacco Smoking	40/48(83)	72/88(86)	1.06(0.49-2.29)	0.89
	Betel quid chewing	29/48(60)	54/88(64)	1.23(0.57-2.68)	0.60
	Alcohol consumption	13/48(27)	17/88(20)	1.50(0.65-3.45)	0.34
Assam	**Age(years±SD)**	52±12	49±10		
	Sex (M/F)	50/18	82/25		
	Tobacco Chewing	38/68(56)	64/107(60)	1.07(0.56-2.06)	0.83
	Tobacco Smoking	31/68(45)	53/107(50)	0.51(0.21-0.87)	**0.02**
	Betel quid chewing	61/68(93)	90/107(84)	3.61(1.06-12.21)	**0.04**
	Alcohol consumption	27/68(39)	/107(32)	0.86(0.42-1.78)	0.69
Sikkim	**Age(years±SD)**	60±11	62±10		
	Sex (M/F)	11/6	50/22		
	Tobacco Chewing	8/17(47)	39/72(53)	0.89(0.24-3.37)	0.86
	Tobacco Smoking	8/17(47)	27/72(37)	1.08(0.29-4.0)	0.90
	Betel quid chewing	6/17(35)	25/72(34)	0.11(0.01-1.03)	**0.05**
	Alcohol consumption	4/17(24)	24/72(33)	2.0(0.22-17.9)	0.53

p-value of <0.05 is considered statistically significant
***Adjusted with all other risk variables under consideration**
OR, odd ratio; 95% CI, 95% Confidencial interval
SD, standard deviation of the mean

Table 4.2B- Region specific distribution of *GSTT1, GSTM1* and *GSTP1* genotypes among gastric cancer cases and controls

	Risk Factors	Cases	Controls	Adjusted	
		n/N (%)	n/N (%)	OR (95% C.I.)	P
Mizoram	T1 Present			1.00	
	T1 Null	24/48(50)	45/88(51)	0.91(0.46-1.79)	0.78
	GSTM1 Present			1.00	
	Null	18/48(39)	47/88(53)	0.65(0.29-1.12)	0.10
	P1 Ile/Ile			1.00	
	P1 Ile/Val or Val/Val	19/48(40)	30/88(34)	1.36(0.61-3.03)	0.45
Assam	T1 Present			1.00	
	T1 Null	18/68 (27)	13/107(12)	3.07(1.33-7.09)	0.009
	GSTM1 Present			1.00	
	Null	20/68(29)	36/107(33)	1.08(0.52-2.27)	0.83
	P1 Ile/Ile			1.00	
	P1 Ile/Val or Val/Val	34/68(50)	39/107(36)	1.58(0.83-2.99)	0.16
Sikkim	T1 Present			1.00	
	T1 Null	8/17(47)	27/72(37)	2.12(0.62-7.3)	0.23
	GSTM1 Present			1.00	
	Null	11/17(62)	37/72(51)	1.64(0.47-5.67)	0.43
	P1 Ile/Ile			1.00	
	P1 Ile/Val or Val/Val	5/17(29)	28/72(38)	0.63(0.19-2.07)	0.44

p-value of <0.05 is considered statistically significant.

***Adjusted with all other risk variables under consideration.**

OR, odd ratio; 95% CI, 95% Confidencial interval

Table 4.3- Combined effects of *GSTT1*, *GSTM1* and *GSTP1* genotypes in the study subjects

Genotype	Cases	Controls	Adjusted	
GSTT1/GSTM1	**N=133**	**N=267**	**OR (95% C.I.)**	**P**
T1(+)/M1(+)	58 (44)	101(38)	1.00	
T1(-)/M1(+)	26 (20)	47(18)	0.98 (0.52-1.83)	0.94
T1(+)/M1(-)	25 (19)	81(30)	0.59 (0.33-1.06)	0.08
T1(-)/M1(-)	24 (39)	38 (61)	1.05 (0.53-2.09)	0.88
GSTT1/GSTP1				
T1(+)/P1(+)	47 (35)	114 (43)	1.00	
T1(-)/P1(+)	28 (21)	57 (21)	1.18 (0.63-2.18)	0.61
T1(+)/P1(-)	36 (27)	68 (26)	1.22 (0.69-2.13)	0.49
T1(-)/P1(-)	22 (17)	28 (11)	1.75 (0.83-3.69)	0.14
GSTM1/GSTP1				
M1(+)/P1(+)	48 (36)	89 (33)	1.00	
M1(-)/P1(+)	27 (20)	82 (31)	0.59 (0.32-1.07)	0.08
M1(+)/P1(-)	36 (27)	59 (22)	1.03 (0.58-1.83)	0.93
M1(-)/P1(-)	22 (16)	37 (14)	1.08 (0.54-2.17)	0.83

A p-value of <0.05 is considered statistically significant.
***Adjusted with all other risk variables under consideration.**
OR, odd ratio; 95% CI, 95% Confidencial interval

Discussion

Lack of *GSTT1* and *GSTM1* isoenzymes activity or differences in the activity and distribution of allelic variants of *GSTP1* have been earlier implicated in increased cancer risk following exposure to environmental carcinogens. Of these, *GSTT1* is responsible for the biotransformation of the constituents of tobacco smoke, such as alkyl halides, and its derivatives, such as monohaloethanes, ethylene oxide, benzo(a)pyrene diol epoxide, and acrolien [223, 346]. *GSTM1* subfamily metabolizes lipid peroxidation products, DNA hydroperoxides, and polyaromatic hydrocarbons such as benzo [alpha] pyrene [347, 348]. The *GSTP1* enzyme is widely expressed in tumor cells and is responsible for the detoxification of benzo(a)pyrene diol epoxide and acrolein present in cigarette smoke. The *GSTP1* isoform is also known to metabolize tobacco-related carcinogens with elimination of the oxidative products of thymidine or uracil propenal [16].

Polymorphism of the *GSTT1* and *GSTM1* genes, which are located on chromosome 22q11.2 and 1p13.3, respectively, results in deletion of their loci with subsequent loss of specific enzymatic functional activity and reduced ability to detoxify potentially toxic substances. Polymorphism of *GSTP1* gene, which is located on chromosome region 11q13, shows a single base pair substitution where adenine is replaced by guanine, resulting in amino acid isoleucine (I105) being replaced by valine (V105) [221, 335]. As GST genes are involved in the detoxification of tobacco constituents, there is a possibility that the genetic

polymorphisms of these enzymes may be a high risk factor for the widespread occurrence of tobacco-associated malignancies in NE Indians.

The association of tobacco consumption with cancer such as oral and lung have been well documented [349]. However, in our study no significant association of tobacco consumption in any form was found to be associated with GC when NE population was considered as one group. This was in contrast to earlier reports where tobacco consumption was found to be significantly associated with GC [10, 349].

Region wise analysis reveals that betel quid chewing increase threefold risk of developing GC in Assam population only. Besides the frequency of betel quid chewers in Assam was observed to be the highest (93%) to that of Aizawl (60%) and Sikkim population (35%). Tobacco chewing have 49% less chance of developing GC in Assam population and it is statistically significanct. This could be due to distribution of chewers being lower in cases than controls.

Earlier studies from different regions of the world have reported a higher risk for the occurrence of several cancers in patients with *GSTT1* and *GSTM1* null genotypes. However, many other studies have reported conflicting results. *GSTM1* null genotype has been reported as a risk factor for oral cancer [350, 351], GC [352-354], and lung cancer [355, 356]. This is in contrast to other reports where no significant association of *GSTM1* null genotype was found with risk of oral cancer [357], GC [358], and lung cancer [359, 360]. In fact, there are reports that have shown *GSTM1* null genotype as a protective factor for some cancers such as breast , oral and skin cancers [345, 361-363].

GSTT1 null genotype has been reported as a risk factor for GC [231, 364], whereas no significant association of *GSTT1* null genotype had been reported with GC in other studies [358]. As reported for *GSTM1* null genotype, *GSTT1* null genotype has been also reported as a protective factor for some cancers such as head and neck cancer [365], bladder cancer [366], and breast cancer [367]. A review of studies done on these cancers in India also showed conflicting results of association with GST polymorphism. In a study by Malik et al both *GSTT1* null and *GSTM1* null genotypes were reported as a significant risk for GC in the Kashmir valley population [368].

In the present study, *GSTT1* null genotype was not found to be associated with risk of GC when the NE population was taken as one group. However, analysis of GST polymorphisms in different geographic regions of NE India showed *GSTT1* genotype to be a significant risk factor for in the Assam region of NE India. However *GSTM1* was found to confer a protective effect to GC when NE population was taken as one group and also in Assam and Mizoram. GSTs have been reported to have dual role [369], beside its function as xenobiotic detoxifying enzyme, *GSTM1* has also been reported to influence DNA damage [370]. Possible explaination for the dual role (protective as well as risk) of *GSTM1* could be, first, due to different substract specificity and, second, due to possible interaction with environmental factors, which modify the risk associated with the gene.

GSTP1 variants have been reported as a risk factor for GC in the Lucknow region of North India [354], whereas no association has been found for GC in the

Kashmir valley [368]**.** Moreover, data from different geographical regions of India show large variation in different ethnic groups in a healthy population [342]. In the present study, the variant *GSTP1* Ile/Val and Val/Val genotypes were not significantly associated with GC when the NE population was taken as one group or when analyzed for different geographical regions though the variant genotypes (Ile/Val and Val/Val) of *GSTP1* were found more frequently in samples from patients with GC (44%) when compared with controls (36%), but the difference was not statistically significant. Subjects carrying the *GSTP1* Ile105Val Val/ Val genotype have a lower ability to detoxify electrophilic compounds than subjects carrying the wildtype genotype, Ile/Ile [222]. Variants in these genes may reduce an individual's ability to detoxify PAHs and could increase risk for various cancers, including GC [210, 223]. Several earlier studies have also reported conflicting results for *GSTP1* polymorphism, with both risk factor and no association having been reported [338, 371, 372]. In addition, epigenetic factors such as hypermethylation of the promoter region of *GSTP1* gene may lead to downregulated gene expression and reduced activity of the enzyme. Methylation of the *GSTP1* promoter region has been earlier found to be associated with some cancers, particularly prostate cancers, where it has been used for its early diagnosis and prognosis [373]. However, no such significant association has been so far reported for GC. The inconsistency in results of association of GST polymorphism with GC may be due to different ethnicity or interaction between different environmental and genetic factors.

The variation in the association of GST genes with various other cancer can also be attributed to the fact that the distribution of GST enzymes in different organs may also vary with the age and sex of different individuals. These factors may lead to variation in the carcinogenic concentration of toxins in different tissues and to a variable role of GST genotypes in different populations exposed to different environmental carcinogens as has been found in our study. To the best of our knowledge, this is the first study on these detoxifying genes involving GC in a high-risk region of India where the local population has peculiar betel quid and tobacco consumption habits.

Chapter 5: Determination of Association of p53 Gene, its Interaction with Tobacco, Betel quid and Alcohol Consumption and Risk of Gastric Cancer: a Case-Control Study in Northeast Population to Understand the Etiology

---

Introduction

The human *p53* tumor suppressor gene plays a central role in many cellular processes, regulating cell growth, DNA maintenance and apoptosis. It is an important component of DNA repair machinery in response to DNA damage induced by radiation or adduct formation [374]. This might explain the occurrence of the *p53* gene mutation and alteration in about 50% of all cancers. Studies have shown a relationship between tobacco smoke exposures, carcinogen-DNA adduct formation, tumor specific mutation of *p53* gene and cancer risk. *p53* gene plays a significant role in the regulation of cellular response to benzo[a]pyrene, one of the most important polycyclic aromatic hydrocarbon (PAH) compounds of tobacco smoke [374]. Moreover, recent studies have indicated that there is a strong coincidence in mutational hotspots and sites of preferential formation of PAH adducts along the *p53* gene in lung, laryngeal and head and neck cancers. Besides mutation, polymorphisms in *p53* gene have also been implicated in the process of tobacco carcinogenesis. Numerous polymorphism in the wild type *p53* have been reported both in coding and non coding regions [38]. Out of the five polymorphisms described in the coding region, polymorphisms in codon 47 and 72 in exon 4 are functionally well characterized. More common of the two, codon 72 polymorphism is a single base

substitution of cytosine for guanine, leading to arginine (A72) being replaced by proline (P72) [38] that has been reported to be associated with the risk of several cancers [32-36]. However, the results are conflicting with Pro/Pro genotype showing association with lung cancer [39-42] breast cancer [43, 44] and GC [45]. whereas Arg/Arg genotype being more prevalent in cervical cancer [46, 47]. However, no association between either genotype and cancer risk have also been reported for head and neck [48] and cervical cancer [49]. Literature available from India is limited and inconsistent. Two different studies have reported both Arg/Arg and Pro/Pro genotypes to be associated with risk of lung cancer [42, 50] whereas no association was reported with oral cancer [32, 51]. However another study done in relation to HPV status showed carriers of Arg/Arg genotype to be more susceptible in HPV positive oral cancer cases [375]. There are no reports on association of *p53* codon 72 polymorphism with GC from NE India.

Studies on codon 72 polymorphism have revealed striking ethnic differences [52, 53] have demonstrated that frequency of *p53* variant allele varies with latitude, increasing in a linear trend as populations near the equator. Thus ethnicity might be related to allelic distribution of the gene and its determinacy in disease involvement; however some studies do refute the ethnicity-risk confounding relationship [54]. NE part of India, due to its unique, strategic geographic location and the presence of linguistically, culturally and demographically diverse populations is a hotspot for population genetics. This area shows the world's highest incidences of cancers of tongue and stomach particularly in men, associated with tobacco use [11]. A high AAR for GC has been

reported from Mizoram (50.6 in males and 23.3 in females) and Sikkim (14.2 in males). The area also reports tobacco use in variety of ways of chewing and smoking that are different from the rest of India. Unlike the Western nations or the urban India, use of tobacco or alcohol in crude forms is more prevalent. High risk to cancer can be an outcome of either environmental and genetic risk factors or a complex interplay of both. Literatures have reported *p53* allelic polymorphisms to be possible predisposing factors for tumor development. The association of *p53* codon 72 polymorphism with cancer susceptibility remains uncertain and varies with ethnicity. NE India represents geographically and culturally an ethnically isolated population. These area reports high rate of tobacco usage in variety of ways of consumption, compared to the rest of Indian population [329, 332]. A number of smoking and smokeless tobacco products are in use all over the world. But unlike other smokeless tobacco products, a unique tobacco smoke–infused water is used in Mizoram and is locally known as tuibur. This product is made locally by passing smoke, generated by burning tobacco, through water until the preparation turns cognac in color and has a pungent smell. In vitro studies using the allium root test show the toxic nature of tuibur [332]. Indigenous crude devices are used for the production of tuibur on small scale. Users take about 5 to 10 mL tuibur orally and keep it in the mouth for some time and then spit it out. Most of the users take it several times a day. Meiziol, a local cigarette made from vaihlo (Nicotiana dadacum) tobacco. The tobacco content of each meiziol is about 0.8 to 1g. Lack of data on *p53* codon 72 polymorphism and high incidence of GC in NE region of India incited us to

explore and evaluate any relevance of this polymorphism in this ethnic population. We carried out a case control study on GC, in NE part of India. The role of *p53* codon 72 polymorphism and its interaction with tobacco, betel quid and alcohol use was also analyzed.

Materials and methods

The present case–control study was performed on 134 (68 from Assam, 18 from Sikkim and 48 from Mizoram) histopathologically diagnosed GC cases and a total of 282 (107 from Assam, 77 from Sikim and 98 from Mizoram) age sex match healthy controls collected during the period of december 2005 to 2008. All incident cases willing to participate in the study were recruited from the collaborating centers in NE India. Five ml of peripheral blood was collected in EDTA vials and stored under -70°C refrigeration. Sample batches were later sent to the working centre under frozen conditions. Only cases with stomach as their primary site of cancer were included. Controls included were healthy relatives of cancer patients, willing to participate in the study. All study subjects provided informed consent for participation in this research which was done under a protocol approved by the institutional ethics committee of Dr. B. Barooah Cancer Institute, Guwahati, Civil Hospital, Aizawl, and Sir T.N.M. Hospital, Gangtok. Information regarding smoking, usage of tobacco, betel quid and alcohol were obtained from subjects in a standard questionnaire used for all the centers.

DNA extraction from peripheral blood lymphocytes

Genomic DNA was extracted and purified using proteinase K phenol–chloroform extraction procedure [344] as described in chapter one.

PCR-RFLP Assay for p53 codon 72 polymorphism analysis

Standard PCRs were performed on GeneAmp PCR system 9700 (Applied biosystems). Each PCR reaction mixture (20µl) contained 0.2µM of each primer, 1.5 mM MgCl2, 0.2mM each dNTP, 0.75 unit of Taq polymerase and 500ng of genomic DNA. Reaction mixtures were preincubated for 10 min at 94°C. PCR conditions were 94C° for 45 s and 60°C for 45s, followed by 72°C for 45s for 40 rounds. The size of the PCR product is 199bp **(Figure 5.1)**. The primer sequence used for amplification in given below **(Table 5.1).**

Table 5.1: Oligonucleotide primer sequences used for amplification of p*53* gene

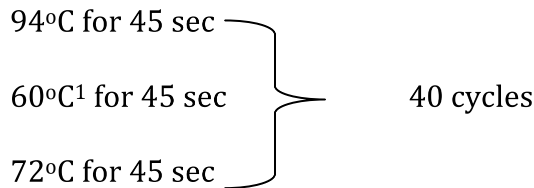
Gene	Primer Sequence	Tm	Amplicon size (bp)
p53	Forward:5'-TTG CCG TCC CAA GCA ATG GAT GA-3'	60°C	199
	Reverse: 5'-TCT GGG AAG GGA CAG AAG ATG AC-3'		

PCR Amplification

Component Concentration	Volume (final concentration)
10 X buffer (10X)	2.0µl (1X)
$MgCl_2$ (25mM)	1.2µl (1.5 mM)
dNTPs (5mM)	0.8µl (0.2mM)
Forward Primer (10µM /µl)	0.4µl (0.2µM /µl)
Reverse Primer (10µM /µl)	0.4µl (0.2µM /µl)
Taq Polymerase (5units/µl)	0.15µl (0.75units/ µl)
DNA Template	2.0µl (500ng/µl)
DNase free water	13.05µl

PCR amplification conditios were as follows

$94^{o}C$ for 10 min



$94^{o}C$ for 45 sec

$60^{o}C$[1] for 45 sec   40 cycles

$72^{o}C$ for 45 sec

$72^{o}C$ for 10 min

[1]Annealing temperatures were primer specific. Specific annealing temperatures are indicated in experiments where appropriate.

Electrophoresis of PCR amplicons

All PCR amplicons were electrophoresed and separated by molecular weight on either 1or 2% (m/v) agarose gels. The agarose was dissolved in 1X TAE (Tris Acetate Ethylene diamine tetra acetate) buffer (i.e. for a 1% gel, 1g of agarose/100 ml of TAE). This was microwaved at high setting for 1-2min, cooled to approximately 50, 1ul of ethidium bromide (10mg/ml) added and the gel poured and left to set. Approximately 1/5 of loading volume of loading dye (0.25% bromophenol blue, 0.25% xylene cyanol, 50% glycerol) was then added to the samples and mixed. The electrophoresis was carried out at 100V in a Bangalore genie System (Genei, India) for approximately 30mins and the image was captured using a Syngene UV system (Syngene, Frederick, MD) **(Figure 5.1)**.

Genotyping: Restriction analysis was performed by digesting the PCR products with 5 units of restriction enzyme BstUI (New England Biolabs, Beverly, MA) at 60°C for 16hrs. The digested products were electrophoresed through a 2.5% agarose gel and stained with ethidium bromide **(Figure 5.2)**. The *Arg/Arg* genotype resulted in the presence of two bands of 113 base pairs and 86 base pairs each and the *Pro/Pro* genotype resulted in a single uncut band of 199 base pairs. PCR results were evaluated without knowledge of case and control status of the sample.

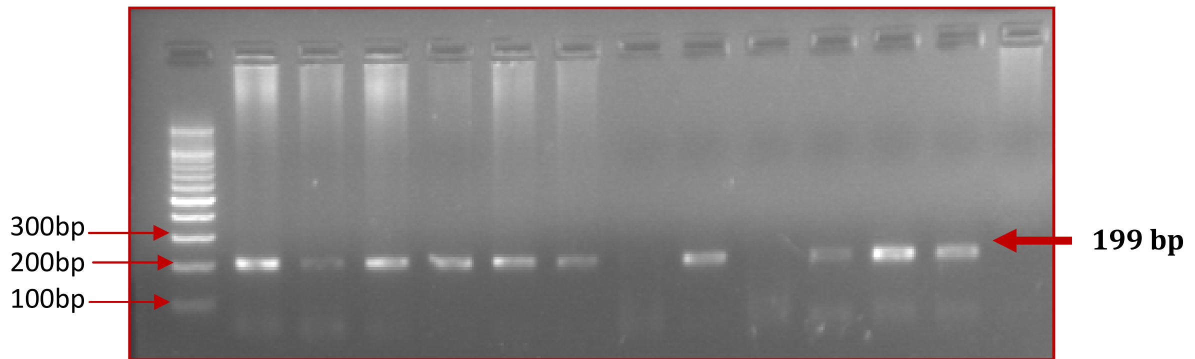


Figure 5.1. Representing PCR amplification of p53 gene (199bp)

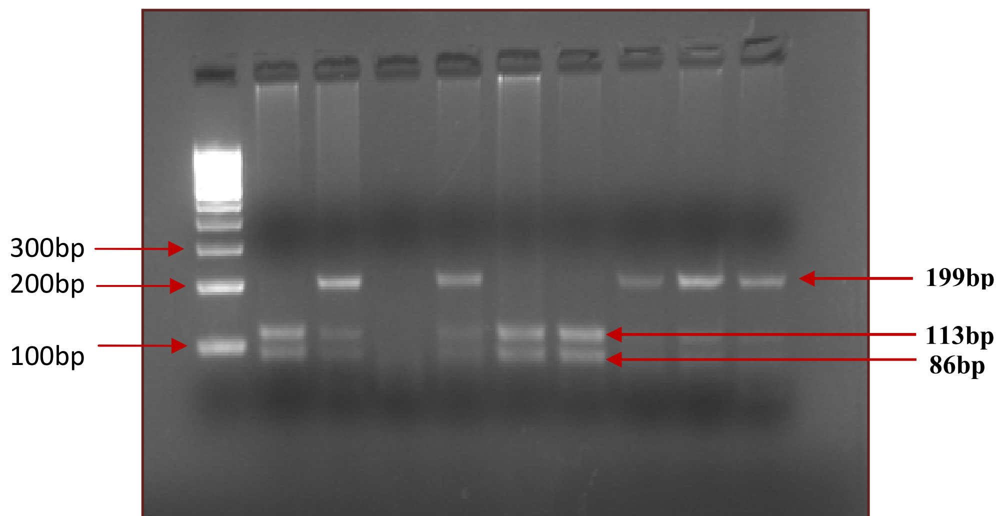


Figure 5.2: RFLP analysis for p53 gene codon 72 polymorphism using 100bp marker (M/s Fermentas, Lithuania). Pro allele was not cleaved by BstUI and had a band with a fragment of 199 bp. Arg allele was cleaved by BstUI and yielded two fragments (113 and 86 bp). The heterozygote had three bands (199bp, 113bp and 86bp.

Sequencing

Genotyping of selected samples were confirmed by sequencing method as described earlier in chapter one (Figure 5.3) . No discrepancies were observed.

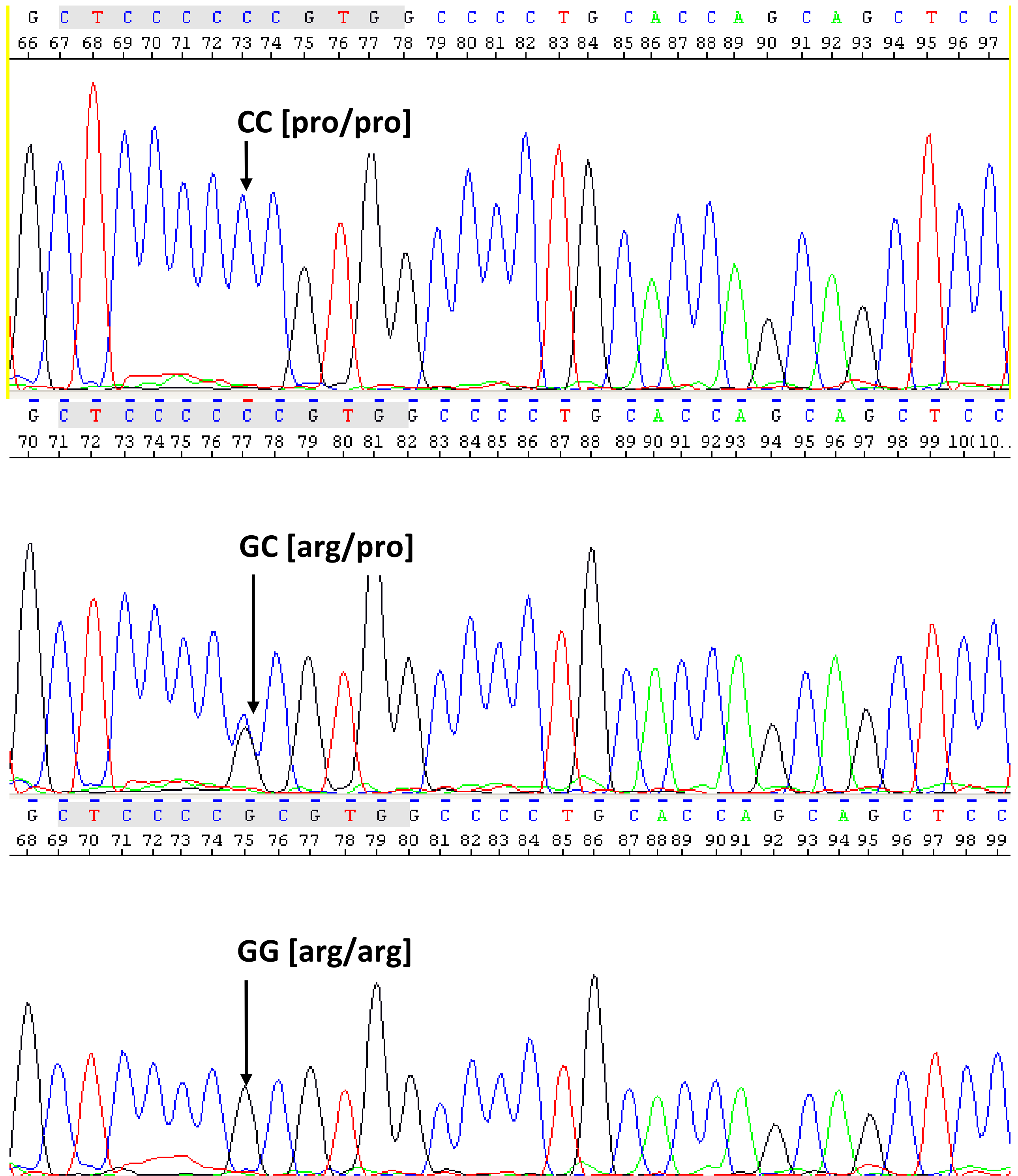


Figure 5.3: Electrogram showing the sequence analysis of different genotypes of p53 codon 72 polymorphism in gastric cancer samples sequencing analysis of the polymorphism

Statistical Analysis

Cases were individually matched with controls from the pool of 282 control samples on the basis of age (±5 years), sex and ethnicity, in a case control ratio of approximately 1:2. Thus a subset of 282 controls was matched against GC cases respectively. The association of *p53* codon 72 genotypes with GC was evaluated by multivariable conditional logistic regression in dominant (Pro/Pro and Arg/Pro versus Arg/Arg) and recessive (Pro/Pro versus Arg/Arg and Arg/Pro) genetic models. Armitage trend test was performed to calculate P for trends in additive model (Pro/Pro, Arg/Pro, versus Arg/Arg) **(Table 5.4)**. Interactions between *p53* genotypes and risk factors were analyzed to look for gene-environment interactions **(Table 5.5)**. The association of tobacco smoking, tobacco chewing, betel quid chewing, alcohol intake with disease outcome was assessed by $\chi 2$/Fisher's exact test. Estimates of risk to cancer, imparted by *p53* genotypes and other covariates as tobacco smoking, chewing, betel quid chewing and alcohol was determined by deriving the odds ratio (ORs) and corresponding 95% confidence intervals (95% CIs) using univariate and multivariable conditional logistic regression models **(Table 5.3)**. To evaluate potential modifying effects of *p53* genotypes on the association between various co-variates and cancer risk, cases and controls were tabulated according to the joint distribution of these factors in various possible combinations **(Table 5.5)**. Tests for studying interactions were performed for each *p53* genotype with all the considered covariates **(Table 5.5)**. For all the estimates two odds ratio were calculated. A crude unadjusted OR (OR1) was calculated using univariate conditional logistic regression

model. OR2 was the adjusted odds ratio using conditional multivariate logistic regression model. For all the tests a two sided $p<0.05$ was considered statistically significant. The data analysis was performed on STATA 8.0 software.

Results

The distributions of demographic characteristics and potential risk factors are summarized in **Table 5.2.** Higher percentage of males was seen in cases as well as in controls. However the gender distribution in GC cases (68.7% in men) was comparable to that in controls (73.4% in men). The age (mean age ±SD) of matched cases and controls respectively was, 56.54±13.5 and 53.34±12.28 years. The distribution of *p53* genotype between cases and controls is shown in **Table 5.2**. Deviation from Hardy–Weinberg equilibrium (HWE) was seen in cases ($\chi2$ =7.68, p=0.005). No significant increase in risk of GC was observed in a univariate or in a multivariable analysis for dominant and recessive models of inheritance **(Table 5.4)**. Frequency of Arg/Arg, Arg/Pro, and Pro/Pro genotypes was 20.1%, 61.9%, and 17.9% in the cases and 22.7%, 51.4%, and 25.9% in controls **(Table 5.2)**. Conditional logistic regression analysis showed higher risk associated with Arg/Pro genotype, whereas the Pro/Pro appeared to be a protective genotype. These observations were represented by both OR1 and OR2 but lacked statistical power **(Table 5.3)**. Tobacco smoking (OR2=0.83, 95% CI=0.52–1.32, p=0.44), Tobacco chewers (OR2=0.83, 95% CI=0.52–1.32, p=0.44); Betel quid chewing (OR2=1.74, 95% CI=1.04–2.92, p=0.03) and Alcohol drinking (OR2=1.09, 95%

CI=0.64–1.86, p=0.73) showed no significant increase of risk of GC **(Table 5.3)**. Analysis for interactions of risk habits with p53 genotypes did not reveal any combination to be significant toward GC. However, interaction of betel quid chewing with Arg/Pro genotype reached a near significance level, with a twofold risk of GC (OR2=2.40, 95% CI=0.91–6.26, p=0.07) **(Table.5.5)**.

Table 5.2: Distribution of demographic variables and genotypes between gastric cancer cases and controls

	Categories	Case (134) n (%)	Control (282) n (%)
Sex			
	Female	42 (31.3)	75 (26.6)
	Male	92 (68.7)	207 (73.4)
Smoking status			
	Non-smokers	59 (44)	141 (50)
	Smokers	75 (56)	141 (50)
Tobacco chewing			
	Non-chewers	79 (59)	143 (50.7)
	Chewers	55 (41)	139 (49.3)
Betel quid chewing			
	Non chewers	46 (34.3)	107 (38.1)
	Chewers	88 (65.7)	174 (61.9)
Alcohol consumption			
	Non-alcoholic	97 (72.4)	200 (70.9)
	Alcoholic	37 (27.6)	82 (29.1)
***p53* genotypes**			
	Arg/Arg	27 (20.1)	64 (22.7)
	Arg/Pro	83 (61.9)	145 (51.4)
	Pro/Pro	24 (17.9)	73 (25.9)

** χ^2 significant; $p<0.05$
Allele probabilities (P-value HWE) for gastric cancer cases: Arg: 0.51, Pro: 0.49 (0.005)
Allele probabilities (P-value HWE) for gastric matched controls: Arg: 0.48, Pro: 0.52 (0.62)

Table 5.3: Risk estimates for *p53* codon 72 polymorphism and co-variates in cancer patients

	Categories	OR1*	$OR2^a$
Tobacco smoking	Non-smokers	1.0	1.0
	Smokers	1.46(0.91-2.33),p=0.11	1.43(0.88-2.32),p=0.14
		OR1*	**$OR2^b$**
Tobacco chewing	Non-chewers	1.0	1.0
	Chewers	0.88(0.56-1.38),p=0.60	0.83(0.52-1.32),p=0.44
		OR1*	**$OR2^c$**
Betel quid chewing	Non-chewers	1.0	1.0
	Chewers		
		OR1*	**$OR2^d$**
Alcohol consumption	Non-drinkers	1.0	1.0
	Drinkers	1.19(0.72-1.98),p=0.48	1.09(0.64-1.86),p=0.73
		OR1*	**$OR2^e$**
***p53* Genotypes**	Arg/Arg	1.0	1.0
	Arg/Pro	1.21(0.68-2.15),p=0.51	1.16(0.64-2.08),p=0.61
		OR1*	**$OR2^e$**
	Pro/Pro	0.68(0.33-1.39),p=0.29	0.66(0.32-1.37),p=0.27

OR_1*: crude odds ratio

OR_2^a : adjusted for tobacco chewing, betel quid chewing, alcohol consumption and *p53* genotypes

OR_2^b : adjusted for tobacco smoking, betel quid chewing, alcohol consumption and *p53* genotypes

OR_2^c : adjusted for tobacco smoking, tobacco chewing, alcohol consumption and *p53* genotypes

OR_2^d : adjusted for tobacco smoking, tobacco chewing, betel quid chewing and *p53* genotypes

OR_2^e : adjusted for tobacco smoking, tobacco chewing, betel quid chewing and alcohol consumption

Table 5.4: Estimate of the effect of the *p53* codon 72 polymorphism on cancer risk modeled with logistic regression

Associations	Genotype	Case/Control n (%)	OR1 (95% CI)	OR2 (95% CI)
Dominant Model	Arg/Arg	27 (20.1) /64 (22.7)	1.00	1.00
	Arg/Pro and Pro/Pro	107 (79.9)/218 (77.3)	1.04 (0.59-1.81), p=0.89	1.00 (0.57-1.76), p=0.98
Recessive Model	Arg/Arg and Arg/Pro	110 (82.1)/209 (74.1)	1.00	1.00
	Pro/Pro	24 (17.9)/73 (25.9)	0.59 (0.33-1.04), p=0.07	0.59 (0.33-1.06), p=0.07

Table 5.5: Gene-Environment interactions odds ratio for *p53* codon 72 genotypes and risk habits for cancer patients

Variables	Interactions	Case (n=134) n(%)	Control (n=282) n(%)	OR_1*	OR_2
Smoking					
	Arg/Arg X Non smoker	11 (8.2)	32 (11.3)	1.0	1.0
	Arg/Arg X smoker	16 (11.9)	32 (1.3)	1.78(0.63-4.98), p=0.26	1.91(0.66-5.46), p=0.22
	Arg/Pro X Non smoker	38 (28.4)	70 (24.8)	1.42(0.59-.46), p=0.42	1.50(0.61-3.71), p=0.37
	Arg/Pro X smoker	45 (33.5)	75 (26.6)	1.87(0.78- .46), p=0.15	1.81(0.75-4.37), p=0.18
	Pro/Pro X Non smoker	10 (7.5)	39 (13.8)	0.69(0.24- .95), p=0.48	0.71(0.24-2.05), p=0.53
	Pro/Pro X smoker	14 (10.4)	34 (12.1)	1.17(0.42- .25), p=0.75	1.21(0.43-3.37), p=0.70
Tobacco chewing					
	Arg/Arg X Non chewer	13 (9.7)	25 (8.9)	1.0	1.0
	Arg/Arg X Chewer	14 (10.4)	39 (13.8)	0.72(0.26- .94), p=0.52	0.71(0.26-1.93), p=0.50
	Arg/Pro X Non chewer	50 (37.3)	76 (27.0)	0.98(0.43- .25), p=0.97	0.97(0.42-2.24), p=0.95
	Arg/Pro X Chewer	33 (24.6)	69 (24.5)	1.07(0.46- .50), p=0.86	1.02(0.43-2.41), p=0.95
	Pro/Pro X Non chewer	16 (11.9)	42 (14.9)	0.68(0.26-.78), p=0.43	0.70(0.26-1.85), p=0.47
	Pro/Pro X Chewer	8 (6.0)	31 (11.0)	0.41(0.13-.26) ,p=0.12	0.37(0.11-.17), p=0.09

Table 5.5 continued					
Betel quid chewing					
	Arg/Arg X Non chewer	7 (5.2)	25 (8.9)	1.0	1.0
	Arg/Arg X Chewer	20 (14.9)	39 (13.8)	2.21(0.76-6.37), p=0.14	2.16(0.74-6.26), p=0.15
	Arg/Pro X Non chewer	27 (20.1)	54 (19.1)	1.36(0.48-3.79), p=0.55	1.31(0.46-3.67), p=0.60
	Arg/Pro X Chewer	56 (41.8)	91 (32.3)	2.47(0.95-6.46), p=0.06	2.40(0.91-6.26), p=0.07
	Pro/Pro X Non chewer	12 (9.0)	29 (10.3)	0.99(0.31-3.19), p=0.99	0.93(0.29-3.02), p=0.91
	Pro/Pro X Chewer	12 (9.0)	44 (15.6)	1.21(0.39-3.70), p=0.73	1.15(0.37-3.51), p=0.80
Alcohol consumption					
	Arg/Arg X Non alcoholic	17 (12.7)	41 (14.5)	1.0	1.0
	Arg/Arg X Alcoholic	10 (7.5)	23 (8.2)	1.45(0.53-3.99), p=0.46	1.34(0.47-3.75), p=0.57
	Arg/Pro X Non alcoholic	60 (44.8)	108 (38.3)	1.23(0.60-2.52), p=0.56	1.18(0.57-2.43), p=0.65
	Arg/Pro XAlcoholic	23 (17.2)	37 (13.1)	2.00(0.85-4.73), p=0.11	1.68(0.70-4.04), p=0.24
	Pro/Pro X Non alcoholic	20 (14.9)	51 (18.1)	0.93(0.40-2.17),p=0.87	0.90(0.38-2.11), p=0.81
	Pro/Pro X Alcoholic	4 (3.0)	22 (7.8)	0.38(0.10-1.42), p=0.15	0.37(0.10-1.33), p=0.13

OR_1*: crude odds ratio

a : OR_2 adjusted for tobacco chewing, betel quid chewing and alcoho consumption

b : OR_2 adjusted for tobacco smoking, betel quid chewing and alcohol consumption

c : OR_2adjusted for tobacco smoking, tobacco chewing and alcohol consumption

Table 5.6: Frequency of p53 Polymorphism in Gastric Cancer: Worldwide Scenario

Tumor Site	Place of Study	N (Case/ Control)	Allele freq (Arg) (Case/Control)	Allele freq (Pro) (Case/Control)	First Author
Gastric Cancer	Korea	292/216	0.56/0.65	0.44/0.35	Yi et al (2006)
	USA	155/134	0.56/0.61	0.44/0.39	J. Sulet al (2006)
	Taiwan	89/192	0.42/0.46	0.58/0.54	Wu et al (2004)
	Iran	92/163	0.40/0.41	0.60/0.59	Mojtahedi et al (2010)

Discussion

P53 is a highly conserved gene with only five polymorphisms being known till date in 11 exons, of which polymorphism of codon 72 is the commonest and most characterized single base substitution of Proline (P72) for Arginine (A72) leading to structural changes in the protein [376, 377]. The polymorphism occurs in the proline rich region of *p53*, which plays a vital role in apoptosis and growth suppression functions, thus indicating that these two polymorphic variants differ in their biological properties [377]. The P72 variant is a stronger inducer of transcription, probably owing to its stronger affinity to bind to transcription factors and the R72 variant is considered to be a better inducer of apoptosis, thus suppressing transformation more efficiently, than the P72 variant [377].

Observations on association of *p53* codon 72 polymorphism and GC are reported to be inconsistent in different ethnic and geographical region with allele frequency varying from 0.40 to 0.56 for Arg and 0.44 to 0.60 for Pro **(Table 5.6)**. Hiyama et al. (2002) reported Pro/Pro genotype association with increased risk of developing GC [45] whereas study by Zhang et al. (2003) showed that risk associated with Arg/Arg genotype was 3.1 times higher as compared to Pro/Pro and Arg/Pro carriers [378]. Considerable number of studies has reported no difference in distribution of any genotype between cases of GC and controls [37, 188, 379, 380]. In the present study, no significant effect of the polymorphism on susceptibility to GC was seen. These findings are concordant with some previous reports spread over different

ethnic populations in different cancer. No association between *p53* variants and lung cancer was observed in African Americans and caucasians in United States by Weston et al. [381] and in northwestern Mediterranean population by To-Figueras et al. [382]. Similar results in GC are reported from Korea, Costa Rica and Iran [383-385] and in oral cancer by Drummond et al. [386] and Kietthubthew et al. [387] in Brazilian and Thai population. However most of these studies on gastric constituted small sample size, from 58 to 97 and lacked stratification by risk factors. In the present study Arg/Pro heterozygous genotype was seen to confer greater risk to cancer than Pro/Pro genotype **(Table 5.2)**. Literature available, report preferential retention of *p53* codon 72 arginine allele in tumors of patients with Arg/Pro heterozygous germline genotype [35]. Furthermore, presence of arginine allele at codon 72 in tumor related with reduced sensitivity to chemotherapy [388] and decreased survival in heterozygous breast cancer [389].

For studying role of gene-environment interaction that might modify susceptibility of cancers, potential interactions of *p53* with known risk factors was analyzed. Perhaps investigating the mutational status of tumor and its correlations could possibly provide better understanding. Major limitation of this study is the small sample size. A reasonable fraction of the control came from friends and accompanying family members of the cancer patients. Such selection of control would reduce any confounding bias by reducing background variations with cases. GC cases showed deviation from Hardy–Weinberg equilibrium. Source population comprising of culturally or

religiously non communicating strata which generally do not inter-marry within other caste or religion can be a reason for this deviation. However our cases were incident and thus the data does not show report or recall bias. Also case control matching was done in reference to age, gender and ethnicity thereby controlling for any confounding effect on account of these variables.

Studies on p53 codon 72 polymorphism available in Indian population have not explored gene-environment interaction [390]; therefore sample size for such an analysis could not be reliably determined and limit our ability to estimate the interaction effects precisely. The results of interaction analysis should therefore be considered empirical observations for further studies on larger number of sample. The current study indicates that there is no significant relationship between *p53* codon 72 polymorphism and GC in the high risk NE population of India. Taking into account other confounding variables such as dietary habits, environment (working environmental exposures, passive smoking etc) and infections (HPV, *H pylori* status etc) can give more conclusive perspective.

Chapter 6: Evaluation of PCR Assays for Detection of the Presence of Helicobacter Pylori in Gastric Cancer Patients of Northeast India

---

Introduction

Over 50% of the world's population is infected, with the highest prevalence in developing countries [391, 392]. *H. pylori* infection is common worldwide [393]. Acquisition of *H. pylori* infection in children results in adulthood disorders. It is estimated that life time risk of gastritis, Mucosa-Associated Lymphatic Tissue Lymphomas and gastric cancer (GC) in *H. pylori* infected population is above 100%, 10%, <1% and 1-2% respectively. On a global scale, GC is the second commonest cancer in the world. There is substantial international variation in GC incidence with the highest rates reported from China, Japan and other East Asian countries [394].

According to Cornea's hypothesis GC develops from chronic inflammation either directly or indirectly leading to different stages such as metaplasia, dysplasia and finally to cancer [395]. The role of infectious agents such as *H. pylori* and chronic inflammation in carcinogenesis is being increasingly recognized [396]. Gastric or duodenal ulcers (commonly referred to as peptic ulcers) are defined as mucosal defects with a diameter of at least 0.5 cm penetrating through the muscularis mucosa which are believe to be the initial stages of gastric carcinoma [397]. Both gastric and duodenal ulcer diseases are strongly related to *H. pylori*. In initial reports from all over the world in the first

decade after the discovery of *H. pylori*, approximately 95% of duodenal ulcers and 85% of gastric ulcers occurred in the presence of *H. pylori* infection [398].

H. pylori is a gram-negative curved or spiral bacterium which colonizes the human gastric mucosa. Work by Warren and Marshall in the early 1980s demonstrated the organism's association with gastritis and peptic ulceration. Chronic *H. pylori* infection has also been linked to the development of GC [143, 399]. Gastric mucosal damage involves both host and *H. pylori* dependent factors [400]. Several prospective serological studies indicate that *H. pylori* infection causes a three to sixfold higher risk for the development of GC [401, 402]. A sero epidemiological study with two large cohorts from different continents has also revealed a sixfold increased risk for the development of gastric carcinoma in patients with anti-*H. pylori* antibodies compared with uninfected controls [403]. Moreover, the World Health Organization has included *H. pylori* as a class I carcinogen due to its strong correlation with GC [404]. For preventive purpose, a simple and non-invasive tests for detecting *H. pylori* infection is required. *H. pylori* antibodies found in human serum are indicators for *H. pylori* infection. Positive *H. pylori* IgG results for diagnosis of *H. pylori* infection has been proven in reports of various countries with high sensitivity and specificity [405]. However, the results are different among countries, especially between developed and developing countries [406].

The most specific way of detecting *H. pylori* in tissue is a combination of culture and histological staining of mucosal biopsy specimens obtained by endoscopies [407]. This procedure is not fully satisfactory because organisms

resembling *H. pylori* can be falsely detected by histological examination [408, 409]. Culture has been for long the method of choice to detect infectious agents. However, for some organisms that are growing slowing like *H. pylori* culture is time consuming [410]. Biopsy urease test has been developed as a rapid technique that is combined with histological examination. Unfortunately, this test is somewhat less specific than culture [411] and urease assays can lead to non-specific results due to the presence of other urease-positive bacteria and false negative results have also been reported in individuals taking proton pump inhibitors [412, 413].

Polymerase chain reaction (PCR) has been shown to be a valuable method for detection of various microorganism [414]. The polymerase chain reaction (PCR) assay opens up the possibility of detection for genomic *H pylori* without the need for microbiological cultures using tissue routinely obtained for diagnostic histopathology.

Tissue samples from biopsy were taken for the study to represent the actual conditions when testing clinical samples (e.g., presence of inhibitors), since experiments using pure *H. pylori* DNA might not represent the definite clinical condition. ELISA has been used as a gold standard. Positive *H. pylori* IgG results for diagnosis of *H. pylori* infection has been proven in reports of various countries with high sensitivity and specificity [415-417]. The primer choosen for the study were for *GlmM*, *UreA* and *16S rRNA*. Although previous studies have used these primers for detection [418-420], systematic studies comparing between them are rare in tissue biopsy specimen which represent actual clinical conditions.

Moreover the sensitivity and specificity of different diagnostic tests for *H. pylori* detection varies widely [421]. The aims of the present study were to compare the accuracy of the reported PCR primer pairs using GC biopsy specimens known to either contain *H. pylori* or to be *H. pylori* negative by serological tests in high risk region of NE India.

Materials and Methods

Study population

A total of 75 confirmed GC patients were included in the study. Patients who received antibiotic therapy, bismuth treatment, or a proton pump inhibitor or H2 blocker within 1 month prior to the study were excluded. Five milliliters of blood was obtained on the day of endoscopy. One – two ml of serum was successfully collected from all these samples. Sera were kept at -20°C until analyzed. For PCR analysis, the biopsy was collected in PBS and frozen at -70° c until processed. The biopsy specimens from the stomach were obtained for extraction of genomic DNA for PCR assays. All the 75 patients have sufficient tissue material for DNA isolation. Written informed consent was obtained from all individuals involved before the study. All patients gave informed consent to be biopsied and participation in this study**.** Detailed questionnaire with specific information regarding diet, smoking, alcohol consumption habits and family history of cancer was completed for all patients as well as controls. Approval for this study had been obtained from the Institutional Human Ethics Committee.

Antibody assay

Sera from 75 were analyzed for presence of IgG antibodies to *H. pylori* by commercially available enzyme-linked immunosorbent assay (ELISA) kit (DRG Instruments GmbH, Germany), according to the manufacturer's instructions.

Genomic DNA extraction from tissue samples

For the detection of the *H. pylori 16SRNA, UreA* and *GlmM* genes in DNA extracted from the tissue, nucleic acids was extracted from 10 mg tissue using a standard phenol-chloroform extraction method. The tissue was first homogenized and digested by lysis buffer I (30mM Tris, 5mM EDTA, and 50mM NaCl) and lyses buffer II (20%SDS, 100µg/ml Prot.K). The mixture was incubated for 3hr at 45°C under agitation. All centrifugation steps were carried out at 10000g, 10min, 4°C. Lysate was centrifuged and the supernatant was collected and stored on ice until the phenol/chloroform/isoamyl alcohol extraction step to prevent DNA degradation that might occur during the physical cell lysis treatment. Then DNA was extracted from this supernatant by adding an equal volume of phenol/chloroform/isoamyl alcohol mixture (25: 4:1 v/v/v) (Sigma, USA). The tube was briefly vortexed to obtain an emulsion then centrifuged. The aqueous layer was collected. Traces of phenol were removed by adding an equal volume of chloroform. After centrifugation, the top layer was collected and DNA was precipitated overnight at -20°C with two volumes of cold isopropanol. The pellet of crude DNA was obtained by centrifugation for 20min and was washed with 1ml of 70% cold ethanol. It was air dried and resuspended in 200µl of 1X TE buffer (10mM Tris-Cl, 1mM; Na_2EDTA, pH 8). This genomic DNA was then used

for detection of the presence of *H. pylori* specific genes i,e *UreA*, *16SrRNA* and *GlmM*.

Quantitativeand Quality estimation of extracted genomic DNA

The estimation of concentration of genomic DNA in solution, obtained from different sources was done using UV spectrophotometer (Shimadzu Corporation) at 260 nm and 280 nm and the concentration of DNA was calculated as mention in chapter 1. A good quality of high molecular weight genomic DNA was obtained having a concentration of 200-500 ηg/µl. The genomic DNA were checked in ethidium bromide stained 1% agarose gel (1g agarose in 1X TBE).

Primers

The Primers for detection were **(Table 6.1)** based on earlier published paper and their specificity were checked by BLAST search, and synthesised by M/s Microsynth, Germany. An NCBI BLAST analysis confirmed that these primers identify the target sequence within the genome of *H. pylori*.

Table 6.1: Oligonucleotide primer sequences for amplification of for *UreA*, *GlmM* and *16SRNA*

Gene	Primer sequence (forward) 5'---3'	Primer sequence (reverse) 5'---3'	Tm	Amplicon size
GlmM	GGATAAGCTTTTAGGGGT GTTAGGG	GCTTACTTTCTAACACTA ACGCGC	62	296 bp
UreA	GCCAATGGTAAATTAGTT	CTCCTTAATTGTTTTTAC	46	411 bp
16SrRNA	GCTATGACGGGTATCC	TCGCAATGAGTATTCCTCTT	53	428 bp

Detection of the GlmM gene

The genomic DNA extracted from seventy five biopsy of GC samples were tested for the presence of *GlmM* gene. The PCR mixture of 20 µl volume was prepared containing 0.2µl of 10x Taq buffer, 0.8µl $MgCl_2$ (50mM), 0.6µl of dNTPmix (10mM) (Invitrogen), 0.6µl of each forward and reverse primers for *GlmM* , 500-700ng of template DNA and 1 Unit of Taq Polymerase (Invitrogen). Positive control was purchase from KPL (Gaithersburg, USA) and included in the PCR run to ensure proper functioning of PCR reaction. To test for contamination, negative controls (distilled water) were included in every PCR run. Programme set in PCR machine (Eppendorf) for amplification include initial activation step at 94^0C for 5 min, 35 cycles of denaturation at 94^0C for 1 min; annealing at 62^0C for 1 min; 72^0C for 1 min 15 sec and in addition a final extension at 72^0C for 10 min.

PCR amplification for *GlmM*

Component concentration	Volume (Final concentration)
PCR buffer (10X)	2 µl (1X)
MgCl2 (50 mM)	0.8 µl (2 mM)
dNTP (10mM)	0.6 µl (0.3 mM)
Forward primer (50 pmol)	0.2 µl (0.5pmol)
Reverse primer (50 pmol)	0.2µl (0.5pmol)
Platinum Taq (5U/µl)	0.2 µl (1U)
DNA	5 µl
DNase free water	Fill up to the total volume
Total	20 µl

The PCR conditions were as followed

Initial Denaturation (94ºC)	:	5 min	
Denaturation (94ºC)	:	1 min	35 cycle
Primer annealing (62 ºC)	:	1 min	
Extension (72ºC)	:	1 min 15 sec	
Final extension (74ºC)	:	10 min	
Hold (4ºC)	:	till the electrophoresis	

Electrophoresis of PCR amplicons

The amplified product (6 µl) was electrophoresed in 2% agarose (USB, USA) containing 0.5 µg/mL ethidium bromide and examined under Transilluminator (Alphaimager, Cell Bioscience) as describe earlier in chapter 1. The presence of 269bp band indicates and the presence of *H. pylori* specific *GlmM* gene **(Figure 6.1B)**.

Detection of the UreA gene

The genomic DNA extracted fron tissue samples from seventy five patients were tested for the presence of *UreA* gene. About 500ng-700ng of DNA was amplified in a total volume of 20 µl reaction mixture containing 0.5 pmol of each primer as described, 0.3mM of dNTPs and 1U Taq polymerase (Invitrogen). PCR was performed with initial denaturation at 94˚C for 5 min, followed by 35 cycles of 94˚C for 1min, 53˚C for 1min, 72˚C for 1min and a final extension at 72˚C for 7.5 min. Presence of *UreA* yielded bands of 411bp. PCR on genomic DNA from *H. pylori* strains ATCC 43504 (KPL, Gaithersburg, USA) and distilled water was used as positive and negative controls respectively.

PCR amplification for *UreA*

Component concentration	Volume (Final concentration)
PCR buffer (10X)	2 µl (1X)
MgCl2 (50 mM)	1µl (2.5 mM)
dNTP (10 mM)	0.6 µl (0.3 mM)
Forward primer (50 pmol)	0.2 µl (0.5 pmol)
Reverse primer (50 pmol)	0.2 µl (0.5 pmol)
Platinum Taq (5U/µl)	0.2 µl (1 U)
DNA	5 µl
DNase free water	Fill up to the total volume
Total	20 µl

The PCR conditions were as followed

Initial Denaturation (94ºC) : 5 min

Denaturation (94ºC) : 1 min

Primer annealing (53ºC) : 1min } 35 cycle

Extension (72ºC) : 1min

Final extension (74ºC) : 7.5 min

Hold (4ºC) : till the electrophoresis

Electrophoresis of PCR ampliconse gel

Agarose gel electrophoresis (2%) resolved amplified DNA fragments as described in chapter 1. The presence of 411bp band indicates and the presence of *H. pylori* specific *UreA* gene **(Figure 6.2B)**.

Detection of the 16S rRNA gene

Seventy five biopsy specimens were also tested for the presence of *16SrRNA* gene. The PCR mixture contained 1U of Taq polymerase (Invitrogen), 50mM $MgCl_2$, 0.2 mM deoxynucleotides (Invitrogen) and 0.4 pmol each oligonucleotide primer (M/s Microsynth, Germany). PCR was performed in DNA thermocycler (Eppendorf). Conditions used were as follows: 94°C for 5 min, then 35 cycles of 94°C for 1 min, 53°C for 1 min, 72°C for 1 min, and, finally, 72°C for 7.5 min. Tube containing distilled water instead of DNA sample (negative control) and DNA samples of strains ATCC 43504 (KPL, Gaithersburg, USA) (Positive control) were used with each batch of amplification mixture.

Analysis of the PCR-amplified products

Amplification products were detected by electrophoresis in 2% agarose gels, stained with ethidium bromide 0.5 g/m1 and visualized under UV illumination (Alphaimager, Cell Biosciences). Electrophoresis was performed in TAE buffer as describe earlier. The presence of 428bp indicate the presence of *16S rRNA* **(Figure 6.3B)**.

PCR amplification for *16SrRNA*

Component concentration	Volume (Final concentration)
PCR buffer (10X)	2 µl (1X)
MgCl2 (50 mM)	1µl (2.5 mM)
dNTP (10 mM)	0.6 µl (0.3 mM)
Forward primer (50 pmol)	0.16 µl (0.4 pmol)
Reverse primer (50 pmol)	0.16 µl (0.4 pmol)
Platinum Taq (5U/µl)	0.2 µl (1 U)
DNA	5 µl
DNase free water	Fill up to the total volume
Total	20 µl

The PCR conditions were as followed

Initial Denaturation (94ºC)	:	5 min	
Denaturation (94ºC)	:	1 min	35 cycle
Primer annealing (53ºC)	:	1min	
Extension (72ºC)	:	1min	
Final extension (72ºC)	:	7.5 min	
Hold (4ºC)	:	till the electrophoresis	

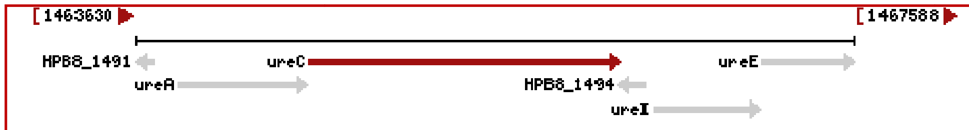
DNA Sequencing

Directly sequencing were carried out on ABI Prism 3,700 Genetic analyzer (Applied Biosystems, Foster City, CA, U.S.A.) using the ABI Prism BigDye Terminator Cycle Sequencing Ready Reaction kit (Applied Biosystems). Products were sequenced in both the directions and eletropherogram **(Figure 6.1C, 6.2C , 6.3C)** are viwed using Sequence Analysis software (Applied Biosystems) and confirmed by aligning with the reference sequence *of H. pylori* from NCBI using SeqMan module of DNASTAR v5.07 software.

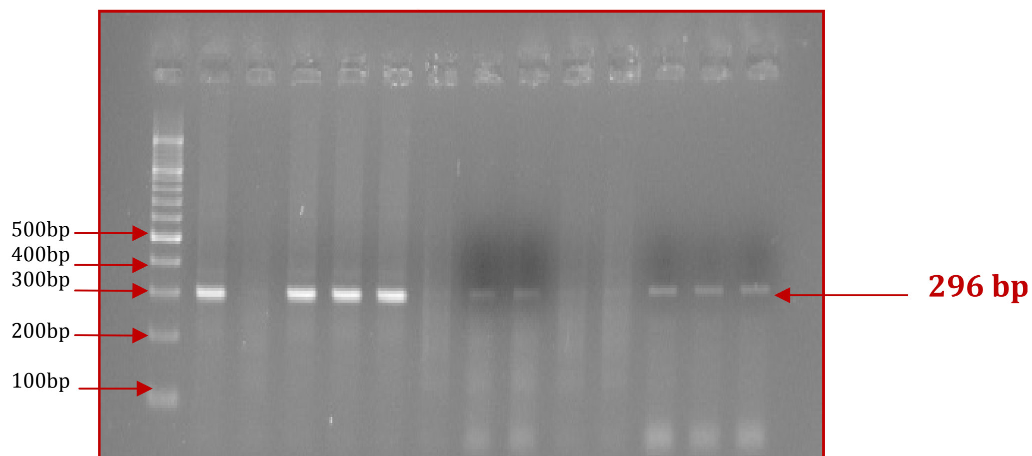
A. **Gene symbol** ***ureC*** **(*GLmM*)**

Gene description **urease alpha subunit**

Locus tag **HPB8_1493**



B.



C.

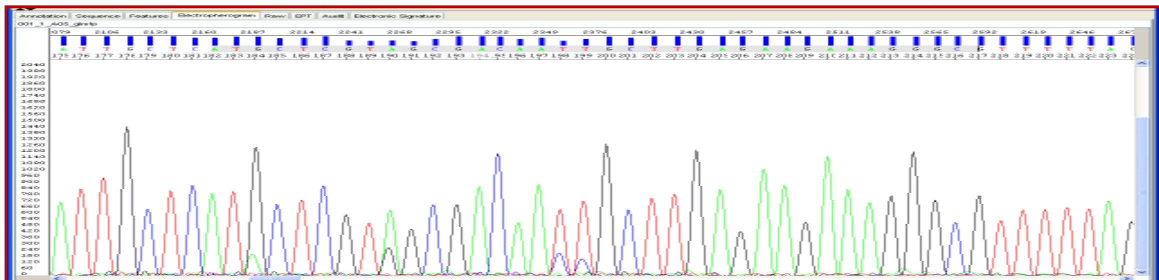
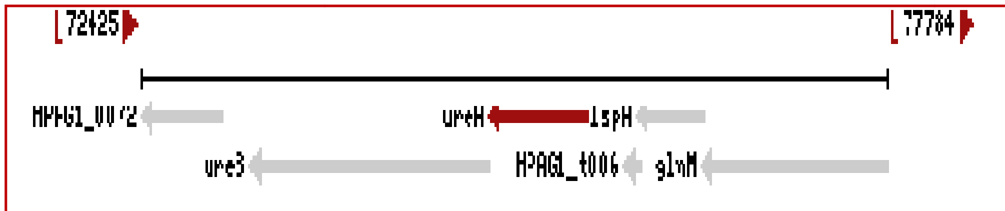


Figure 6.1: (A) Schematic representation of gene orientation in GlmM. (B) 2 % agarose gel electrophoresis of 296 bp amplicon. Lane 1 is 100 bp DNA Ladder, Lane 2 is positive control and Lane 3 is negative control. (C) Representative picture showing sequencing of the 296 bp GlmM gene

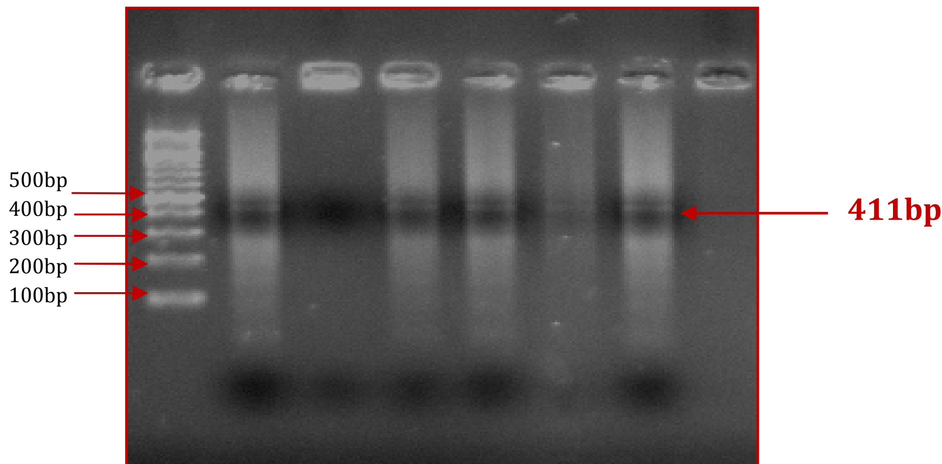
A. **Gene Symbol** **UreA**

Gene description **urease subunit alpha**

Locus tag **HPAG1_0074**



B.



C.

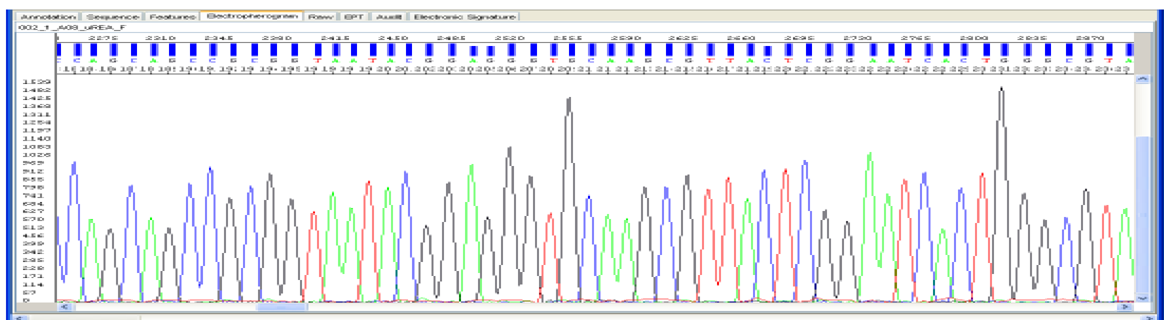


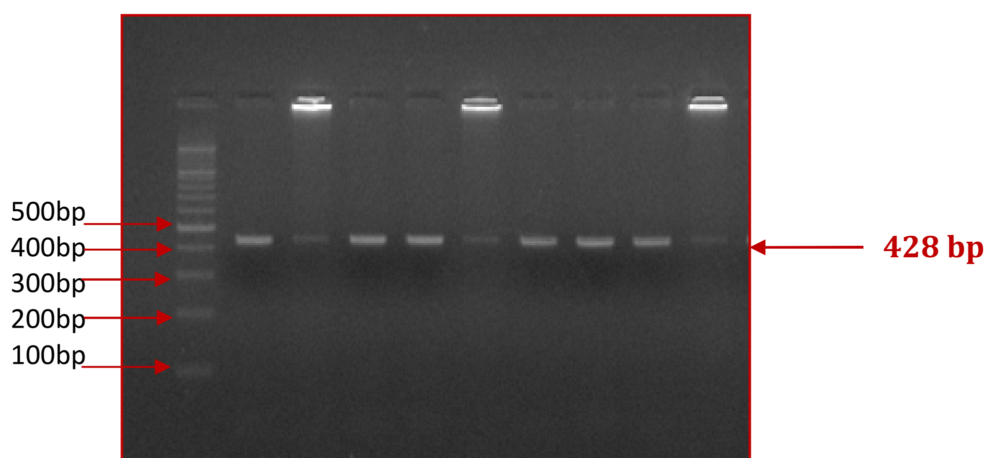
Figure 6.2: (A) Schematic representation of gene orientation in UreA. (B) 2 % agarose gel electrophoresis of 411bp amplicon. Lane 1 is 100 bp DNA Ladder, Lane 2 is positive control and Lane 3 is negative control. (C) Representative picture showing sequencing of the 411 bp UreA gene

A. **Gene symbol** ***HPSH_r08358***

Gene description **16S ribosomal RNA**

Locus tag **HPSH_r08358**

B.



C.

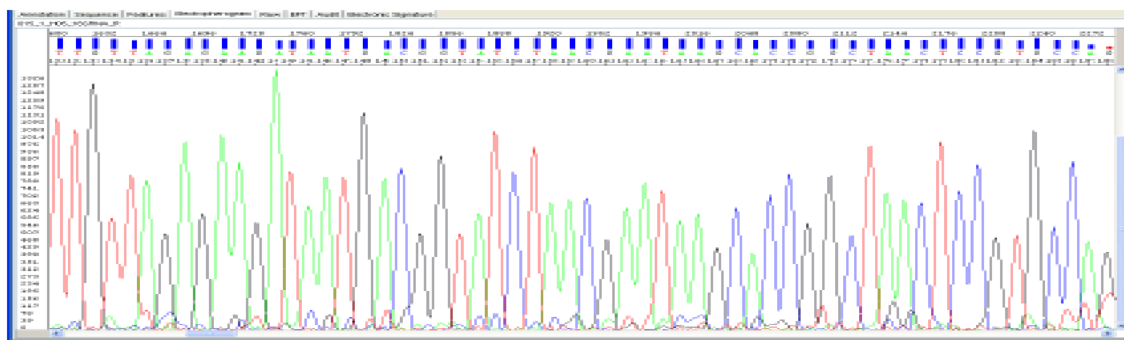


Figure 6.3: (A) Schematic representation of gene orientation in 16S rRNA. (B) 2 % agarose gel electrophoresis of 428 bp amplicon. Lane 1 is 100 bp DNA Ladder, Lane 2 is positive control (C) Representative picture showing sequencing of the 428 bp 16S RNA gene

Statistical analysis

The efficiency of *Glm M, Ure A* and *16SRNA* and their combination were estimated in term of sensitivity and specificity, livelihood ratio of positive and negative test, taking immunoassay test for Ig G antibody to *H. pylori* as the gold standard **(Table 6.3, Table 6.4)**. The probability of true positivity and true negativity of the each primer alone and also in different combinations were also calculated using Clinical Calculator 1 **(Table 6.5, Table 6.6)**.

Results

Out of the 75 GC tissue screened, a total of forty were found positive for IgG antibodies against *H. pylori* in sera. Twenty six (35%) were positive in all three PCR methods (*GlmM, UreA, 16SRNA*). Samples positive for *UreA* is 38 (51%), of these 6 (8%) were positive for *UreA* alone. Samples positive for *GlmM* is 45 (60%), of these 13 (17%) were positive for *GlmM* alone. Samples positive for *16SRNA* were 32 (43%), of these 2 (3%) were positive for *16S RNA* alone. Five were positive for *GlmM* and *UreA*, 3 for *GlmM*/*16SRNA* and 2 for *UreA/16SRNA*. Eight ELISA negative samples were positive in all three genes screened. Eighteen (24%) of the tissue samples were negative for all three genes screened. All the negative samples from the three genes were negative by ELISA. There were two samples which were negative for *UreA*, *GlmM* and also by ELISA but positive for *16SRNA* alone. Only one sample each was positive by *UreA* and *GlmM* alone while negative by the other three methods used for detection **(Table 6.3)**. On the basis of these values we have calculated the specificity, sensitivity as well as

probability of true positivity and negativity of the primers signally and in combination also.

The PCR methods using *GlmM, UreA* and *16SRNA* genes have shown a sensitivity of 85%, 65% and 53% and specificity of 63%, 63%, 66% respectively **(Table 6.3)**. The positive likelihood ratio of *GlmM, UreA* and *16SRNA* are 2.29, 1.75 and 1.53 and the negative likelihood ratio are 0.24, 0.56 and 0.72 respectively **(Table 6.3)**. The sensitivities of different combination of primers i,e of *GlmM+ve/16SrRNA+ve GlmM+ve/UreA+ve, UreA+ve/16SrRNA+ve* and are 50%, 49%, 48% and whereas their specificities are 74%, 69% and 73% respectively **(Table 6.4)**. The positive likelihood ratio of *Glm+ve/16SrRNA+ve*, *GlmM+ve/UreA+ve* and *UreA+ve/16SrRNA+ve* are 1.94, 1.59 and 1.85 respectively. The negative likelihood ratios are 0.67, 0.73 and 0.71 respectively **(Table 6.4)**. The probability of true positive for *GlmM*, *UreA* and *16srRNA* were 0.72, 0.64 and 0.61 and true negative were 0.79, 0.64 and 0.55 respectively **(Table 6.5)** whereas probability of true positive of *GlmM/16SrRNA, GlmM/UreA* and *UreA/16SrRNA* were 0.69, 0.65 and 0.57 respectively and probability of true negative were 0.57, 0.55 and 0.55 respectively **(Table 6.6)**.

Table 6.2: Results of three PCR methods and ELISA for the detection of *H. pylori* from GC cases

Total Biopsy n=75 n (%)	***Glm M***	***Ure A***	***16S RNA***	**ELISA**
26(35)	**+**	**+**	**+**	18+, 8-
13(17)	**+**	**-**	**-**	12+, 1-
6(8)	**-**	**+**	**-**	5+ , 1-
2(3)	**-**	**-**	**+**	2-
5(7)	**+**	**+**	**-**	2+, 3-
3(4)	**+**	**-**	**+**	2+, 1-
2(3)	**-**	**+**	**+**	1+, 1-
18(24)	**-**	**-**	**-**	18-

Table 6.3: Performance of PCR methods for the detection of *H.pylori*

PCR methods	Sensitivity	Specificity	Positive likelihood ratio	Negative likelihood ratio
Glm M	0.85 (CI: 0.69-0.94)	0.63 (CI: 0.45-0.78)	2.29 (CI: 1.46-3.59)	0.24 (CI: 0.11-0.52)
Ure A	0.65 (CI: 0.4820.79)	0.63 (CI: 0.45-0.78)	1.75 (CI: 1.075-2.848)	0.56 (CI: 0.353-0.878)
16S RNA	0.53 (CI: 0.36-0.68)	0.66 (CI: 0.48-0.8)	1.53 (CI: 0.89-2.64)	0.72 (CI: 0.51-1.0)

Table 6.4: Performance of combination of PCR methods for the detection of *H. pylori*

PCR methods	Sensitivity	Specificity	Positive likelihood ratio	Negative likelihood ratio
GlmM+ve/ UreA+ve	0.49 (CI:0.34-0.66)	0.69 (CI:0.51-0.83)	1.59 (CI:0.89-2.84)	0.73 (CI:0.52-1.019)
UreA+ve/ 16SrRNA+ve	0.48 (CI:0.32-0.64)	0.74 (CI:0.56-0.87)	1.85 (CI-0.96-3.54)	0.71 (0.52-0.98)
GlmM+ve/ 16SrRNA+ve	0.5 (CI:0.34-0.66)	0.74 (CI: 0.56-0.87)	1.94 (CI:1.02-3.69)	0.67 (CI:0.48-0.94)

Table 6.5: Probability of positivity and negativity for *GlmM, UreA* and *16S RNA*

Glm M		**95% Confidence Interval**	
	Estimated Value	**Lower Limit**	**Upper Limit**
For any particular positive test result, the probability that it is:			
True Positive	0.72	0.57	0.84
False Positive	0.28	0.16	0.43
For any particular negative test result, the probability that it is:			
True Negative	0.79	0.59	0.91
False Negative	0.21	0.09	0.41

Ure A		**95% Confidence Interval**	
	Estimated Value	**Lower Limit**	**Upper Limit**
For any particular positive test result, the probability that it is:			
True Positive	0.64	0.49	0.81
False Positive	0.33	0.19	0.51
For any particular negative test result, the probability that it is:			
True Negative	0.61	0.44	0.76
False Negative	0.39	0.24	0.56

16S RNA		**95% Confidence Interval**	
	Estimated Value	**Lower Limit**	**Upper Limit**
For any particular positive test result, the probability that it is:			
True Positive	0.64	0.45	0.79
False Positive	0.36	0.21	0.55
For any particular negative test result, the probability that it is:			
True Negative	0.55	0.39	0.69
False Negative	0.45	0.31	0.61

Table 6.6: Probability of positivity and negativity for combination of *GlmM+ve- UreA+ve, UreA+ve-16S rRNA+ve*, and GlmM+ve-*16S RNA+ve*

GlmM/16SrRNA		**95% Confidence Interval**	
	Estimated Value	**Lower Limit**	**Upper Limit**
For any particular positive test result, the probability that it is:			
True Positive	0.69	0.49	0.84
False Positive	0.31	0.16	0.51
For any particular negative test result, the probability that it is:			
True Negative	0.57	0.41	0.71
False Negative	0.43	0.29	0.59

Glm M/UreA		**95% Confidence Interval**	
	Estimated Value	**Lower Limit**	**Upper Limit**
For any particular positive test result, the probability that it is:			
True Positive	0.65	0.45	0.81
False Positive	0.35	0.19	0.55
For any particular negative test result, the probability that it is:			
True Negative	0.55	0.39	0.69
False Negative	0.45	0.31	0.61

UreA/16SrRNA		**95% Confidence Interval**	
	Estimated Value	**Lower Limit**	**Upper Limit**
For any particular positive test result, the probability that it is:			
True Positive	0.37	0.47	0.83
False Positive	0.32	0.17	0.52
For any particular negative test result, the probability that it is:			
True Negative	0.55	0.41	0.69
False Negative	0.45	0.31	0.59

Discussion

The association between chronic *H. pylori* infection and development of GC is well established [422]. Several prospective serological studies indicate that *H. pylori* infection causes a three to sixfold higher risk for the development of GC [401, 402]. However, in spite of the high prevalence of *H. pylori* infections only a minority of infected individuals will develop malignancies. In addition to a variety of host and environmental factors, there are specific bacterial virulence genes that might determine the possible sequelae of infection by *H. pylori* [423].

The prevalence of *H. pylori* shows large geographical variations. In various developing countries, more than 80% of the population is *H. pylori* positive, even at young ages [424]. The prevalence of *H. pylori* in industrialized countries generally remains under 40% [425]. Within geographical areas, the prevalence of *H. pylori* inversely correlates with socioeconomic status [426]. In Western countries, the prevalence of this bacterium is often considerably higher among first- and second-generation immigrants from the developing world [427, 428]. While the prevalence of *H. pylori* infection in developing countries remains relatively constant, it is rapidly declining in the industrialized world [429]. The latter is thought to be caused by the reduced chances of childhood infection due to improved hygiene and sanitation and the active elimination of carriership via antimicrobial treatment. In developing countries, *H. pylori* infection rates rise rapidly in the first 5 years of life and remain constantly high [430]. Such different prevalence is caused by age of acquisition at initial *H. pylori* infection [431] and technically by different method of examination with different

sensitivity and specificity for *H. pylori* detection [432]. For diagnostic purpose, ideally there should be a test with high sensitivity and specificity, reaching up to 100% for each. However, such test has not been present yet [433]. The PCR is the most sensitive of the existing rapid methods to detect microbial pathogens in clinical specimens. In particular, when specific pathogens that are difficult to culture in vitro or require a long cultivation period are expected to be present in specimens, the diagnostic value of PCR is known to be significant. In some non routine studies, conventional PCR was found to be a very sensitive method. The superiority of conventional PCR over other routine tests has already been reported [434-436]. The feasibility of PCR assay is good, with a duration of two to three hours and its cost is reasonable. It can be performed on any Light Cycler apparatus available in the laboratory.

Our study of the three choosen primer s viz, *GlmM*, *UreA* and *16SrRNA* revealed that the test positive samples of *GlmM* and *UreA* were 13 whereas test positive samples were 22 in the 35 ELISA negative samples **(Table 6.2)** and as such they have same sensitivity i,e 0.63 (CI:0.449-0.78)**(Table 6.3)**. The sample which were positive by *GLmM* alone was thirteen of which twelve is also positive by ELISA and that of UreA were 6 out of which 5 is also positive by ELISA whereas that of 16SrRNA were 2, none of which were positive by ELISA **(Table 6.2)**. The overall sensitivity of *GlmM* was found to be the highest (85%) as compare to that of *UreA* (65%) and *16SrRNA* (53%). Among the three primers choosen for the study that of *GlmM* seems to a promising one for detection of *H. pylori* with higher sensitivity and high positive likelihood ratio of 2.29 as compare to that of *UreA* (1.75) and *16SRNA* (1.53) **(Table 6.3)**. We also searched for combinations of primer pairs that

would improve the results, the combination of *GlmM/16SrRNA* shows a higher probability of true positivity and true negativity of 0.69 and 0.57 respectively to that of *GlmM/UreA* (0.65 and 0.55) and *UreA/16SrRNA* (0.37 and 0.55) **(Table 6.6)** but the combination of *GlmM* and*16SrRNA* shows only a slight higher sensitivity (50%) and specificity (74%) to that of *GlmM/UreA* (49% and 69%) and *UreA/16SrRNA* (48% and 73%) **(Table 6.4)**. Our study of the three choosen primers reveal *GlmM* to be promising for detection of *H. pylori having* high specificity and sensitivity for detection of the presence of *H.pylori* in clinical tissue samples directly obtained from endoscopy or surgery in the population under study. However the various combinations of primers shows no increase in sensitivity or specificity. To the best our knowledge this is the first study in NE India which determines the sensitive and specificity of PCR for the *H. pylori* detection.

Chapter 7: Gene Expression Profiling of Gastric Cancer in Northeast India: Role of Helicobacter Pylori and Tobacco

Introduction

Gastric cancer (GC) is the fourth most common malignancy and the second leading cause of cancer death worldwide, following cancers of the lung, breast and colo-rectum. Many disparities exist in the prevalence and incidence of GC. More than 70% of cases occur in developing countries, and half of the total cases in the world occur in Eastern Asia, mainly in China [57, 437, 438]. In Southeast Asian countries, including India, low incidence of GC has been reported [439]. However, a very high Age Adjusted incidence rate (AAR) of $47.7/10^5$ in males and $25.7/10^5$ in females [440] has been reported from Mizoram state in Northeast (NE) region of India. In fact, Mizoram can be grouped under high-risk region within the low prevalence area for GC in India. Risk factors predisposing to the formation of GC include a combination of environmental risks, such as diet and infection (*H.* pylori), and, in some cases, genetic predisposition [57]. *H. pylori* is the most important carcinogen for gastric adenocarcinoma [441]. Mizoram has distinct tobacco and diet habits with homogenous population where individuals rarely marry outside their community. In this region, tobacco is widely used and has been reported to increase the risk of GC [10]. In addition to tobacco, consumption of diet high in salt and high incidence of *H. pylori* infection has also been reported in this population [89]. Besides, diet high in salt has been suggested to be risk factor for GC [442, 443].

Traditional clinicopathologic factors and several interesting molecules, including cell cycle regulation factors such as p27 or cyclin E, cell adhesion molecules such as E-cadherin, angiogenic factors such as vascular endothelial growth factor and placental growth factor, oncogenes such as *c-erbB2* and *c-myc* and tumor supressor genes such as *p53*, have been reported to correlate with the prognosis of GC [444-447]. However, there exists inconsistency among different studies, and the reported parameters provided limited information about prognosis of individual patients due to complex biology of the disease. So far most of the studies have used candidate gene approach yielding different results among different studies [448, 449]. Thus, the potential use of combinations of multiple markers instead of a single marker has been previously recommended upon for the understanding of pathobiology or identification of biomarkers for diagnosis, prognosis and therapeutic response [450-452]. The past decade has seen a revolution in high-throughput technologies for molecular profiling in cancer research. Such analyses of tumour tissues have provided unique molecular signature that can distinguish, identify, and classify discrete subsets of disease, predict the disease outcome, and even predicts the response to therapy [453, 454]. Inspite of the high incidence, the molecular study to understand the mechanism underlying the development and progression of GC in Mizoram is lacking. In the present study, the gene expression profiles of GC patients and correlation with *H.pylori* infection were examined in this region. This study attempts towards better understanding of the candidate players for gastric carcinogenesis.

Materials and methods

Selection of patients and collection of samples

One hundred and twenty four patients with GC were registered at Civil Hospital in Aizawl, Mizoram from 2007 to 2009. Among them 12 cases were either already operated or having malignancy in other organ, gastrointestinal hemorrhage, or receiving aspirin or other nonsteroidal anti-inflammatory drugs, hence were excluded from the study leaving total of 112 patients registered for this study. Five ml of blood samples was also collected from 112 patients and from 66 voluntary, age and sex matched healthy controls. Controls were ruled out for previous history of any malignancy or gastroenterological disorders. Serums were obtained for these blood samples to be used in assayed for IgG antibodies to *H.pylori* in cases as well as controls.

Of the 112 cases only 43 cases had adequate tissue collected in RNA later for research purpose. Multiple endoscopic biopsies were obtained from the tumour site as well as from normal appearing adjacent site for parallel assessment of gene expression and histopathology. One bit of the biopsy sample was stored in RNA Later (Ambion, USA) at 4^0C and the rest of the tissue sample was stored in buffered formalin for histopathogic examination. Samples stored in RNA Later were frozen at -70^0C till further processed. Of these 37 tissue samples were included in the expression study as six samples had degraded RNA hence could not be used for further experiments. All 37 samples had a confirmed histopathologic diagnosis of gastric adenocarcinoma. Detailed questionnaire with specific information regarding diet, smoking, alcohol consumption habits and family history of cancer was completed for all

patients. Informed consent was also obtained from all the patients for participation in this study. Approval for this study had been obtained from the Institutional Human Ethics Committee.

Antibody assay

Sera from 112 patients and 66 control individuals were analyzed for presence of IgG antibodies to *H. pylori* by commercially available enzyme-linked immunosorbent assay (ELISA) kit (DRG Instruments GmbH, Germany), according to the manufacturer's instructions.

Microarray experiments

Five samples with RNA quantity > 1μg and RNA integrity > 8 in the tumour tissue and corresponding normal appearing tissue distant from the tumour site were selected for microarray experiments. Only samples from patients with confirmed diagnosis of gastric adenocarcinoma who gave history of tobacco, betel nut chewing and alcohol consumption were selected for microarray experiments to maintain uniformity of the experimental design. All these cases showed positivity for IgG antibodies to *H. pylori* on ELISA assay. Demographic and clinical characteristics of gastric adenocarcinoma cases were given in **Table 7.I**. All other samples were used for validation by real-time RT-PCR assay.

Total RNA was isolated from snap-frozen biopsies using RNeasy mini kit (Qiagen, Valencia, CA) and its quantity was determined by the NanoDrop® ND-1000 UV-Vis spectrophotometer. RNA integrity was determined using the RNA 6000 Nano LabChip on the Agilent 2100 Bioanalyzer (Agilent

Technologies, Palo Alto, CA). The Agilent 2100 bioanalyzer is a small bench top system to integrate sample separation, detection, quantification, and data analysis. Each disposable RNA chip is used to determine the concentration and purity/integrity of 12 RNA samples with a total analysis time of about 25 minutes. The 18S and 28S ribosomal RNA peaks are identified by the Agilent 2100 bioanalyzer software and dominate the electropherogram. The Agilent 2100 biosizing software includes data collection, presentation and interpretation functions. Data can be displayed as a gel-like image and/or as electropherogram **(Figure 7.1)**. For each sample, the ratio of the two major ribosomal RNA bands is automatically determined and displayed with the RNA quantitation data on the electropherogram. For the RNA applications, the instrument uses fluorescence detection, monitoring the fluorescence between 670 nm and 700 nm. In order to standardize the process of RNA integrity interpretation, Agilent Technologies has introduced a new tool for RNA quality assessment. The RNA Integrity Number (RIN) was developed to remove individual interpretation in RNA quality control. It takes the entire electrophoretic trace into account. The RIN software algorithm allows for the classification of eukaryotic total RNA, based on a numbering system from 1 to 10, with 1 being the most degraded profile and 10 being the most intact. The RIN software algorithm was developed for samples acquired with the Eukaryote Total RNA Nano assay on the Agilent 2100 bioanalyzer.Equal quantities of RNA extracted from corresponding normal gastric mucosa were pooled as shown in **Figure 7.2**. ExpressArt® Amino Allyl mRNA amplification

Kits (Ocimum Biosolution, Hyderabad, India) was used for labeling cRNA following manufacturer's protocol. Gene expression profile of all the five

tumors were compared with that of pooled normal control individually. For this, cRNA from each tumour samples and pooled controls were labeled with cyanine 3 and hybridized on 'OciChip Human A' chip (Ocimum Biosolution, Hyderabad, India) which contained 20160 genes. The labeled and fragmented cRNAs were hybridized at 65˚C for 17h.

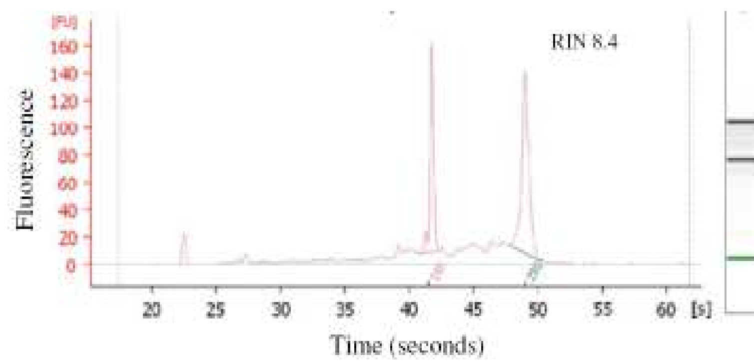


Figure 7.1: Representative picture of electropherogram showing RNA quality

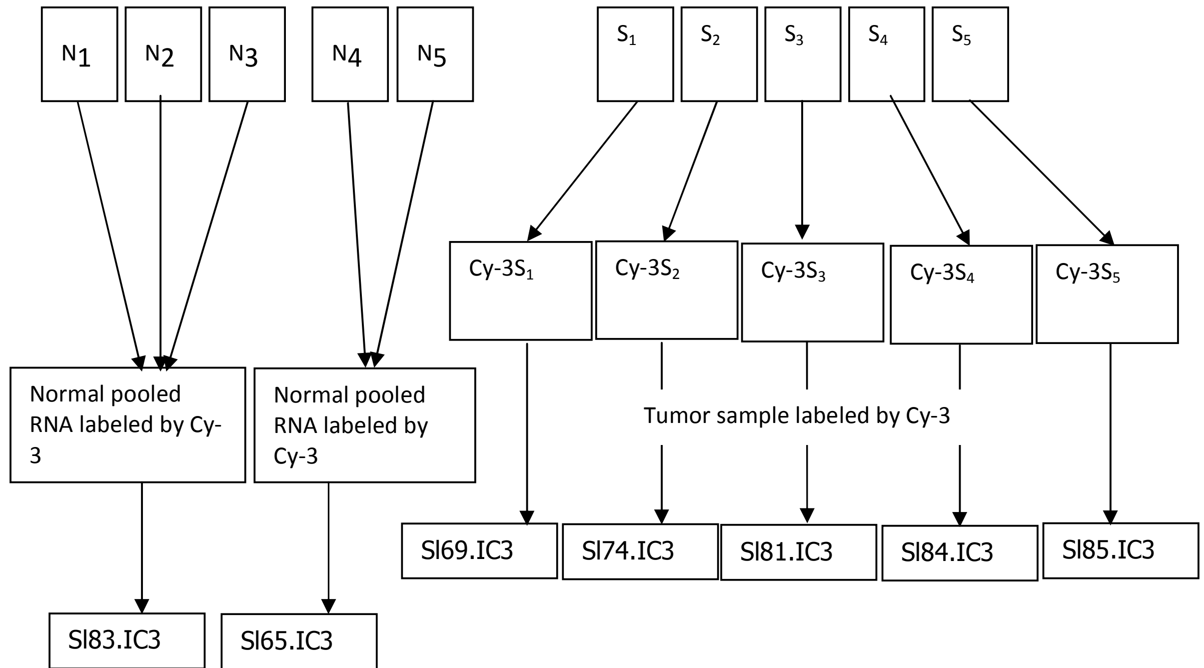


Figure 7.2: Experimental design: S_{1-5} ***and*** N_{1-5} ***indicate samples and their corresponding normal tissue RNA respectively.*** $Cy\text{-}3_{1-5}$ ***indicates the Cy-3 labeled tumor RNA samples. Sl83.IC3 Normal, Sl65.IC3 Normal, Sl69.IC3 Tumor, Sl74.IC3 Tumor, Sl81.IC3 Tumor, Sl84.IC3 Tumor, Sl85.IC3 Tumor indicated barcode of microarray chips***

Microarray image acquisition and data analysis

Hybridized arrays were scanned at 5 µm resolution on an AFFYMETRIX 428™ Array Scanner at 100% laser power and 30% PMT at 532 nm for Cy3-labeled samples. The resulting TIFF images were analyzed by R package and Genowiz™ Software (Ocimun BIosolution, India). The expression data was filtered by removing the blank spots and controls spots on the chip and 19700 probes of the 20160 were used for further analysis. The data obtained by image processing showed positively skewed distribution for each array. In order to have across array comparison, the data was normalized using $\log^2$ transformation. The normalization of each array could be visualized

through **Figure 7.3**. The left panel shows box plots of raw data for different arrays, while the right panel shows the box plots for the log transformed data.

It is evident from the figure that the transformation could normalize the data for each array; however, mean intensity level across arrays was different. Also, the scatter for each array was found varying. To overcome this, median centering and median absolute deviation (MAD) scaling was performed on each array. The resulting array wise distribution is shown in **Figure 7.4**. The simple pre-processing adjusted the mean intensity levels of each array to zero.

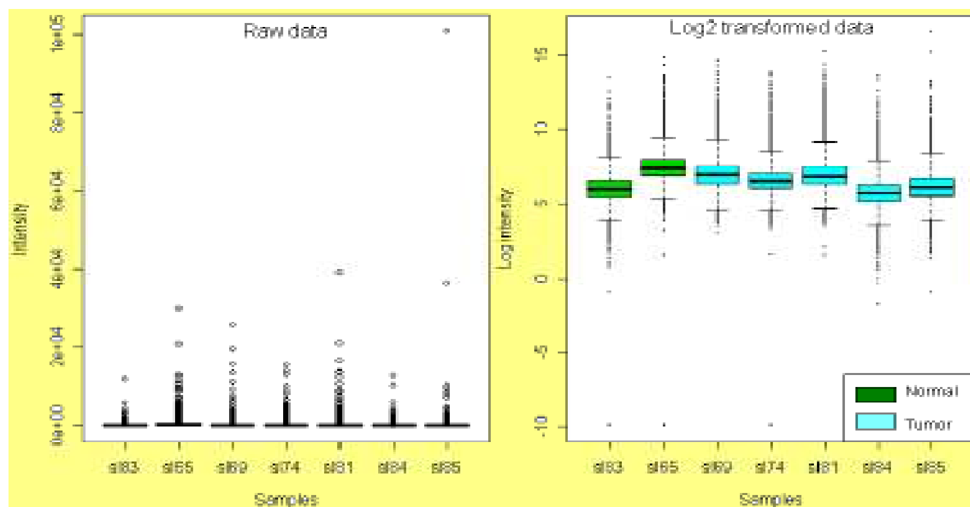


Figure 7.3: Box plot for raw and log2 transformed data for each array

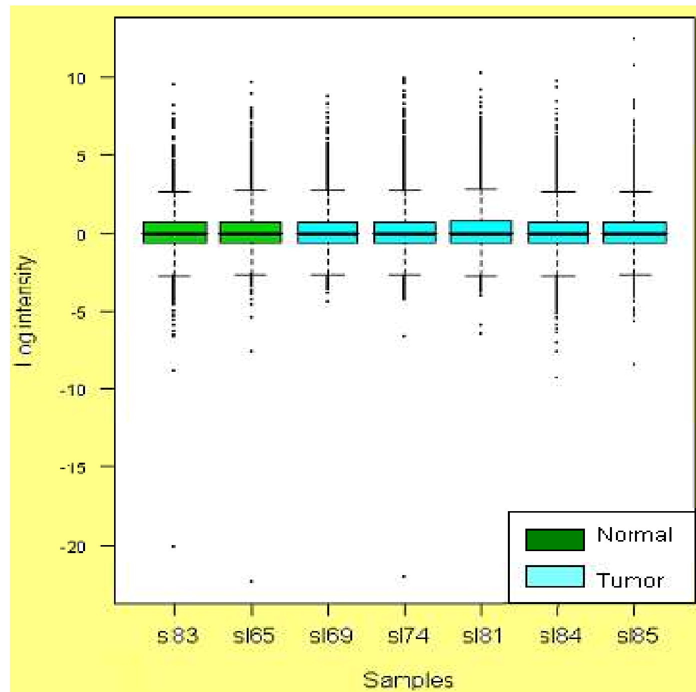


Figure 7.4: Box plot for each sample after median centering and MAD scaling

Also, the uniformity of scatter across arrays was achieved through MAD scaling. This data was further used for differential expression analysis.A threshold p-value of 0.05 was fixed so that all probes having p-value less than the threshold were declared significantly expressed across the two conditions. Moreover, the fold change for each probe was calculated as the difference between the mean log transformed intensity levels, thereby resulting into log fold change. The probes with log fold change either greater than +1 or less than -1 (equivalent to 2-fold change) were considered to be biologically significant. The probes having log fold change greater than 1 were treated as up-regulated, while those with log fold change less than -1 were considered as down-regulated. A volcano plot showing the scatter of probes based on log fold change and p-value is shown in **Figure 7.5**. The up and down regulated

probes are indicated with green and blue color codes in the figure. The expression data on the up and down regulated probes on seven experimental samples (five from tumour samples and two from normal controls) was considered for hierarchical clustering. Two-way hierarchical clustering was used with Euclidean distance as a measure of proximity and average linkage method were used to determine clustering. Both the samples and probes were clustered simultaneously and visualized through a heatmap. **Figure 7.6** shows the heatmap with samples clustered horizontally and probes clustered vertically. Dendrograms are shown on the left and top showing the relatedness of probes and samples respectively. The mapping of color codes and intensity levels is shown through 'Color key'. The microarray data was submitted to the GEO repository (GSE 20143) at http://www.ncbi.nlm.nih.gov/geo.

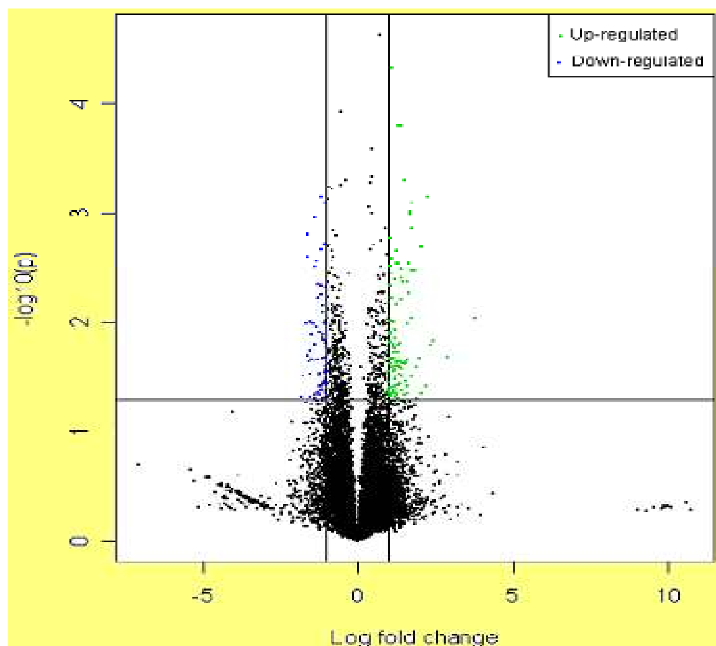


Figure 7.5: Volcano plot showing up and down regulated genes.

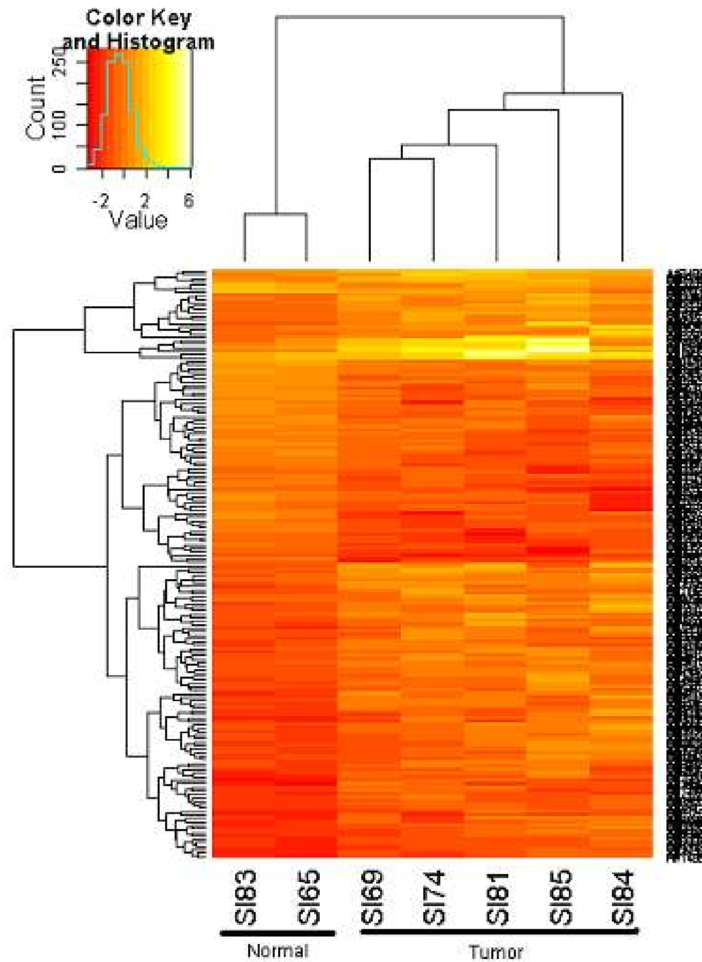


Figure 7.6: Heatmap showing the clustering of samples and probes as a measure of proximity. Samples clustered horizontally and probes clustered vertically. Dendrograms are shown on the left and top showing the relatedness of probes and samples respectively

Gene enrichment analysis

The probes obtained through the comparison were studied for their over abundance in different Gene Ontology (GO) terms as well as Pathways. The terms could be categorized into biological process, molecular function and cellular component. Fisher's exact test was used to determine the significance of the GO term. If a term was significant with say $p < 0.05$, then it was implied that it was enriched with genes. Accordingly, the biological relevance of the term and the associated genes could be explored.

Validation of microarray results by quantitative real-time RT-PCR analysis

Thirty Seven of the tumor tissues and counter normal parts (including those used in microarray experiment) were used in the experiment. One microgram of tumor and normal RNA was reverse transcribed into cDNA with random primers (High Capacity cDNA archive kit, Applied Biosystems, Foster City, CA). Real-time RT-PCR reactions were performed using an ABI Prism 7000 sequence detection system (Applied Biosystems). Primers and TaqMan probes of eight target genes and an internal control gene *TBP* were purchased as assays-on-demand from Applied Biosystems**.** The thermal cycling conditions included an initial denaturation step at 95˚C for 10 min, 40 cycles at 95˚C for 15 sec and 60˚C for one min. The 2-ΔΔCT method was used to calculate relative changes in gene expression determined from real-time quantitative PCR experiments. Validation of microarray results was done in 37 gastric adenocarcinoma cases including the samples where microarray experiments were carried out. Wilcoxon signed rank tests (paired) were used to determine the statistical significance of expression difference for each test gene in the 37 gastric adenocarcinoma cases and controls **(Table 7.4A, 7.4B)**. The statistical analyses were performed with Graphpad Prism Version 5 (La Jolla, CA, USA).

Statistical analysis

Correlation study of *H. pylori* positive and negative patients with the expression of *LDHB*, *PXN*, *RYK*, *RET*, *ANG*, *BMP7*, *RAD51L3* and *CXCL10* was performed by Karl Pearson's correlation coefficient **(Table 7.5A, 7.5B)**. The association of tobacco smoking, tobacco chewing, betel quid chewing, alcohol consumption and *H. pylori* infection with disease outcome was assessed using univariate and multivariate conditional logistic regression models **(Table 7.2)**. Test for studying interaction were performed for *H. pylori* infection with all the considered covariates **(Table 7.6)**. For all the estimates, two ORs were calculated. A crude unadjusted OR (OR1) was calculated using univariate conditional logistic regression model. OR2 was the adjusted OR using conditional multivariate logistic regression model. For all the tests, a two-sided $p<0.05$ was considered statistically significant. The data analysis was performed on STRATA 8.0 software.

Results

Clinical and epidemiological information

The mean age of the 112 patients included in the study for *H. pylori* detection study was 59 ± 12.2 SD and the males to females ratio was 3:1. The mean age of the 66 healthy controls used for the same study was 58.5 ± 12.04 SD and males to females ratio was 2.6:1. Three of the cases included in the study had first-degree relatives with esophageal, lung and breast cancer. Of the 112 samples included in the study, 54% (60 of 112) had well differentiated, 8% (9 of 112) had moderately differentiated and 38% (43 of 112) had poorly differentiated adenocarcinoma. The tumour location was as follows: 54 in cardia (48%), 25 in body (22%), 12 in GE junction (11%), 9 in pylorus (8%), 12 in fundus (11%). Of the 112 samples 69% (77 of 112) were tobacco chewers, 56% (63 of 112) were smokers, 54% (60 of 112) were betel quid chewers and 26% (29 of 112) were alcohol consumer. Fifty three (47%) were smokers as well as chewers, 52 (46%) were smokers as well as betel quid consumers, 33 (29%) were smokers as well as alcohol consumer, 54 (48%) were chewers as well as betel quid consumers, 25 (22%) were chewers as well as alcohol consumers, 22 (19%) were betel quid as well as alcohol consumers. Twenty (16%) of them has all the three habits. Seventy five of the 112 (67%) patients were found positive for IgG antibodies against *H. pylori* in sera. *H. pylori* specific IgG antibodies were also found positive in the sera of 42 of 66 (64%) age and sex matched voluntary healthy control individuals. Of the 37 patient studied for real time validation 26 patients were positive for *H. pylori* specific IgG antibodies.

In the total 112 patients, frequency of tobacco chewers was higher in cases (77 of 112, 69%) than controls (34 of 66, 52%) (x^2=4.55, *p*=0.032) and conferred a significant increase risk of upto two fold (OR2=2.11, 95% CI=1.09-4.05, *p*=0.026) of developing GC. The frequency of *H. pylori* positive cases did not show significant difference among cases and controls (OR2=1.19, 95% CI=0.59-2.39, *p*=0.63) **(Table 7.2)**. For studying the role of *H.pylori*-environment interaction that might modify susceptibility of GC, potential interactions of *H. pylori* with known risk factors were analysed. Although, betel quid chewing do not emerged as a risk factor, interestingly on interaction with *H. pylori* it causes a significant threefold risk of GC (OR2=3.52, 95% CI=1.16-10.68, *p*=0.026) **(Table 7.6)**. However, interaction of *H. pylori* status with tobacco smoking (OR2=1.4, 95% CI=0.42-4.64, *p*=1.4) and alcohol consumption (OR2=1.01, 95% CI=0.37-2.76, *p*=0.98) does not conveyed any significant interaction. When relationship between tobacco chewers were analysed with *H. pylori*, significant interaction were observed with all the three combination **(Table 7.6)** but the significant level was enhanced in those individuals who were *H. pylori* positive as well as chewers (OR2=7.12, 95% CI=5.02-36.31, *p*≤0.0001).

Gene expression profiling by oligonucleotide microarray

Gene expression profile of five gastric adenocarcinoma samples were compared with normal appearing pooled gastric mucosal tissue. All five samples were obtained from patients with a history of tobacco consumption, alcohol consumption and betel quid chewing. None of them had family history

of cancer **(Table 7.1)**. Sera of all the five cases showed positivity for IgG antibody against *H. pylori* by ELISA.

Using stringent criteria (*P≤0.05* and ≥1-fold change), 170 genes were found differentially expressed, 108 upregulated and 62 down-regulated and are categorized using the GO database into known or probable functional categories on the basis of biological processes and molecular function.

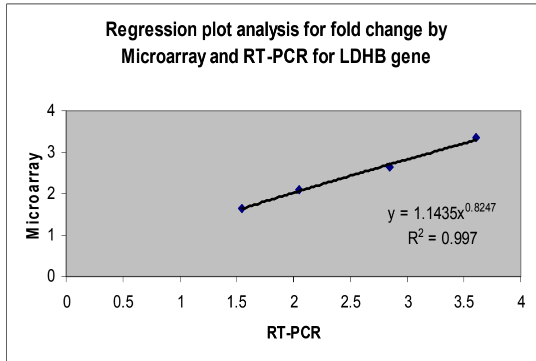
The heatmap revealed that the two normal samples were clustered together and were separated from tumor samples. This was also evident through color profiles **(Figure 7.6)**. Of the 108 significantly up-regulated genes, genes involved in apoptosis (*MALT1*, *HBXIP, RYK, SH2D2A, CSNK1E, LRP12, CHP*) proinflammatory response (*IL9*), angiogenesis (*SH2D2A, ANG*), cell adhesion (*CD151, PXN, WASF1)*, cell proliferation *(CDC2L1, JAG2, IL9),* regulators of I-kappa B kinase / NF-kappa B cascade (*CXXC5*), Wnt signaling (*CHP, CSNK1E, RYK*), Notch signaling (*JAG2*), Hedgehog signaling (*CSNK1E, BMP7*), metobolic pathway (*LDHB*), thyroid cancer (*RET)* and epithelial cell signaling in *H. pylori* infection (*ATP6V1G2*) were found to be biologically relevant in tumorigenesis **(Table 7.3A)**. Of the 62 significantly down-regulated genes, genes involved in base-excision repair (*RAD51L3*), inhibition of cellular proliferation (*INSM1),* immune response (*RFX1, CXCL10, C3AR1*), viral response (*ATP6V1G2, HBXIP, ACE2*), epithelial cell signaling in *H. pylori* infection (*JAM3*) were found to be relevant in tumorigenesis

(Table 7.3B).

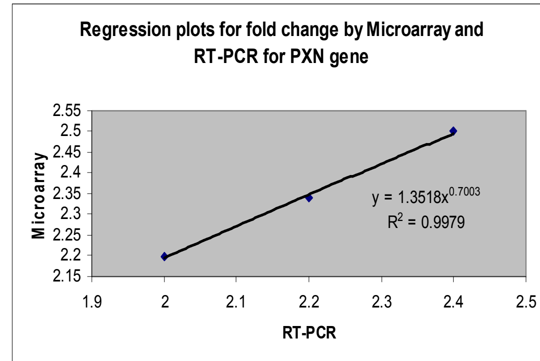
Validation of Selected Genes

To validate the microarray data, real-time RT-PCR was done and the expression levels of *LDHB, PXN, RYK, RET, ANG, BMP7, RAD51L3* and *CXCL10* genes were analyzed in 37 cases of GC **(Table 7.4A)**. Up-regulation of *LDHB, PXN, RYK, RET, ANG* and down-regulation of *BMP7, RAD51L3* and *CXCL10* in tumour tissue relative to controls were confirmed, supporting the microarray results. An internal control gene *TBP* purchased as assays-on-demand from Applied Biosystems was used for normalization. Wilcoxon signed rank test revealed that the difference in the expression levels of each gene as relative messenger RNA to the reference samples normalized to *TBP* were statistically significant. The mRNA expression levels of *LDHB, PXN, RYK, RET, ANG, BMP7, RAD51L3* and *CXCL10* were compared with normal controls ($p= 0.046$, $p= 0.026$, $p= 0.0071$, $p<0.0001$, $p=0.0035$, $p<0.0001$, $p= 0.023$, $p= 0.027$) respectively **(Table 7.4A)**. Regression plot analyses for the seven genes showed positive correlation between the gene expression measured by microarray and real-time RT-PCR **(Figure 7.7)**.

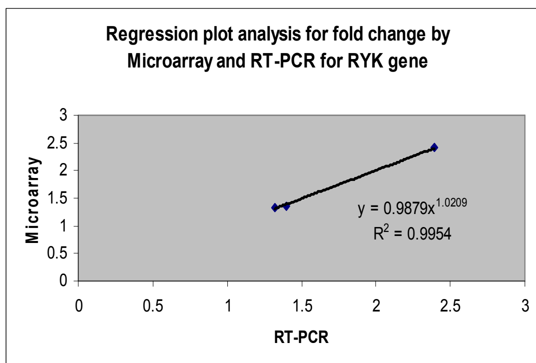
A



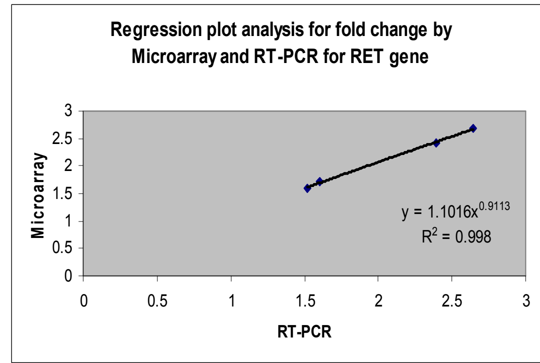
B



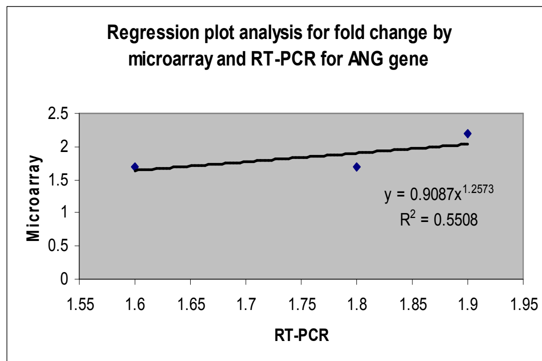
C



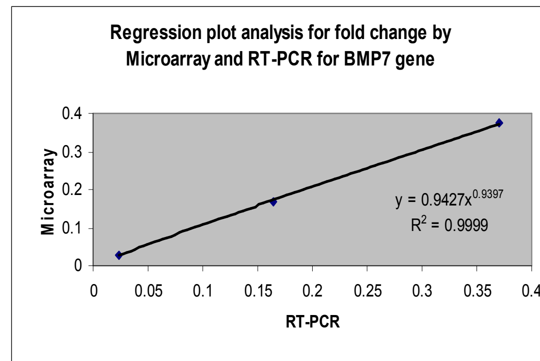
D



E



F



G

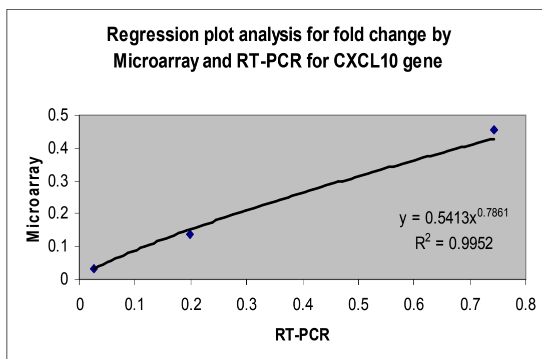


Figure 7.7: Regression plots for fold change by microarray (Y-axis) and quantitative real-time PCR assay (X-axis) for LDHB (A), PXN (B), RYK (C), RET (D), ANG (E), BMP7 (F) and CXCL10 (G)

In *H. pylori* positive cases significant down-regulation in expression of *RAD51L3 (p=0.0013), CXCL10 (p=0.0038) and BMP7 (p=0.0002)* genes whereas significantly up-regulation in expression of *PXN (p=0.0077), ANG (p=0.0015), LDHB (p=0.019), RYK (p=0.044)* and *RET (p<0.0001)* genes were observed as compared to adjacent normal tissue. The *H. pylori* negative cases showed a significant up-regulation of expression of *RET* (*p=0.024*) and a significant down-regulation of expression of *BMP7* (*p*=0.018) as compared to its adjacent normal tissue. The gene expression of *PXN (p=0.039)* was alone found to be significantly up-regulated in *H. pylori* positive GC cases on comparison to *H. pylori* negative GC cases **(Table 7.4B)**.

A significant direct correlation of expression of *PXN* with *RET* (0.523, *P=0.05*) was observed in *H. pylori* positive patients **(Table 7.5A)** while *H. pylori* negative cases did not show any correlation in the expression of any of the genes studied **(Table 7.5B).**

Table 7.1. Demographic and clinical characteristics of gastric adenocarcinoma cases for microarray and real time experiments.

Patient ID	Age	Gender	Tobacco chewing	Smoking	Alcohol	Betel	FH	Site	Histological Grade	Experiments
ST-204	59	M	Yes	Yes	Yes	Yes	No	Cardia	Moderately diff.	MA, RT
ST-269	55	M	Yes	Yes	Yes	Yes	No	Cardia	Well diff.	MA, RT
ST-258	57	M	Yes	Yes	Yes	Yes	No	Cardia	Well diff.	MA, RT
ST-262	60	M	Yes	Yes	Yes	Yes	No	Cardia	Poorly diff.	MA, RT
ST-276	62	M	Yes	Yes	Yes	Yes	No	Cardia	Well diff.	MA, RT
ST-199	54	M	No	Yes	Yes	No	No	Body	Moderately diff.	RT
ST-203	47	F	Yes	No	Yes	Yes	No	Body	Poorly diff.	RT
ST-205	72	F	No	Yes	Never	Yes	No	Body	Well diff.	RT
ST-206	38	F	No	Yes	Never	Yes	No	GE junction	Poorly diff.	RT
ST-207	65	M	No	Yes	Never	No	Yes	Cardia	Well diff.	RT
ST-208	54	M	No	Yes	Yes	Yes	No	Pylorus	Moderately diff.	RT
ST-209	73	M	No	Yes	Yes	Yes	No	Cardia	Poorly diff.	RT
ST-210	52	M	Yes	Yes	Never	No	No	Fundus	Well diff.	RT
ST-211	54	F	Yes	Yes	Never	No	No	Cardia	Well diff.	RT
ST-213	50	F	No	Yes	Never	No	No	GE junction	Well diff.	RT
ST-214	83	M	Yes	Yes	Never	No	No	Cardia	Well diff.	RT
ST-281	57	F	Yes	Yes	Yes	No	No	Cardia	Well diff.	RT
ST-259	52	M	No	No	Never	No	No	Cardia	Poorly diff.	RT
ST-260	50	M	No	Yes	Yes	Yes	No	Cardia	Well diff.	RT
ST-261	75	M	No	Yes	Never	Yes	No	Fundus	Well diff.	RT
ST-264	58	M	No	Yes	Never	Yes	No	Cardia	Well diff.	RT
ST-265	52	F	No	Yes	Never	No	No	Cardia	Well diff.	RT
ST-266	65	M	No	Yes	Never	Yes	No	Cardia	Well diff.	RT
ST-267	67	AM	No	Yes	Yes	Yes	No	Cardia	Poorly diff.	RT

ST-268	77	F	No	Yes	Never	Yes	No	Pylorus	Poorly diff.	RT
ST-271	65	M	No	Yes	Never	No	No	Cardia	Well diff.	RT
ST-272	54	F	No	Yes	Never	No	No	Body	Poorly diff.	RT
ST-273	75	M	Yes	Yes	Never	No	Yes	GE Junction	Poorly diff.	RT
ST-274	49	M	Yes	Yes	Never	No	No	Fundus	Well diff.	RT
ST-275	47	M	Yes	Yes	Never	Yes	No	Body	Poorly diff.	RT
ST-277	75	F	No	No	Never	No	Yes	Cardia	Poorly diff.	RT
ST-278	80	M	Yes	Yes	Never	Yes	No	Body	Well diff.	RT
ST-279	40	F	Yes	Yes	Never	Yes	No	GE Junction	Poorly diff.	RT
ST-304	65	M	No	Yes	Never	Yes	No	Body	Poorly diff.	RT
ST-306	58	M	No	Yes	Never	No	No	Fundus	Well diff.	RT
ST-307	70	M	No	No	Never	No	No	Pylorus	Poorly diff.	RT
ST-311	58	M	No	Yes	Never	No	No	Body	Well diff.	RT

diff. – Differentiated **MA –** Microarray **RT –** Real Time **FH** – Family history

Table 7.2: Risk estimates for *H pylori* status, tobacco, betel quid and alcohol consumption habits in GC patients

Risk factors	Cases (n=112)	Controls (n=66)	Univariate		Multivariate	
	n (%)	n (%)	OR1 (95% CI)	*P* value	OR2 (95% CI)	*P* value
H.pylori Positive[a]	75 (67)	42 (64)	1.26 (0.66-2.39)	0.48	1.19 (0.59-2.39)	0.63
Tobacco Smoking[b]	63 (56)	32 (48)	1.37 (0.74-2.52)	0.32	1.54 (0.79-3.02)	0.21
Tobacco Chewing[c]	77 (69)	34 (52)	2.07 (1.11-3.88)	0.023	2.11 (1.09-4.05)	**0.026**
Betel-quid Chewing[d]	60 (54)	30 (45)	1.23 (0.67-2.25)	0.51	1.27 (0.68-2.38)	0.45
Alcohol Consumption[e]	29 (26)	18 (27)	0.93 (0.47-1.85)	0.84	0.89 (0.42-1.81)	0.72

OR1*: crude odds ratio
a : OR2 adjusted for tobacco smoking, tobacco chewing, betel quid chewing and alcohol consumption
b : OR2 adjusted for *H. pylori* positive, tobacco chewing, betel quid chewing and alcohol consumption
c : OR2 adjusted for *H. pylori* positive, tobacco smoking, betel quid chewing and alcohol consumption
d : OR2 adjusted for *H. pylori* positive, tobacco smoking, tobacco chewing and alcohol consumption
e : OR2 adjusted for *H. pylori* positive, tobacco smoking, tobacco chewing and betel quid chewing

Table 7.3. Biologically relevant and statistically significant up-regulated and down-regulated genes in gastric cancer patients
A, Up-regulated genes

Genes	Gene symbol	Gene bank ID	Chromosomal location	fold change	*p-value* (gene specific)[a]	GO category and pathway
Angiotensin I converting enzyme (peptidyl-dipeptidase A) 2	*ACE2*	NM_021804.1	Xp22	2	0.02	Proteolysis, Viral receptor activity, Renin-angiotensin system pathway
Heat shock transcription factor 2	*HSF2*	NM_004506.2	6q22.31	1.7	0.00	Response to stress
Low density lipoprotein-related protein 12	*LRP12*	NM_013437.2	8q22.2-q23.1	1.7	0.001	Caspase activation, Cell cycle
Angiogenin, ribonuclease, RNase A family, 5	*ANG*	NM_001145.2	14q11.1-q11.2	2.2	0.04	Angiogenesis
Ret proto-oncogene	*RET*	X12949.1	10q11.2	2.8	0.03	Endocytosis, Pathways in cancer, Thyroid cancer
Interleukin 9	*IL9*	NM_000590.1	5q31.1	1.5	0.02	Regulation of cell proliferation, Proinflammatory response, Cytokine-cytokine receptor interaction pathway
Sperm associated antigen 6	*SPAG6*	NM_012443.2	10p12.2	1.5	0.005	Microtubule, Cell organization and biogenesis
Cullin 4B	*CUL4B*	NM_003588.2	Xq23	1.1	0.03	Cell cycle, Cellular component
Paxillin	*PXN*	AK128712.1	12q24.31	2	0.03	Cell-matrix adhesion, Cell adhesion

WAS protein family, member 1	*WASF1*	BC068546.1	6q21-q22	1	0.04	Cell Communication; Adherens junction, Bacterial invasion of epithelial cells pathway
RNA binding protein S1, serine-rich domain	*RNPS1*	BC001659.2	16p13.3	1	0.04	mRNA catabolic process
CD151 antigen	*CD151*	BT020132.1	11p15.5	1	0.009	Cell adhesion
RYK receptor-like tyrosine kinase	*RYK*	S59184.1	3q22	2.7	0.003	Caspase activation
SH2 domain protein 2A	*SH2D2A*	NM_003975.2	1q21	1.6	0.02	Angiogenesis, caspase activation, VEGF signaling pathway
Calcium binding protein P22	*CHP*	NM_007236.3	15q13.3	1.2	0.04	Small GTPase mediated signal transduction, Wnt signaling, VEGF signalling pathway
Casein kinase 1, epsilon	*CSNK1E*	BT019831.1	22q13.1	1.03	0.04	Caspase activation, Wnt and Hedgehog signaling pathway
Fatty acid desaturase 2	*FADS2*	NM_004265.2	11q12-q13.1	1.2	0.02	Lipid metabolic process, PPAR signaling pathway
Neuregulin 1	*NRG1*	AF491780.1	8p12	1.06	0.01	Nervous system development, ErbB signaling pathway
Myosin VI	*MYO6*	NM_004999.2	6q13	1.2	0.03	Regulation of DNA replication initiation, Calmodulin binding
ATPase, H+ transporting, lysosomal 13kDa, V1 subunit G isoform 2	*ATP6V1G2*	BC068023.1	6p21.3	1.3	0.02	ATP biosynthetic process, Vibrio cholera infection, Epithelial cell signaling in *H. pylori* infection pathway
Jagged 2	*JAG2*	Y14330.1	14q32	1.4	0.04	Cell proliferation, Cell migration, Cell cycle, Notch signaling pathway
Formin 2	*FMN2*	NM_020066.3	1q43	1.3	0.03	Intracellular signaling cascade

Lactate dehydrogenase B	*LDHB*	NM_002300.3	12p12.2-p12.1	2.05	0.05	Metabolic pathway (glycolysis, gluconeogenesis)
Hepatitis B virus x interacting protein	*HBXIP*	NM_006402.2	1p13.3	1.1	0.02	Viral genome replication, Anti-apoptosis
CDC37 cell division cycle 37 homolog (S. cerevisiae)-like 1	*CDC37L1*	NM_017913.2	9p24.1	1.3	0.002	Cell division
Dynein, axonemal, heavy polypeptide 5	*DNAH5*	NM_001369.1	5p15.2	1.1	0.009	Cellular components
CXXC finger 5	*CXXC5*	NM_016463.5	5q31.2	1.2	0.007	Positive regulation of I-kappa B kinase/ NF-kappa B cascade
Mucosa associated lymphoid tissue lymphoma translocation gene 1	*MALT1*	BC030143.2	18q21	-1.1	0.005	Anti- apoptosis, Signal transducer activity, B cell receptor signaling pathway

B. Down-regulated genes

Genes	**Gene symbol**	**Gene bank ID**	**Chromosomal location**	**fold change**	***p-value* (gene specific)[a]**	**GO category and pathway**
RAD51-like 3	*RAD51L3*	NM_133628.1	17q11	-2.4	0.02	Base-excision repair
RAB8B, member RAS oncogene family	*RAB8B*	NM_016530.2	15q22.2	-1.4	0.01	[b]NA
Signal transducer and activator of transcription 4	*STAT4*	NM_003151.2	2q32.2-q32.3	-1.4	0.02	Signal tranducer activity
Bone morphogenetic protein 7	*BMP7*	NM_001719.1	20q13	-2.7	0.04	Signaling Molecules and Interaction, Hedgehog signaling pathway, cellular processes
Chemokine (C-X-C motif) ligand 10	*CXCL10*	NM_001565.1	4q21	-1.1	0.05	Immune response
Sulfotransferase family 1E, estrogen-preferring,member 1	*SULT1E1*	NM_005420.2	4q13.1	-1.1	0.002	Estrogen metabolic process, Sulfur metabolism pathway
Insulinoma-associated 1	*INSM1*	NM_002196.2	20p11.2	-1.2	0.029	Inhibit cellular proliferation, Cell differentiation

Corticotropin releasing hormone binding protein	*CRHBP*	NM_001882.3	5q11.2-q13.3	-1.1	0.013	Protein binding, hormone metabolic process
Regulatory factor X, 1	*RFX1*	NM_002918.3	19p13.1	-1.2	0.04	Immune response
TAF7-like RNA polymerase II	*TAF7L*	NM_024885.2	Xq22.1	-1.2	0.002	Cell differentiation
Cysteine rich transmembrane BMP regulator 1	*CRIM1*	NM_016441.1	2p21	-1.2	0.0007	Regulation of cell growth
Junctional adhesion molecule 3	*JAM3*	NM_032801.3	11q25	-1.3	0.004	Epithelial cell signaling in *H. pylori* infection pathway
Complement component 3a receptor 1	*C3AR1*	NM_004054.2	12p13.31	-1.1	0.008	Complement component C3a receptor activity, G-protein coupled receptor protein signaling pathway, Neuroactive ligand-receptor interaction pathway

[a]Biological significance of differentials was computed and functionally classified using the Genowiz™ software on the basis of gene ontology. Pathways were obtained using enrichment analysis based on gene ontology categories. [b]NA, No information available.

Table 7.4A. Locations, function, assay ID and amplicon size of eight genes that were validated by Real-Time RT-PCR

Gene	Gene Bank ID	Location	Gene Expression status	Putative function	p[a]	Assay ID[b]	Amplicon size (bp)
LDHB	NM_002300.3	12p12.2-p12.1	Up	Metabolic pathways	0.046	Hs00929953_m1	90
PXN	AK128712.1	12q24.31	Up	Cell-matrix adhesion, Cell adhesion	0.026	Hs01104424_m1	89
RYK	S59184.1	3q22	Up	Caspase activation	0.0071	Hs00243196_m1	84
RET	X12949.1	10q11.2	Up	Proliferation	p<0.0001	Hs01120027_m1	92
ANG	NM_001145.2	14q11.1-q11.2	Up	Protein synthesis	0.0035	Hs01590076_m1	156
BMP7	NM_001719.1	20q13	Down	Cytokine-cytokine interaction	p<0.0001	Hs00233477_m1	81
RAD51L3	NM_133628.1	17q11	Down	Base-excision repair	0.023	Hs00172529_m1	69
CXCL10	NM_001565.1	4q21	Down	Apoptosis, Immune response	0.027	Hs00171042_m1	98

[a]Wilcoxon signed rank tests were used to determine the statistical significance of expression difference for each test gene in 37 samples. Statistical significance was defined as $p<0.05$. [b]m1 denotes that assay's probe spans an exon junction and will not detect genomic DNA.

Table 7.4B: Differential expression between *H. pylori* positive (HP+) patients and *H. pylori* negative (HP-) patients

Gene symbols	HP+ve cases Mean±SE n=26	HP+ve cases adjacent normals Mean±SE n=26	*p-value*[a]	HP-ve case Mean±SE n=11	HP-ve cases adjacent normals Mean±SE n=11	*p-value*[a]
PXN	2.59 ± 0.50	0.46 ± 0.29	0.0077	0.97 ± 0.71	1.45 ± 0.49	0.56
RAD51L3	1.83 ± 0.64	3.39 ± 0.36	0.0013	2.65 ± 0.67	2.56 ± 0.65	0.59
ANG	1.88 ± 0.39	0.71 ± 0.29	0.0015	1.39 ± 0.65	1.61 ± 0.46	0.59
LDHB	0.61 ± 0.41	0.68 ± 0.34	0.019	0.14 ± 0.61	0.13 ± 0.57	0.89
RYK	2.30 ± 0.26	1.58 ± 0.25	0.044	1.96 ± 0.49	0.91 ± 0.22	0.97
CXCL10	0.41 ± 0.24	1.23 ± 0.23	0.0038	0.83 ± 0.45	1.44 ± 0.33	0.64
BMP7	-1.86 ± 0.32	0.89 ± 0.32	0.0002	2.28 ± 0.59	0.63 ± 0.73	0.018
RET	1.29 ± 0.20	-0.89 ± 0.24	$p<0.0001$	1.36 ± 0.66	-0.59 ± 0.41	0.024

[a]Wilcoxon signed rank tests were used to determine the statistical significance of expression difference for each test gene in HP+ve and HP-ve cases with their corresponding normal counter part. Significance difference was defined as $p<0.05$.

Table 7.5A: Matrix of gene expression of *RAD51L3*, *PXN*, *ANG*, *BMP7*, *CXCL10*, *RET*, *RYK* and *LDHB* in *H. pylori* positive patients

Variable	***RAD51L3***	***PXN***	***ANG***	***BMP7***	***CXCL10***	***RET***	***RYK***	***LDHB***
RAD51L3	1							
PXN	-0.078	1						
ANG	0.122	0.367	1					
BMP7	-0.08	-0.044	-0.108	1				
CXCL10	0.006	0.042	0.316	0.339	1			
RET	-0.223	**0.523**[a]	0.318	0.347	0.389	1		
RYK	-0.088	0.216	-0.258	-0.111	-0.07	0.045	1	
LDHB	-0.072	-0.043	-0.410	0.157	-0.183	0.026	0.381	1

a-Correlation is significant at the 0.05 level (two-tailed)

Table 7.5B: Matrix of gene expression of *RAD51L3*, *PXN*, *ANG*, *BMP7*, *CXCL10*, *RET*, *RYK* and *LDHB* in *H. pylori* negative patients

Variable	***RAD51L3***	***PXN***	***ANG***	***BMP7***	***CXCL10***	***RET***	***RYK***	***LDHB***
RAD51L3	1							
PXN	-0.076	1						
ANG	0.210	-0.148	1					
BMP7	0.139	0.047	0.142	1				
CXCL10	0.301	-0.101	0.029	-0.087	1			
RET	0.408	0.080	0.153	0.072	-0.063	1		
RYK	0.058	0.301	-0.423	0.274	-0.217	0.515	1	
LDHB	0.103	0.129	0.410	0.428	-0.382	0.470	0.195	1

Correlation is significant at the 0.05 level (two-tailed)

Table 7.6: Distribution of interactions terms and odds ratio for *H. pylori* and risk habits

Variables	Interaction	OR1*(95% CI)	*p*-value	OR2(95% CI)	*p*-value
Smoking[a]					
	H. pylori -ve X Non smoker	1.0		1.0	
	H. pylori -ve X Smoker	0.73(0.23-2.3)	0.59	0.81(0.24-2.68)	0.81
	H. pylori +ve X Non smoker	0.74(0.24-2.24)	0.6	0.67(0.22-2.08)	0.67
	H. pylori +ve X Smoker	1.52(0.47-4.89)	0.48	1.4(0.42-4.64)	1.4
Tobacco chewing[b]					
	H. pylori -ve X Non chewer	1.0		1.0	
	H. pylori -ve X Chewer	3.00(1.29-6.97)	0.01	3.7(1.52-9.03)	0.004
	H. pylori +ve X Non chewer	3.19(1.18-8.68)	0.02	4.06(1.41-11.69)	0.009
	H. pylori +ve X Chewer	6.08(3.91-31.12)	0.001	7.12(5.02-36.31)	≤0.0001
Betel-quid chewing[c]					
	H. pylori -ve X Non chewer	1.0		1.0	
	H. pylori -ve X Chewer	2.13(0.84-5.45)	0.11	2.01(0.76-5.31)	0.16
	H. pylori +ve X Non chewer	2.94(1.14-7.6)	0.026	2.67(0.98-7.25)	0.06
	H. pylori +ve X Chewer	3.64(1.22-10.84)	0.02	3.52(1.16-10.68)	0.026
Alcohol consumption[d]					
	H. pylori -ve X Non alcoholic	1.0		1.0	
	H. pylori -ve X Alcoholic	1.11(0.36-3.46)	0.85	1.07(0.34-3.4)	0.917
	H. pylori +ve X Non alcoholic	1.35(0.63-2.88)	0.44	1.31(0.57-3.00)	0.52
R 1	*H. pylori* +ve X Alcoholic	1.16(0.44-3.07)	0.77	1.01(0.37-2.76)	0.98

*: crude odds ratio
a : OR2adjusted for tobacco chewing, betel quid chewing and alcohol consumption
b : OR2adjusted for tobacco smoking, betel quid chewing and alcohol consumption
c : OR2adjusted for tobacco smoking, tobacco chewing and alcohol consumption
d : OR2adjusted for tobacco smoking, tobacco chewing and betel quid chewing

Discussion

It is now evident that multiple etiologic factors associated with environmental, geographical and ethnic differences are responsible for the development of cancer [455-457]. NE region of India have distinct life style and food habits. Earlier studies from this region have showed risk factors such as betel quid chewing or tobacco smoking to be significantly associated with cancers of the breast, oral and lung. In addition genes involved in the detoxification pathways of tobacco constituents are also reported to be associated with the risk of developing oral and GC in NE region [458, 459]. Although several factors including *H. pylori* infection and exposure to tobacco constituents has earlier been reported to be associated with GC [460], very few molecular studies has been carried out to understand its role in the pathogenesis of GC in this high risk region. Studies continue to show that smoking is associated with cancer risk including not only lung but also for oral, larynx, esophageal, gastric and colorectal [461-463]. In concordant with it, the present study showed tobacco chewing to conferred a significant risk (OR2=2.11, 95% CI=1.09-4.05, *p*=0.026) for the development of GC. Tobacco consumption plays an important role in tumorigenesis as it is linked to cellular processes such as apoptosis, inflammation and cell proliferation [464-466]. Earlier studies have established several cancer related molecular targets for tobacco smoke including the *TP53*, *KRAS*, *FHIT*, *RB1* and *HPRT* genes [467-470] supporting the potential role of tobacco as risk factor. Further, our interaction study revealed that tobacco chewers who were also *H. pylori* positive had a seven fold increase risk (OR2=7.12, 95% CI=5.02-36.31,

$p \leq 0.0001$) of developing GC suggesting of a strong role of *H. pylori* and tobacco chewing in combination. A significant interaction of *H. pylori* was also observed with betel quid chewers (OR2=3.52, 95% CI=1.16-10.68, *p*=0.026) whereas interaction with tobacco smoking and alcohol consumption were not significant **(Table 7.6)**. Thus the results also indicate a carcinogen-specific modulation of cancer risk by *H. pylori.* Our results suggest the role of environmental risk factors in modifying the influence between *H. pylori* and GC*. H. pylori* infection might act as a co-carcinogen or promoter of GC risk in association with other factors. However, mechanistic explanation of the association is not clear. Furthermore, studies of *H. pylori* infection status and environmental interaction are scare, therefore, sample size for such an analysis could not be reliably determined and limit our ability to estimate the interaction effect precisely. However, a regionable fraction of the controls came from friends and accompanying family members of the cancer patients. Such selection of control would reduce any confounding bias by reducing background variations with cases. Moreover, our data indicated the area in which search for underlying mechanism. The results of interaction analysis should be considered empirical observations for further studies on larger number of samples to evaluate the relationship between *H. pylori* and environmental risk in the development of GC.

Recent studies have also focused attention on the importance of chronic inflammation, cell proliferation and impaired host immune response due to *H. pylori* infection in the development of GC [471-473]. It has been reported previously that recruitment of immune cells to the site of infection causes

inflammation, manifesting as chronic gastritis which in turn initiates epithelial hyperproliferation [474]. Concordantly, our microarray analysis in *H. pylori* positive cases revealed genes which stimulate inflammation and proliferation viz *IL9*, *CDC2L1* and *JAG2* to be significantly upregulated. Other genes involved in immune response (*RFX1*, *CXCL10, C3AR1*) and viral response *(ATP6V1G2, HBXIP, ACE2)* were significantly down regulated suggesting an inhibition of immune response and might be associated with *H. pylori* as its infection is reported to inhibit the immune system of the host in earlier studies [472].

In addition several other genes such as those governing regulation of cell cycle (*CDC37L1*, *LRP12*, *JAG2*, *CUL4B*), apoptosis related genes (*HBXIP, MALT1, SH2D2A, RYK, CSNK1E, LRP12, FGB, CHP*) were also found significantly up-regulated. Of the apoptotic genes, *MALT1* and *HBXIP* are anti-apoptotic and relevant for the process of tumorigenesis. Upregulation of other apoptotic genes could be explain as these genes also play important role in carcinogenesis related function. For instance, *SH2D2A* have role in angiogenesis. *CSNK1E* in Wnt signaling pathway, *CHP* in GTPase mediated signal transduction, *LRP12* in cell cycle and *RYK* at . Moreover, Florence Lerebours et al has also reported upregulation of apoptotic gene in breast cancer. Moreover this genes are also involved in other cellular processes for instance, , RYK in Beside genes involved in inflammation and immune response, cell cycle and apoptosis, genes involved in cell adhesion (*PXN, CD151, WASF1*) were also significantly up-regulated in our study. An earlier study has reported cell adhesion molecules pathway to be the most significantly deregulated pathway in human gastric epithelial

adenocarcinoma cells and this may contribute to invasion and cell proliferation in GC [475].

Analysis of the molecular functional pathways by Genowiz™ [Ocimun Biosolution, India] identified different potential relevant molecular functional pathways to be deregulated. These included Hedgehog signalling (*CSNK1E, BMP7)*, Wnt signalling (*CHP, CSNK1E, RYK*) and Notch signalling (*JAG2).* Balance between Wnt-FGF-Notch and BMP-Hedgehog signaling network is important to regulate the homoestasis among stem and progenitor cells and disruption of the signaling network results in cancer [476]. In few studies it has earlier been reported that activation of Wnt, Notch and Hedgehog signaling activity may contribute to the development of GC [477-482]. Wnt signalling is involved in a variety of developmental and cellular processes, and aberrant activation of Wnt signaling is linked to several known cancers [483]. Notch signalling pathway was also reported to be associated with the aggressiveness of GC. The activated form of Notch1 receptor promotes the colony-forming ability and xenografted tumor growth of human gastric adenocarcinoma cells [484]. Hedgehog pathway plays an important role in gastric carcinogenesis and has also been linked with chronic gastric inflammation in response to *H. pylori* infection in the mesenchymal stem cell [485]. Besides, genes involved in epithelial cell signalling in *H. pylori* infection (*JAM3, ATP6V1G2)* were also found significantly deregulated supporting the involvement of *H. pylori* in development of GC. The expression of *JAM3* was down-regulated whereas that *of ATP6V1G2* was up-regulated. Both these genes play an important role in epithelial cell signalling in *H. pylori* infection. Deregulation of

this pathway has been reported to lead to abnormal proliferation and movement of gastric epithelial cells [486].

According to Cornea's hypothesis GC develops from chronic inflammation either directly or indirectly leading progressively to metaplasia, dysplasia and finally to cancer [395]. Chronic inflammation in the gastric mucosa may be induced by tobacco consumption and certain dietary factors such as high salt intake in addition to *H. pylori* infection [465, 473, 487]. The people of Mizoram have been reported to have distinct tobacco consumption and food habits that includes high amount of smoked (sa-um) and other salted food stuffs. Epidemiologic studies from this region have earlier shown that theses habits were associated with the high incidence of GC. In addition, a high prevalence of *H. pylori* infection has also been reported from this population [89]. In our study, 67% (75 of 112) GC patients were positive for IgG antibodies against *H. pylori*, however 64% (42 of 66) age and sex matched voluntary healthy control individuals from the same region were also found to be positive for IgG antibodies against *H. pylori*. However, *H. pylori* was not found to be independently associated with the risk of developing GC in our study (p=0.63) (**Table 7.2**). An independent association might have been masked due to high incidence of infection in both cases and control.

Our results revealed different expression level of the validated genes in *H.pylori* infected and non-infected cases **(Table 7.4B)** suggesting the involvement of different molecular processes of carcinogenesis in the two groups. The existence of different molecular mechanism in the two groups have been

supported by other studies such as those of Park DW et al where they have shown the prevalence of different histopathological type of GC in *H.pylori* infected and non-infected patients [488]. Further, Kato S et al has reported the existence of GC risk in the absence of *H.pylori* infection as a fraction of GC patients in their study displayed multifactorial carcinogenesis without *H.pylori* infection [489].

The genes (*PXN, RAD51L3, ANG, LDHB, RYK, CXCL10, BMP7, RET*) choosen for real time RT-PCR were found to be significantly different as compared to their adjacent counter controls in *H. pylori* positive GC cases whereas only *BMP7* and *RET* genes were found to be significantly different in *H. pylori* negative GC cases **(Table 7.4B)**. Moreover a *positive* correlation was also observed between *PXN* and *RET* **(Table 7.5A)** in *H. pylori* positive GC patients suggesting the important role of *PXN* and *RET* in the process of tumorigenesis related with *H. pylori* infection. Both *PXN* and *RET* undergo tyrosine phosphorylation upon stimulation by ligand. *RET*, encodes one of the receptor tyrosine kinases, which are cell-surface molecules that transducer signals for cell growths and differentiation [490]. They have critical role in the development and progression of many types of cancer. *PXN*, encodes a focal adhesion protein, paxillin that could be involved in the progression of cancers through its interaction with the actin cytoskeleton and key signal transduction oncogenes. Phosphorylation of paxillin induced disassembly of adhesion suggesting its role in the stimulation of cell invasiveness in the progression of cancer [491, 492].

Our study set light on the possible molecular mechanism that could be involved in gastric carcinogenesis based on the outcome of *H. pylori* infection and

its interaction with other environmental factors and also on the basis of differentially expression significant genes involved with a number of cellular processes such as apoptosis, immune response, inflammation, cell proliferation and cell-to-cell contacts and tumor relevant pathways such as Wnt, Notch and Hedgehog suggesting of the involvement of a complex mechanism. Our study supported the previous assumption that the process of gastric carcinogenesis is triggered by *H. pylori* and is further characterized by its complexity of interaction with other risk factors [441]. Formation of gastric tumors may be a result of an imbalance between bacterial attack, dietary factors and the complex cellular processes of the host. This study provides data for identifying the possible molecular mechanisms driving gastric carcinogenesis. It is possible that inflammation induced by *H. pylori* infection and high tobacco consumption promotes Wnt, Notch and Hedgehog signaling activating cellular processes leading GC. This is a novel protumorigenic mechanism of inflammation in GC. Continued investigation in these areas will yield novel insight and help to elucidate the mechanisms of gastric carcinogenesis. This is the first gene expression profiling study of GC in Mizoram population of India where a high incidence of GC has been reported. Our work identifies a limited set of pathways that might play significant role in association with earlier reported risks factors such as *H. pylori* infection and dietary factors.

Chapter 8: Assessing Matched Normal and Tumor Pools in Next-Generation Sequencing Studies to Indentify Genomic Alteration in Gastric Adenocarcinoma in High Risk Northeastern Region of India

---

Introduction

Genome-wide association study (GWAS), is a new approach widely applied to genomics following completion of the human genome project. Several early GWAS studies reported potentially promising results, however the majority of GWAS studies were disappointing because of constraint of arrays for certain genetic variations and insufficient sample size and heterogeneity in phenotype [493]. These obstacles may be overcome by new genomic technology, i.e., next-generation sequencing (NGS), also known as massively parallel sequencing or multiplex cyclic sequencing. In the last few years, NGS has emerged as a revolutionary genomic tool and has profoundly impacted biological research.

Like other new genomic technologies, NGS techniques will provide radical insights and change the landscape of genomics. Since many genetic variants which contribute to many human cancer conditions are still not known, NGS will help to identify these genetic variants, including single nucleotide variants (SNVs) or single nucleotide polymorphisms (SNPs), small insertions and deletions, and structural and genomic variants [494]. Previously, DNA sequencing was performed almost exclusively by the Sanger method, which has excellent accuracy and reasonable read length but very low throughput [495]. The limited data due to low throughput have so far lead to a bottom neck condition in cancer

research. A techniques to match the evergrowing need to provide high throughput genomic variants to oncologist to better understand the underlying mechanism of carcinogenesis is the demand of the day. Several NGS methods recently developed allow larger-scale DNA sequencing [495]. Furthermore, this approach is not restricted to few mutation considered relevant. NGS provide data on various SNPs and indels, both known and unknown. A new generation of non-Sanger-based sequencing technologies has delivered on its promise of sequencing DNA at unprecedented speed, thereby enabling impressive scientific achievements and novel biological applications. Finally, NGS has applications that are immediately relevant to the medical field. In cancer genetics, for example, specific cancer alleles can now be detected in tissues through ultra-deep sequencing of genomic DNA, in instances where previous Sanger-based trails have failed. Short read length, initially deemed a major drawback of NGS, becomes a blessing when the Sanger-based 700-bp read length is traded for a much larger number of sequence reads [324, 496].

Mutations in multiple genes are required for cancer to occur [497]. A mutation limited to one oncogene would be suppressed by normal mitosis control and tumor suppressor genes, first hypothesised by the Knudson hypothesis [290]. A mutation to only one tumor suppressor gene would not cause cancer either, due to the presence of many back up genes that duplicate its functions. It is only when enough proto-oncogenes have mutated into oncogenes, and enough tumor suppressor genes deactivated or damaged, that the signals for cell growth overwhelm the signals to regulate it, that cell growth quickly spirals

out of control. It has been postulated that cancer arises as the result of slow accumulation of multiple mutations.

GC is the fourth most common malignancy and second leading cause of cancer death in the world. More than 70% of cases (713 000 cases) occur in developing countries (467 000 in men, 246 000 in women), and half of the total cases in the world occurs in Eastern Asia (mainly in China) [437].In Aizawl district of Mizoram state in Northeast (NE) India, one out of every 13 males carries the risk of developing cancer [72]. Also, very high age-adjusted rates (AAR) of GC ($47.7/10^5$ in males and $25.7/10^5$ in females) [498] has been reported from Mizoram. Mizoram can be grouped under high-risk region within the low prevalence area for GC in India.

Mutations are hallmark of cancers and identification of the mutations is imperative in our understanding of the disease [327]. Traditional clinicopathologic factors and several interesting molecules, including cell cycle regulation factors such as p27 or cyclin E, cell adhesion molecules such as E-cadherin, angiogenic factors such as vascular endothelial growth factor and placental growth factor, oncogenes such as *c-erbB2* and *c-myc* and tumor supressor genes such as *p53*, have been reported to correlate with the prognosis of GC [444-447]. However, there exists inconsistency among different studies, and the reported parameters provided limited information about prognosis of individual patients due to complex biology of the disease. So far most of the studies have used candidate gene approach yielding different results among different studies [448, 449]. NGS provide the plateform where a large number of

genes can be analysed for mutation detection in a short period of time with an effective cost. NGS technologies have increased the speed and throughput capacities of DNA sequencing and, as a result, dramatically reduced overall sequencing costs [499-502]. High throughput and lower per base cost in sequencing allows us to be more focused in our approach to design sequencing projects to fulfill the aim of our research. NGS approaches has moved the expression analysis to the genomic level, which is leading to new discoveries and understandings at much faster pace. With NGS we are also able to correct errors in earlier sequenced genomes, making it in an invaluable genomics based tool in discovering and understanding of complex biological phenomena.

In the current study we have used the solexy plateform for NGS of relevant cancer related genes with high mutation rate earlier reported in various cancers with an attempt to reveal the genomic alteration involved in the carcinogenesis of GC with enhanced pace which is critically required to identify the genetic modification and variation involved in the process of carcinogenesis. For this study of Targeted Re-sequencing of Human, Illumina-Solexa platform has been used, more specifically on the GAIIx instrument. This platform involves Sequencing by Synthesis approach using reversible dye terminator chemistry. Seventy two bp singleton sequence reads were generated and data analyzed including alignment, assembly, and Variation Discovery. One of the most important underlying objective in our approach is the identification of relatively specific mutations involved in GC to develop individualised diagnostics for cancer detection in this region.

Materials and methods

Normal and cancerous tissue samples were taken from twenty GC patients, and next generation sequencing technology was used to look for differences. The samples were collected at Civil Hospital in Aizawl, Mizoram in 2009. Multiple endoscopic biopsies were obtained from the tumour site. One bit of the biopsy sample was stored in PBS at 4^0C for NGS experiment and the rest of the tissue sample was stored in buffered formalin for histopathogic examination. Samples stored in PBS were frozen at -70^0C till further processed. All the samples had a confirmed histopathologic diagnosis of gastric adenocarcinoma. Total genomic DNA were extracted using Qiagen Dneasy blood and tissue kit. Detailed questionnaire with specific information regarding diet, smoking, alcohol consumption habits and family history of cancer was completed for all patients. Informed consent was also obtained from all the patients for participation in this study. Approval for this study had been obtained from the Institutional Human Ethics Committee.

Next generation sequencing

Specific regions of one hundred and sixty nine genes were analyzed by solexa sequencing based on NGS technology. The region of the genes were those where a high prevalence of mutations distribution has been observed in different cancers. These regions were choosen on the basis of the data available on dbSNP database, NCBI.

Capture array design:

The 1 x 244 K Agilent capture array comprising 60 mer tiling probes was designed for exonic regions of interest. The repetitive regions were removed by RepeatMasker filtering and unique probes selected to avoid nonspecific binding.

Library construction:

The GC DNA samples and normal DNA samples were each pooled in equal concentrations to generate two different sets of pooled samples. 10 micrograms of genomic DNA from each pooled sample was made up to 200 µl with nuclease free water and sonicated using a VibraCell (12 pulses of 10s on and 10s off @ 20% amplitude) to fragment size ranging between 100 to 800 bp **(Figure 8.1)**. The resulting fragmented DNA was cleaned up using QIAquick columns (QIAGEN). The size distribution was checked by running aliquots of the samples on Agilent Bioanalyzer 7500 Nano chips. Subsequently, DNA was subjected to a series of enzymatic reactions that repair frayed ends, phosphorylate the fragments, and add a single nucleotide A overhang **(Figure 8.2)** using Illumina recommended reagents for multiplexed paired end library preparation. After ligating Illumina adaptors, ~300 bp fragments were size selected by gel electrophoresis and purified. Multiple PCR amplifications were performed for the ligated products (enrichment PCR) so as to obtain ~10 micrograms of amplicons per sample.

Pooled Gastric Cancer samples ***Pooled Normal Samples***

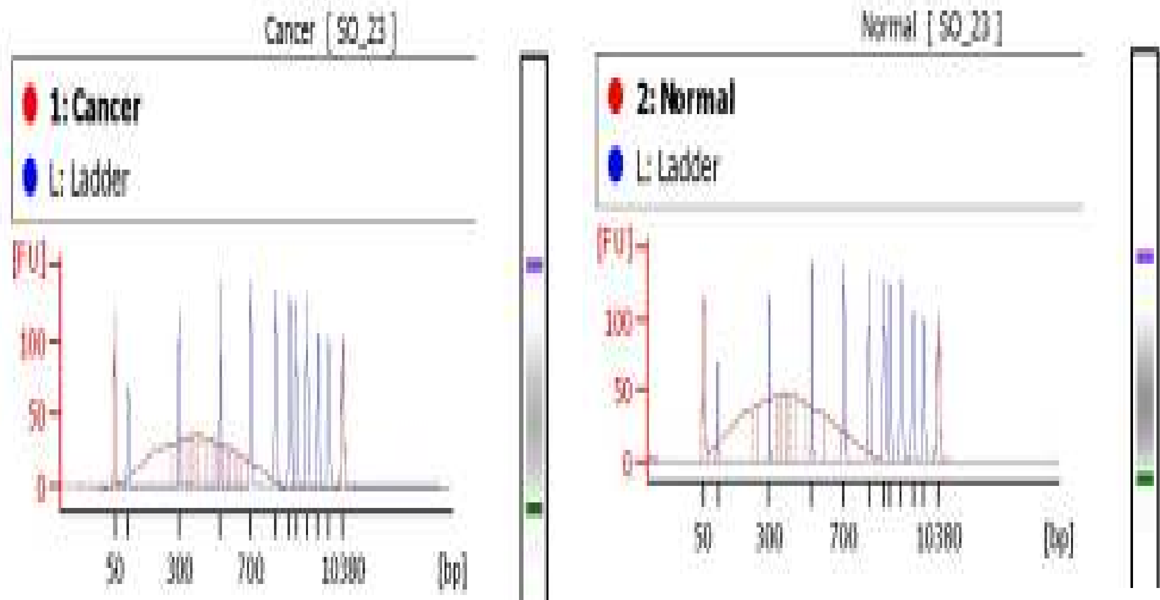


Figure 8. 1: Bioanalyzer profiles of sonicated samples

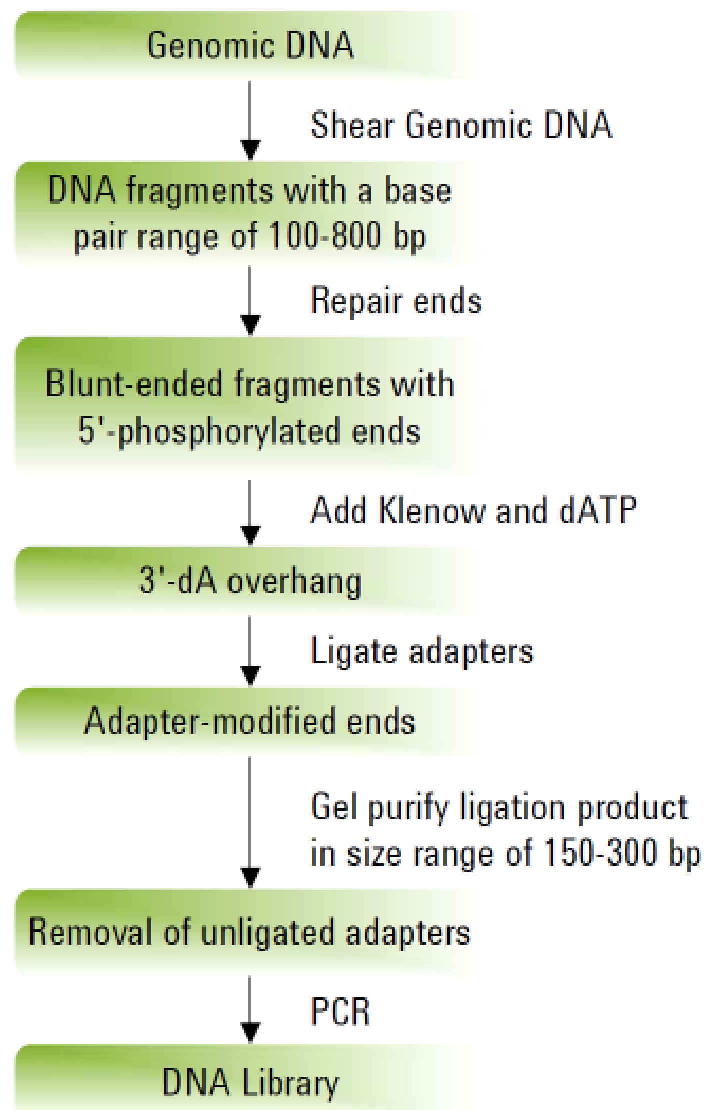


Figure 8.2: Process of library preparation

Hybridization and Elution:

The GC and normal libraries were each pooled in equal amounts to a total of 20 micrograms and hybridized on Agilent 244k Microarray (AMADID: 027271 and AMADID: 027271) following standard protocol recommended by Agilent (Hodges et al., 2009) for 65h at 65°C (Hodges et al., 2009). After standard washing procedures, the slides were reassembled with nuclease free water (Ambion) and exposed to high temperature (95°C for 10 min). DNA eluted in nuclease free water was recovered using a syringe. PCR was carried out with the eluted DNA in several replicates and cleaned up using QIAquick columns (QIAGEN).

Validation of target enrichment by quantitative PCR using targeted primers and non-targeted primers were performed. Comparison of ePCR1 (PCR amplification with common adapter primers) and ePCR2 products for target regions showed enrichment of target region. Captured samples showed early amplification over ePCR1 samples signifying enrichment. On the other hand the non-target regions were not detected in ePCR2 samples when compared with ePCR1.

Illumina Sequences:

Seventy two base pair single end reads were sequenced using Illumina GAIIx Analyzer ~25x coverage against the reference sequence. Quality control was perfromed for the Illumina sequencing using Genotypic's inhouse QC tool SeqQC. Using 20 phred quality for high quality cutoff, we found that 88-97% of the bases were of high quality and 88-97% of reads were of high quality across

the twenty samples. Only High Quality Reads were passed for further analysis. Human Genome (HG19) was used as reference sequence **(Figure 8.3)**. Sequence was downloaded from UCSC, *http://hgdownload.cse.ucsc.edu/goldenPath/hg19/.*

Gapped alignment:

Alignment software BWA version 0.5.7 was used to perform gapped alignment of Illumina sequences against the reference sequence. Parameters used for alignment were a maximum number of gap opens of 2 and a maximum number of gap extension of 10.

SNP Calling and Indel detection:

Samtools version 0.1.11 was used for calling SNPs. Parameters used for SNP calling were a minimum read depth of 5, maximum read depth of 255, minimum mapping quality of 20, minimum neighboring quality of 20, window size around potential indels of 5, maximum number of SNPs in a window of 2.

Indel detection:

Indels are detected using Samtools version 0.1.11 from the gapped alignment peformed using BWA3. Read depth at Indel bases >=8 and % of reads representing Indel >= 50 are the parameters used for Indel filtering.

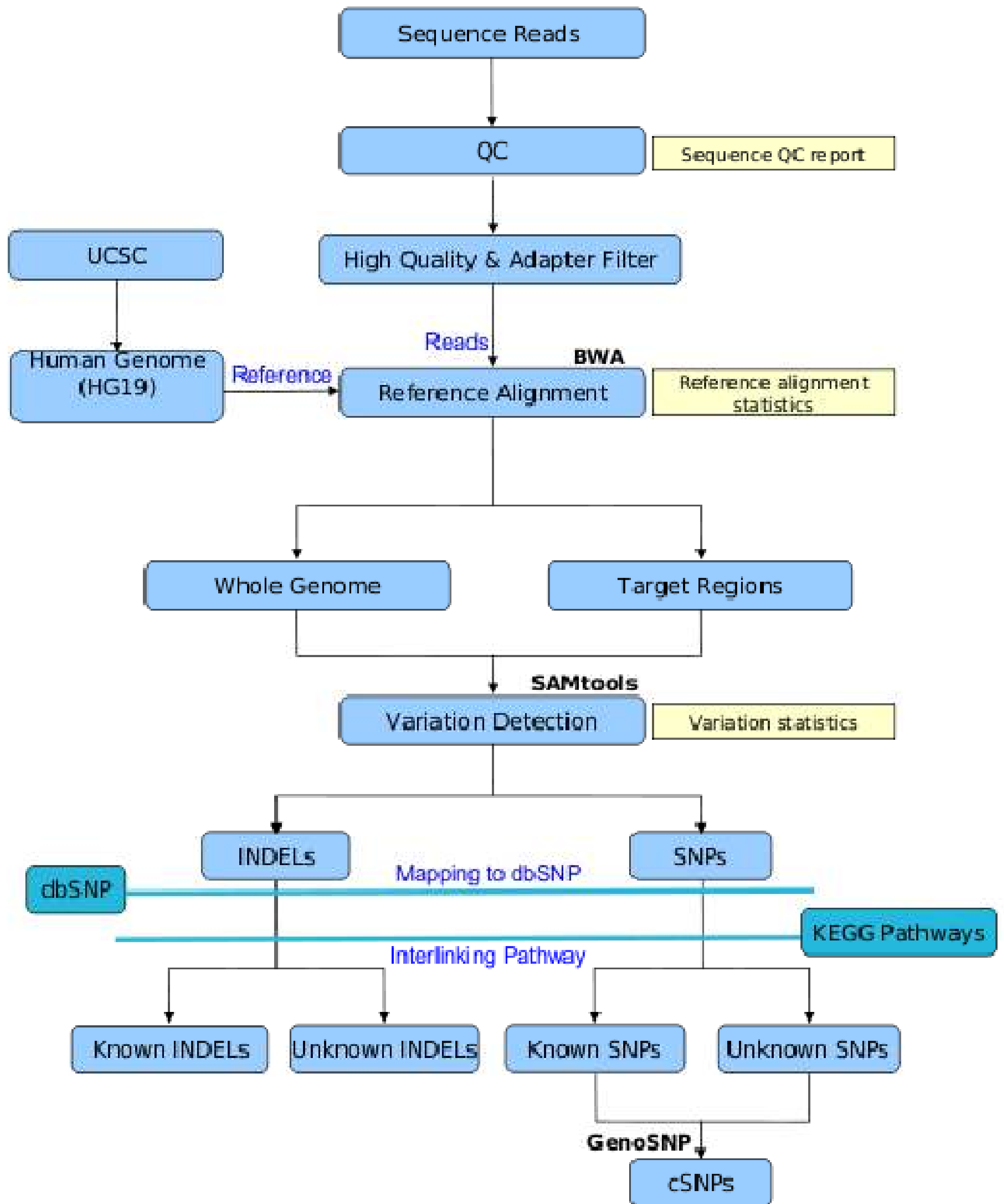


Figure 8.3: Data analysis flow for human targeted resequencing

Results

Clinical and epidemiological information

The mean age of the twenty patients included in the study was 59 ± 12.2 SD and males to females ratio was 2.6:1. None of the cases included in the study had first-degree relatives. Of the twenty samples included in the study, 54% (11 of 20) had well differentiated, 8% (2 of 20) had moderately differentiated and 38% (8 of 20) had poorly differentiated adenocarcinoma. The tumour location was as follows: 10 in cardia (48%), 4 in body (22%), 2 in GE junction (11%), 2 in pylorus (8%), 2 in fundus (11%). Of the 20 samples 69% (14 of 20) were tobacco chewers, 56% (11 of 20) were smokers, 54% (11 of 20) were betel quid chewers and 26% (5 of 20) were alcohol consumer **(Table 8.1)**.

Next generation sequencing

The total number of SNPs observed were 178 in cases whereas those of indels were 23 (12 insertion and 11 deletion). Of the 178 SNPs detected in normal samples 160 were known and 18 were unknown.

The *ACTL6A*, *PARK2*, *PTPN1*, *TP63*, *TSC2*, *VHL* and *NGFR* were the exclusive genes with SNPs only in cases among the 169 highly mutated genes choosen for the study **(Table 8.3)**. No SNPs were observed in these genes in the normal tissue samples.

Known SNPs

Further we have concentrated on variations that were detected on the cancer group in particular and not in the control group as they were of more importance. Knowns SNPs observed in cases only were at chromosomes no 1, 2,

3, 4, 6, 7, 10, 11, 13, 15, 16, 17, 18, 19 and 20 **(Figure 8.4, 8.6)** with dbSNP ID rs951715, rs4272, rs3743251, rs664143, rs538118, rs11640206, rs2287251, rs779804, rs17135764, rs1998291, rs16935840, rs7630340, rs309497, rs79072548, rs741071, rs1800935, rs2292044, rs3816885, rs41737, rs7628293, rs1122470, rs1132429, rs10932374, rs3733413 and rs6790167 targeting *IGF1R, CDK6, ATM, TGFA, BCAR1, CUL4A, VHL, TSC2, RASSF2, CUL2, DLEC1, EPHB2, ACTN4;CAPN12, NGFR, MSH6, DCC, FLNC, MET, MLF1, PARK2, ACTL6A, ERBB4, FAT1* and *TP63* genes **(Table 8.4)**. Of these rs664143, rs2287251, rs951715, rs309497, rs17135764, rs1998291 and rs6790167 were located at intron region. The rs79072548, rs779804, rs3743251, rs11640206, rs229204, rs538118, rs10932374, rs7628293, rs1122470, rs4272 and rs741071 were located at UTRs whereas the rest viz rs3816885, rs3733413, rs1132429, rs16935840, rs1800935, rs7630340 and rs41737 were located at coding regions.

Novel SNPs

Novel SNPs observed in cases only were at Chromosome no 1, 2, 4, 5, 6, 7, 11, 15, 17 and 22 with reference position at 18548356, 30061107, 32792687, 40352138, 41023714, 110908810, 112102905, 128488058, 152424347, 176516714, 176516720, 187549238, 202073970. Four of the SNPs were located in the coding region of *TSG101, APC, FLNC and CASP10* genes **(Table 8.5)** and the other SNPs were in the non-coding regions.

Known indels

Known indels observed in cases only were at chromosomes 2, 3, 7, 9, 17 and 22 **(Figure 8.5, 8.6)** with dbSNP ID rs67396500, rs17879991, rs16935545,

rs2234731, rs72220100, rs11274444, rs79930460 and rs66653070 involving *GAS7*, *CHEK2*, *FANCG*, *MSH6*, *MET*, *ERBB4*, *BARD1* and *DLG1*. Of these rs2234731, rs72220100, rs17879991 and rs79930460 were located at intron region and the rs67396500, rs11274444, rs66653070 and rs16935545 were located at UTRs whereas none of them were located at coding region **(Table 8.6)**.

Novel indels

The novel indels observed in case group alone was at chromosomes 11 with ref Seq ID NM_003682 involving MADD gene **(Table 8.6)**. It is an insertion of guanine base at position 47291817.

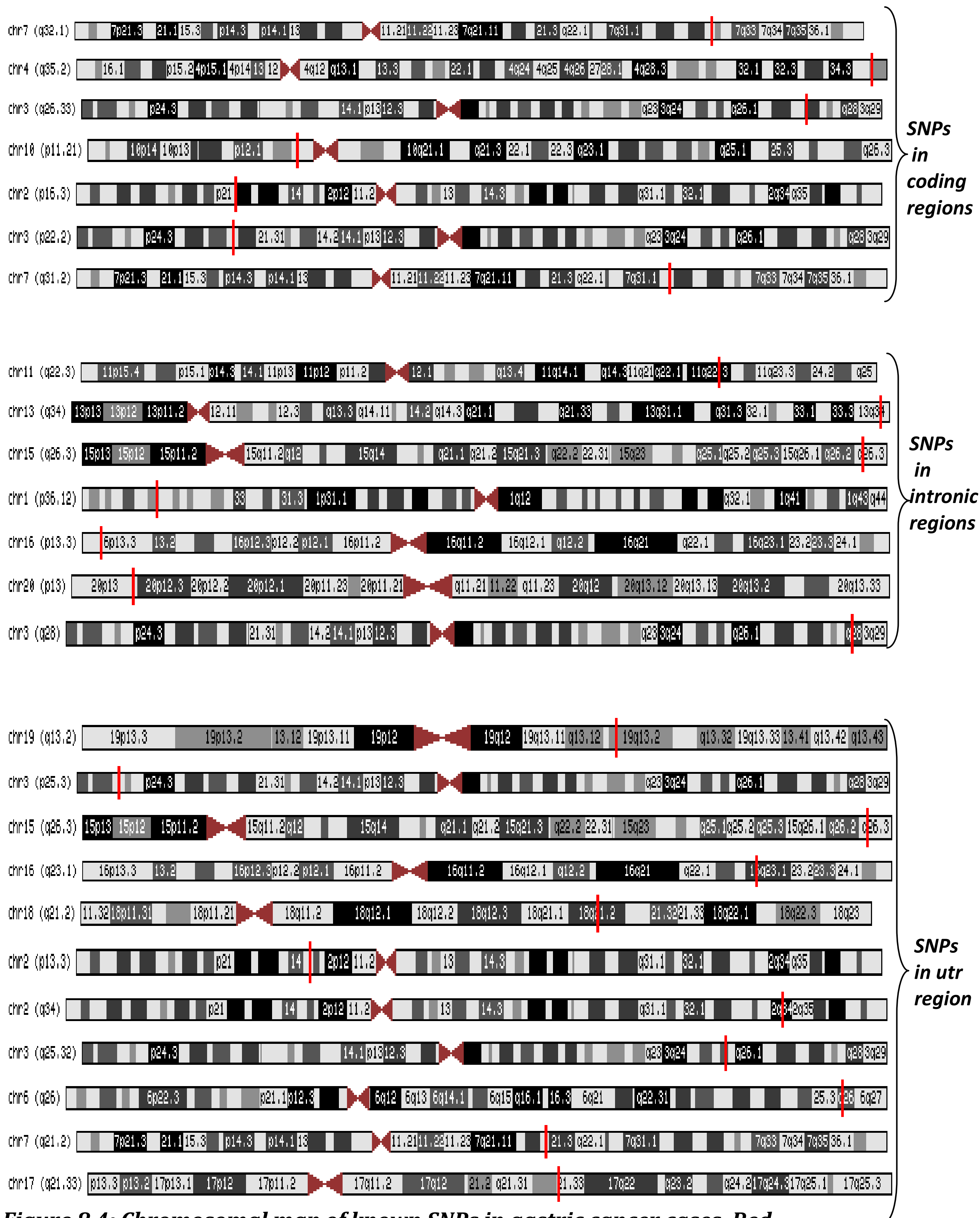


Figure 8.4: Chromosomal map of known SNPs in gastric cancer cases. Red colour indicate the location of the SNPs

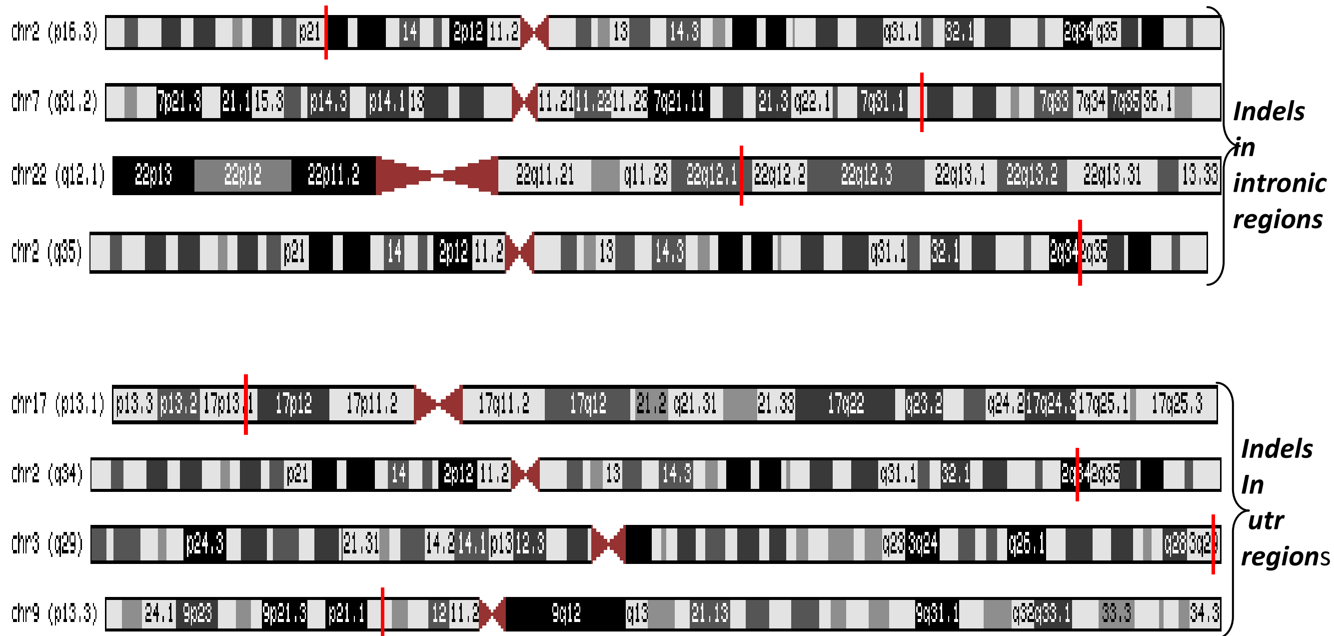


Figure 8.5: Chromosomal map of known indels in gastric cancer cases. Red colour indicate the location of the SNPs

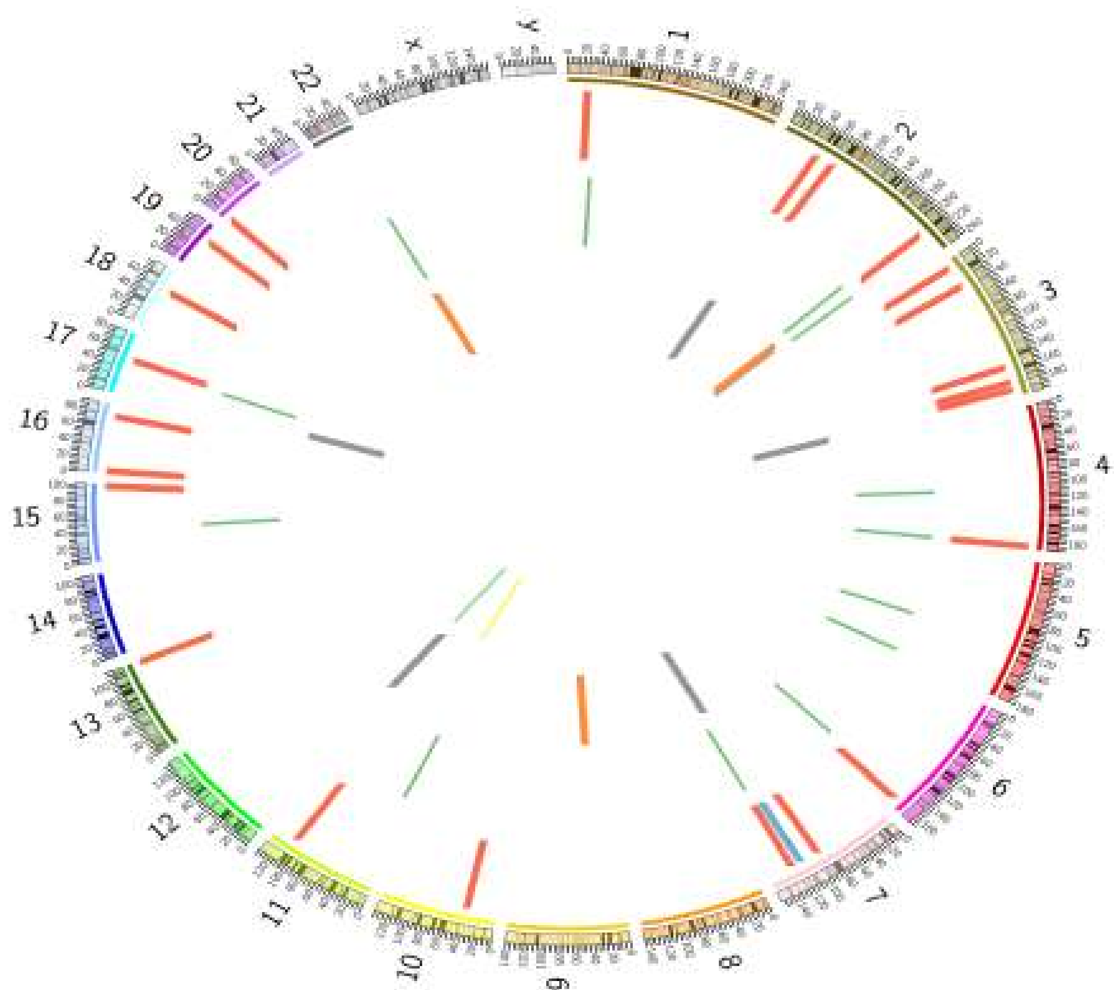


Figure 8.6: Figurative representation of the catalogue of mutations in gastric cancer samples. Chromosome ideograms are shown around the outer ring and are oriented pter–qter in a clockwise direction with centromeres indicated in red. Other tracks contain alterations (from outside to inside): Known SNP: heterozygous(red); Novel SNP: heterozygous(green); Known indel: deletion(grey), insertion(orange); Novel indel: insertion(yellow). (Source:Circos, http//www.circos.ca)

Table 8.1: Demographic and clinical characteristics of gastric cancer

FACTORS	CATEGORIES	CASES
Sex	Male	13
	Female	7
Histopathology(adenocarcinoma)		
	well differentiated	11
	moderately differentiated	2
	poorly differentiated	8
Tumor location	Cardia	10
	Body	4
	GE junction	2
	Pylorus	2
	Fundus	2
Smoking status	Non-smokers	9
	Smokers	11
Tobacco chewing	Non-chewers	6
	Chewers	14
Betel quid chewing	Non chewers	9
	Chewers	11
Alcohol consumption	Non-alcoholic	15
	Alcoholic	5

Table 8.2: Total number of SNPs and indels in gastric cancer cases

	Gastric Cancer
Total SNPs	178
Known among them	160
Unhnkown	18
Inserts	12
Deletions	11
Total indels	23
Known among them	22
Unknown	1

Table 8.3: List of genes with SNPs in gastric cancer

Gene Symbol	SNPs	Chromosome	Position	Read Depth	Zygosity	rsID	Type of variation	Amino acid
ACTL6A	A/C/G/T	chr3	179298999	50	Heterozygous	rs1132429	Coding non-synonymous	E; D
NGFR	C/T	chr17	47591690	37	Heterozygous	rs741071	utr-3	NA
PARK2	C/T	chr6	161769110	47	Heterozygous	rs1122470	utr-3	NA
TSC2	C/T	chr16	2111779	45	Heterozygous	rs17135764	Intron	NA
TP63	A/G	chr3	189587274	10	Heterozygous	rs6790167	Intron	NA
PTPN1	C/T	chr20	49195248	9	Heterozygous	rs968701	Intron	NA
VHL	C/T	chr3	10183274	5	Homozygous	rs779804	utr-5	NA

Table 8.4: Known SNPs in gastric cancer cases

Chromosome	Position	SNP	Read Depth	SNP Ratio	DBSNP ID	Gene Symbol	Function class	Type of variation
chr7	128495338	Y	45	C(29)/T(16)	rs3816885	*FLNC*	coding	Synonymous
chr4	187629538	Y	71	C(46)/T(25)	rs3733413	*FAT1*	coding	non-synonymous
chr3	179298999	Y	50	T(36)/C(14)	rs1132429	*ACTL6A*	coding	Synonymous
chr10	35321414	Y	13	C(9)/T(4)	rs16935840	*CUL2*	coding	Synonymous
chr2	48023115	Y	31	T(26)/C(5)	rs1800935	*MSH6*	coding	Synonymous
chr3	38087123	Y	45	T(34)/C(11)	rs7630340	*DLEC1*	coding	Synonymous
chr7	116436097	R	44	A(23)/G(21)	rs41737	*MET*	coding	Synonymous
chr11	108225661	R	5	G(4)/A(1)	rs664143	*ATM*	intron	NA
chr13	113889499	R	35	A(22)/G(13)	rs2287251	*CUL4A*	intron	NA
chr15	99456553	R	26	A(20)/G(6)	rs951715	*IGF1R*	intron	NA
chr1	23236826	Y	64	C(44)/T(20)	rs309497	*EPHB2*	intron	NA
chr16	2111779	Y	45	C(42)/T(3)	rs17135764	*TSC2*	intron	NA
chr20	4773338	R	22	A(14)/G(8)	rs1998291	*RASSF2*	intron	NA
chr3	189587274	R	10	G(6)/A(3)	rs6790167	*TP63*	intron	NA
chr19	39221183	K	12	T(6)/G(6)	rs79072548	*ACTN4;CAPN2*	utr-3	NA
chr3	10183274	T	5	T(5)	rs779804	*VHL*	utr-5	NA
chr15	99504129	R	17	G(11)/A(6)	rs3743251	*IGF1R*	utr-3	NA
chr16	75262929	R	7	A(4)/G(3)	rs11640206	*BCAR1*	utr-3	NA
chr18	51059341	S	7	C(4)/G(3)	rs2292044	*DCC*	utr-3	NA
chr2	70676639	R	10	A(6)/G(4)	rs538118	*TGFA*	utr-3	NA
chr2	212244403	R	18	G(10)/A(8)	rs10932374	*ERBB4*	utr-3	NA
chr3	158323665	R	15	A(10)/G(5)	rs7628293	*MLF1*	utr-3	NA
chr6	161769110	Y	47	C(33)/T(14)	rs1122470	*PARK2*	utr-3	NA
chr7	92236829	R	32	G(20)/A(12)	rs4272	*CDK6*	utr-3	NA
chr17	47591690	Y	37	C(24)/T(13)	rs741071	*FLNC, NGFR*	utr-3	NA

Table 8.5: Novel SNPs in gastric cancer cases

Chromosome	Position	SNP	Read Depth	SNP Ratio	Gene Symbol	Changed AA residue	Type of variation	Codon	Variation Codon	RefSeq ID
chr11	18548356	M	5	A(3)/C(2)	*TSG101*	E,A	Non-Synonymous	GTG	GMG	NM_006292
chr22	30061107	S	8	G(7)/C(1)				NA	NA	Non-coding
chr1	32792687	R	39	G(33)/A(6)				NA	NA	Non-coding
chr17	40352138	Y	7	C(6)/T(1)				NA	NA	Non-coding
chr15	41023714	K	8	T(7)/G(1)				NA	NA	Non-coding
chr4	110908810	Y	10	T(8)/C(2)				NA	NA	Non-coding
chr5	112102905	Y	5	T(4)/C(1)	*APC*	S,S	Synonymous	AGT	AGY	NM_000038 NM_001127510 NM_001127511
chr7	128488058	M	22	A(12)/C(10)	*FLNC*	T,P	Non-Synonymous	ACC	MCC	NM_001127487 NM_001458
chr6	152424347	R	5	G(4)/A(1)				NA	NA	Non-coding
chr5	176516714	R	11	G(7)/A(4)				NA	NA	Non-coding
chr5	176516720	K	7	G(4)/T(3)				NA	NA	Non-coding
chr4	187549238	M	6	A(5)/C(1)				NA	NA	Non-coding
chr2	202073970	M	17	C(15)/A(2)	*CASP10*	Stop,S	Non-Synonymous	TCG	TMG	NM_001230 NM_032974 NM_032977
chr2	225335540	R	7	A(6)/G(1)				NA	NA	Non-coding

Gene symbol *signifies those genes where these SNPs are located

Table 8.6: Known and novel indels in gastric cancer cases (Excluding those common to controls)

Reference Name	Reference Position	Type of variation	InDels	Read Depth	DBSNP ID	Gene Symbol	Function class
Known indels							
chr2	48032875	Deletion	CTAT	6	rs2234731	***MSH6***	Intron
chr7	116409676	Deletion	T	10	rs72220100	***MET***	Intron
chr22	29130347	Insertion	T	7	rs17879991	***CHEK2***	Intron
chr2	215632125	Deletion	TT	10	rs79930460	***BARD1***	Intron
chr17	9813903	Deletion	A	8	rs67396500	***GAS7***	utr-3
chr2	212245090	Insertion	TGAAAATAGGAT	8	rs11274444	***ERBB4***	utr-3
chr3	196770356	Deletion	AAGG	6	rs66653070	***DLG1***	utr-3
chr9	35079973	Insertion	A	7	rs16935545	***FANCG***	utr-5
Novel indels							
chr11	47291817	Insertion	G	21	NA	***MADD***	Intron

Table 8.7: GO analysis of genes identified in gastric cancer

Gene Symbol	GO functions
ACTL6A	GO:0043968~histone H2A acetylation; development; GO:0045449~regulation of transcription; GO:0040008~regulation of growth; GO:0007165~signal transduction; GO:0006338~chromatin remodeling; GO:0043967~histone H4 acetylation
ACTN4	GO:0015031~protein transport; GO:0001666~response to hypoxia; GO:0051017~actin filament bundle assembly; GO:0051272~positive regulation of cellular component movement; GO:0042981~regulation of apoptosis; GO:0051271~negative regulation of cellular component movement
APC	GO:0031116~positive regulation of microtubule polymerization; GO:0007094~mitotic cell cycle spindle assembly checkpoint;; GO:0006915~apoptosis; GO:0009954~proximal/distal pattern formation; GO:0051781~positive regulation of cell division; GO:0045785~positive regulation of cell adhesion; GO:0051726~regulation of cell cycle; GO:0060070~canonical Wnt receptor signaling pathway; GO:0008285~negative regulation of cell proliferation; GO:0035019~somatic stem cell maintenance; GO:0042493~response to drug; GO:0007155~cell adhesion; GO:0030335~positive regulation of cell migration; GO:0030858~positive regulation of epithelial cell differentiation
ATM	GO:0043065~positive regulation of apoptosis; GO:0007165~signal transduction; GO:0007094~mitotic cell cycle spindle assembly checkpoint; GO:0007131~reciprocal meiotic recombination; GO:0000075~cell cycle checkpoint; GO:0043066~negative regulation of apoptosis; GO:0000724~double-strand break repair via homologous recombination; GO:0006975~DNA damage induced protein phosphorylation; GO:0008630~DNA damage response; GO:0008219~cell death; GO:0031572~G2/M transition DNA damage checkpoint; GO:0007050~cell cycle arrest

BARD1	GO:0006974~response to DNA damage stimulus; GO:0001894~tissue homeostasis; GO:0043066~negative regulation of apoptosis; GO:0042325~regulation of phosphorylation; GO:0006281~DNA repair; GO:0031441~negative regulation of mRNA 3'-end processing; GO:0043065~positive regulation of apoptosis; GO:0007050~cell cycle arrest; protein export from nucleus
BCAR1	GO:0051301~cell division; GO:0007155~cell adhesion; GO:0008283~cell proliferation; GO:0007015~actin filament organization; GO:0007229~integrin-mediated signaling pathway; GO:0008286~insulin receptor signaling pathway; GO:0030335~positive regulation of cell migration; GO:0007173~epidermal growth factor receptor signaling pathway; GO:0050853~B cell receptor signaling pathway; GO:0042981~regulation of apoptosis; GO:0001558~regulation of cell growth; GO:0050852~T cell receptor signaling pathway; GO:0016477~cell migration
CASP10	GO:0008624~induction of apoptosis by extracellular signals; GO:0042981~regulation of apoptosis; GO:0006917~induction of apoptosis; GO:0043123~positive regulation of I-kappaB kinase/NF-kappaB cascade; GO:0006915~apoptosis
CDK6	GO:0048146~positive regulation of fibroblast proliferation; GO:0051301~cell division; GO:0000278~mitotic cell cycle; GO:0000080~G1 phase of mitotic cell cycle; GO:0043697~cell dedifferentiation; GO:0045786~negative regulation of cell cycle; GO:0001954~positive regulation of cell-matrix adhesion; GO:0050680~negative regulation of epithelial cell proliferation
CHEK2	GO:0000077~DNA damage checkpoint; GO:0008630~DNA damage response, signal transduction resulting in induction of apoptosis; GO:0006974~response to DNA damage stimulus; GO:0007049~cell cycle; GO:0000075~cell cycle checkpoint
CUL2	GO:0030163~protein catabolic process; GO:0000082~G1/S transition of mitotic cell cycle; GO:0008285~negative regulation of cell proliferation; GO:0008629~induction of apoptosis by intracellular signals; GO:0007050~cell cycle arrest

CUL4A	GO:0007050~cell cycle arrest; GO:0006281~DNA repair; GO:0008285~negative regulation of cell proliferation; GO:0000082~G1/S transition of mitotic cell cycle; GO:0008629~induction of apoptosis by intracellular signals
DCC	GO:0006917~induction of apoptosis; GO:0001975~response to amphetamine; GO:0042981~regulation of apoptosis; GO:0070374~positive regulation of ERK1 and ERK2 cascade; GO:0006915~apoptosis; GO:0001764~neuron migration
DLEC1	GO:0008285~negative regulation of cell proliferation
DLG1	GO:0001935~endothelial cell proliferation; GO:0007015~actin filament organization; GO:0016337~cell-cell adhesion; GO:0030866~cortical actin cytoskeleton organization; GO:0031575~mitotic cell cycle G1/S transition checkpoint; GO:0045930~negative regulation of mitotic cell cycle
EPHB2	GO:0000902~cell morphogenesis; GO:0048013~ephrin receptor signaling pathway; GO:0007399~nervous system development; GO:0016310~phosphorylation;
ERBB4	GO:0007169~transmembrane receptor protein tyrosine kinase signaling pathway; GO:0030335~positive regulation of cell migration; GO:0008283~cell proliferation; GO:0045449~regulation of transcription; GO:0007165~signal transduction; GO:0045768~positive regulation of anti-apoptosis; GO:0050679~positive regulation of epithelial cell proliferation; GO:0042542~response to hydrogen peroxide;
FANCG	GO:0009314~response to radiation; GO:0006281~DNA repair; GO:0000075~cell cycle checkpoint
FAT1	GO:0007155~cell adhesion; GO:0007267~cell-cell signaling; GO:0009653~anatomical structure morphogenesis; GO:0007156~homophilic cell adhesion
FLNC	GO:0034329~cell junction assembly

GAS7	GO:0007050~cell cycle arrest; GO:0030041~actin filament polymerization; GO:0051017~actin filament bundle assembly
IGF1R	GO:0007165~signal transduction; GO:0007409~axonogenesis; GO:0043409~negative regulation of MAPKKK cascade; GO:0030335~positive regulation of cell migration; GO:0051898~negative regulation of protein kinase B signaling cascade; GO:0048009~insulin-like growth factor receptor signaling pathway; GO:0006955~immune response; GO:0008284~positive regulation of cell proliferation; GO:0043410~positive regulation of MAPKKK cascade; GO:0045768~positive regulation of anti-apoptosis; GO:0051897~positive regulation of protein kinase B signaling cascade; GO:0032467~positive regulation of cytokinesis
MET	GO:0008283~cell proliferation; GO:0007166~cell surface receptor linked signaling pathway; GO:0006468~protein amino acid phosphorylation; GO:0007165~signal transduction
MLF1	GO:0007049~cell cycle; GO:0002318~myeloid progenitor cell differentiation; GO:0006350~transcription; GO:0030154~cell differentiation; GO:0007050~cell cycle arrest
MSH6	GO:0009411~response to UV; GO:0008630~DNA damage response, signal transduction resulting in induction of apoptosis; GO:0006298~mismatch repair; GO:0000710~meiotic mismatch repair; GO:0045910~negative regulation of DNA recombination; GO:0043570~maintenance of DNA repeat elements; GO:0016447~somatic recombination of immunoglobulin gene segments; GO:0006281~DNA repair; GO:0008629~induction of apoptosis by intracellular signals
PARK2	GO:0043123~positive regulation of I-kappaB kinase/NF-kappaB cascade; GO:0001933~negative regulation of protein amino acid phosphorylation; GO:0060548~negative regulation of cell death; GO:0019538~protein metabolic process; GO:0090201~negative regulation of release of cytochrome c from mitochondria
RASSF2	GO:0007049~cell cycle; GO:0007165~signal transduction

TGFA	GO:0008283~cell proliferation; GO:0051781~positive regulation of cell division; GO:0045741~positive regulation of epidermal growth factor receptor activity; GO:0045840~positive regulation of mitosis; GO:0050679~positive regulation of epithelial cell proliferation; GO:0000187~activation of MAPK activity
TP63	GO:0006978~DNA damage response, signal transduction by p53 class mediator resulting in transcription of p21 class mediator; GO:0034644~cellular response to UV; GO:0006915~apoptosis; GO:0030308~negative regulation of cell growth; GO:0002347~response to tumor cell; GO:0043523~regulation of neuron apoptosis; GO:0007050~cell cycle arrest; GO:0045747~positive regulation of Notch signaling pathway; GO:0006916~anti-apoptosis; GO:0006917~induction of apoptosis; GO:0031571~mitotic cell cycle G1/S DNA damage checkpoint
TSC2	GO:0007050~cell cycle arrest; GO:0032007~negative regulation of TOR signaling cascade; GO:0051893~regulation of focal adhesion assembly; GO:0043407~negative regulation of MAP kinase activity; GOGO:0032956~regulation of actin cytoskeleton organization; GO:0030178~negative regulation of Wnt receptor signaling pathway; GO:0051726~regulation of cell cycle; kinase B signaling cascade; GO:0016192~vesicle-mediated transport; GO:0023034~intracellular signaling pathway
TSG101	GO:0043687~post-translational protein modification; GO:0008285~negative regulation of cell proliferation; GO:0045892~negative regulation of transcription, GO:0007049~cell cycle; GO:0001558~regulation of cell growth
VHL	GO:0050821~protein stabilization; GO:0006916~anti-apoptosis; GO:0000902~cell morphogenesis; GO:0008285~negative regulation of cell proliferation; GO:0043534~blood vessel endothelial cell migration; GO:0045449~regulation of transcription; GO:0001525~angiogenesis; GO:0001666~response to hypoxia; GO:0045597~positive regulation of cell differentiation

Discussion

NGS is a rapid and robust assay that allows the simultaneous screening of genomic alteration, DNA copy number alterations and LOH of the entire genome. This technology can provide valuable informations for the identification of pathobiologically relevant cancer genes. We have used this comprehensive approach in a panel of 20 gastric adenocarcinoma and 20 controls tissues. The results of the study expand the view of the genomic complexity of gastric tumors with the recoznization of known and novel SNPs, deletions and insertions.

Our NGS analysis has detected genes with SNPs in GC group alone and not in the normal tissue samples **(Table 8.3)**. The SNPs in these genes can serve as biomarkers for diagnosis of GC and are of potential for therapeutic intervention. Among them the SNPs rs1132429 targeting *ACTL6A* is the only SNP located in the coding region, the dbSNP allele of this SNP was A/C/T. The T allele is the ancestral allele and it is located in GAT codon and code for Asp (D) but when it is substituted with A/C results in a non-synonymous variation which code for Glu (E). The read depth observed during SNP calling for this SNP is 50 suggesting rs1132429 to be of high importance for the pathogenesis of GC in this region. So far no clinical data of this SNP is reported in patients or controls.

Although introns have sometimes been loosely called "junk DNA," the fact that they are so common and have been preserved during evolution leads many researchers to believe that they serve some function. Interestingly, several recent studies have also shown that SNPs located in the intronic regions served significant functions [503]. Of the known SNPs, rs951715, rs664143, rs2287251,

rs17135764, rs1998291, rs309497 and rs6790167 were found in the intronic regions **(Table 8.4)**. The SNPs in this region targeted four known tumor suppressor genes (*ATM*, *TSC2*, *RASSF2, EPHB2*). Mutations in the tumor suppressor genes *TSC2* and *ATM* caused Tuberous sclerosis and Ataxia telangiectasia. Tuberous sclerosis is a disorder that causes tumors formation in many organs [504]. Likewise, approximately one third of Ataxia telangiectasia patients are reported to develop cancer [505]. Several studies have been carried out illustrating the ability of TSC2 to stimulate specific GTPases and regulate mTOR signaling pathway by forming complex with proto-oncogene Rheb [504, 506, 507]. However, so far no studies have been carried out examining its role in specific cancer. The tumor suppressor gene *ATM*, codes for protein that plays a critical role in cell cycle arrest, apoptosis, and DNA repair. Studies have reported a high prevalence of *ATM* gene alterations in breast, brain, leukemia and lymphoid tumours adding evidence to the postulated contribution of *ATM* in the pathogenesis of these tumours [508-512]. Conversely, for the SNP (rs664143) observed in *ATM* gene, Lee SA et al has reported no association of it with the risk of breast cancer [513], though studies by Malmer BS et al showed that haplotype including rs664143 to be higher in frequency in meningioma cases as compare to control suggesting of a positive association [509]. Inspite of the role of *ATM* in various cancers, studies of its role in GC is lacking. Considering these facts it seems that association studies of rs664143 polymorphism detected in our with the risk of gastric will provide new data that could add to our understanding of GC susceptibility. Also *RASSF2*, a member of the *RASSF1* family is a potential

tumour suppressor and has been reported to be specificantly hypermethylated at the transcription start site in GC [514] and its tumor suppressive activity has been extended to other cancers such as those of colorectal [515], nasopharyngeal [516], endometrial [517] and lung [518]. Likewise, *EPHB2*, has been reported to be a tumor suppressor gene in colorectal cancer [519]. Inactivation of it has been shown to correlate with progression of colorectal tumorigenesis [519] and somatic mutations have been reported in both colorectal and prostate tumors [519]. This gene is localized at 1p35-p36.1, a frequently deleted region in colon and other cancer. Contrastingly its overexpression is also reported in several gastrointestinal tumors [520]. Thus suggesting of a conflicting role. Moreover, increase expression of Eph receptor tyrosine kinases and their ephrin ligands has been implicated in tumor progression in a number of malignancies [521]. Mutational study by Davalos V et al has further particularize the role of *EPHB2* in gastric tumor progression [522]. Thus studies of theses tumor suppressor genes in relation to GC will be of high implication.

Beside these tumor suppressor genes three other genes (*IGF1R*, CUL4A, *TP63*) were targeted by the SNPs of the intronic region. *IGF1R* is a potent mitogen and postulated to exert autocrine and paracrine effects on growth regulation in human gastric cancer [523]. Clinical data of the rs951715 observed in our study which is located in this gene was available and reported to be associated with the risk of a subgroup of breast cancer [524].

The genes i,e *CUL4A* and *TP63* also play role in aspects relevant to the process of carcinogenesis. CUL4A is an E3 ubiquitin ligase that has been shown to

be involved in the DNA damage response and it mediates activation p16INK4A during oncogenic checkpoint response [525, 526]. As for *TP63*, it is a member of the *TP53* gene family. In contrast with *TP53*, this gene is not frequently inactivated by mutation in cancer [527, 528]. Nevertheless, there is growing evidence of the involment of *TP63* in oncogenesis through several mechanisms [529]. Therefore, we see that the genes targeted by the SNPs of the intronic regions are more or less involved in functions applicable to carcinogenesis but beside rs664143 and rs951715 located in the *ATM* and *IGF1R* genes at chr11 and chr15 with reference positions 108225661 and 99456553, no other variants in the intronic regions have clinical report on patients and controls. However because of the functional characteristic of the targeted genes studying them is pertinent with their role in GC.

The function of 3'UTRs and 5'UTRs is still rather poorly studied still there is evidence that elements in 3'UTRs and 5'UTRs control post-transcriptional gene regulation, efficiency of expression and the stability of mRNAs [530-532]. At the same time reports from several studies showed association of polymorphisms in 3'UTR as well as 5'UTR regions with the risk of GC [533-536]. The known SNPs bearing the dbSNP ID rs3743251, rs11640206, rs2292044, rs538118, rs10932374, rs7628293, rs1122470, rs4272, rs741071, rs79072548, rs779804 were located in the UTR regions. Of these only one SNP targeting the *VHL* gene was located in the 5'UTR other were located at 3'UTRs **(Table 8.4)**. *VHL*, tumor-suppressor gene mutation leads to VHL syndrome a dominantly inherited familial

cancer syndrome predisposing to a variety of malignant and benign tumors [537].

Beside *VHL* in 5'UTR, the other genes targeted by the SNPs in 3'UTR regions were *PARK2, CDK6, IGF1R, BCAR1, DCC, TGFA, ERBB4, MLF1, NGFR, ACTN4* and *CAPN12.* Although the molecular functions of these genes did not provide informations of the variants characteristic to the predisposition of GC in particular as none of the SNPs in the UTRs have previous data on clinical samples. However, GO functional analysis of the genes have revealed their relevant in the process of carcinogenesis **(Table 8.7)**.

In addition to the variations in the intronic and UTRs, seven of the known SNPs were located in the coding regions of various genes. Among *FLNC, FAT1, ACTL6A, CUL2, MSH6, DLEC1* and *MET* which were targeted by these SNPs, the variation in *FAT1* were of high importance leading to non synonymous changes of amino acid at codon position 482 whereas those in the others genes were of synonymous type. The functional aspect of these particular codons is not known, however changes in the amino acid sequence will modify the functioning of these gene and hamper the normal performance of the cells as *FAT1* is a well known tumor suppressor gene. Reports were available of *FAT1 i*n various other cancers [538-540] but not with GC. Nevertheless, our findings suggest the possibility of the existence of its role in the development of GC.

Of the novel SNPs observed , that at chr11 with reference position 18548356, chr5 with reference position 112102905, chr7 with reference position 128488058 and at chr7 with reference position 128488058 were found to be

located in the coding region of *APC, TSG101, FLNC* and *CASP10* respectively **(Table 8.5)**. APC is a well known tumor suppressor genes. It is one of the extensively studied gene and defects in this gene is confirmed to cause familial adenomatous polyposis in various studies [541-544]. *APC* mutation is reported to be involved in carcinogenesis of GC [545, 546]. Beside gastric, mutations and polymorphisms in *APC* is associated with various other cancers such as that of colon, breast, prostate, thyroid, pancreatic, melanoma and epithelial odontogenic tumors [547-554]. Promoter hypermethylation of *APC* is observed in cancers of the breast, esophagous, prostate, lung, colorectal, endometrial and gastric [555-560]. Inspite of its relevant functions, paradoxically, the novel SNP observed in *APC* coding region is synonymous and hence code for the same amino acid S, thus might have less functional significance unlike other genes viz *TSG101, FLNC* and *CASP10* that have a non-synonymous type of variation. The gene product of *TSG101* contains a coiled-coil domain that interacts with stathmin, a cytosolic phosphoprotein implicated in tumorigenesis. The variation observed alter a codon of *TSG101* from GTG which code for amino acid E to GMG which code for A. The change in the amino acid coded have towering chances of structural alteration leading to significant defect in the functional aspect of the *TSG101* protein product. Moreover, high frequency of mutations in it have been reported in breast cancer. Unlikely, the gene *FLNC* has not been associated with any of the cancer in earlier studies, though this gene codes for filamin proteins that crosslink actin filaments into orthogonal networks in cortical cytoplasm and participate in the anchoring of membrane proteins for the actin cytoskeleton and

as a result perceptibly is of high structural importance and modification of amino acid sequence will alter its function to a great magnitude thereby affecting its normal role and hence have high possibility to play significant part in cancer mechanism. The novel SNP detected in *FLNC* lead to change of the codon from ACC to MCC altering the amino acid coded by it from T to P. Also, the novel SNPs in *CASP10* observed in our study bring above a stop codon leading to truncated protein product. *CASP10* have been reported in association with GC in previous studies and it is well known that activation of caspases play central role in the execution phase of cell apoptosis. Apoptosis being a major process concerned with tumorigenesis, *CASP10* has a clear position in cancer related cellular process although not precise for GC.

All the known indels detected in our studies were located at introns and UTR regions at chr 2, 3, 7, 9, 17 and 22. No citation results were obtained for the indels observed and noticeably association studies of these indels with GC is lacking. The novel indels detected exclusively in cases was located in chr11 and targeted the *MADD* and it is an insertion of guanine base at position 47291817 **(Table 8.6)**. GO analysis of *MADD* has revealed it involvement in GO:0042981~regulation of apoptosis; GO:0007166~cell surface receptor linked signaling pathway; GO:0000187~activation of MAPK activity and GO:0051726~regulation of cell cycle. These GO biological functions are of significant importance in the development of cancer. Compared to studies on other cancers the rate of mutations observed in our study were low.

Concordantly studies carried by Liang Goh et al on GC samples have also revealed a low rate of mutation in GC [327].

The NGS data have thus revealed a limited number of novel and known genomic alterations in patients with gastric tumors in this population and provided a useful information which narrow down the search for biomarkers. In addition to the novel alterations, majority of the known SNPs observed in our study lacked clinical data and as such exploring the role of these SNPs in GC may disclosed new insight in the functional aspect of the genes in relation to the process of gastric carcinogenesis. Further, it is to be mention that NGS can served as a potential technology for reporting rare mutations in various cancers, many previously undetected. To our knowledge this is the first data on NGS involving GC samples in Northeast population of India where the incidence is reported to be very high. These findings illustrate the potential for NGS to provide unprecedented insights into mutational processes underlying GC. It has produced valuable information concerning the molecular and cellular biology of GC and provided a tool to investigate the process of carcinogenesis. Although whole genome sequencing is currently more expensive than CT scanning, given the falling costs and development of ever faster and cheaper NGS technologies, this situation is likely to be reversed within the next few years. This method represents an important foray into offering truly personalised medicine. Moreover, because it is based on the analysis of somatic genetic changes, rather than an individual's inherited genetic code, many of the complex ethical, legal and social issues raised by personal genome profiling are avoided.

Chapter 9: Conclusion and Future Scope

Although our results do not show consistent significant association of the candidate gene polymorphism in the risk of GC, we cannot totally rule out the possibility of GSTs gene effect on the carcinogenesis of GC as they are important with the detoxification of tobacco related carcinogens. Moreover a significant association of *GSTT1* null genotype in Assam population with GC is observed. Notably one of the major advances in our understanding of the risk factors and *H. pylori* interaction is that the bacterial factors are influence by the risk factors such as tobacco, betelquid and alcohol in determining disease outcome, it is also the attempts of the environmental and genetic system of the host that influence the chronic infection with *H. pylori* that might have elicited an overly aggressive response resulting in a series of morphological changes that lead to cancer. Pathways identified in our study i,e wnt, notch and hedgehog and genes involved in different cellular processes such as proliferation, inflammation, cell adhesion and apoptosis may have their specific important role during the process constituting the genetic decomposition of the host. The mechanism of GC carcinogenesis in this region might involve other genetic effect such as those in *ACTL6A, PARK2, PTPN1, TP63, TSC2, VHL*, *NGFR*, *FAT1*, *APC*, *TSG101, FLNC, CASP10* and *MADD* as evident in NGS study. These supported the multiplicative risk factor hypothesis of gastric carcinogenesis. It has been speculated that many risk factors such as *H. pylori* and environmental factors may work in part through the genetic factors. Intervention to eradicate *H. pylori* and control over betelquid and

tobacco consumption early in this pathological process might prevent progressively more severe disease.

The development of bioinformatics in biological research is developing rapidly; and the amplification procedures could be pushed further. The identification of the linked between abnormal gene expression and mutation in diseased tissues will lead to the discovery of new drug targets as well as advances in diagnosis. If in the future these elements can be successfully integrated, it's a realistic possibility that we will be able to create a window through which to view gene expression and candidate gene at the cellular level in normal and diseased tissues.

References

---

1. Parkin, D.M., et al., *Global cancer statistics, 2002.* CA Cancer J Clin, 2005. **55**(2): p. 74-108.
2. http://www.who.int, *Cancer (Fact sheet N°297)". World Health Organization, February 2009.* Retrieved 2009-05-11.
3. http://www.cancerhelp.org.uk, *Stomach cancer symptoms: Cancer Research UK. CancerHelp UK. 2007-06-15.* Retrieved 2009-05-11.
4. Nouraie, M., et al., *Fruits, vegetables, and antioxidants and risk of gastric cancer among male smokers.* Cancer Epidemiol Biomarkers Prev, 2005. **14**(9): p. 2087-92.
5. Normark, S., et al., *Persistent infection with Helicobacter pylori and the development of gastric cancer.* Adv Cancer Res, 2003. **90**: p. 63-89.
6. Correa, P. and B.G. Schneider, *Etiology of gastric cancer: what is new?* Cancer Epidemiol Biomarkers Prev, 2005. **14**(8): p. 1865-8.
7. *Biennial Report, International Agency For Research on Cancer, WHO, Lyon, France,*2000. **41**.
8. Rao, D.N. and B. Ganesh, *Estimate of cancer incidence in India in 1991.* Indian J Cancer, 1998. **35**(1): p. 10-8.
9. R. K. PHUKAN, E.Z., N. C. HAZARIKA, D. BARUAH, J. MAHANTA, *High prevalence of stomach cancer among the people of Mizoram, India.* CURRENT SCIENCE, 2004. **87**(3).
10. Phukan, R.K., et al., *Tobacco use and stomach cancer in Mizoram, India.* Cancer Epidemiol Biomarkers Prev, 2005. **14**(8): p. 1892-6.
11. Mudur, G., *India has some of the highest cancer rates in the world.* BMJ, 2005. **330**(7485): p. 215.
12. Neber, D.W. and A.L. Roe, *Ethnic and genetic differences in metabolism genes and risk of toxicity and cancer.* Sci Total Environ, 2001. **274**(1-3): p. 93-102.
13. Lucas, D., et al., *Cytochrome CYP2E1 phenotyping and genotyping in the evaluation of health risks from exposure to polluted environments.* Toxicol Lett, 2001. **124**(1-3): p. 71-81.
14. Wang, L.D., et al., *CYP1A1, GSTs and mEH polymorphisms and susceptibility to esophageal carcinoma: study of population from a high- incidence area in north China.* World J Gastroenterol, 2003. **9**(7): p. 1394-7.
15. Morita, S., et al., *CYP1A1, CYP2E1 and GSTM1 polymorphisms are not associated with susceptibility to squamous-cell carcinoma of the esophagus.* Int J Cancer, 1997. **71**(2): p. 192-5.
16. Matthias, C., et al., *The glutathione S-transferase GSTP1 polymorphism: effects on susceptibility to oral/pharyngeal and laryngeal carcinomas.* Pharmacogenetics, 1998. **8**(1): p. 1-6.
17. La Torre, G., S. Boccia, and G. Ricciardi, *Glutathione S-transferase M1 status and gastric cancer risk: a meta-analysis.* Cancer Lett, 2005. **217**(1): p. 53-60.

18. Lavender, N.A., et al., *Examination of polymorphic glutathione S-transferase (GST) genes, tobacco smoking and prostate cancer risk among men of African descent: a case-control study.* BMC Cancer, 2009. **9**: p. 397.
19. Ali-Osman, F., et al., *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, 1997. **272**(15): p. 10004-12.
20. Sundberg, K., et al., *Differences in the catalytic efficiencies of allelic variants of glutathione transferase P1-1 towards carcinogenic diol epoxides of polycyclic aromatic hydrocarbons.* Carcinogenesis, 1998. **19**(3): p. 433-6.
21. Coughlin, S.S. and I.J. Hall, *A review of genetic polymorphisms and prostate cancer risk.* Ann Epidemiol, 2002. **12**(3): p. 182-96.
22. Lee, E., et al., *Genetic polymorphism of conjugating enzymes and cancer risk: GSTM1, GSTT1, NAT1 and NAT2.* J Toxicol Sci, 1998. **23 Suppl 2**: p. 140-2.
23. Sreelekha, T.T., et al., *Genetic polymorphism of CYP1A1, GSTM1 and GSTT1 genes in Indian oral cancer.* Oral Oncol, 2001. **37**(7): p. 593-8.
24. Katoh, T., et al., *Effects of glutathione S-transferase (GST) M1 and GSTT1 genotypes on urothelial cancer risk.* Cancer Lett, 1998. **132**(1-2): p. 147-52.
25. Salagovic, J., et al., *Genetic polymorphism of glutathione S-transferases M1 and T1 as a risk factor in lung and bladder cancers.* Neoplasma, 1998. **45**(5): p. 312-7.
26. 1.Garcia M, J.A., Ward EM, Center MM, Hao Y, Siegel RL, et al., *Global Cancer Facts & Figures 2007. American Cancer Society. Available at http://www.cancer.org/downloads/STT/Global_Facts_and_Figures_2007_rev2.pdf. Accessed October 19, 2009.*
27. Zhu, L.B.a.W.-G., *P53: Structure, Function and Therapeutic Applications. .* Journal of Cancer Molecules 2006. **2**(4): p. 141-153.
28. Schmale, H. and C. Bamberger, *A novel protein with strong homology to the tumor suppressor p53.* Oncogene, 1997. **15**(11): p. 1363-7.
29. Kaghad, M., et al., *Monoallelically expressed gene related to p53 at 1p36, a region frequently deleted in neuroblastoma and other human cancers.* Cell, 1997. **90**(4): p. 809-19.
30. Irwin, M.S. and W.G. Kaelin, *p53 family update: p73 and p63 develop their own identities.* Cell Growth Differ, 2001. **12**(7): p. 337-49.
31. Whibley, C., P.D. Pharoah, and M. Hollstein, *p53 polymorphisms: cancer implications.* Nat Rev Cancer, 2009. **9**(2): p. 95-107.
32. Tandle, A.T., V. Sanghvi, and D. Saranath, *Determination of p53 genotypes in oral cancer patients from India.* Br J Cancer, 2001. **84**(6): p. 739-42.
33. Rogounovitch, T.I., et al., *TP53 codon 72 polymorphism in radiation-associated human papillary thyroid cancer.* Oncol Rep, 2006. **15**(4): p. 949-56.
34. Mitra, S., et al., *Association of specific genotype and haplotype of p53 gene with cervical cancer in India.* J Clin Pathol, 2005. **58**(1): p. 26-31.
35. Papadakis, E.D., N. Soulitzis, and D.A. Spandidos, *Association of p53 codon 72 polymorphism with advanced lung cancer: the Arg allele is preferentially*

retained in tumours arising in Arg/Pro germline heterozygotes. Br J Cancer, 2002. **87**(9): p. 1013-8.
36. Wu, M.T., M.C. Chen, and D.C. Wu, *Influences of lifestyle habits and p53 codon 72 and p21 codon 31 polymorphisms on gastric cancer risk in Taiwan.* Cancer Lett, 2004. **205**(1): p. 61-8.
37. Shen, H., et al., *P53 codon 72 polymorphism and risk of gastric cancer in a Chinese population.* Oncol Rep, 2004. **11**(5): p. 1115-20.
38. Pietsch, E.C., O. Humbey, and M.E. Murphy, *Polymorphisms in the p53 pathway.* Oncogene, 2006. **25**(11): p. 1602-11.
39. Wang, Y.C., et al., *Prognostic significance of p53 codon 72 polymorphism in lung carcinomas.* Eur J Cancer, 1999. **35**(2): p. 226-30.
40. Birgander, R., et al., *p53 polymorphisms and haplotypes in nasopharyngeal cancer.* Hum Hered, 1996. **46**(1): p. 49-54.
41. Kawajiri, K., et al., *Germ line polymorphisms of p53 and CYP1A1 genes involved in human lung cancer.* Carcinogenesis, 1993. **14**(6): p. 1085-9.
42. Jin, X., et al., *Higher lung cancer risk for younger African-Americans with the Pro/Pro p53 genotype.* Carcinogenesis, 1995. **16**(9): p. 2205-8.
43. Papadakis, E.N., D.N. Dokianakis, and D.A. Spandidos, *p53 codon 72 polymorphism as a risk factor in the development of breast cancer.* Mol Cell Biol Res Commun, 2000. **3**(6): p. 389-92.
44. Sjalander, A., et al., *p53 polymorphisms and haplotypes in breast cancer.* Carcinogenesis, 1996. **17**(6): p. 1313-6.
45. Hiyama, T., et al., *p53 Codon 72 polymorphism in gastric cancer susceptibility in patients with Helicobacter pylori-associated chronic gastritis.* Int J Cancer, 2002. **100**(3): p. 304-8.
46. Dokianakis, D.N. and D.A. Spandidos, *P53 codon 72 polymorphism as a risk factor in the development of HPV-associated cervical cancer.* Mol Cell Biol Res Commun, 2000. **3**(2): p. 111-4.
47. Storey, A., et al., *Role of a p53 polymorphism in the development of human papillomavirus-associated cancer.* Nature, 1998. **393**(6682): p. 229-34.
48. Hamel, N., et al., *No association between P53 codon 72 polymorphism and risk of squamous cell carcinoma of the head and neck.* Br J Cancer, 2000. **82**(4): p. 757-9.
49. Minaguchi, T., et al., *No evidence of correlation between polymorphism at codon 72 of p53 and risk of cervical cancer in Japanese patients with human papillomavirus 16/18 infection.* Cancer Res, 1998. **58**(20): p. 4585-6.
50. Sreeja, L., et al., *p53 Arg72Pro polymorphism predicts survival outcome in lung cancer patients in Indian population.* Cancer Invest, 2008. **26**(1): p. 41-6.
51. Nagpal, J.K., S. Patnaik, and B.R. Das, *Prevalence of high-risk human papilloma virus types and its association with P53 codon 72 polymorphism in tobacco addicted oral squamous cell carcinoma (OSCC) patients of Eastern India.* Int J Cancer, 2002. **97**(5): p. 649-53.
52. Sjalander, A., et al., *p53 polymorphisms and haplotypes in different ethnic groups.* Hum Hered, 1995. **45**(3): p. 144-9.

53. Beckman, G., et al., *Is p53 polymorphism maintained by natural selection?* Hum Hered, 1994. **44**(5): p. 266-70.
54. Fan, R., et al., *The p53 codon 72 polymorphism and lung cancer risk.* Cancer Epidemiol Biomarkers Prev, 2000. **9**(10): p. 1037-42.
55. Zhou, Y., et al., *p53 Codon 72 polymorphism and gastric cancer risk in a Chinese Han population.* Genet Test Mol Biomarkers. **14**(6): p. 829-33.
56. Stolte, M. and A. Meining, *Helicobacter pylori and Gastric Cancer.* Oncologist, 1998. **3**(2): p. 124-128.
57. Kim, K.E., *Gastric Cancer in Korean Americans: Risks and Reductions.* Korean Korean Am Stud Bull, 2003. **13**(1/2): p. 84-90.
58. Hippo, Y., et al., *Global gene expression analysis of gastric cancer by oligonucleotide microarrays.* Cancer Res, 2002. **62**(1): p. 233-40.
59. Cui, D.X., et al., *A microarray-based gastric carcinoma prewarning system.* World J Gastroenterol, 2005. **11**(9): p. 1273-82.
60. Kim, J.M., et al., *Identification of gastric cancer-related genes using a cDNA microarray containing novel expressed sequence tags expressed in gastric cancer cells.* Clin Cancer Res, 2005. **11**(2 Pt 1): p. 473-82.
61. Wang, K., et al., *Exome sequencing identifies frequent mutation of ARID1A in molecular subtypes of gastric cancer.* Nat Genet. **43**(12): p. 1219-23.
62. Bao, S., et al., *Evaluation of next-generation sequencing software in mapping and assembly.* J Hum Genet. **56**(6): p. 406-14.
63. Katsios, C., et al., *Translating cancer genomes sequencing revolution into surgical oncology practice.* J Surg Res. **173**(2): p. 365-9.
64. Hamilton, J.P. and S.J. Meltzer, *A review of the genomics of gastric cancer.* Clin Gastroenterol Hepatol, 2006. **4**(4): p. 416-25.
65. Boccia, S., Gianfagna, F., La Torre, G., Persiani, R., D'Ugo, D., and C.M. van Duijn, Ricciardi, G, *Genetic Susceptibility to Gastric Cancer: A review of the Published Meta-Analyses. In Cardinni, D.C. (ed.), Research Focus on Gastric Cancer. Hauppauge, .* NY: Nova Science Publishers, 2008: p. 137-163.
66. Yin, M., et al., *Molecular epidemiology of genetic susceptibility to gastric cancer: focus on single nucleotide polymorphisms in gastric carcinogenesis.* Am J Transl Res, 2009. **1**(1): p. 44-54.
67. Zacho, A., J. Nielsen, and V. Larsen, *On the consumption of unburned tobacco in patients with cancer of the stomach.* Acta Chir Scand, 1968. **134**(3): p. 272-4.
68. *Vumson, N. T. and Thawagn, V. (eds), Zo History, India, Vumson Publisher, Aizawl.* 1986.
69. Mahanta J, C.M., Hazarika NC, *Toxicity of tuibur, a unique form of tobacco smoke extract used in Mizoram, India.* Curr Sci, 1998. **75**: p. 381 - 4.
70. *International Agency for Research on Cancer (IARC). 1999. Monographs on the..carcinogenicity..of..chemical..substances..*http://193.51.164.11/monoeval/grlist.html.
71. Tredaniel, J., et al., *Tobacco smoking and gastric cancer: review and meta-analysis.* Int J Cancer, 1997. **72**(4): p. 565-73.
72. Satyanarayana, L., S. Asthana, and K.C. Sharma, *Tobacco related cancer risk in north eastern region of India.* Indian J Med Res, 2008. **128**(3): p. 318-9.

73. Satyanarayana L, A.S., Sharma KC., *Trends in cumulative risk for life time development of tobacco related cancers in India during 1982 to 2000.* Indian J Pediatr Oncol, 2007. **28**: p. 26-32.
74. Howe, G.R., *Use of computerized record linkage in cohort studies.* Epidemiol Rev, 1998. **20**(1): p. 112-21.
75. Ferlay J, B.F., Pisani P, Parkin DM, *Cancer incidence, mortality and prevalence worldwide, version 1.0. Lyon: IARC Press, IARC Cancer Base No. 5; 2001.2.*
76. Levi, F., et al., *Trends in mortality from major cancers in the European Union, including acceding countries, in 2004.* Cancer, 2004. **101**(12): p. 2843-50.
77. Henson, D.E., et al., *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, 2004. **128**(7): p. 765-70.
78. Ando, T., et al., *Causal role of Helicobacter pylori infection in gastric cancer.* World J Gastroenterol, 2006. **12**(2): p. 181-6.
79. Crew, K.D. and A.I. Neugut, *Epidemiology of gastric cancer.* World J Gastroenterol, 2006. **12**(3): p. 354-62.
80. Jemal, A., et al., *Cancer statistics, 2006.* CA Cancer J Clin, 2006. **56**(2): p. 106-30.
81. Nakamura, T., et al., *A clinicopathological study in young patients with gastric carcinoma.* J Surg Oncol, 1999. **71**(4): p. 214-9.
82. Theuer, C.P., et al., *Gastric adenocarcinoma in patients 40 years of age or younger.* Am J Surg, 1996. **172**(5): p. 473-6; discussion 476-7.
83. Brenner, H., D. Rothenbacher, and V. Arndt, *Epidemiology of stomach cancer.* Methods Mol Biol, 2009. **472**: p. 467-77.
84. Sant, M., et al., *EUROCARE-3: survival of cancer patients diagnosed 1990-94--results and commentary.* Ann Oncol, 2003. **14 Suppl 5**: p. v61-118.
85. Dickman, P.W., et al., *Survival of cancer patients in Finland 1955-1994.* Acta Oncol, 1999. **38 Suppl 12**: p. 1-103.
86. Teppo, L., et al., *Cancer patient survival--patterns, comparisons, trends--a population-based Cancer Registry study in Finland.* Acta Oncol, 1999. **38**(3): p. 283-94.
87. Sinha, R., et al., *Cancer risk and diet in India.* J Postgrad Med, 2003. **49**(3): p. 222-8.
88. ICMR, *Consolidated report of population based cancer registries 2001-2004 , Incidence and distribution of cancer. National cancer registry programme. .* Indian council of medical research (ICMR), Bangalore, 2006.
89. Phukan, R.K., et al., *Dietary habits and stomach cancer in Mizoram, India.* J Gastroenterol, 2006. **41**(5): p. 418-24.
90. Vumson NT, Z.h., Thawagn V, *Aizawl, Mizoram: India.* Vumson Publisher, 1986: p. 26–39.
91. *World Cancer Research Fund, American Institute for Cancer Research. Food, nutrition, and the prevention of cancer: a global perspective / World Cancer Research Fund, in association with American Institute for Cancer Research.* Washington, DC: American Institute for Cancer Research,, 1997.

92. Mahady, G.B., et al., *Turmeric (Curcuma longa) and curcumin inhibit the growth of Helicobacter pylori, a group 1 carcinogen.* Anticancer Res, 2002. **22**(6C): p. 4179-81.
93. Shanmugasundaram, K.R., S. Ramanujam, and E.R. Shanmugasundaram, *Amrita Bindu--a salt-spice-herbal health food supplement for the prevention of nitrosamine induced depletion of antioxidants.* J Ethnopharmacol, 1994. **42**(2): p. 83-93.
94. Rao, D.N., et al., *A case-control study of stomach cancer in Mumbai, India.* Int J Cancer, 2002. **99**(5): p. 727-31.
95. Mathew, A., et al., *Diet and stomach cancer: a case-control study in South India.* Eur J Cancer Prev, 2000. **9**(2): p. 89-97.
96. Chyou, P.H., et al., *Lung cancer: a prospective study of smoking, occupation, and nutrient intake.* Arch Environ Health, 1993. **48**(2): p. 69-72.
97. Tandon, M., et al., *Role of micro-nutrients and trace elements in carcinoma of larynx.* J Assoc Physicians India, 2000. **48**(10): p. 995-8.
98. Rajkumar, T., et al., *Oral cancer in Southern India: the influence of body size, diet, infections and sexual practices.* Eur J Cancer Prev, 2003. **12**(2): p. 135-43.
99. Krishnaswamy, K. and K. Polasa, *Diet, nutrition & cancer--the Indian scenario.* Indian J Med Res, 1995. **102**: p. 200-9.
100. Hebert, J.R., et al., *Dietary exposures and oral precancerous lesions in Srikakulam District, Andhra Pradesh, India.* Public Health Nutr, 2002. **5**(2): p. 303-12.
101. Nayar, D., et al., *Nutritional risk factors in esophageal cancer.* J Assoc Physicians India, 2000. **48**(8): p. 781-7.
102. Malkan, G. and K.M. Mohandas, *Epidemiology of digestive cancers in India. I. General principles and esophageal cancer.* Indian J Gastroenterol, 1997. **16**(3): p. 98-102.
103. Phukan, R.K., et al., *Role of dietary habits in the development of esophageal cancer in Assam, the north-eastern region of India.* Nutr Cancer, 2001. **39**(2): p. 204-9.
104. Glade, M.J., *Food, nutrition, and the prevention of cancer: a global perspective. American Institute for Cancer Research/World Cancer Research Fund, American Institute for Cancer Research, 1997.* Nutrition, 1999. **15**(6): p. 523-6.
105. Hisada, M., et al., *Prospective study of antibody to human papilloma virus type 16 and risk of cervical, endometrial, and ovarian cancers (United States).* Cancer Causes Control, 2001. **12**(4): p. 335-41.
106. Fenley J, B.F., Pisani DMe, *World Health Organization. GLOBOCAN 2000: Cancer incidence, mortality and prevalence worldwide. Lyon, France: IARC Press; 2001.*
107. Nagayo, T., *Precursors of human gastric cancer: their frequencies and histological characteristics. Pathophysiology of carcinogenesis in digestive organs p 151 (editor E Farber). .* Univ. Tokyo Press,Tokyo 1977.

108. Imai, T., T. Kubo, and H. Watanabe, *Chronic gastritis in Japanese with reference to high incidence of gastric carcinoma.* J Natl Cancer Inst, 1971. **47**(1): p. 179-95.
109. Cheli, R., et al., *Atrophic gastritis and intestinal metaplasia in asymptomatic Hungarian and Italian populations.* Endoscopy, 1980. **12**(3): p. 105-8.
110. Siurala, M., J. Lehtola, and T. Ihamaki, *Atrophic gastritis and its sequelae. Results of 19-23 years' follow-up examinations.* Scand J Gastroenterol, 1974. **9**(5): p. 441-6.
111. Walker, I.R., et al., *Simple atrophic gastritis and gastric carcinoma.* Gut, 1971. **12**(11): p. 906-11.
112. Correa, P., C. Cuello, and E. Duque, *Carcinoma and intestinal metaplasia of the stomach in Colombian migrants.* J Natl Cancer Inst, 1970. **44**(2): p. 297-306.
113. Lauren, P., *The Two Histological Main Types of Gastric Carcinoma: Diffuse and So-Called Intestinal-Type Carcinoma. An Attempt at a Histo-Clinical Classification.* Acta Pathol Microbiol Scand, 1965. **64**: p. 31-49.
114. Rubio, C.A., et al., *Intestinal metaplasia of the stomach. I: Quantitative analysis in gastric peptic ulcer and in incipient adenocarcinoma in Japanese subjects.* Anticancer Res, 1985. **5**(4): p. 435-40.
115. Matsukura, N., et al., *Distribution of marker enzymes and mucin in intestinal metaplasia in human stomach and relation to complete and incomplete types of intestinal metaplasia to minute gastric carcinomas.* J Natl Cancer Inst, 1980. **65**(2): p. 231-40.
116. Morson, B.C., *Carcinoma arising from areas of intestinal metaplasia in the gastric mucosa.* Br J Cancer, 1955. **9**(3): p. 377-85.
117. Kang, J.Y., et al., *Risk of gastric carcinoma in patients with atrophic gastritis and intestinal metaplasia.* Gut, 2002. **51**(6): p. 899.
118. Rugge, M. and R.M. Genta, *Staging and grading of chronic gastritis.* Hum Pathol, 2005. **36**(3): p. 228-33.
119. Morson, B.C., et al., *Precancerous conditions and epithelial dysplasia in the stomach.* J Clin Pathol, 1980. **33**(8): p. 711-21.
120. Cuello, C., et al., *Histopathology of gastric dysplasias: correlations with gastric juice chemistry.* Am J Surg Pathol, 1979. **3**(6): p. 491-500.
121. Jass, J.R., *A classification of gastric dysplasia.* Histopathology, 1983. **7**(2): p. 181-93.
122. Lauren, P.A. and T.J. Nevalainen, *Epidemiology of intestinal and diffuse types of gastric carcinoma. A time-trend study in Finland with comparison between studies from high- and low-risk areas.* Cancer, 1993. **71**(10): p. 2926-33.
123. Munoz, N. and R. Connelly, *Time trends of intestinal and diffuse types of gastric cancer in the United States.* Int J Cancer, 1971. **8**(1): p. 158-64.
124. Roukos, D., M. Lorenz, and C. Hottenrott, *[Prognostic significance of the Lauren classification of patients with stomach carcinoma. A statistical analysis of long-term results following gastrectomy].* Schweiz Med Wochenschr, 1989. **119**(21): p. 755-9.

125. Fuchs, C.S. and R.J. Mayer, *Gastric carcinoma.* N Engl J Med, 1995. **333**(1): p. 32-41.
126. el-Rifai, W. and S.M. Powell, *Molecular and biologic basis of upper gastrointestinal malignancy. Gastric carcinoma.* Surg Oncol Clin N Am, 2002. **11**(2): p. 273-91, viii.
127. El-Rifai, W. and S.M. Powell, *Molecular biology of gastric cancer.* Semin Radiat Oncol, 2002. **12**(2): p. 128-40.
128. La Vecchia, C., *Mediterranean diet and cancer.* Public Health Nutr, 2004. **7**(7): p. 965-8.
129. Palli, D., et al., *Red meat, family history, and increased risk of gastric cancer with microsatellite instability.* Cancer Res, 2001. **61**(14): p. 5415-9.
130. Palli, D., et al., *O6-alkylguanines, dietary N-nitroso compounds, and their precursors in gastric cancer.* Nutr Cancer, 2001. **39**(1): p. 42-9.
131. Sjodahl, K., et al., *Smoking and alcohol drinking in relation to risk of gastric cancer: a population-based, prospective cohort study.* Int J Cancer, 2007. **120**(1): p. 128-32.
132. Gao, C., et al., *Interaction between cytochrome P-450 2E1 polymorphisms and environmental factors with risk of esophageal and stomach cancers in Chinese.* Cancer Epidemiol Biomarkers Prev, 2002. **11**(1): p. 29-34.
133. Koizumi, Y., et al., *Cigarette smoking and the risk of gastric cancer: a pooled analysis of two prospective studies in Japan.* Int J Cancer, 2004. **112**(6): p. 1049-55.
134. Nishikawa, A., et al., *Cigarette smoking, metabolic activation and carcinogenesis.* Curr Drug Metab, 2004. **5**(5): p. 363-73.
135. Iishi, H., et al., *Promotion by ethanol of gastric carcinogenesis induced by N-methyl-N'-nitro-N-nitrosoguanidine in Wistar rats.* Br J Cancer, 1989. **59**(5): p. 719-21.
136. Kapadia, C.R., *Gastric atrophy, metaplasia, and dysplasia: a clinical perspective.* J Clin Gastroenterol, 2003. **36**(5 Suppl): p. S29-36; discussion S61-2.
137. Hoskins, L.C., et al., *Distribution of ABO blood groups in patients with pernicious anemia, gastric carcinoma and gastric carcinoma associated with pernicious anemia.* N Engl J Med, 1965. **273**(12): p. 633-7.
138. Callender, S., et al., *ABO blood groups in patients with gastric carcinoma associated with pernicious anaemia.* Gut, 1971. **12**(6): p. 465-7.
139. Kusters, J.G., A.H. van Vliet, and E.J. Kuipers, *Pathogenesis of Helicobacter pylori infection.* Clin Microbiol Rev, 2006. **19**(3): p. 449-90.
140. De Luca, A., et al., *Effects of Helicobacter pylori infection on cell cycle progression and the expression of cell cycle regulatory proteins.* J Cell Physiol, 2004. **200**(3): p. 334-42.
141. *Schistosomes, liver flukes and Helicobacter pylori. IARC Working Group on the Evaluation of Carcinogenic Risks to Humans. Lyon, 7-14 June 1994.* IARC Monogr Eval Carcinog Risks Hum, 1994. **61**: p. 1-241.
142. Salih, B.A., *Helicobacter pylori infection in developing countries: the burden for how long?* Saudi J Gastroenterol, 2009. **15**(3): p. 201-7.

143. *Unidentified curved bacilli on gastric epithelium in active chronic gastritis.* Lancet, 1983. **1**(8336): p. 1273-5.
144. Maconi, G., G. Manes, and G.B. Porro, *Role of symptoms in diagnosis and outcome of gastric cancer.* World J Gastroenterol, 2008. **14**(8): p. 1149-55.
145. Maconi, G., et al., *Gastric cancer in young patients with no alarm symptoms: focus on delay in diagnosis, stage of neoplasm and survival.* Scand J Gastroenterol, 2003. **38**(12): p. 1249-55.
146. Stephens, M.R., et al., *Prognostic significance of alarm symptoms in patients with gastric cancer.* Br J Surg, 2005. **92**(7): p. 840-6.
147. Bowrey, D.J., et al., *Use of alarm symptoms to select dyspeptics for endoscopy causes patients with curable esophagogastric cancer to be overlooked.* Surg Endosc, 2006. **20**(11): p. 1725-8.
148. Dikken, J.L., et al., *Treatment of resectable gastric cancer.* Therap Adv Gastroenterol. **5**(1): p. 49-69.
149. Graziano, F., B. Humar, and P. Guilford, *The role of the E-cadherin gene (CDH1) in diffuse gastric cancer susceptibility: from the laboratory to clinical practice.* Ann Oncol, 2003. **14**(12): p. 1705-13.
150. Huscher, C.G., et al., *Laparoscopic versus open subtotal gastrectomy for distal gastric cancer: five-year results of a randomized prospective trial.* Ann Surg, 2005. **241**(2): p. 232-7.
151. Benthin F, M.H., Aigner K, *Celiac axis infusion (C.A.I)induction chemotherapy prior to surgery for gastric cancer.* Regional Ca Treat, 1993. **6**.
152. Stephens, F.O., *The role of regional chemotherapy in gastric cancer.* Eur J Surg Oncol, 1994. **20**(2): p. 187-8.
153. Fujimoto, S., et al., *A study of survival in patients with stomach cancer treated by a combination of preoperative intra-arterial infusion therapy and surgery.* Cancer, 1976. **37**(4): p. 1648-53.
154. Iida, T., et al., *Preoperative intraarterial infusion chemotherapy for advanced gastric cancer--a retrospective review of four cases.* Radiat Med, 2003. **21**(4): p. 172-7.
155. JC, M., *Carcinomas of the Gastrointestinal Tract. In: Skeel RT, editor. Handbook of Cancer Chemotherapy.* Philadelphia: Lippincott Williams & Wilkins, 1999: p. 214-237.
156. Gong, S.J., et al., *Growth inhibitory effects of trastuzumab and chemotherapeutic drugs in gastric cancer cell lines.* Cancer Lett, 2004. **214**(2): p. 215-24.
157. Scartozzi, M., et al., *Molecular biology of sporadic gastric cancer: prognostic indicators and novel therapeutic approaches.* Cancer Treat Rev, 2004. **30**(5): p. 451-9.
158. Jung, Y.D., et al., *Effects of combination anti-vascular endothelial growth factor receptor and anti-epidermal growth factor receptor therapies on the growth of gastric cancer in a nude mouse model.* Eur J Cancer, 2002. **38**(8): p. 1133-40.
159. Toffoli, G. and E. Cecchin, *Pharmacogenomics and stomach cancer.* Pharmacogenomics, 2004. **5**(6): p. 627-41.

160. Ottini, L., et al., *Microsatellite instability in gastric cancer is associated with tumor location and family history in a high-risk population from Tuscany.* Cancer Res, 1997. **57**(20): p. 4523-9.
161. Leung, S.Y., et al., *hMLH1 promoter methylation and lack of hMLH1 expression in sporadic gastric carcinomas with high-frequency microsatellite instability.* Cancer Res, 1999. **59**(1): p. 159-64.
162. Motomura, K., et al., *Loss of alleles at loci on chromosome 13 in human primary gastric cancers.* Genomics, 1988. **2**(2): p. 180-4.
163. Kim, C.J., et al., *Detection of 17p loss in gastric carcinoma using polymerase chain reaction.* Lab Invest, 1995. **72**(2): p. 232-6.
164. Semba, S., et al., *Frequent microsatellite instability and loss of heterozygosity in the region including BRCA1 (17q21) in young patients with gastric cancer.* Int J Oncol, 1998. **12**(6): p. 1245-51.
165. Guilford, P., et al., *E-cadherin germline mutations in familial gastric cancer.* Nature, 1998. **392**(6674): p. 402-5.
166. Brooks-Wilson, A.R., et al., *Germline E-cadherin mutations in hereditary diffuse gastric cancer: assessment of 42 new families and review of genetic screening criteria.* J Med Genet, 2004. **41**(7): p. 508-17.
167. Morris, L.E., et al., *Nucleotide variants within the IQGAP1 gene in diffuse-type gastric cancers.* Genes Chromosomes Cancer, 2005. **42**(3): p. 280-6.
168. Weber, G.F., et al., *Receptor-ligand interaction between CD44 and osteopontin (Eta-1).* Science, 1996. **271**(5248): p. 509-12.
169. Ue, T., et al., *Co-expression of osteopontin and CD44v9 in gastric cancer.* Int J Cancer, 1998. **79**(2): p. 127-32.
170. Lotan, R., et al., *Expression of a 31-kDa lactoside-binding lectin in normal human gastric mucosa and in primary and metastatic gastric carcinomas.* Int J Cancer, 1994. **56**(4): p. 474-80.
171. Kawanishi, J., et al., *Loss of E-cadherin-dependent cell-cell adhesion due to mutation of the beta-catenin gene in a human cancer cell line, HSC-39.* Mol Cell Biol, 1995. **15**(3): p. 1175-81.
172. Caca, K., et al., *Beta- and gamma-catenin mutations, but not E-cadherin inactivation, underlie T-cell factor/lymphoid enhancer factor transcriptional deregulation in gastric and pancreatic cancer.* Cell Growth Differ, 1999. **10**(6): p. 369-76.
173. Shibata, T., et al., *Dominant negative inhibition of the association between beta-catenin and c-erbB-2 by N-terminally deleted beta-catenin suppresses the invasion and metastasis of cancer cells.* Oncogene, 1996. **13**(5): p. 883-9.
174. Tahara, E., *Genetic pathways of two types of gastric cancer.* IARC Sci Publ, 2004(157): p. 327-49.
175. Yoshida, K., et al., *Expression of TGF-beta and procollagen type I and type III in human gastric carcinomas.* Int J Cancer, 1989. **44**(3): p. 394-8.
176. Akagi, M., et al., *Induction of neuropilin-1 and vascular endothelial growth factor by epidermal growth factor in human gastric cancer cells.* Br J Cancer, 2003. **88**(5): p. 796-802.

177. Kitadai, Y., et al., *Regulation of disease-progression genes in human gastric carcinoma cells by interleukin 8.* Clin Cancer Res, 2000. **6**(7): p. 2735-40.
178. Kitadai, Y., et al., *Expression of interleukin-8 correlates with vascularity in human gastric carcinomas.* Am J Pathol, 1998. **152**(1): p. 93-100.
179. Kuniyasu, H., et al., *Frequent amplification of the c-met gene in scirrhous type stomach cancer.* Biochem Biophys Res Commun, 1992. **189**(1): p. 227-32.
180. Hattori, Y., et al., *K-sam, an amplified gene in stomach cancer, is a member of the heparin-binding growth factor receptor genes.* Proc Natl Acad Sci U S A, 1990. **87**(15): p. 5983-7.
181. Yokota, J., et al., *Genetic alterations of the c-erbB-2 oncogene occur frequently in tubular adenocarcinoma of the stomach and are often accompanied by amplification of the v-erbA homologue.* Oncogene, 1988. **2**(3): p. 283-7.
182. Yonemura, Y., et al., *Expression of c-erbB-2 oncoprotein in gastric carcinoma. Immunoreactivity for c-erbB-2 protein is an independent indicator of poor short-term prognosis in patients with gastric carcinoma.* Cancer, 1991. **67**(11): p. 2914-8.
183. Lee, K.H., et al., *Clinicopathologic significance of the K-ras gene codon 12 point mutation in stomach cancer. An analysis of 140 cases.* Cancer, 1995. **75**(12): p. 2794-801.
184. Sano, T., et al., *Frequent loss of heterozygosity on chromosomes 1q, 5q, and 17p in human gastric carcinomas.* Cancer Res, 1991. **51**(11): p. 2926-31.
185. Schabath, M.B., et al., *Combined effects of the p53 and p73 polymorphisms on lung cancer risk.* Cancer Epidemiol Biomarkers Prev, 2006. **15**(1): p. 158-61.
186. Toyama, T., et al., *Association of TP53 codon 72 polymorphism and the outcome of adjuvant therapy in breast cancer patients.* Breast Cancer Res, 2007. **9**(3): p. R34.
187. Shepherd, T., et al., *Alterations in exon 4 of the p53 gene in gastric carcinoma.* Gastroenterology, 2000. **118**(6): p. 1039-44.
188. Zhou, Y., et al., *P53 codon 72 polymorphism and gastric cancer: a meta-analysis of the literature.* Int J Cancer, 2007. **121**(7): p. 1481-6.
189. Yokozaki, H., et al., *p53 point mutations in primary human gastric carcinomas.* J Cancer Res Clin Oncol, 1992. **119**(2): p. 67-70.
190. Sugimura, T., S. Fujimura, and T. Baba, *Tumor production in the glandular stomach and alimentary tract of the rat by N-methyl-N'-nitro-N-nitrosoguanidine.* Cancer Res, 1970. **30**(2): p. 455-65.
191. Mirvish, S.S., *Kinetics of nitrosamide formation from alkylureas, N-alkylurethans, and alkylguanidines: possible implications for the etiology of human gastric cancer.* J Natl Cancer Inst, 1971. **46**(6): p. 1183-93.
192. Yokozaki, H., et al., *Alterations of p73 preferentially occur in gastric adenocarcinomas with foveolar epithelial phenotype.* Int J Cancer, 1999. **83**(2): p. 192-6.
193. Lefebvre, O., et al., *Gastric mucosa abnormalities and tumorigenesis in mice lacking the pS2 trefoil protein.* Science, 1996. **274**(5285): p. 259-62.

194. Rowley, P.T., *Inherited susceptibility to colorectal cancer.* Annu Rev Med, 2005. **56**: p. 539-54.
195. Kinzler, K.W., et al., *Identification of FAP locus genes from chromosome 5q21.* Science, 1991. **253**(5020): p. 661-5.
196. Nakatsuru, S., et al., *Somatic mutations of the APC gene in precancerous lesion of the stomach.* Hum Mol Genet, 1993. **2**(9): p. 1463-5.
197. Nakatsuru, S., et al., *Somatic mutation of the APC gene in gastric cancer: frequent mutations in very well differentiated adenocarcinoma and signet-ring cell carcinoma.* Hum Mol Genet, 1992. **1**(8): p. 559-63.
198. Li, Y.L., et al., *Loss of heterozygosity on 10q23.3 and mutation of tumor suppressor gene PTEN in gastric cancer and precancerous lesions.* World J Gastroenterol, 2005. **11**(2): p. 285-8.
199. Deng, H., et al., *Significance of Survivin and PTEN expression in full lymph node-examined gastric cancer.* World J Gastroenterol, 2006. **12**(7): p. 1013-7.
200. Li, Q.L., et al., *Causal relationship between the loss of RUNX3 expression and gastric cancer.* Cell, 2002. **109**(1): p. 113-24.
201. Hayashi, K., et al., *Inactivation of retinoic acid receptor beta by promoter CpG hypermethylation in gastric cancer.* Differentiation, 2001. **68**(1): p. 13-21.
202. Boccia, S., Brand, A., Brand, H., Ricciardi, G, *The integration of gene environment interactions into healthcare and disease prevention as a major challenge for Public Health Genomics. Mut. Res.*
203. Roberts-Thomson, I.C. and W.J. Butler, *Polymorphism and gastric cancer.* J Gastroenterol Hepatol, 2005. **20**(5): p. 793-4.
204. Gonzalez, C.A., N. Sala, and G. Capella, *Genetic susceptibility and gastric cancer risk.* Int J Cancer, 2002. **100**(3): p. 249-60.
205. Saadat, M., *Genetic polymorphisms of glutathione S-transferase T1 (GSTT1) and susceptibility to gastric cancer: a meta-analysis.* Cancer Sci. , 2006. **97** p. 505-509.
206. Zintzaras, E., *Association of methylenetetrahydrofolate reductase (MTHFR) polymorphisms with genetic susceptibility to gastric cancer: a meta-analysis.* J Hum Genet, 2006. **51**(7): p. 618-24.
207. Boccia, S., et al., *CYP2E1PstI/RsaI polymorphism and interaction with tobacco, alcohol and GSTs in gastric cancer susceptibility: A meta-analysis of the literature.* Carcinogenesis, 2007. **28**(1): p. 101-6.
208. Kato, S., et al., *Cytochrome P4502E1 (CYP2E1) genetic polymorphism in a case-control study of gastric cancer and liver disease.* Pharmacogenetics, 1995. **5 Spec No**: p. S141-4.
209. Suzuki, S., et al., *Relationship between genetic polymorphisms of drug-metabolizing enzymes (CYP1A1, CYP2E1, GSTM1, and NAT2), drinking habits, histological subtypes, and p53 gene point mutations in Japanese patients with gastric cancer.* J Gastroenterol, 2004. **39**(3): p. 220-30.
210. Hirvonen, A., *Polymorphisms of xenobiotic-metabolizing enzymes and susceptibility to cancer.* Environ Health Perspect, 1999. **107 Suppl 1**: p. 37-47.

211. Hung, R.J., et al., *Perspectives on the molecular epidemiology of aerodigestive tract cancers.* Mutat Res, 2005. **592**(1-2): p. 102-18.
212. Gonzalez, F.J., *Genetic polymorphism and cancer susceptibility: fourteenth Sapporo Cancer Seminar.* Cancer Res, 1995. **55**(3): p. 710-5.
213. Nebert, D.W., R.A. McKinnon, and A. Puga, *Human drug-metabolizing enzyme polymorphisms: effects on risk of toxicity and cancer.* DNA Cell Biol, 1996. **15**(4): p. 273-80.
214. Anttila, S., et al., *Immunohistochemical localization of glutathione S-transferases in human lung.* Cancer Res, 1993. **53**(23): p. 5643-8.
215. Ketterer, B., et al., *The human glutathione S-transferase supergene family, its polymorphism, and its effects on susceptibility to lung cancer.* Environ Health Perspect, 1992. **98**: p. 87-94.
216. Martin, J.L., *Thioredoxin - a fold for all reasons. Structure 3, 245±250.* 1995
217. Koonin, E.V., Mushegian, A. R., Tatusov, R. L., Altschul, S. F., Bryant, S. H., Bork, P. and Valencia, A. , *Eukaryotic translation elongation factor 1c contains a glutathione transferase domain - Study of a diverse, ancient protein superfamily using motif search and structural modeling.* Protein Sci. , 1994. **3**: p. 2045±2054.
218. Rossjohn, J., Board, P. G., Parker, M. W. and Wilce, M. C. J. , *A structurallyderived consensus pattern for theta class glutathione transferases* Protein Eng. , 1996 **9** p. 327±332.
219. Armstrong, R.N., *Mechanistic imperatives for the evolution of glutathione transferases.* Curr Opin Chem Biol, 1998. **2**(5): p. 618-23.
220. Hansson, L.O., et al., *Evolution of differential substrate specificities in Mu class glutathione transferases probed by DNA shuffling.* J Mol Biol, 1999. **287**(2): p. 265-76.
221. Coles, B.F. and F.F. Kadlubar, *Detoxification of electrophilic compounds by glutathione S-transferase catalysis: determinants of individual response to chemical carcinogens and chemotherapeutic drugs?* Biofactors, 2003. **17**(1-4): p. 115-30.
222. Bolt, H.M. and R. Thier, *Relevance of the deletion polymorphisms of the glutathione S-transferases GSTT1 and GSTM1 in pharmacology and toxicology.* Curr Drug Metab, 2006. **7**(6): p. 613-28.
223. Rebbeck, T.R., *Molecular epidemiology of the human glutathione S-transferase genotypes GSTM1 and GSTT1 in cancer susceptibility.* Cancer Epidemiol Biomarkers Prev, 1997. **6**(9): p. 733-43.
224. Egan, K.M., et al., *Genetic polymorphisms in GSTM1, GSTP1, and GSTT1 and the risk for breast cancer: results from the Shanghai Breast Cancer Study and meta-analysis.* Cancer Epidemiol Biomarkers Prev, 2004. **13**(2): p. 197-204.
225. Saadat, I. and M. Saadat, *Glutathione S-transferase M1 and T1 null genotypes and the risk of gastric and colorectal cancers.* Cancer Lett, 2001. **169**(1): p. 21-6.
226. Strange, R.C., et al., *The human glutathione S-transferases: a case-control study of the incidence of the GST1 0 phenotype in patients with adenocarcinoma.* Carcinogenesis, 1991. **12**(1): p. 25-8.

227. Deakin, M., et al., *Glutathione S-transferase GSTT1 genotypes and susceptibility to cancer: studies of interactions with GSTM1 in lung, oral, gastric and colorectal cancers.* Carcinogenesis, 1996. **17**(4): p. 881-4.
228. Gao, C.M., et al., *Glutathione-S-transferases M1 (GSTM1) and GSTT1 genotype, smoking, consumption of alcohol and tea and risk of esophageal and stomach cancers: a case-control study of a high-incidence area in Jiangsu Province, China.* Cancer Lett, 2002. **188**(1-2): p. 95-102.
229. Chen, B., et al., *Glutathione S-transferase T1 (GSTT1) gene polymorphism and gastric cancer susceptibility: a meta-analysis of epidemiologic studies.* Dig Dis Sci. **55**(7): p. 1831-8.
230. Ott, K., et al., *Glutathione-S-transferase P1, T1 and M1 genetic polymorphisms in neoadjuvant-treated locally advanced gastric cancer: GSTM1-present genotype is associated with better prognosis in completely resected patients.* Int J Colorectal Dis, 2008. **23**(8): p. 773-82.
231. Boccia, S., et al., *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, 2007. **7**: p. 206.
232. Al-Moundhri, M.S., et al., *Combined polymorphism analysis of glutathione S-transferase M1/G1 and interleukin-1B (IL-1B)/interleukin 1-receptor antagonist (IL-1RN) and gastric cancer risk in an Omani Arab Population.* J Clin Gastroenterol, 2009. **43**(2): p. 152-6.
233. Nguyen, T.V., et al., *Genetic polymorphisms in GSTA1, GSTP1, GSTT1, and GSTM1 and gastric cancer risk in a Vietnamese population.* Oncol Res. **18**(7): p. 349-55.
234. Piao, J.M., et al., *Glutathione-S-transferase (GSTM1, GSTT1) and the risk of gastrointestinal cancer in a Korean population.* World J Gastroenterol, 2009. **15**(45): p. 5716-21.
235. Goto, S., et al., *Overexpression of glutathione S-transferase pi enhances the adduct formation of cisplatin with glutathione in human cancer cells.* Free Radic Res, 1999. **31**(6): p. 549-58.
236. Nakagawa, K., et al., *Glutathione-S-transferase pi as a determinant of drug resistance in transfectant cell lines.* J Biol Chem, 1990. **265**(8): p. 4296-301.
237. Hu, X., et al., *Mechanism of differential catalytic efficiency of two polymorphic forms of human glutathione S-transferase P1-1 in the glutathione conjugation of carcinogenic diol epoxide of chrysene.* Arch Biochem Biophys, 1997. **345**(1): p. 32-8.
238. Butkiewicz, D., et al., *Polymorphisms of the GSTP1 and GSTM1 genes and PAH-DNA adducts in human mononuclear white blood cells.* Environ Mol Mutagen, 2000. **35**(2): p. 99-105.
239. Guengerich, F.P., *Catalytic selectivity of human cytochrome P450 enzymes: relevance to drug metabolism and toxicity.* Toxicol Lett, 1994. **70**(2): p. 133-8.
240. Gonzalez, F.J., T. Aoyama, and H.V. Gelboin, *Activation of promutagens by human cDNA-expressed cytochrome P450s.* Prog Clin Biol Res, 1990. **340B**: p. 77-86.

241. Eaton, D.L., et al., *Role of cytochrome P4501A2 in chemical carcinogenesis: implications for human variability in expression and enzyme activity.* Pharmacogenetics, 1995. **5**(5): p. 259-74.
242. Gajecka, M., et al., *CYP1A1, CYP2D6, CYP2E1, NAT2, GSTM1 and GSTT1 polymorphisms or their combinations are associated with the increased risk of the laryngeal squamous cell carcinoma.* Mutat Res, 2005. **574**(1-2): p. 112-23.
243. Nelson, D.R., et al., *P450 superfamily: update on new sequences, gene mapping, accession numbers and nomenclature.* Pharmacogenetics, 1996. **6**(1): p. 1-42.
244. Shimada, T., et al., *Interindividual variations in human liver cytochrome P-450 enzymes involved in the oxidation of drugs, carcinogens and toxic chemicals: studies with liver microsomes of 30 Japanese and 30 Caucasians.* J Pharmacol Exp Ther, 1994. **270**(1): p. 414-23.
245. Nouso, K., S.S. Thorgeirsson, and N. Battula, *Stable expression of human cytochrome P450IIE1 in mammalian cells: metabolic activation of nitrosodimethylamine and formation of adducts with cellular DNA.* Cancer Res, 1992. **52**(7): p. 1796-800.
246. Koop, D.R., *Oxidative and reductive metabolism by cytochrome P450 2E1.* FASEB J, 1992. **6**(2): p. 724-30.
247. Vondracek, M., et al., *Cytochrome P450 expression and related metabolism in human buccal mucosa.* Carcinogenesis, 2001. **22**(3): p. 481-8.
248. McBride OW, U.M., Gelboin HV, Gonzalez and FJ., *A Taql polymorphism in the human P45011E1 gene on chromosome 10 (CYP2E). .* NucI Acid Res, 1987. **15**.
249. Uematsu F, K.H., Ohmachi T, Sagami 1,, K.T. Motomiya M, Komori M, Watanabe, and M., *Two common RFLPs of the human CYP2E1 gene.* NucI Acid Res, 1991. **19**: p. 2803.
250. Uematsu, F., et al., *Association between restriction fragment length polymorphism of the human cytochrome P450IIE1 gene and susceptibility to lung cancer.* Jpn J Cancer Res, 1991. **82**(3): p. 254-6.
251. Watanabe, J., S. Hayashi, and K. Kawajiri, *Different regulation and expression of the human CYP2E1 gene due to the RsaI polymorphism in the 5'-flanking region.* J Biochem, 1994. **116**(2): p. 321-6.
252. Lind, C., et al., *DT-diaphorase: purification, properties, and function.* Methods Enzymol, 1990. **186**: p. 287-301.
253. Traver, R.D., et al., *NAD(P)H:quinone oxidoreductase gene expression in human colon carcinoma cells: characterization of a mutation which modulates DT-diaphorase activity and mitomycin sensitivity.* Cancer Res, 1992. **52**(4): p. 797-802.
254. Ross, D. and D. Siegel, *NAD(P)H:quinone oxidoreductase 1 (NQO1, DT-diaphorase), functions and pharmacogenetics.* Methods Enzymol, 2004. **382**: p. 115-44.
255. Gaedigk, A., et al., *NAD(P)H:quinone oxidoreductase: polymorphisms and allele frequencies in Caucasian, Chinese and Canadian Native Indian and Inuit populations.* Pharmacogenetics, 1998. **8**(4): p. 305-13.

256. Kelsey, K.T., et al., *Ethnic variation in the prevalence of a common NAD(P)H quinone oxidoreductase polymorphism and its implications for anti-cancer chemotherapy.* Br J Cancer, 1997. **76**(7): p. 852-4.
257. Kiyohara, C., et al., *NQO1, MPO, and the risk of lung cancer: a HuGE review.* Genet Med, 2005. **7**(7): p. 463-78.
258. Gasdaska, P.Y., H. Fisher, and G. Powis, *An alternatively spliced form of NQO1 (DT-diaphorase) messenger RNA lacking the putative quinone substrate binding site is present in human normal and tumor tissues.* Cancer Res, 1995. **55**(12): p. 2542-7.
259. Pan, S.S., et al., *NAD(P)H:quinone oxidoreductase expression and mitomycin C resistance developed by human colon cancer HCT 116 cells.* Cancer Res, 1995. **55**(2): p. 330-5.
260. Akkiz, H., et al., *No association of NAD(P)H: quinone oxidoreductase 1 (NQO1) C609T polymorphism and risk of hepatocellular carcinoma development in Turkish subjects.* Asian Pac J Cancer Prev. **11**(4): p. 1051-8.
261. Evans, D., *N-Acetyltransferase. In: Pharmacogenetics of Drug Metabolism (Kalow W, ed).* New York:Pergamon Press, 1992: p. 95-178.
262. Hein, D.W., et al., *Metabolic activation and deactivation of arylamine carcinogens by recombinant human NAT1 and polymorphic NAT2 acetyltransferases.* Carcinogenesis, 1993. **14**(8): p. 1633-8.
263. Hearse, D.J. and W.W. Weber, *Multiple N-acetyltransferases and drug metabolism. Tissue distribution, characterization and significance of mammalian N-acetyltransferase.* Biochem J, 1973. **132**(3): p. 519-26.
264. Coroneos, E. and E. Sim, *Arylamine N-acetyltransferase activity in human cultured cell lines.* Biochem J, 1993. **294 (Pt 2)**: p. 481-6.
265. Hirvonen, A., *Polymorphic NATs and cancer predisposition.* IARC Sci Publ, 1999(148): p. 251-70.
266. Grant, D.M., et al., *Monomorphic and polymorphic human arylamine N-acetyltransferases: a comparison of liver isozymes and expressed products of two cloned genes.* Mol Pharmacol, 1991. **39**(2): p. 184-91.
267. Grant DM, V.P., Avis Y, Ima A. , *Detection of a new polymorphism of human arylamine Nacetyltransferase NAT1 using p-aminosalicylic acid as an in vivo probe* J Basic Clin Physiol Pharmacol, 1992. **3**: p. S244.
268. Grant DM, V.K., Meyer UA., *In vitro metabolism of dinaline and acetyidinaline by human liver. Abstracts of 12th European Workshop on Drug Metabolism.* Basel, Switzerland, 1990. **147**.
269. Grant, D.M., et al., *Human acetyltransferase polymorphisms.* Mutat Res, 1997. **376**(1-2): p. 61-70.
270. Lin, H.J., et al., *Ethnic distribution of slow acetylator mutations in the polymorphic N-acetyltransferase (NAT2) gene.* Pharmacogenetics, 1994. **4**(3): p. 125-34.
271. Lin, H.J., et al., *Slow acetylator mutations in the human polymorphic N-acetyltransferase gene in 786 Asians, blacks, Hispanics, and whites: application to metabolic epidemiology.* Am J Hum Genet, 1993. **52**(4): p. 827-34.

272. Cascorbi, I., et al., *Arylamine N-acetyltransferase (NAT2) mutations and their allelic linkage in unrelated Caucasian individuals: correlation with phenotypic activity.* Am J Hum Genet, 1995. **57**(3): p. 581-92.
273. Mrozikiewicz, P.M., N. Drakoulis, and I. Roots, *Polymorphic arylamine N-acetyltransferase (NAT2) genes in children with insulin-dependent diabetes mellitus.* Clin Pharmacol Ther, 1994. **56**(6 Pt 1): p. 626-34.
274. Hirvonen, A., et al., *Inherited GSTM1 and NAT2 defects as concurrent risk modifiers in asbestos-related human malignant mesothelioma.* Cancer Res, 1995. **55**(14): p. 2981-3.
275. Mashimo M, S.T., Abe M, Deguchi T., *Molecular genotyping of N-acetylation polymorphism to predict phenotype.* Hum Genet 1992. **90**: p. 139-142.
276. Rothman, N., et al., *Correlation between N-acetyltransferase activity and NAT2 genotype in Chinese males.* Pharmacogenetics, 1993. **3**(5): p. 250-5.
277. WW, W., *The Acetylator Genes and Drug Response.* New York:Oxford University Press, 1987.
278. Hein, D.W., et al., *Molecular genetics and epidemiology of the NAT1 and NAT2 acetylation polymorphisms.* Cancer Epidemiol Biomarkers Prev, 2000. **9**(1): p. 29-42.
279. Risch, A., et al., *Slow N-acetylation genotype is a susceptibility factor in occupational and smoking related bladder cancer.* Hum Mol Genet, 1995. **4**(2): p. 231-6.
280. Golka, K., et al., *Occupational history and genetic N-acetyltransferase polymorphism in urothelial cancer patients of Leverkusen, Germany.* Scand J Work Environ Health, 1996. **22**(5): p. 332-8.
281. Minchin, R.F., et al., *N-and O-acetylation of aromatic and heterocyclic amine carcinogens by human monomorphic and polymorphic acetyltransferases expressed in COS-1 cells.* Biochem Biophys Res Commun, 1992. **185**(3): p. 839-44.
282. Yanagawa, Y., et al., *Stable expression of human CYP1A2 and N-acetyltransferases in Chinese hamster CHL cells: mutagenic activation of 2-amino-3-methylimidazo[4,5-f]quinoline and 2-amino-3,8-dimethylimidazo[4,5-f]quinoxaline.* Cancer Res, 1994. **54**(13): p. 3422-7.
283. Ambrosone, C.B., et al., *Cigarette smoking, N-acetyltransferase 2 genetic polymorphisms, and breast cancer risk.* JAMA, 1996. **276**(18): p. 1494-501.
284. Martinez, C., et al., *Lung cancer and mutations at the polymorphic NAT2 gene locus.* Pharmacogenetics, 1995. **5**(4): p. 207-14.
285. Cascorbi, I., et al., *Homozygous rapid arylamine N-acetyltransferase (NAT2) genotype as a susceptibility factor for lung cancer.* Cancer Res, 1996. **56**(17): p. 3961-6.
286. Saarikoski, S.T., et al., *Role of NAT2 deficiency in susceptibility to lung cancer among asbestos-exposed individuals.* Pharmacogenetics, 2000. **10**(2): p. 183-5.
287. Hirvonen A, S.S., linnainmaa K, and H.-P.K. Koskinen K, Vainio H., *GST and NAT genotypes and asbestos-associated pulmonary disorders.* J NatI Cancer Inst 1996. **88**: p. 1853-1856.

288. Agundez, J.A., et al., *Identification and prevalence study of 17 allelic variants of the human NAT2 gene in a white population.* Pharmacogenetics, 1996. **6**(5): p. 423-8.
289. Zhang, Y.W., et al., *Effects of dietary factors and the NAT2 acetylator status on gastric cancer in Koreans.* Int J Cancer, 2009. **125**(1): p. 139-45.
290. Knudson, A.G., *Two genetic hits (more or less) to cancer.* Nat Rev Cancer, 2001. **1**(2): p. 157-62.
291. Zhou, J., et al., *Gene expression profiles at different stages of human esophageal squamous cell carcinoma.* World J Gastroenterol, 2003. **9**(1): p. 9-15.
292. Mohr, S., et al., *Microarrays as cancer keys: an array of possibilities.* J Clin Oncol, 2002. **20**(14): p. 3165-75.
293. Bullinger, L., et al., *Use of gene-expression profiling to identify prognostic subclasses in adult acute myeloid leukemia.* N Engl J Med, 2004. **350**(16): p. 1605-16.
294. Valk, P.J., et al., *Prognostically useful gene-expression profiles in acute myeloid leukemia.* N Engl J Med, 2004. **350**(16): p. 1617-28.
295. Mao, X., B.D. Young, and Y.J. Lu, *The application of single nucleotide polymorphism microarrays in cancer research.* Curr Genomics, 2007. **8**(4): p. 219-28.
296. Cahan, P., et al., *Meta-analysis of microarray results: challenges, opportunities, and recommendations for standardization.* Gene, 2007. **401**(1-2): p. 12-8.
297. Murtaza, I., et al., *A study on p53 gene alterations in esophageal squamous cell carcinoma and their correlation to common dietary risk factors among population of the Kashmir valley.* World J Gastroenterol, 2006. **12**(25): p. 4033-7.
298. Schulze, A. and J. Downward, *Navigating gene expression using microarrays--a technology review.* Nat Cell Biol, 2001. **3**(8): p. E190-5.
299. Russo, G., C. Zegar, and A. Giordano, *Advantages and limitations of microarray technology in human cancer.* Oncogene, 2003. **22**(42): p. 6497-507.
300. Macgregor, P.F. and J.A. Squire, *Application of microarrays to the analysis of gene expression in cancer.* Clin Chem, 2002. **48**(8): p. 1170-7.
301. Kurian, K.M., C.J. Watson, and A.H. Wyllie, *DNA chip technology.* J Pathol, 1999. **187**(3): p. 267-71.
302. Glanzer, J.G., P.G. Haydon, and J.H. Eberwine, *Expression profile analysis of neurodegenerative disease: advances in specificity and resolution.* Neurochem Res, 2004. **29**(6): p. 1161-8.
303. Khan, J., et al., *Expression profiling in cancer using cDNA microarrays.* Electrophoresis, 1999. **20**(2): p. 223-9.
304. Sotiriou, C., et al., *Gene expression profiles derived from fine needle aspiration correlate with response to systemic chemotherapy in breast cancer.* Breast Cancer Res, 2002. **4**(3): p. R3.

305. Kim, I.J., et al., *RET oligonucleotide microarray for the detection of RET mutations in multiple endocrine neoplasia type 2 syndromes.* Clin Cancer Res, 2002. **8**(2): p. 457-63.
306. Kim, I.J., et al., *Development and applications of a beta-catenin oligonucleotide microarray: beta-catenin mutations are dominantly found in the proximal colon cancers with microsatellite instability.* Clin Cancer Res, 2003. **9**(8): p. 2920-5.
307. Andreyev, H.J., et al., *Kirsten ras mutations in patients with colorectal cancer: the 'RASCAL II' study.* Br J Cancer, 2001. **85**(5): p. 692-6.
308. Lockhart, D.J., et al., *Expression monitoring by hybridization to high-density oligonucleotide arrays.* Nat Biotechnol, 1996. **14**(13): p. 1675-80.
309. Schmidt, U. and C.G. Begley, *Cancer diagnosis and microarrays.* Int J Biochem Cell Biol, 2003. **35**(2): p. 119-24.
310. Lee, S., et al., *Identification of genes differentially expressed between gastric cancers and normal gastric mucosa with cDNA microarrays.* Cancer Lett, 2002. **184**(2): p. 197-206.
311. Raetz, E.A., et al., *Gene expression profiling. Methods and clinical applications in oncology.* Hematol Oncol Clin North Am, 2001. **15**(5): p. 911-30, ix.
312. Boussioutas, A. and D. Taupin, *Towards a molecular approach to gastric cancer management.* Intern Med J, 2001. **31**(5): p. 296-303.
313. Yasui, W., et al., *Molecular diagnosis of gastric cancer: present and future.* Gastric Cancer, 2001. **4**(3): p. 113-21.
314. Lockhart, D.J. and E.A. Winzeler, *Genomics, gene expression and DNA arrays.* Nature, 2000. **405**(6788): p. 827-36.
315. Ji, J., et al., *Comprehensive analysis of the gene expression profiles in human gastric cancer cell lines.* Oncogene, 2002. **21**(42): p. 6549-56.
316. Inoue, H., et al., *Prognostic score of gastric cancer determined by cDNA microarray.* Clin Cancer Res, 2002. **8**(11): p. 3475-9.
317. Hasegawa, S., et al., *Genome-wide analysis of gene expression in intestinal-type gastric cancers using a complementary DNA microarray representing 23,040 genes.* Cancer Res, 2002. **62**(23): p. 7012-7.
318. Dicken, B.J., et al., *Gastric adenocarcinoma: review and considerations for future directions.* Ann Surg, 2005. **241**(1): p. 27-39.
319. Liu, L.X., et al., *Profiling of differentially expressed genes in human gastric carcinoma by cDNA expression array.* World J Gastroenterol, 2002. **8**(4): p. 580-5.
320. Yasui, W., et al., *Search for new biomarkers of gastric cancer through serial analysis of gene expression and its clinical implications.* Cancer Sci, 2004. **95**(5): p. 385-92.
321. El-Rifai, W., et al., *Expression profiling of gastric adenocarcinoma using cDNA array.* Int J Cancer, 2001. **92**(6): p. 832-8.
322. Shao, Y., et al., *[Gene expression profile of human adenocarcinoma by cDNA microarray and clustering].* Zhonghua Yi Xue Yi Chuan Xue Za Zhi, 2004. **21**(2): p. 110-5.

323. Zhang, X.Q., et al., *[Gene expression profiling of diffuse-type gastric cancer by cDNA microarray].* Zhonghua Zhong Liu Za Zhi, 2006. **28**(2): p. 116-9.
324. Schuster, S.C., *Next-generation sequencing transforms today's biology.* Nat Methods, 2008. **5**(1): p. 16-8.
325. Mardis, E.R., *The impact of next-generation sequencing technology on genetics.* Trends Genet, 2008. **24**(3): p. 133-41.
326. Thomas, R.K., et al., *High-throughput oncogene mutation profiling in human cancer.* Nat Genet, 2007. **39**(3): p. 347-51.
327. Goh, L., et al., *Assessing matched normal and tumor pairs in next-generation sequencing studies.* PLoS One. **6**(3): p. e17810.
328. Stratton, M.R., P.J. Campbell, and P.A. Futreal, *The cancer genome.* Nature, 2009. **458**(7239): p. 719-24.
329. Phukan RK, Z.E., Hazarika N, Baruah D, Mahanta J, *High prevalence of stomach cancer among the people of Mizoram, India.* Current Science, 2004. **87**: p. 285-286
330. Bhattacharjee A, C.A., Purkaystha P. , *Prevalance of head and neck cancers in the north east- an institutional study.* Indian Journal of otolaryngology and head and neck surgery 2006. **58**: p. 15-19.
331. ICMR, *Consolidated report of population based cancer registries 2001-2004, Incidence and distribution of cancer*, in *Indian council of medical research (ICMR), Bangalore* 2006.
332. Phukan RK, Z.E., Narain K, Hazarika NC, Mahanta J, *Tobacco use and stomach cancer in Mizoram, India.* Cancer Epidemiol Biomarkers Prev 2005. **14**: p. 1892-1896.
333. Guengerich, F.P., *Enzymatic oxidation of xenobiotic chemicals.* Crit Rev Biochem Mol Biol, 1990. **25**(2): p. 97-153.
334. Sheehan, D., et al., *Structure, function and evolution of glutathione transferases: implications for classification of non-mammalian members of an ancient enzyme superfamily.* Biochem J, 2001. **360**(Pt 1): p. 1-16.
335. Watson, M.A., et al., *Human glutathione S-transferase P1 polymorphisms: relationship to lung tissue enzyme activity and population frequency distribution.* Carcinogenesis, 1998. **19**(2): p. 275-80.
336. Jain, M., et al., *GSTT1, GSTM1 and GSTP1 genetic polymorphisms and interaction with tobacco, alcohol and occupational exposure in esophageal cancer patients from North India.* Cancer Lett, 2006. **242**(1): p. 60-7.
337. Buch, S.C., P.N. Notani, and R.A. Bhisey, *Polymorphism at GSTM1, GSTM3 and GSTT1 gene loci and susceptibility to oral cancer in an Indian population.* Carcinogenesis, 2002. **23**(5): p. 803-7.
338. Sobti, R.C., et al., *Combined effect of GSTM1, GSTT1 and GSTP1 polymorphisms on histological subtypes of lung cancer.* Biomarkers, 2008. **13**(3): p. 282-95.
339. Anantharaman, D., et al., *Susceptibility to oral cancer by genetic polymorphisms at CYP1A1, GSTM1 and GSTT1 loci among Indians: tobacco exposure as a risk modulator.* Carcinogenesis, 2007. **28**(7): p. 1455-62.
340. Soya, S.S., et al., *Genetic polymorphisms of glutathione-S-transferase genes (GSTM1, GSTT1 and GSTP1) and upper aerodigestive tract cancer risk*

among smokers, tobacco chewers and alcoholics in an Indian population. Eur J Cancer, 2007. **43**(18): p. 2698-706.

341. Singh, M., et al., *Association of genetic polymorphisms in glutathione S-transferases and susceptibility to head and neck cancer.* Mutat Res, 2008. **638**(1-2): p. 184-94.
342. Thoudam, R.D., et al., *Distribution of glutathione S-transferase T1 and M1 genes polymorphisms in North East Indians: a potential report.* Genet Test Mol Biomarkers. **14**(2): p. 163-9.
343. Chatterjee, S., et al., *Prevalence of CYP1A1 and GST polymorphisms in the population of northeastern India and susceptibility of oral cancer.* Oncol Res, 2009. **17**(9): p. 397-403.
344. Sambrook Joseph, D.W.R., *Molecular Cloning, A Laboratory Manual.* Third Edition, 2001. **Volume 1**(www.MolecularCloning.com): p. 6.4 - 6.11.
345. Hatagima, A., et al., *Glutathione S-transferase polymorphisms and oral cancer: a case-control study in Rio de Janeiro, Brazil.* Oral Oncol, 2008. **44**(2): p. 200-7.
346. Pemble, S., et al., *Human glutathione S-transferase theta (GSTT1): cDNA cloning and the characterization of a genetic polymorphism.* Biochem J, 1994. **300 (Pt 1)**: p. 271-6.
347. Lear, J.T., et al., *Detoxifying enzyme genotypes and susceptibility to cutaneous malignancy.* Br J Dermatol, 2000. **142**(1): p. 8-15.
348. Ye, Z., H. Song, and Y. Guo, *Glutathione S-transferase M1, T1 status and the risk of head and neck cancer: a meta-analysis.* J Med Genet, 2004. **41**(5): p. 360-5.
349. Cho CH, V.P.e., *Alcohol, tobacco and cancer. Vol, 2006. Karger, Basel, Switzerland.* 2006.
350. Gattas, G.J., et al., *Genetic polymorphisms of CYP1A1, CYP2E1, GSTM1, and GSTT1 associated with head and neck cancer.* Head Neck, 2006. **28**(9): p. 819-26.
351. Duarte, E.C., et al., *Genetic polymorphisms of carcinogen metabolizing enzymes are associated with oral leukoplakia development and p53 overexpression.* Anticancer Res, 2008. **28**(2A): p. 1101-6.
352. Lai, K.C., et al., *Glutathione S-transferase M1 gene null genotype and gastric cancer risk in Taiwan.* Hepatogastroenterology, 2005. **52**(66): p. 1916-9.
353. Martinez, C., et al., *Glutathione S-transferases mu 1, theta 1, pi 1, alpha 1 and mu 3 genetic polymorphisms and the risk of colorectal and gastric cancers in humans.* Pharmacogenomics, 2006. **7**(5): p. 711-8.
354. Tripathi, S., et al., *Gastric carcinogenesis: Possible role of polymorphisms of GSTM1, GSTT1, and GSTP1 genes.* Scand J Gastroenterol, 2008. **43**(4): p. 431-9.
355. Lan, Q., et al., *Indoor coal combustion emissions, GSTM1 and GSTT1 genotypes, and lung cancer risk: a case-control study in Xuan Wei, China.* Cancer Epidemiol Biomarkers Prev, 2000. **9**(6): p. 605-8.
356. Chen, H.C., et al., *Genetic polymorphisms of phase II metabolic enzymes and lung cancer susceptibility in a population of Central South China.* Dis Markers, 2006. **22**(3): p. 141-52.

357. Losi-Guembarovski, R., et al., *Lack of association among polymorphic xenobiotic-metabolizing enzyme genotypes and the occurrence and progression of oral carcinoma in a Brazilian population.* Anticancer Res, 2008. **28**(2A): p. 1023-8.
358. Wideroff, L., et al., *GST, NAT1, CYP1A1 polymorphisms and risk of esophageal and gastric adenocarcinomas.* Cancer Detect Prev, 2007. **31**(3): p. 233-6.
359. Sorensen, M., et al., *Interactions between GSTM1, GSTT1 and GSTP1 polymorphisms and smoking and intake of fruit and vegetables in relation to lung cancer.* Lung Cancer, 2007. **55**(2): p. 137-44.
360. Honma, H.N., et al., *Influence of p53 codon 72 exon 4, GSTM1, GSTT1 and GSTP1*B polymorphisms in lung cancer risk in a Brazilian population.* Lung Cancer, 2008. **61**(2): p. 152-62.
361. Heagerty, A.H., et al., *Glutathione S-transferase GSTM1 phenotypes and protection against cutaneous tumours.* Lancet, 1994. **343**(8892): p. 266-8.
362. Roodi, N., et al., *Association of homozygous wild-type glutathione S-transferase M1 genotype with increased breast cancer risk.* Cancer Res, 2004. **64**(4): p. 1233-6.
363. Ramachandran, S., et al., *Presentation with multiple cutaneous basal cell carcinomas: association of glutathione S-transferase and cytochrome P450 genotypes with clinical phenotype.* Cancer Epidemiol Biomarkers Prev, 1999. **8**(1): p. 61-7.
364. Lan, Q., et al., *Glutathione S-transferase genotypes and stomach cancer in a population-based case-control study in Warsaw, Poland.* Pharmacogenetics, 2001. **11**(8): p. 655-61.
365. Evans, A.J., et al., *Polymorphisms of GSTT1 and related genes in head and neck cancer risk.* Head Neck, 2004. **26**(1): p. 63-70.
366. Kim, W.J., et al., *GSTT1-null genotype is a protective factor against bladder cancer.* Urology, 2002. **60**(5): p. 913-8.
367. Garcia-Closas, M., et al., *Glutathione S-transferase mu and theta polymorphisms and breast cancer susceptibility.* J Natl Cancer Inst, 1999. **91**(22): p. 1960-4.
368. Malik, M.A., et al., *Role of xenobiotic-metabolizing enzyme gene polymorphisms and interactions with environmental factors in susceptibility to gastric cancer in Kashmir Valley.* J Gastrointest Cancer, 2009. **40**(1-2): p. 26-32.
369. Keating, A.F., et al., *Dual protective role for glutathione S-transferase class pi against VCD-induced ovotoxicity in the rat ovary.* Toxicol Appl Pharmacol. **247**(2): p. 71-5.
370. Singh, S., et al., *Influence of CYP2C9, GSTM1, GSTT1 and NAT2 genetic polymorphisms on DNA damage in workers occupationally exposed to organophosphate pesticides.* Mutat Res. **741**(1-2): p. 101-8.
371. Syamala, V.S., et al., *Influence of germline polymorphisms of GSTT1, GSTM1, and GSTP1 in familial versus sporadic breast cancer susceptibility and survival.* Fam Cancer, 2008. **7**(3): p. 213-20.

372. Reszka, E., et al., *Glutathione S-transferase M1 and P1 metabolic polymorphism and lung cancer predisposition.* Neoplasma, 2003. **50**(5): p. 357-62.
373. Wu, T., et al., *Measurement of GSTP1 promoter methylation in body fluids may complement PSA screening: a meta-analysis.* Br J Cancer. **105**(1): p. 65-73.
374. Xiao, H. and S.V. Singh, *p53 regulates cellular responses to environmental carcinogen benzo[a]pyrene-7,8-diol-9,10-epoxide in human lung cancer cells.* Cell Cycle, 2007. **6**(14): p. 1753-61.
375. Katiyar, S., et al., *Polymorphism of the p53 codon 72 Arg/Pro and the risk of HPV type 16/18-associated cervical and oral cancer in India.* Mol Cell Biochem, 2003. **252**(1-2): p. 117-24.
376. Matlashewski, G.J., et al., *Primary structure polymorphism at amino acid residue 72 of human p53.* Mol Cell Biol, 1987. **7**(2): p. 961-3.
377. Thomas, M., et al., *Two polymorphic variants of wild-type p53 differ biochemically and biologically.* Mol Cell Biol, 1999. **19**(2): p. 1092-100.
378. Zhang, Z.W., et al., *Age-associated increase of codon 72 Arginine p53 frequency in gastric cardia and non-cardia adenocarcinoma.* Clin Cancer Res, 2003. **9**(6): p. 2151-6.
379. Song, H.R., et al., *p53 codon 72 polymorphism in patients with gastric and colorectal cancer in a Korean population.* Gastric Cancer. **14**(3): p. 242-8.
380. Gomes de Souza, L., et al., *P53 Arg72Pro polymorphism in gastric cancer patients.* J Gastrointest Cancer, 2009. **40**(1-2): p. 41-5.
381. Weston, A., et al., *Allelic frequency of a p53 polymorphism in human lung cancer.* Cancer Epidemiol Biomarkers Prev, 1992. **1**(6): p. 481-3.
382. To-Figueras, J., et al., *Glutathione-S-Transferase M1 and codon 72 p53 polymorphisms in a northwestern Mediterranean population and their relation to lung cancer susceptibility.* Cancer Epidemiol Biomarkers Prev, 1996. **5**(5): p. 337-42.
383. Kim, J.M., et al., *[p53 Codon 72 and 16-bp duplication polymorphisms of gastric cancer in Koreans].* Korean J Gastroenterol, 2007. **50**(5): p. 292-8.
384. Alpizar-Alpizar, W., et al., *[Association of the p53 codon 72 polymorphism to gastric cancer risk in a hight risk population of Costa Rica].* Rev Biol Trop, 2005. **53**(3-4): p. 317-24.
385. Mojtahedi, Z., et al., *p 53 codon 72 polymorphism in stomach and colorectal adenocarcinomas in Iranian patients.* Indian J Cancer. **47**(1): p. 31-4.
386. Drummond, S.N., et al., *TP53 codon 72 polymorphism in oral squamous cell carcinoma.* Anticancer Res, 2002. **22**(6A): p. 3379-81.
387. Kietthubthew, S., et al., *The p53 codon 72 polymorphism and risk of oral cancer in Southern Thailand.* Asian Pac J Cancer Prev, 2003. **4**(3): p. 209-14.
388. Bergamaschi, D., et al., *p53 polymorphism influences response in cancer chemotherapy via modulation of p73-dependent apoptosis.* Cancer Cell, 2003. **3**(4): p. 387-402.
389. Bonafe, M., et al., *Retention of the p53 codon 72 arginine allele is associated with a reduction of disease-free and overall survival in arginine/proline*

heterozygous breast cancer patients. Clin Cancer Res, 2003. **9**(13): p. 4860-4.
390. Ihsan, R., et al., *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): p. 163-71.
391. Parsonnet, J., *Helicobacter pylori.* Infect Dis Clin North Am, 1998. **12**(1): p. 185-97.
392. Taylor, D.N. and M.J. Blaser, *The epidemiology of Helicobacter pylori infection.* Epidemiol Rev, 1991. **13**: p. 42-59.
393. Ernst, P.B. and B.D. Gold, *Helicobacter pylori in childhood: new insights into the immunopathogenesis of gastric disease and implications for managing infection in children.* J Pediatr Gastroenterol Nutr, 1999. **28**(5): p. 462-73.
394. Singh, K. and U.C. Ghoshal, *Causal role of Helicobacter pylori infection in gastric cancer: an Asian enigma.* World J Gastroenterol, 2006. **12**(9): p. 1346-51.
395. Lam, S.K., *9th Seah Cheng Siang Memorial Lecture: gastric cancer--where are we now?* Ann Acad Med Singapore, 1999. **28**(6): p. 881-9.
396. Piazuelo, M.B., M. Epplein, and P. Correa, *Gastric cancer: an infectious disease.* Infect Dis Clin North Am. **24**(4): p. 853-69, vii.
397. Kuipers, E.J., J.C. Thijs, and H.P. Festen, *The prevalence of Helicobacter pylori in peptic ulcer disease.* Aliment Pharmacol Ther, 1995. **9 Suppl 2**: p. 59-69.
398. Van Zanten, S.J., M.F. Dixon, and A. Lee, *The gastric transitional zones: neglected links between gastroduodenal pathology and helicobacter ecology.* Gastroenterology, 1999. **116**(5): p. 1217-29.
399. Czinn, S.J., *Serodiagnosis of Helicobacter pylori in pediatric patients.* J Pediatr Gastroenterol Nutr, 1999. **28**(2): p. 132-4.
400. Koehler, C.I., et al., *Helicobacter pylori genotyping in gastric adenocarcinoma and MALT lymphoma by multiplex PCR analyses of paraffin wax embedded tissues.* Mol Pathol, 2003. **56**(1): p. 36-42.
401. *An international association between Helicobacter pylori infection and gastric cancer. The EUROGAST Study Group.* Lancet, 1993. **341**(8857): p. 1359-62.
402. Hansson, L.E., et al., *Helicobacter pylori infection: independent risk indicator of gastric adenocarcinoma.* Gastroenterology, 1993. **105**(4): p. 1098-103.
403. Parsonnet, J., et al., *Helicobacter pylori infection and gastric lymphoma.* N Engl J Med, 1994. **330**(18): p. 1267-71.
404. Wong, B.C., et al., *Helicobacter pylori eradication to prevent gastric cancer in a high-risk region of China: a randomized controlled trial.* JAMA, 2004. **291**(2): p. 187-94.
405. Cutler, A.F., et al., *Accuracy of invasive and noninvasive tests to diagnose Helicobacter pylori infection.* Gastroenterology, 1995. **109**(1): p. 136-41.
406. Khanna, B., et al., *Use caution with serologic testing for Helicobacter pylori infection in children.* J Infect Dis, 1998. **178**(2): p. 460-5.

407. Mishra, K.K., et al., *UreC PCR based diagnosis of Helicobacter pylori infection and detection of cag A gene in gastric biopsies.* Indian J Pathol Microbiol, 2002. **45**(1): p. 31-7.
408. Fabre, R., et al., *Polymerase chain reaction assay for the detection of Helicobacter pylori in gastric biopsy specimens: comparison with culture, rapid urease test, and histopathological tests.* Gut, 1994. **35**(7): p. 905-8.
409. Dye, K.R., et al., *Ultrastructure of another spiral organism associated with human gastritis.* Dig Dis Sci, 1989. **34**(11): p. 1787-91.
410. Wahlfors, J., et al., *Development of a rapid PCR method for identification of Helicobacter pylori in dental plaque and gastric biopsy specimens.* Eur J Clin Microbiol Infect Dis, 1995. **14**(9): p. 780-6.
411. McNulty, C.A., et al., *Detection of Campylobacter pylori by the biopsy urease test: an assessment in 1445 patients.* Gut, 1989. **30**(8): p. 1058-62.
412. Mobley, H.L., L.T. Hu, and P.A. Foxal, *Helicobacter pylori urease: properties and role in pathogenesis.* Scand J Gastroenterol Suppl, 1991. **187**: p. 39-46.
413. Graham, D.Y., *Helicobacter pylori infection is the primary cause of gastric cancer.* J Gastroenterol, 2000. **35 Suppl 12**: p. 90-7.
414. Yamamoto, Y., *PCR in diagnosis of infection: detection of bacteria in cerebrospinal fluids.* Clin Diagn Lab Immunol, 2002. **9**(3): p. 508-14.
415. Jensen, A.K., L.P. Andersen, and C.H. Wachmann, *Evaluation of eight commercial kits for Helicobacter pylori IgG antibody detection.* APMIS, 1993. **101**(10): p. 795-801.
416. Locatelli, A., et al., *Detection of anti-Helicobacter pylori antibodies in serum and duodenal fluid in peptic gastroduodenal disease.* World J Gastroenterol, 2004. **10**(20): p. 2997-3000.
417. Misawa, K., et al., *[Clinical and etiological studies of IgG antibodies to Helicobacter pylori detected by enzyme-linked immunosorbent assay].* Rinsho Byori, 1995. **43**(4): p. 375-80.
418. *cagA-Positive Helicobacter pylori Populations in China and the Netherlands Are Distinct.* INFECTION AND IMMUNITY, 1998. **66**(5): p. 1822–1826.
419. Clayton, C., K. Kleanthous, and S. Tabaqchali, *Detection and identification of Helicobacter pylori by the polymerase chain reaction.* J Clin Pathol, 1991. **44**(6): p. 515-6.
420. Smith, S.I., et al., *Comparison of three PCR methods for detection of Helicobacter pylori DNA and detection of cagA gene in gastric biopsy specimens.* World J Gastroenterol, 2004. **10**(13): p. 1958-60.
421. Ho, S.A., et al., *Direct polymerase chain reaction test for detection of Helicobacter pylori in humans and animals.* J Clin Microbiol, 1991. **29**(11): p. 2543-9.
422. Kato, M. and M. Asaka, *Recent knowledge of the relationship between Helicobacter pylori and gastric cancer and recent progress of gastroendoscopic diagnosis and treatment for gastric cancer.* Jpn J Clin Oncol. **40**(9): p. 828-37.
423. Covacci, A., et al., *Helicobacter pylori virulence and genetic geography.* Science, 1999. **284**(5418): p. 1328-33.

424. Perez-Perez, G.I., D. Rothenbacher, and H. Brenner, *Epidemiology of Helicobacter pylori infection.* Helicobacter, 2004. **9 Suppl 1**: p. 1-6.
425. Pounder, R.E. and D. Ng, *The prevalence of Helicobacter pylori infection in different countries.* Aliment Pharmacol Ther, 1995. **9 Suppl 2**: p. 33-9.
426. Malaty, H.M. and D.Y. Graham, *Importance of childhood socioeconomic status on the current prevalence of Helicobacter pylori infection.* Gut, 1994. **35**(6): p. 742-5.
427. Perez-Perez, G.I., et al., *Seroprevalence of Helicobacter pylori in New York City populations originating in East Asia.* J Urban Health, 2005. **82**(3): p. 510-6.
428. Tsai, C.J., et al., *Helicobacter pylori infection in different generations of Hispanics in the San Francisco Bay Area.* Am J Epidemiol, 2005. **162**(4): p. 351-7.
429. Genta, R.M., *Review article: after gastritis--an imaginary journey into a Helicobacter-free world.* Aliment Pharmacol Ther, 2002. **16 Suppl 4**: p. 89-94.
430. Fiedorek, S.C., et al., *Factors influencing the epidemiology of Helicobacter pylori infection in children.* Pediatrics, 1991. **88**(3): p. 578-82.
431. HM., M., *Epidemiology of Infection.Helicobacter pylori: physiology and genetics. .* Washington DC: ASM Press, 2001: p. 7-18.
432. Asrat, D., et al., *Prevalence of Helicobacter pylori infection among adult dyspeptic patients in Ethiopia.* Ann Trop Med Parasitol, 2004. **98**(2): p. 181-9.
433. Kaplan, C., *Use of the laboratory walker.* Butterworth Publishers, 1990: p. 40-81990.
434. Goodwin, C.S. and J.A. Armstrong, *Microbiological aspects of Helicobacter pylori (Campylobacter pylori).* Eur J Clin Microbiol Infect Dis, 1990. **9**(1): p. 1-13.
435. Gulbis, B. and P. Galand, *Immunodetection of the p21-ras products in human normal and preneoplastic tissues and solid tumors: a review.* Hum Pathol, 1993. **24**(12): p. 1271-85.
436. Goodwin, C.S., et al., *Unusual cellular fatty acids and distinctive ultrastructure in a new spiral bacterium (Campylobacter pyloridis) from the human gastric mucosa.* J Med Microbiol, 1985. **19**(2): p. 257-67.
437. Ferlay, J., et al., *Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008.* Int J Cancer. **127**(12): p. 2893-917.
438. Boussioutas, A., et al., *Distinctive patterns of gene expression in premalignant gastric mucosa and gastric cancer.* Cancer Res, 2003. **63**(10): p. 2569-77.
439. Sharma, A. and V. Radhakrishnan, *Gastric cancer in India.* Indian J Med Paediatr Oncol. **32**(1): p. 12-6.
440. Dr S.K. Bhattacharya, D.N.K.G., *North East Population Based Cancer Registries, Incidence and Distribution of Cancer*. SECOND REPORT : 2005 - 2006. NATIONAL CANCER REGISTRY PROGRAMME. Indian Council of Medical Research Bangalore, India.

441. Bornschein, J. and P. Malfertheiner, *Gastric carcinogenesis.* Langenbecks Arch Surg. **396**(6): p. 729-42.
442. Tsugane, S., *Salt, salted food intake, and risk of gastric cancer: epidemiologic evidence.* Cancer Sci, 2005. **96**(1): p. 1-6.
443. Zhang, Z. and X. Zhang, *Salt taste preference, sodium intake and gastric cancer in china.* Asian Pac J Cancer Prev. **12**(5): p. 1207-10.
444. Sanz-Ortega, J., et al., *Comparative study of tumor angiogenesis and immunohistochemistry for p53, c-ErbB2, c-myc and EGFr as prognostic factors in gastric cancer.* Histol Histopathol, 2000. **15**(2): p. 455-62.
445. So, J.B., et al., *Expression of cell-cycle regulators p27 and cyclin E correlates with survival in gastric carcinoma patients.* J Surg Res, 2000. **94**(1): p. 56-60.
446. Chu, Y.Q., et al., *Relationship between cell adhesion molecules expression and the biological behavior of gastric carcinoma.* World J Gastroenterol, 2008. **14**(13): p. 1990-6.
447. Shimada, S., et al., *Synergistic tumour suppressor activity of E-cadherin and p53 in a conditional mouse model for metastatic diffuse-type gastric cancer.* Gut.
448. Zhu, M. and S. Zhao, *Candidate gene identification approach: progress and challenges.* Int J Biol Sci, 2007. **3**(7): p. 420-7.
449. Srivastava, K., et al., *Candidate gene studies in gallbladder cancer: A systematic review and meta-analysis.* Mutat Res. **728**(1-2): p. 67-79.
450. Miyaki, K., et al., *High throughput multiple combination extraction from large scale polymorphism data by exact tree method.* J Hum Genet, 2004. **49**(9): p. 455-62.
451. Wang, J.Y., et al., *Multiple molecular markers as predictors of colorectal cancer in patients with normal perioperative serum carcinoembryonic antigen levels.* Clin Cancer Res, 2007. **13**(8): p. 2406-13.
452. Lu, K.H., et al., *Selection of potential markers for epithelial ovarian cancer with gene expression arrays and recursive descent partition analysis.* Clin Cancer Res, 2004. **10**(10): p. 3291-300.
453. Pittman, J., et al., *Integrated modeling of clinical and gene expression information for personalized prediction of disease outcomes.* Proc Natl Acad Sci U S A, 2004. **101**(22): p. 8431-6.
454. Sotiriou, C., et al., *Gene expression profiling in breast cancer: understanding the molecular basis of histologic grade to improve prognosis.* J Natl Cancer Inst, 2006. **98**(4): p. 262-72.
455. Lynch, H.T., et al., *Gastric cancer: new genetic developments.* J Surg Oncol, 2005. **90**(3): p. 114-33; discussion 133.
456. Yamaoka, Y., M. Kato, and M. Asaka, *Geographic differences in gastric cancer incidence can be explained by differences between Helicobacter pylori strains.* Intern Med, 2008. **47**(12): p. 1077-83.
457. Aragones, N., et al., *The striking geographical pattern of gastric cancer mortality in Spain: environmental hypotheses revisited.* BMC Cancer, 2009. **9**: p. 316.

458. Yadav, D.S., et al., *Polymorphisms of glutathione-S-transferase genes and the risk of aerodigestive tract cancers in the Northeast Indian population.* Genet Test Mol Biomarkers. **14**(5): p. 715-23.
459. Kaushal, M., et al., *Betel quid chewing as an environmental risk factor for breast cancer.* Mutat Res. **703**(2): p. 143-8.
460. Yeh, J.M., et al., *Effects of Helicobacter pylori infection and smoking on gastric cancer incidence in China: a population-level analysis of trends and projections.* Cancer Causes Control, 2009. **20**(10): p. 2021-9.
461. Lubin, J.H., et al., *Cigarette smoking and cancer risk: modeling total exposure and intensity.* Am J Epidemiol, 2007. **166**(4): p. 479-89.
462. Ji, B.T., et al., *Tobacco smoking and colorectal hyperplastic and adenomatous polyps.* Cancer Epidemiol Biomarkers Prev, 2006. **15**(5): p. 897-901.
463. *Tobacco smoke and involuntary smoking.* IARC Monogr Eval Carcinog Risks Hum, 2004. **83**: p. 1-1438.
464. Nibau, C., et al., *Arabidopsis and Tobacco superman regulate hormone signalling and mediate cell proliferation and differentiation.* J Exp Bot. **62**(3): p. 949-61.
465. Fernandes, A.M., et al., *Tobacco and inflammation effects in immunoexpression of hMSH2 and hMLH1 in epithelium of oral mucosa.* Anticancer Res, 2007. **27**(4B): p. 2433-7.
466. Ramage, L., A.C. Jones, and C.J. Whelan, *Induction of apoptosis with tobacco smoke and related products in A549 lung epithelial cells in vitro.* J Inflamm (Lond), 2006. **3**: p. 3.
467. Gealy, R., et al., *Comparison of mutations in the p53 and K-ras genes in lung carcinomas from smoking and nonsmoking women.* Cancer Epidemiol Biomarkers Prev, 1999. **8**(4 Pt 1): p. 297-302.
468. Husgafvel-Pursiainen, K. and A. Kannio, *Cigarette smoking and p53 mutations in lung cancer and bladder cancer.* Environ Health Perspect, 1996. **104 Suppl 3**: p. 553-6.
469. Wistuba, II, A.F. Gazdar, and J.D. Minna, *Molecular genetics of small cell lung carcinoma.* Semin Oncol, 2001. **28**(2 Suppl 4): p. 3-13.
470. Massion, P.P. and D.P. Carbone, *The molecular basis of lung cancer: molecular abnormalities and therapeutic implications.* Respir Res, 2003. **4**: p. 12.
471. Asaka, M., et al., *What role does Helicobacter pylori play in gastric cancer?* Gastroenterology, 1997. **113**(6 Suppl): p. S56-60.
472. Algood, H.M. and T.L. Cover, *Helicobacter pylori persistence: an overview of interactions between H. pylori and host immune defenses.* Clin Microbiol Rev, 2006. **19**(4): p. 597-613.
473. Wilson, K.T. and J.E. Crabtree, *Immunology of Helicobacter pylori: insights into the failure of the immune response and perspectives on vaccine studies.* Gastroenterology, 2007. **133**(1): p. 288-308.
474. Roth, K.A., et al., *Cellular immune responses are essential for the development of Helicobacter felis-associated gastric pathology.* J Immunol, 1999. **163**(3): p. 1490-7.

475. Chen, X., et al., *Variation in gene expression patterns in human gastric cancers.* Mol Biol Cell, 2003. **14**(8): p. 3208-15.
476. Katoh, M., *Comparative integromics on non-canonical WNT or planar cell polarity signaling molecules: transcriptional mechanism of PTK7 in colorectal cancer and that of SEMA6A in undifferentiated ES cells.* Int J Mol Med, 2007. **20**(3): p. 405-9.
477. Zhang, H. and Y. Xue, *Wnt pathway is involved in advanced gastric carcinoma.* Hepatogastroenterology, 2008. **55**(84): p. 1126-30.
478. Oguma, K., H. Oshima, and M. Oshima, *Inflammation, tumor necrosis factor and Wnt promotion in gastric cancer development.* Future Oncol. **6**(4): p. 515-26.
479. Katoh, M., *Notch signaling in gastrointestinal tract (review).* Int J Oncol, 2007. **30**(1): p. 247-51.
480. Kim, T.H. and R.A. Shivdasani, *Notch signaling in stomach epithelial stem cell homeostasis.* J Exp Med. **208**(4): p. 677-88.
481. Yoo, Y.A., et al., *Sonic hedgehog signaling promotes motility and invasiveness of gastric cancer cells through TGF-beta-mediated activation of the ALK5-Smad 3 pathway.* Carcinogenesis, 2008. **29**(3): p. 480-90.
482. Lee, S.Y., et al., *Sonic hedgehog expression in gastric cancer and gastric adenoma.* Oncol Rep, 2007. **17**(5): p. 1051-5.
483. Peng, S., et al., *Effects of Wnt5a protein on proliferation and apoptosis in JAR choriocarcinoma cells.* Mol Med Report. **4**(1): p. 99-104.
484. Yeh, T.S., et al., *The activated Notch1 signal pathway is associated with gastric cancer progression through cyclooxygenase-2.* Cancer Res, 2009. **69**(12): p. 5039-48.
485. Martin, J., et al., *The role of sonic hedgehog reemergence during gastric cancer.* Dig Dis Sci. **55**(6): p. 1516-24.
486. Naumann, M. and J.E. Crabtree, *Helicobacter pylori-induced epithelial cell signalling in gastric carcinogenesis.* Trends Microbiol, 2004. **12**(1): p. 29-36.
487. Fiore, M.C., et al., *Statins reverse renal inflammation and endothelial dysfunction induced by chronic high salt intake.* Am J Physiol Renal Physiol. **301**(2): p. F263-70.
488. Park, D.W., et al., *Phenotypic Differences of Gastric Cancer according to the Helicobacter pylori Infection in Korean Patients.* J Gastric Cancer. **10**(4): p. 168-74.
489. Kato, S., et al., *Helicobacter pylori infection-negative gastric cancer in Japanese hospital patients: incidence and pathological characteristics.* Cancer Sci, 2007. **98**(6): p. 790-4.
490. Crockett, D.K., et al., *Predicting phenotypic severity of uncertain gene variants in the RET proto-oncogene.* PLoS One. **6**(3): p. e18380.
491. Thomas, S.M., M. Hagel, and C.E. Turner, *Characterization of a focal adhesion protein, Hic-5, that shares extensive homology with paxillin.* J Cell Sci, 1999. **112 (Pt 2)**: p. 181-90.

492. Jagadeeswaran, R., et al., *Paxillin is a target for somatic mutations in lung cancer: implications for cell growth and invasion.* Cancer Res, 2008. **68**(1): p. 132-42.
493. Daly, A.K., *Genome-wide association studies in pharmacogenomics.* Nat Rev Genet. **11**(4): p. 241-6.
494. Daly, A.K., *Pharmacogenetics and human genetic polymorphisms.* Biochem J. **429**(3): p. 435-49.
495. Zhang, J., et al., *The impact of next-generation sequencing on genomics.* J Genet Genomics. **38**(3): p. 95-109.
496. Wall, P.K., et al., *Comparison of next generation sequencing technologies for transcriptome characterization.* BMC Genomics, 2009. **10**: p. 347.
497. Peltomaki, P., *Mutations and epimutations in the origin of cancer.* Exp Cell Res.
498. *Three-Year Reports of Population Based Cancer Registries* National cancer Registry Programme, 2006-2008.
499. Mardis, E.R., *Anticipating the 1,000 dollar genome.* Genome Biol, 2006. **7**(7): p. 112.
500. Mardis, E.R., *New strategies and emerging technologies for massively parallel sequencing: applications in medical research.* Genome Med, 2009. **1**(4): p. 40.
501. Tucker, T., M. Marra, and J.M. Friedman, *Massively parallel sequencing: the next big thing in genetic medicine.* Am J Hum Genet, 2009. **85**(2): p. 142-54.
502. Metzker, M.L., *Sequencing technologies - the next generation.* Nat Rev Genet. **11**(1): p. 31-46.
503. Eom, S. and C. Lee, *Functions of intronic nucleotide variants in the gene encoding pleckstrin homology like domain beta 2 (PHLDB2) on susceptibility to vascular dementia.* World J Biol Psychiatry.
504. Lacher, M.D., et al., *Rheb activates AMPK and reduces p27Kip1 levels in Tsc2-null cells via mTORC1-independent mechanisms: implications for cell proliferation and tumorigenesis.* Oncogene. **29**(50): p. 6543-56.
505. Mavrou, A., et al., *The ATM gene and ataxia telangiectasia.* Anticancer Res, 2008. **28**(1B): p. 401-5.
506. Nobukini, T. and G. Thomas, *The mTOR/S6K signalling pathway: the role of the TSC1/2 tumour suppressor complex and the proto-oncogene Rheb.* Novartis Found Symp, 2004. **262**: p. 148-54; discussion 154-9, 265-8.
507. Lacher, M.D., R.J. Pincheira, and A.F. Castro, *Consequences of interrupted Rheb-to-AMPK feedback signaling in tuberous sclerosis complex and cancer.* Small Gtpases. **2**(4): p. 211-216.
508. Broeks, A., et al., *ATM-heterozygous germline mutations contribute to breast cancer-susceptibility.* Am J Hum Genet, 2000. **66**(2): p. 494-500.
509. Malmer, B.S., et al., *Genetic variation in p53 and ATM haplotypes and risk of glioma and meningioma.* J Neurooncol, 2007. **82**(3): p. 229-37.
510. Kheirollahi, M., et al., *Expression of cyclin D2, P53, Rb and ATM cell cycle genes in brain tumors.* Med Oncol. **28**(1): p. 7-14.
511. Boultwood, J., *Ataxia telangiectasia gene mutations in leukaemia and lymphoma.* J Clin Pathol, 2001. **54**(7): p. 512-6.

512. Gumy-Pause, F., P. Wacker, and A.P. Sappino, *ATM gene and lymphoid malignancies.* Leukemia, 2004. **18**(2): p. 238-42.
513. Lee, S.A., et al., *Antioxidant vitamins intake, ataxia telangiectasia mutated (ATM) genetic polymorphisms, and breast cancer risk.* Nutr Cancer. **62**(8): p. 1087-94.
514. Endoh, M., et al., *RASSF2, a potential tumour suppressor, is silenced by CpG island hypermethylation in gastric cancer.* Br J Cancer, 2005. **93**(12): p. 1395-9.
515. Akino, K., et al., *The Ras effector RASSF2 is a novel tumor-suppressor gene in human colorectal cancer.* Gastroenterology, 2005. **129**(1): p. 156-69.
516. Zhang, Z., et al., *Inactivation of RASSF2A by promoter methylation correlates with lymph node metastasis in nasopharyngeal carcinoma.* Int J Cancer, 2007. **120**(1): p. 32-8.
517. Liao, X., et al., *Hypermethylation of RAS effector related genes and DNA methyltransferase 1 expression in endometrial carcinogenesis.* Int J Cancer, 2008. **123**(2): p. 296-302.
518. Kaira, K., et al., *Epigenetic inactivation of the RAS-effector gene RASSF2 in lung cancers.* Int J Oncol, 2007. **31**(1): p. 169-73.
519. Kokko, A., et al., *EPHB2 germline variants in patients with colorectal cancer or hyperplastic polyposis.* BMC Cancer, 2006. **6**: p. 145.
520. Oba, S.M., et al., *Genomic structure and loss of heterozygosity of EPHB2 in colorectal cancer.* Cancer Lett, 2001. **164**(1): p. 97-104.
521. Herath, N.I. and A.W. Boyd, *The role of Eph receptors and ephrin ligands in colorectal cancer.* Int J Cancer. **126**(9): p. 2003-11.
522. Davalos, V., et al., *High EPHB2 mutation rate in gastric but not endometrial tumors with microsatellite instability.* Oncogene, 2007. **26**(2): p. 308-11.
523. Yi, H.K., et al., *Expression of the insulin-like growth factors (IGFs) and the IGF-binding proteins (IGFBPs) in human gastric cancer cells.* Eur J Cancer, 2001. **37**(17): p. 2257-63.
524. Deming, S.L., et al., *Genetic variation in IGF1, IGF-1R, IGFALS, and IGFBP3 in breast cancer survival among Chinese women: a report from the Shanghai Breast Cancer Study.* Breast Cancer Res Treat, 2007. **104**(3): p. 309-19.
525. Higa, L.A., et al., *Radiation-mediated proteolysis of CDT1 by CUL4-ROC1 and CSN complexes constitutes a new checkpoint.* Nat Cell Biol, 2003. **5**(11): p. 1008-15.
526. Kotake, Y., Y. Zeng, and Y. Xiong, *DDB1-CUL4 and MLL1 mediate oncogene-induced p16INK4a activation.* Cancer Res, 2009. **69**(5): p. 1809-14.
527. Miki, D., et al., *Variation in TP63 is associated with lung adenocarcinoma susceptibility in Japanese and Korean populations.* Nat Genet. **42**(10): p. 893-6.
528. Wang, Y., et al., *Variation in TP63 is associated with lung adenocarcinoma in the UK population.* Cancer Epidemiol Biomarkers Prev. **20**(7): p. 1453-62.
529. Petitjean, A., P. Hainaut, and C. Caron de Fromentel, *TP63 gene in stress response and carcinogenesis: a broader role than expected.* Bull Cancer, 2006. **93**(12): p. E126-35.

530. Mignone, F., et al., *Untranslated regions of mRNAs.* Genome Biol, 2002. **3**(3): p. REVIEWS0004.
531. Cenik, C., et al., *Genome-wide functional analysis of human 5' untranslated region introns.* Genome Biol. **11**(3): p. R29.
532. Chung, B.Y., et al., *Effect of 5'UTR introns on gene expression in Arabidopsis thaliana.* BMC Genomics, 2006. **7**: p. 120.
533. Tahara, T., et al., *Effect of polymorphisms in the 3' untranslated region (3'-UTR) of vascular endothelial growth factor gene on gastric cancer and peptic ulcer diseases in Japan.* Mol Carcinog, 2009. **48**(11): p. 1030-7.
534. Liu, X., et al., *The hOGG1 gene 5'-UTR variant c.-53G>C contributes to the risk of gastric cancer but not colorectal cancer in the Chinese population: the functional variation of hOGG1 for gastric cancer risk.* J Cancer Res Clin Oncol. **137**(10): p. 1477-85.
535. Shibata, T., et al., *A polymorphism in the 3'-utr of cyclooxygenase-2 and the risk of gastric cancer.* Mol Med Report, 2008. **1**(4): p. 561-4.
536. Zhang, Z., et al., *Polymorphisms of thymidylate synthase in the 5'- and 3'-untranslated regions associated with risk of gastric cancer in South China: a case-control analysis.* Carcinogenesis, 2005. **26**(10): p. 1764-9.
537. Huang, Y., et al., *Germline mutations of the VHL gene in seven Chinese families with von.* Int J Mol Med. **29**(1): p. 47-52.
538. Nishikawa, Y., et al., *Human FAT1 cadherin controls cell migration and invasion of oral squamous cell carcinoma through the localization of beta-catenin.* Oncol Rep. **26**(3): p. 587-92.
539. Chosdol, K., et al., *Frequent loss of heterozygosity and altered expression of the candidate tumor suppressor gene 'FAT' in human astrocytic tumors.* BMC Cancer, 2009. **9**: p. 5.
540. Nakaya, K., et al., *Identification of homozygous deletions of tumor suppressor gene FAT in oral cancer using CGH-array.* Oncogene, 2007. **26**(36): p. 5300-8.
541. Merg, A., et al., *Hereditary colon cancer--part I.* Curr Probl Surg, 2005. **42**(4): p. 195-256.
542. Han, S.H., et al., *Mutation analysis of the APC gene in unrelated Korean patients with FAP: four novel mutations with unusual phenotype.* Fam Cancer. **10**(1): p. 21-6.
543. Fostira, F., et al., *Mutational spectrum of APC and genotype-phenotype correlations in Greek FAP patients.* BMC Cancer. **10**: p. 389.
544. Sheng, J.Q., et al., *APC gene mutations in Chinese familial adenomatous polyposis patients.* World J Gastroenterol. **16**(12): p. 1522-6.
545. Fang, D.C., et al., *Mutation analysis of APC gene in gastric cancer with microsatellite instability.* World J Gastroenterol, 2002. **8**(5): p. 787-91.
546. Lee, J.H., et al., *Inverse relationship between APC gene mutation in gastric adenomas and development of adenocarcinoma.* Am J Pathol, 2002. **161**(2): p. 611-8.
547. Kauh, J. and J. Umbreit, *Colorectal cancer prevention.* Curr Probl Cancer, 2004. **28**(5): p. 240-64.

548. Mukherjee, N., et al., *Association of APC and MCC polymorphisms with increased breast cancer risk in an Indian population.* Int J Biol Markers. **26**(1): p. 43-9.
549. Wang, X., et al., *Association of genetic variation in genes implicated in the beta-catenin destruction complex with risk of breast cancer.* Cancer Epidemiol Biomarkers Prev, 2008. **17**(8): p. 2101-8.
550. Huang, S.P., et al., *Association analysis of Wnt pathway genes on prostate-specific antigen recurrence after radical prostatectomy.* Ann Surg Oncol. **17**(1): p. 312-22.
551. Truta, B., et al., *Genotype and phenotype of patients with both familial adenomatous polyposis and thyroid carcinoma.* Fam Cancer, 2003. **2**(2): p. 95-9.
552. Couch, F.J., et al., *Association of mitotic regulation pathway polymorphisms with pancreatic cancer risk and outcome.* Cancer Epidemiol Biomarkers Prev. **19**(1): p. 251-7.
553. Castiglia, D., et al., *Concomitant activation of Wnt pathway and loss of mismatch repair function in human melanoma.* Genes Chromosomes Cancer, 2008. **47**(7): p. 614-24.
554. Siriwardena, B.S., et al., *Aberrant beta-catenin expression and adenomatous polyposis coli gene mutation in ameloblastoma and odontogenic carcinoma.* Oral Oncol, 2009. **45**(2): p. 103-8.
555. Zhang, J.J., et al., *[Detection and significance of APC gene promoter hypermethylation in serum of breast cancer patients].* Ai Zheng, 2007. **26**(1): p. 44-7.
556. Kim, Y.T., et al., *Aberrant promoter CpG island hypermethylation of the adenomatosis polyposis coli gene can serve as a good prognostic factor by affecting lymph node metastasis in squamous cell carcinoma of the esophagus.* Dis Esophagus, 2009. **22**(2): p. 143-50.
557. Richiardi, L., et al., *Promoter methylation in APC, RUNX3, and GSTP1 and mortality in prostate cancer patients.* J Clin Oncol, 2009. **27**(19): p. 3161-8.
558. Saito, K., et al., *Long interspersed nuclear element 1 hypomethylation is a marker of poor prognosis in stage IA non-small cell lung cancer.* Clin Cancer Res. **16**(8): p. 2418-26.
559. Kamory, E., J. Olasz, and O. Csuka, *Somatic APC inactivation mechanisms in sporadic colorectal cancer cases in Hungary.* Pathol Oncol Res, 2008. **14**(1): p. 51-6.
560. Ignatov, A., et al., *APC promoter hypermethylation is an early event in endometrial tumorigenesis.* Cancer Sci. **101**(2): p. 321-7.

Appendix

Preparation of reagents

Stock solution of commonly used reagents

1M Tris

121.1 gm of Tris base was dissolved in 800 ml of double distilled water and pH set (6.8, 7.4, 8.0) with concentrated HCl. Volume was made up to 1 liter and autoclaved.

0.5M EDTA

186.1 gm of disodium EDTA-2H2O was added in 800 ml of double distilled water, stirred vigorously on a stirrer, pH set to 8.0 with NaOH (~20 gm of NaOH pellets) and volume made up to 1 liter and autoclaved.

10% SDS

10 gm of electrophoresis grade SDS was dissolved in 70 ml of ddH2O, heated at 60°C to dissolve and the volume made up to 100 ml.

Ethidium Bromide (10 mg/ml)

10 mg of ethidium bromide was dissolved in 1 ml ddH2O, stored in a opaque bottle.

DEPC water

0.1% diethylpyrocarbonate was added to 1 liter ddH2O in a fume hood and mixed well. After incubating it for 1 hr at 37°C it was autoclaved.

Phosphate Buffer Saline (PBS)

8 gm of NaCl, 2 gm of KCl, 1.44 gm of Na2HPO4 and 0.2 gm of KH2PO4 were dissolved in 800 ml of dw. pH was set to 7.4 with HCl. Final volume was made up to 1 liter and sterilized by autoclaving at 15 lb/sq.in for 20 min. and stored at room temperature.

10 X TAE buffer (Tris acetate, EDTA)

4.84 gm of Tris base in 80 ml of dw was dissolved and 1.2 ml of glacial acetic acid and 2 ml of 0.5 EDTA pH 8.0 were added. Final volume was made up to 100 ml.

10X TBE buffers (Tris borate, EDTA)

8 gm of Tris base, 55 gm of boric acid and 9.3 gm Na2EDTA. H_2O were dissolved in 700 ml dH_2O and the final volume made up to 1 liter.

10 X Tris EDTA (TE), pH8.0

100 mM Tris-Cl & 10mM EDTA. Sterilize solutions by autoclaving. Store the buffer at room temperature.

B. 1M Tris-Cl, pH 8.0

Dissolve 121.1 gm of Tris base in 800 ml of autoclaved distilled water. Adjust the pH 8.0 by adding 42ml of concentrated HCl. Sterilize solutions by autoclaving. Store the buffer at room temperature.

DNA loading dye (6X)

0.2 gm bromophenol blue, 0.2 gm of xylene cyanol and 30 ml of glycerol were dissolved and volume set to 100 by autoclaved dH_2O.

10X Formaldehyde Gel-loading Buffer

50% (v/v) glycerol diluted in DEPC treated water, 10mM EDTA, pH8.0, 0.25% (w/v) bromophenol blue and 0.25% (w/v) xylene cyanol FF were added

10X MOPS Electrophoresis Buffer

Dissolve 41.8 gm of MOPS (3'N Morpholino Propanesulfonic acid) in 700 ml of sterile 0.1% DEPC (Diethyl pyrocarbonate) water. Adjust the pH to 7.0 with 2N NaOH. Add 20 ml of DEPC-treated 1M sodium acetate and 20 ml of DEPC treated 0.5M EDTA, pH 8.0. Adjust the volume of the solution to 1L with DEPC treated water. Sterilize solutions by autoclaving. Store the buffer at room temperature in dark condition.

List of Publications

Publications in International journals

1. **Distribution of Glutathione S-transferase T1 and M1 genes polymorphisms in North East Indians: A potential report.** ***Thoudam RD***, Yadav DS, Mishra AK, Kaushal M, Ihsan R, Chattopadhyay I, Chauhan P, Sarma J, Zomawia E, Verma Y, Nandkumar A, Mahanta J, Phukan RK, Kapur S, Saxena S. ***Genet Test Mol Biomarkers. Volume 14, Number 2, 2010. (thesis related)***

2. **Polymorphisms of Glutathione-S-transferase (GST) genes and the risk of aerodigestive cancers in Northeast Indian population.** Dhirendra Singh. Yadav, ***Thoudam Regina Devi***, Rakhshan Ihsan, AK Mishra, Mishi Kaushal, Indranil Chattopadhyay, Pradeep Singh Chauhan, Jagannath Sharma, Eric Zomawia, Yogesh Verma, A. Nandkumar, Jagadish Mahanta, Rupkumar Phukan, Sunita Saxena., Sujala Kapur. ***Genetic Testing and Molecular Biomarker. Volume 14, Number 5, 2010. (thesis related)***

3. **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.** Rakhshan Ihsan, ***Thoudam Regina Devi***, Dhirendra Singh Yadav, Ashwani Kumar Mishra, Jagannath Sharma, Eric Zomawia, Yogesh Verma, Rupkumar Phukan, Jagadish Mahanta, Amal Chandra Kataki, Sujala Kapur, Sunita Saxena. ***DNA and Cell Biology.* 2010 Nov 2. *[Epub ahead of print]. (thesis related)***

4. **VDR gene polymorphism(s) and breast cancer risk in North Indians**. Anurupa Chakraborty, A.K Mishra, Abha, ***Regina***, A Bhatnagar, P.C Sharma, Sunita Saxena. ***Cancer Detect Prev. 2009;32(5-6):386-94.***

5. **Association of Glutathione S-Transferase, EPHX, and p53 codon 72 Gene Polymorphisms with Adult Acute Myeloid Leukemia.** Pradeep Singh Chauhan, Rakhshan Ihsan, Dhirendra Singh Yadav, Ashwani Kumar Mishra, Bharat Bhushan, Abha Soni, Mishi Kaushal, ***Thoudam Regina Devi***, Sumita Saluja, Dipendra Kumar Gupta, Vishakha Mittal, Sunita Saxena, and Sujala Kapur. ***DNA and Cell Biology. 2010 Aug 23. [Epub ahead of print].***

6. **Gene Expression Profile and Mutational Analysis of DNA Mismatch Repair Genes in Carcinoma Prostate in the Indian Population.** Abha Soni, Sunita Saxena, Anju Bansal, L.C. Singh, Ashwani Mshra, Majinak Majumdar, ***Thoudam Regina***, and Nayan Mohanty. ***OMICS A Journal of Integrative Biology. 2010 Nov 30 [Epub ahead of print].***

7. **Study on predictive role of AR and EGFR family genes with response to neoadjuvant chemotherapy in locally advanced breast cancer in Indian women**. L. C. Singh, Anurupa Chakraborty , Ashwani K. Mishra, ***Thoudam Regina Devi***, Nidhi Sugandhi, Chintamani, Dinesh Bhatnagar, Sujala Kapur, Sunita Saxena. ***Medical Oncology. 2011 April 8 [Epub ahead of print].***

Manuscript Communicated

1. **Study of molecular carcinogenesis of gastric cancer in North East India: Role of *Helicobacter pylori* and tobacco. *Regina Thoudam***, L C Singh, Indranil Chattopadhyay, Ashwani Kumar Mishra, Dhirendra Singh Yadav, Abha Soni, Anurupa Chakraborty, Eric Zomawia, Sunita Saxena, Sujala Kapur. (Manucript communicated)

2. **Genetic polymorphisms of CYP1A1, NQO1 and NAT2 and risk of oral cancer in northeast India.** Dhirendra S Yadav, ***Regina Thoudam***, Ashwani K Mishra, Sobhit Saxena, Abha Soni, Jagannath Sarma, Eric Zamoawia, Yogesh Verma, Rupkumar Phukan, Amal C Kataki, Jagadish Mahanta, Sunita Saxena, Sujala Kapur. (Manuscript communicated)

Abstracts in Proceedings

1. ***Th. Regina Devi***, D.S. Yadav, A.C. Kataki, E. Zamoawia, Yogesh Verma, S. Kapur, S. Saxena "**Polymorphisms of tumour protein P53 genes and the risk of developing gastric cancer in Northeast India."** in "International symposium on Ethics Culture and population genomics' & 34th annual conference of the Indian society of Human Genetics" organized by ISHG and ASI in New Delhi from March 17-20, 2009. **[poster presentation]**

2. ***Regina D Thoudam***, Dhirendra S Yadav, I Chattopadhyay, AC Kataki, E Zamoawia, S Kapur, S Saxena. "**Differential gene expression profile of stomach and oral cancer in high risk region of India"** in 13th Human genome meeting (HGM 2008) on "Genomics and the Future of Medicine" held in Hyderabad on September 27-30, 2008. **[poster presentation]**

3. ***Th. Regina Devi***[1], D.S. Yadav[1], A.C. Kataki[2], E. Zamoawia[3], Y. Verma[4], S. Kapur[1], S. Saxena. **Detoxifying enzyme genotypes and susceptibility to Gastric cancer**. National Conference on Emerging Trends in Life Sciences Research. March 6-7, 2009. Birla Institute of Technology & Science, Pilani, Rajasthan. **[poster presentation]**

4. D.S. Yadav, ***Th. Regina Devi***, J. Sharma, Y. Verma, E. Jamoivia, S. Kapur, S. Saxena. **Genetic Polymorphisms of *CYP1A1* genotypes in patients with oral cancer**. National Conference on Emerging Trends in Life Sciences Research. March 6-7, 2009. Birla Institute of Technology & Science, Pilani, Rajasthan.

5. D.S. Yadav, ***Th. Regina***, R. Ihsan, P. S. Chauhan, I. Chattopadhyay, A. C. Kataki, J. Sharma, E. Zamoawia, Y. Verma, S. Kapur and S. Saxena. **Prevalence of Glutathione S-Transferase (GST) Polymorphisms in Tobacco-Associated Malignancies in High Risk Northeast Indian Population.** International Symposium on cancer Biology, November 14-16 2007, National Institute of Immunology (NII), New Delhi.

6. D.S. Yadav, ***RD Thoudam***, AC Kataki, E Zamoawia, Y Verma, S. Kapur, S. Saxena. **Codon 72 of *TP53* gene Polymorphism in oral cancer and stomach cancer in high-risk region of India.** Human Genome Meeting, Hyderabad, September 27-30, 2008.

7. D.S. Yadav, ***Th. Regina Devi***, A.K.Mishra, J. Sharma, Y. Verma, E. Jamoivia, S. Kapur, S.Saxena. **Genetic polymorphisms of *CYP1A1* genotypes in various ethnic groups of India**. International symposium on Ethics Culture and population genomics & 34th annual conference of the Indian society of Human Genetics" organized by ISHG and ASI, March 17-20, 2009, New Delhi.

8. Ihsan Rakhshan, Kaushal M., ***Th. R. Devi***, Yadav D. S., Soni A., Mishra A. K ., Kataki A. C., Sharma J., Behera D., Jaiswal A., Gupta K., Kapur Sujala, Saxena Sunita. **Study of interactions between Glutathione-S-transferase metabolic enzymes and smoking in lung cancer.** International Symposium on Frontiers in Functional Genomics, 27th Annual Convention, February 7-6, IACRCON 2008, Ahmedabad.

9. Kaushal Mishi, Chakraborty A, Bagadi S A Raju, Ihsan R, ***Regina T***, Yadav D S,Chatterjee I, Zomawia E, Kataki A C, Sharma J, Verma Y, Mishra A K, Kapur Sujala, Saxena Sunita**. Assessment of Breast cancer Risk: Contribution of genetic Polymorphisms in Estrogen-synthesizing and Metabolizing Genes**. International Symposium on Frontiers in Functional Genomics, 27th Annual Convention, IACRCON, Feb 7-9, 2008, Ahmedabad. (Awarded first prize in poster presentation)

10. Pradeep Singh Chauhan, Rakhshan Ihsan, Ashwani Kumar Mishra, Bharat Bhushan, Sumita Saluja, Mishi Kaushal, Dhirendra Singh Yadav, Abha Soni, ***Thoudam Regina Devi***, Indranil Chattopadhyay, Sunita Saxena, Sujala Kapur. **"Glutathione S-transferase and Microsomal Epoxide Hydrolase Gene Polymorphisms and Risk of Acute myeloid leukemia"** in proceedings of

conference entitled "*Hematologic Malignancies: Bridging the Gap 2010*" held in Singapore City, Singapore on Feb 5-7, 2010.

11. Singh Yadav Dhirendra, Thoudam Regina, Mishra Ashwani, Saxena Sunita, Kapur Sujala. **Genetic Polymorphisms of CY1A1, NQO1 and NAT2 and Risk of Oral Cancer in Northeast India: Tobacco Consumption as Risk Modulator.** Human Genome Meeting 2011, 14-17 March 2011, Dubai.

12. ***Regina Devi Thoudam***, Dhirendra Singh Yadav, E. Zomawai, L.C. Singh, Indranil C., A.K. Mishra, Sunita Saxena, Sujala Kapur. **Gene expression profiling of gastric adenocarcinoma in high-risk northeastern region of India.** An AACR (American Association for Cancer Research) International Conference on New Horizons in Cancer Research. December 13-16, 2011. **[Poster Presentation]**

13. Dhirendra singh Yadav , ***Regina Thoudam D***, Sujala Kapur, Sunita Saxena et al. **"*CYP1A1*, *NQO1*, and *NAT2* gene polymorphisms and risk of oral cancer in Northeast India: Tobacco consumption as risk modulator"** in proceeding of conference entitled "New Horizons in Cancer Research: Biology to Prevention to Therapy" organized by American Association for Cancer Research. December 13-16, 2011, New Delhi, India

Biography of Candidate

Thoudam Regina Devi

Email: thoudamnanao@yahoo.co.in, thoudamregina@gmail.com
Phone: Cell: +91-9811681214

Personal details

Age: 30
Date of birth: March 28, 1981 **Sex:** Female
Marital status: unmarried **Citizenship:** Indian

Correspondence:

Present address:

Tumour Biology
National Institute of Pathology (ICMR)
Safdurjung hospital Campus
New Delhi
INDIA 110029

Permanent address:

Nagamapal Soram Leirak
Opp. Nityananda temple
Imphal-West
Manipur 795001
INDIA

Career at a glance

	Name of employer	***Institute***	***Topic of research project***	***Period From -To***
Junior Research Fellow	*Dr Sunita Saxena*	*National Institute of pathology*	*Role of tobacco in the causation of cancer in north east India*	*Oct 2005-Oct 2007*
Senior Research Fellow	*Dr Sunita Saxena*	*National Institute of pathology*	*Role of tobacco in the causation of cancer in north east India*	*Sep 2007-May 2009*
Senior Research Fellow (Individual ICMR)	*Dr Sunita Saxena*	*National Institute of pathology*	*Study the molecular signature of..gastric cancer in high prevalence region*	*June 2009 till date*

Hands on experience: Microarray technique and analysis, RNA extraction, Real Time RT-PCR, DNA Sequencing, DNA Extraction from Blood samples and tissue samples, PCR (Polymerase Chain Reaction)-Touch-Up, Touch Down, Single Strand Conformation Polymorphisms (SSCP), Heteroduplex analysis, RFLP, RT-PCR, PAGE, ELISA, DHPLC, Western blotting.

Educational Qualifications

Ph.D. pursuing (Thesis under correction) at National Institute of Pathology (ICMR) in Biological Sciences, Safdurjung Hospital campus, New Delhi, India.

Master of Science in Zoology (Specialization in cell and molecular biology), Jiwaji University, Gwalior, India 2005 securing 70 % in aggregate.

Bachelor of Science in Zoology, D.M. College of Science, Manipur, India in 2001, securing 63% in aggregate.

Higher Secondary, Tamphasana Girls Higher Sec School, Manipur, India in 1998, securing 62% in aggregate.

Secondary, Nirmalabas High School, Manipur, India in 1996, securing 62% in aggregate.

Work Shop Attended

1. Fourth Workshop on "**Genetic Epidemiological methods for the Dissection of Complex Traits".** TCG- ISI Centre for population Genomics and University of Pittsburgh. Kolkata, India. 23-28 February 2009.

2. National Workshop on Molecular Cytogenetics Cancer Cytogenetics (hematological malignancy) by FISH, Department of Reproductive Biology**,** All India Institute of Medical Sciences, New Delhi, India. 29 November to 04 December 2010.

3. International Workshop on Molecular and GIS Based Epidemiology of Leprosy. Institute of Pathology, Safdarjung Hospital Campus, New Delhi, (INDIA). 4 -9 March 2010.

4. Hands on Training on Microarray technology with analysis and interpretation of data. **National Workshop on Microarray Technology**, April 2007. Institute of Pathology (Indian Council of Medical Research), New Delhi.

Publications in International journals

1. Distribution of Glutathione S-transferase T1 and M1 genes polymorphisms in North East Indians: A potential report. ***Thoudam RD***, Yadav DS, Mishra AK, Kaushal M, Ihsan R, Chattopadhyay I, Chauhan P, Sarma J, Zomawia E, Verma Y, Nandkumar A, Mahanta J, Phukan RK, Kapur S, Saxena S. ***Genet Test Mol Biomarkers. Volume 14, Number 2, 2010.***

2. Polymorphisms of Glutathione-S-transferase (GST) genes and the risk of aerodigestive cancers in Northeast Indian population. Dhirendra Singh. Yadav, ***Thoudam Regina Devi***, Rakhshan Ihsan, AK Mishra, Mishi Kaushal, Indranil Chattopadhyay, Pradeep Singh Chauhan, Jagannath Sharma, Eric Zomawia, Yogesh Verma, A. Nandkumar, Jagadish Mahanta, Rupkumar Phukan, Sunita Saxena., Sujala Kapur**.** ***Genetic Testing and Molecular Biomarker. Volume 14, Number 5, 2010.***

3. 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. Rakhshan Ihsan, ***Thoudam Regina Devi***, Dhirendra Singh Yadav, Ashwani Kumar Mishra, Jagannath Sharma, Eric Zomawia, Yogesh Verma, Rupkumar Phukan, Jagadish Mahanta, Amal Chandra Kataki, Sujala Kapur, Sunita Saxena. ***DNA and Cell Biology.*** **2010 Nov 2.** ***[Epub ahead of print].***

4..VDR gene polymorphism(s) and breast cancer risk in North Indians. Anurupa Chakraborty, A.K Mishra, Abha, ***Regina***, A Bhatnagar, P.C Sharma, Sunita Saxena**.** ***Cancer Detect Prev. 2009;32(5-6):386-94.***

5. Association of Glutathione S-Transferase, EPHX, and p53 codon 72 Gene Polymorphisms with Adult Acute Myeloid Leukemia. Pradeep Singh Chauhan, Rakhshan Ihsan, Dhirendra Singh Yadav, Ashwani Kumar Mishra, Bharat Bhushan, Abha Soni, Mishi Kaushal, ***Thoudam Regina Devi***, Sumita Saluja, Dipendra Kumar Gupta, Vishakha Mittal, Sunita Saxena, and Sujala Kapur. ***DNA and Cell Biology. 2010 Aug 23. [Epub ahead of print].***

6. Gene Expression Profile and Mutational Analysis of DNA Mismatch Repair Genes in Carcinoma Prostate in the Indian Population. Abha Soni, Sunita Saxena,

Anju Bansal, L.C. Singh, Ashwani Mshra, Majinak Majumdar, ***Thoudam Regina***, and Nayan Mohanty. ***OMICS A Journal of Integrative Biology. 2010 Nov 30 [Epub ahead of print].***

7. **Study on predictive role of AR and EGFR family genes with response to neoadjuvant chemotherapy in locally advanced breast cancer in Indian women** L. C. Singh, Anurupa Chakraborty , Ashwani K. Mishra, ***Thoudam Regina Devi***, Nidhi Sugandhi, Chintamani, Dinesh Bhatnagar, Sujala Kapur, Sunita Saxena. ***Medical Oncology. 2011 April 8 [Epub ahead of print].***

8. **Gene Expression Profiling of Gastric Adenocarcinoma in High Risk North East Region of India.** ***Thoudam RD***, Singh LC, Chattopadhyay I, Yadav DS, Abha S, Anurupa C, Zomawia E, Verma Y, Nandkumar A, Mahanta J, Phukan RK, Kapur S, Saxena S. (Manucripts communicated).

9. **Genetic polymorphisms of CYP1A1, NQO1 and NAT2 and risk of gastric cancer.** ***Thoudam RD***, Yadav DS, Zomawia E, Verma Y, Nandkumar A, Mahanta J, Phukan RK, Kapur S, Saxena S. (Manuscripts under preparation).

10. **Assessing matched normal and tumor pools in next-generation sequencing studies to indentify genomic alteration in gastric adenocarcinoma in high risk northeastern region of India**. **Thoudam Regina,** Anand Verma, Sunita Saxena, Sujala Kapur (Manuscript under preparation).

Abstracts in Proceedings

1.Th. Regina Devi, D.S. Yadav, A.C. Kataki, E. Zamoawia, Yogesh Verma, S. Kapur, S. Saxena "**Polymorphisms of tumour protein P53 genes and the risk of developing gastric cancer in Northeast India."** in "International symposium on Ethics Culture and population genomics' & 34th annual conference of the Indian society of Human Genetics" organized by ISHG and ASI in New Delhi from March 17-20, 2009. **[poster presentation]**

2. Regina D Thoudam, Dhirendra S Yadav, I Chattopadhyay, AC Kataki, E Zamoawia, S Kapur, S Saxena. "**Differential gene expression profile of stomach and oral cancer in high risk region of India"** in 13th Human genome meeting (HGM

2008) on "Genomics and the Future of Medicine" held in Hyderabad on September 27-30, 2008. **[poster presentation]**

3. Th. Regina Devi[1], D.S. Yadav[1], A.C. Kataki[2], E. Zamoawia[3], Y. Verma[4], S. Kapur[1], S. Saxena. **Detoxifying enzyme genotypes and susceptibility to Gastric cancer**. National Conference on Emerging Trends in Life Sciences Research. March 6-7, 2009. Birla Institute of Technology & Science, Pilani, Rajasthan. **[poster presentation]**

4. D.S.Yadav, ***Th. Regina Devi***, J. Sharma, Y. Verma, E. Jamoivia, S. Kapur, S. Saxena. **Genetic Polymorphisms of *CYP1A1* genotypes in patients with oral cancer**. National Conference on Emerging Trends in Life Sciences Research. March 6-7, 2009. Birla Institute of Technology & Science, Pilani, Rajasthan.

5. D. S. Yadav, ***Th. Regina***, R. Ihsan, P. S. Chauhan, I. Chattopadhyay, A. C. Kataki, J. Sharma, E. Zamoawia, Y. Verma, S. Kapur and S. Saxena. **Prevalence of Glutathione S-Transferase (GST) Polymorphisms in Tobacco-Associated Malignancies in High Risk Northeast Indian Population.** International Symposium on cancer Biology, November 14-16 2007, National Institute of Immunology (NII), New Delhi.

6. D. S. Yadav, ***RD Thoudam***, AC Kataki, E Zamoawia, Y Verma, S. Kapur, S. Saxena. **Codon 72 of *TP53* gene Polymorphism in oral cancer and stomach cancer in high-risk region of India.** Human Genome Meeting, Hyderabad, September 27-30, 2008.

7. D.S.Yadav, ***Th. Regina Devi***, A.K.Mishra, J. Sharma, Y. Verma, E. Jamoivia, S. Kapur, S.Saxena. **Genetic polymorphisms of *CYP1A1* genotypes in various ethnic groups of India**. International symposium on Ethics Culture and population genomics & 34th annual conference of the Indian society of Human Genetics" organized by ISHG and ASI, March 17-20, 2009, New Delhi.

8. Ihsan Rakhshan, Kaushal M., ***Th. R. Devi***, Yadav D. S., Soni A., Mishra A. K ., Kataki A. C., Sharma J., Behera D., Jaiswal A., Gupta K., Kapur Sujala, Saxena Sunita. **Study of interactions between Glutathione-S-transferase metabolic enzymes and smoking in lung cancer.** International Symposium on Frontiers in

Functional Genomics, 27th Annual Convention, February 7-6, IACRCON 2008, Ahmedabad.

9. Kaushal Mishi, Chakraborty A, Bagadi S A Raju, Ihsan R, ***Regina T***, Yadav D S,Chatterjee I, Zomawia E, Kataki A C, Sharma J, Verma Y, Mishra A K, Kapur Sujala, Saxena Sunita**. Assessment of Breast cancer Risk: Contribution of genetic Polymorphisms in Estrogen-synthesizing and Metabolizing Genes**. International Symposium on Frontiers in Functional Genomics, 27th Annual Convention, IACRCON, Feb 7-9, 2008, Ahmedabad. **(Awarded first prize in poster presentation)**

10. Pradeep Singh Chauhan, Rakhshan Ihsan, Ashwani Kumar Mishra, Bharat Bhushan, Sumita Saluja, Mishi Kaushal, Dhirendra Singh Yadav, Abha Soni, ***Thoudam Regina Devi***, Indranil Chattopadhyay, Sunita Saxena, Sujala Kapur. **"Glutathione S-transferase and Microsomal Epoxide Hydrolase Gene Polymorphisms and Risk of Acute myeloid leukemia"** in proceedings of conference entitled "*Hematologic Malignancies: Bridging the Gap 2010*" held in Singapore City, Singapore on Feb 5-7, 2010.

11. ***Regina Devi Thoudam,*** Eric Zomawia, Sunita Saxena, Sujala Kapur et al. "Gene expression profiling of gastric adenocarcinoma in high-risk Northeast region of India" in proceeding of conference entitled "New Horizons in Cancer Research: Biology to Prevention to Therapy" organized by American Association for Cancer Research. December 13-16, 2011, New Delhi, India **[poster presentation]**.

12. Dhirendra singh Yadav , ***Regina Thoudam D***, Sujala Kapur, Sunita Saxena et al. **"*CYP1A1*, *NQO1*, and *NAT2* gene polymorphisms and risk of oral cancer in Northeast India: Tobacco consumption as risk modulator"** in proceeding of conference entitled "New Horizons in Cancer Research: Biology to Prevention to Therapy" organized by American Association for Cancer Research. December 13-16, 2011, New Delhi, India.

Biography of Supervisor

Curriculum Vitae

Name	:	**Dr. (Mrs.) Sunita Saxena**
Date of Birth	:	16th September, 1952
Designation	:	**Director**
Address	:	Institute of Pathology-ICMR Safdarjang Hospital Campus, Post Box No.4909, New Delhi – 110029
Academic Qualifications	:	**M.B.B.S., D.C.P., M.D.(Path)**

Degree	**Institute**	**Year**	**Remarks**
M.B.B.S.	M.L.N.Medical College, Allahabad	1974	
D.C.P. (Clinical Pathology M.D. (Pathology)	L.L.R.M.Medical College. Meerut	1978	**Received Gold Medal & Merit Certificate**
	-do-	1981	

Details of Employment

Post	Duration	Institute
Research Officer	April 1981 to Dec.1985	Institute of Pathology, New Delhi.
Sr. Research Officer	Jan.1986 to Jan.1991	-do
Asst. Director	9th Jan.1991-2nd May, 1994	-do-
Deputy Director	2nd May, 1994 to 2nd May,'99	-do-
Deputy Director (Sr.Gr)	3rdMay, 99 to 8thApril, 2002	
Deputy Director (Sr.Gr) & Officer In charge	9thApril, 2002to 13th Dec., 2004	-do-
Director	14th Dec., 2004 till date	-do-

Areas of Specialization : Molecular Oncology, Oncopathology

Areas of Interest : Breast Tumors, Tobacco Associated cancers Genito urinary cancers

Membership of National and International bodies

International : Life Member International Union against Cancer (U.I.C.C.)

National : Life Member: Indian Association of Cancer Research (IACR).
Life Member: Indian Association of Pathologists and Microbiologists .(IAPM).
Life Member: Association for the promotion of DNA fingerprinting and other DNA technologies (ADNAT)
Life Member: Proteomic Society of India
Life member: Human Genomic Organization (HUGO)

Trainings Received:-

1. Trained for **'Culture of fastidious cells and modern techniques of cell manipulation'** at *National Facility for animal Tissue and cell culture, Pune* in Dec., 1991.
2. Trained in **'Genetic mutation detection techniques for BRCA 1 and BRCA2 genes in genomic DNA of Breast cancer patients'** at *Unit Genetic Epidemiology* at *International Agency for Research on Cancer, Lyon, France* in 1998.
3. Attended **Hands-on Training Course on Proteomics and DNA Micro arrays** held from 25th February to 10th March 2003 at CCMB, Hyderabad.

Fellowships:-

1. Awarded **WHO fellowship** to work on **Renal Pathology** at **St. Vincent's Hospital, Melbourne, Australia**, 1987.
2. Awarded **Yamigawa Yoshida Fellowship of U.I.C.C.** to work on *"Mutation analysis of BRCA1 and BRCA2 genes in Breast Cancer Patients*" at unit of Genetic Epidemiology **International Agency for Research on Cancer, Lyon, France** in 1998.

Awards and Honors Received:-

1. Awarded **Gold Medal** and **Merit Certificate** for securing highest marks in **Diploma in Clinical Pathology** (D.C.P.).
2. **K. C. Basu Mullick award** for best research work by **Indian Association of Pathologists and Microbiologists for year 2008.**
3. Received '***NOVARTIS ORATION AWARD 2006***" of Indian Council of Medical Research for her work on Breast cancer on 18th Sept. 2009.

4. Elected ***Fellow of National Academy of Medical Sciences in 2010.***

5. Elected ***Fellow of Indian College of Pathologist in 2010***

6. Paper entitled **"Role of p53 Tumor suppressor gene and Estrogen receptor status of Breast Cancer and its association with Clinical and Histopathologic parameters"** was awarded **T.J. Aggarwal Memorial Gold Medal** in Annual conference of Association of Surgeons of India in 1999.

7. "Role of apoptotic markers in assessing the response to neo-adjuvant chemotherapy (NACT) and its correlation with clinical parameters in patients with carcinoma breast"**, Adjudged BEST PAPER at the Annual Conference of Surgery (Delhi State Chapter) held in December, 2002.**

8. Paper entitled "Flow cytometric analyses of Th1 and Th2 cytokine production as a parameter of immunologic dysfunction in patients with superficial transitial cell Carcinoma" **received special appreciation award for the Best Poster presentation at 24th Annual Convention of Indian Association for Cancer Research & International Symposium of Human Papilloma virus and cervical cancer held at ICPO from 9th – 12th Feb., 2005.**

9. **Best Poster Award for poster entitled** "What Androgen Receptor CAG repeats polymorphism and p53 mutations/polymorphisms have to do with prostate cancer risk and progression? at "XXXIII Annual Conference of the ISHG and International Symposium on Genetics Revisited: the Genomics and Proteomics Advantage", Vishakhapatnam, India from 11-13 Feb 2008.

Research Grants received:-

1. Awarded research grant to work on **"Genetics of Breast Cancer in Indian Women"** by *Indo-French Centre for Promotion of Advanced Research,* **New Delhi (1999-2002).**
2. Awarded Research grant from *Deptt. of Science and Technology* for proposal entitled **" A new approach to the management of superficial bladder cancer. Role of in vitro Cytotoxicity assessment and immunologic enhancement"** (2000-2003).
3. **"Comprehensive study of carcinoma esophagus at North-East India-Multidiscipline approach"**. Collaborative, Multicentric ICMR Task Force project (2004-2007).
4. **"Microsatellite instability in androgen receptor gene and mismatch repair system in Prostate cancer in Indian males"** – Department of Science Technology (2005-2008).
5. **"Role of tobacco use in causation of cancer in north-east India"** – Indian Council of Medical Research Task Force project (2005-2008).

6. **“Effects of pesticide exposure in causation of cancer in north-east India”** – Indian Council of Medical Research Task Force project (2005-2008).
7. **“Establishment of Cell lines from Primary Breast Cancer”** – Indian Council of Medical Research. Task force project (2007-2010).
8. **“Study on Gene Expression and Hypermethylation Profiles in Early Onset Breast Cancer”** Department of Biotechnology (2008-2011)
9. “**Characterization of host immune factors associated with progression of superficial TCC of bladder by microarray analysis”** Indian Council of Medical Research (2009-2012)
10. “**Immunogenetic profile of Nasopharyngeal Cancer in a high prevalence region of Northeast India”** Department of Biotechnology (2010-2013)
11. **“Comparative study of Genetic, Clinical and Epidemiological Factors of Breast Cancer in Rural and Urban Area of India”** Indian Council of Medical Research Task force project (2009-2012).
12. **“Epigenetic studies in esophageal cancer in high risk region of Northeast India”** Department of Biotechnology, Twining Program for NER (2011-2013)
13. “**Genome wide Analysis of Genetic alterations in patients with Esophageal Cancer from Northeast India using Single Nucleotide Polymorphism arrays”** Indian Council of Medical Research (2011-2013)
14. **“Study on miRNA signatures associated with Breast cancer stem like cells (CSC) and their role in drug response”** Indian Council of Medical Research (2012-2015)

Academic Experience:-

a. Diplomat of National Board Examinations (N.B.E.) – Institute is accredited for running this program since 1993.
    - ***Core member, Supervisor & Co-ordinator***.
    - Guided 20 DNB dissertation as supervisor and cosupervisor.
    - Nominated as Inspector, Examiner for theory & Practical examinations and paper setter.

b. Ph.D. - ***Supervisor/Mentor***
    - ***Supervisor*** of student registered under GGSIPS University, Delhi, BITS, Pilani.
    - External Examiner for Ph.D. candidates of AIIMS, PGI, Chandigadh, Agra University, BITS, Pilani.

c. Member of project Review Committee of Divisions of ***Non Communicable Diseases of Indian Council of Medical Research.***
d. Member of project Review Committee of Divisions of ***Basic Medical Sciences of Indian Council of Medical Research.***
e. Member of ***Scientific Advisory committee of Institute Of Cytology and Preventive Oncology, Noida***
f. Reviewer of papers of ***Indian Journal of Medical Research, Human Mutation, Cancer Immunology Immunotherapy, World Journal of Surgical Oncology***, **BMC Cancer*, Cancer Detection and Prevention, British Journal of Urology International, Indian Journal of Medical Research***
g. Member of ***Scientific Advisory committee of National Jalma Institute for Leprosy and other Mycobacterial diseases, Agra.***
h. Member of ***Scientific Advisory committee of Regional Medical Research Centre, Dibrugadh***
i. Member of ***Scientific Advisory Group of Publication & Information Division Of Indian Council Of Medical Research***
j. ***Chairperson of Institutional Ethical Committee of Safdarjang Hospital, New Delhi.***
k. Member of ***Technical committee of Indian Council of Medical Research***
l. Member of **DBT sponsored DSMB on Curcumin Trial in Cancer Cervix**
m. **Senate member of BITS, Pilani.**
n. Nominated as core member of selection committee **for the award of ICMR Post Doctoral Fellowships**
o. Nominated as the expert member of **Task force on Leprosy at ICMR.**
p. Nominated as the member of Scientific Advisory Group for creating a new centre for Environmental Health and Bhopal Gas Tragedy at Bhopal.
q. Nominated as the member of the Data Safety Monitoring Board (DSMB) on "BASANT Clinical Trial" of DBT, New Delhi.
r. Appointed as **Appraiser and Inspector** by National Board of Examination for assessment of DNB students and institutes.

q. Dr. Sunita Saxena has been nominated as the expert member of " ICMR-ICAR Joint Task force on the Epidemiology of Human and Animal Brucellosis".

r. Dr. Sunita Saxena has been nominated as nodal officer for getting ICMR university status.

International Conferences attended.

1. Presented a paper on **"Pattern of lymphokines in minimal change Nephrotic syndrome"** in 5th *Asia Pacific Congress of Nephrology* held in New Delhi during 9-12th Dec., 1992.

2. Presented paper on **'Role of Proto-oncogene, Growth Factor Receptor and Steroid Hormones on Malignant Human Mammary Epithelial Cancer Cells in vitro and vivo"** in XVI, *International Cancer Congress (U.I.C.C.)* at New Delhi, 30th-5th Nov., 1994.

3. **'Stage A carcinoma of Prostate'** paper presented at *first conference of Nephrology, Urology and Transplantation Society of SAARC Countries* held at A.I.I.M.S. , New Delhi during 24th-26th March, 1995.

4. **'Mutation profile of BRCA 1 / 2 genes in Indian patients'** paper presented at *XV Asia Pacific Cancer Congress* held at Chennai during Dec. 12-15, 1999.

5. **'Mutation profile of BRCA 1 / 2 mutations in worldwide population. The MAGIC project'** paper presented in meeting of *American Association of Cancer Research AACR*, 2001, LA, USA.

6. Attended the 7th International Symposium on Molecular Basis of Predictive Oncology and Intervention Strategies' and presented a paper **BRCA1 and BRCA2 Genes in Indian Breast Cancer Patients** held at Nice, France from 7th 10th Feb. 2004.

7. Attended the **"UICC World Cancer Congress and Centre for Disease Control and Prevention (CDC)"** held during 8th to 13th July, 2006 at Washington DC, U.S.A and presented paper **"Study of candidate genes associated with Breast Cancer Susceptibility in the Indian Women".**

8. Attended the NCRI Cancer Conference held at International Convention Centre in Birmingham, UK from 30th September - 3rd October 2007 and presented paper (oral and poster) entitled **"Gene expression profile of esophageal cancer in North East India by cDNA microarray analysis"**

9. Attended World Cancer Congress-2008 held in Shanghai, China during 12th-17th June 2008 and presented paper **"Differential gene expression in familial and tobacco associated esophageal cancers in north-east region of India".**

10. Attended HUGO's 13th Human Genome Meeting, September 27 - 30 2008, Hyderabad. and presented papers-

- **Prognostic value of TP53 Codon 72 polymorphism in oral cancer and stomach cancer in high risk region of India**
- **Differential gene expression profile of stomach and oral cancer in high risk region of India.**
- **Differential expression of MAPK and GPCR pathway in esophageal cancer of North-east region of India**
- **Significance of TP53 codon 72 polymorphism in breast and lung cancer showing different xenobiotic potential spectrum**

11. Attended **First Symposium on HPV Vaccination in the Asia Pacific and Middle East Region held at Seoul, Korea** during1st to 3rd June'09.
12. Visited **University of Minnesota, USA as member of expert team of Indian Scientists on Cancer** for collaborative research projects.
13. Presented papers in **Conference of the Organisation for Oncology and Translational Research (OOTR), 6th Annual Conference on 26 and 27 February, 2010 at Kyoto Japan**
    - Genetic alterations in patients with esophageal cancer from high-risk region in India by SNP array. *Sujala Kapur, Indranil Chattopadhyay, Rupkumar Phukan, Joydeep Purkayastha, Vikki Marshal, Amal Kataki, Jagdish Mohanta, David Bowtell, Sunita Saxena*
    - Genome-wide analysis of genetic alterations in breast cancer patients from Northeast India using 10K SNP arrays. *Sunita Saxena, Mishi Kaushal, Indranil Chatterjee, A. Bhatnagar, Chintamani, D. Bhatnagar, Sujala Kapur*
14. Presented papers in **20th Asia Specific Cancer Conference, November 12-14, 2009,Japan.**
    - GENOME-WIDE ANALYSIS OF GENETIC ALTERATIONS IN ESOPHAGEAL SQUAMOUS CELL CARCINOMA BY SNP ARRAY. Sujala Kapur, Indranil Chattopadhyay, Rupkumar Phukan, Joydeep Purkayastha, Vikki Marshal, Amal Kataki, Jagdish Mohanta, David Bowtell, Sunita Saxena.
    - GENOME-WIDE ANALYSIS OF DNA COPY NUMBER VARIATIONS IN INDIAN BREAST CANCER PATIENTS USING HIGH-DENSITY SNP ARRAYS. Sunita Saxena, Mishi Kaushal Wasson, Indranil Chatterjee, A Bhatnagar, Dr Chintamani, D Bhatnagar, Sujala Kapur.
15. Presented paper entitled "**Molecular Profile of Esophageal Cancer in High Risk Region of India"** in 10th World Congress of OESO held at Boston, USA during 28-31 Aug'2010. (Abstract J. Clin. Gastroenterology. Vol: 45 (2), 2011.
16. Presented paper entitled "**Genomic alterations in breast cancer patients from Northeast India using 10K SNP arrays"** in BMC group conference **Beyond the Genome: The true gene count, human evolution and disease genomics** at Harvard Medical School, Boston, USA during11th-13th Oct'2010

17. Presented paper entitled "**Betel Quid Chewing A Risk Factor For Breast Cancer: Study Of Genomic Alterations"** at 16th Human Genome Meeting 2012 held at Sydney, Australia during 11th-14th March, 2012

Publications: -

Research Articles:

1. Chandra M., **Saxena, S.**, Dave P.K., Kaza R. M. and Saxena, H.M.K.: **Proliferative myositis (a pseudosarcomatous lesion of skeletal muscle) occurring in an infant**. *Ind. J. Path. & Microbiol :* 26, 213, 1983.

2. Nagar P., **Saxena S.** (nee Sinha). Pratap V.K. and Mehrotra M.L.: **Malignant haemangiopericytoma of the uterus**. **A case report with brief review of literture**. *J. of Obst. Gynaec. Of India : 34,753, 1984.*

3. **Saxena S.**, Andal, A. and Saxena H.M.K.: **Ultrastructure study of minimal change nephrotic syndrome – a clinico-morphologic correlation**. *Ind. J. Med. Res.* 82: 171, 1985.

4. **Saxena S.,** Andal A, and Saxena H.M.K. **Stereomicroscopic examination of kidney tissue for rapid identification of glomerulus**. *Nephron:* 45: 249, 1987.

5. **Saxena S.,** Mehrotra M.L.: **Host tissue response in soft tissue sarcomas.** *Ind. J. Path. & Microbiol*. 30:97, 1987.

6. **Saxena S.,** Andal A, and Saxena H.M.K. **Idiopathic nephrotic syndrome of childhood: Ultrastructural immunohistologic and Clinicocomorphologic correlation.** *Ind. J. Path. & Microbiol.* 31 (3) 195, 1988.

7. Andal A, **Saxena S**, Chellani H.K. and Sharma S. **Pure Mesangioproliferative Glomerulonephritis. A Clinicomorphologic analysis and its possible role in morphological transition of minimal change lesion to Focal glomerulosclerosis**. *Nephron:* 51(3): 314, 1989.

8. **Saxena S,** Davies D.J., Krisner R.L.G. **Thin basement membrane in minimally abnormal glomeruli.** *J. Clin. Pathol*. 43: 32, 1990.

9. **Saxena S.**, Andal A, Saxena R.K., Sharma S, Chandra M, Saxena H.M.K. **Immune status of children suffering from Minimal change nephrotic syndrome**. *Ind. J. Path. & Microbiol.* 35(3) 171, 1992.

10. **Saxena. S.**, Davies D.J., **Glomerular alterations in Idiopathic haematuria– Ultrastructural and Morphometric analysis**. *Ind. J. Path. & Microbiol.* 35(4), 326-332, 1992.

11. **Saxena. S.**, Andal. A., Sharma. S, Saxena H.M.K., Chandra M. **Immunomodulation by measles vaccine in children with Minimal change nephrotic syndrome**. *Indian J. of Nephrology* 2, 141-146, 1992

12. Verma. A.K., Tandon, R.., **Saxena. S.**, Pandy, J., Talib. V.H. **Aspiration Cytology of maxillary myxoma**. *Diagnostic Cytopathology* 9(2), 202-204, 1993.

13. **Saxena S**, Mital. A, Andal A,: **Pattern of interleukins in MCNS of childhood**. *Nephron* 65(1) 56-61, 1993

14. **Saxena S.**, Bhargawa R., Mohanty N.K., Talwar M: **Primary adenocarcinoma of the urinary bladder. A case report with review of literature** *Ind J Pathol and Microbiol.* 37(4), 453, 1994.

15. Saha T.K., Jolly B B., Mohanty N.K., **Saxena S.**, Dawson. L. **Multiple stones in Ectopic megaureter with Dysgenetic kidney – A case report**. *Ind. J. Nephrol*. 4(2). 61, 1994.

16. **Saxena S**. **Cytokine growth factors and childhood nephrotic syndrome**. *Jr. of Nephrol*. Vol. 8(6), 287, 1995.

17. Mohanty NK, Jolly BB, **Saxena S**, Dawson L. **Squamous cell carcinoma of peripheral urethrostomy**. *Urol. Int.* 1995, 55: 118-119.

18. Mohanty NK, Jolly BB, Talwar M, **Saxena S**, Dawson L. **Aspergillosis kidney. A case report**. *Indian Jr. of Nephrol*. 6(2), 56-58, 1996.

19. **Saxena S**, Jain A K, Pandey K K, Dewan A K. **Study on role of Steroid Hormone Receptors, Growth factor/receptors and Proto-oncogenes on behavior of Human Mammary Epithelial cancer cells in vitro.** *Pathobiology* 65(2), 75-82, 1997.

20. **Saxena S**, Mohanty N K, Talwar M, Jain A K. **Screening of Prostate Cancer in males with prostatism**. *Ind. J. of Path & Microbiol*. 40(4), 441-450, 1997.

21. Mohanty N K, Gulati P, **Saxena S**. **Role of interferon α-2b in the prevention of superficial carcinoma of bladder recurrence**. *Urol. Intern*.59: 194-196, 1997.

22. Mohanty N.K., Jha AK, **Saxena S**, Kumar S., Arora RP. **Ten years experience with Adujuvant Intravesical Immunotherapy in management of superficial transitional cell carcinoma of Urinary bladder – A review**. *Ind. J. of urology* 2001, 17, 127.

23. **Saxena S**, Jain A.K., Bhatnagar D. **Study of events leading to cellular Senescence to Human Mammary epithelial cancer cells in vitro**. *Indian. J. Cancer*. 38: 103-116, 2001

24. **Saxena S**., Beena KR, Bansal A, Bhatnagar A. **Emperipolesis: Significance of an unusual phenomenon in common breast malignancy**. *Acta Cytologica.* 46: 883-886, 2002

25. **Saxena S**, Szabo C, Barjhoux H, Chopin S, Sinilinova O, Lenoir G, Goldgar D, Bhatnagar D**. BRCA 1 and BRCA 2 in Indian Breast Cancer Patients.** *Human mutation.* 20 (6): 473-74, 2002.

26. Chintamani, Sharma R D, Bardan R, Singhal V, **Saxena S,** Bansal A **Sweat gland Adenocarcinoma – a rare clinical dilemma**. *World J Sur. Oncol.* 1: 13, 2003

27. Bharat R, Burra U, Vidyadharan G, **Saxena S**. **Morphogical spectrum of cysticercus cellulose on cytology in case of malnutrished child**. *J Cytol*. 21 (2): 95-06, 2004

28. **Saxena S,** Bansal A, Mohil R S, Bhatnagar D. **Metaplastic carcinoma of the breast-A rare breast tumor.** *Ind J Pathol and Microbiol.* 47(2): 217-220, 2004

29. **C**hintamani, Shankar M, Singhal V, Singh J P, **Saxena S**. **Squamous cell carcinoma developing in the scar of fournier's gangrene-case report.** *BMC Cancer*. 4:16, 2004.

30. Bharat R, **Saxena S,** Burra U. **Fine needle aspiration cytology of Dermato fibrosarcoma protuberans**. *J Cytol*. 21(3), 2004

31. **C**hintamani, Singhal V, Singh J P, Bansal A, **Saxena S**, Lyall A. **Is drug induced cytotoxicity a good predictor of response to new adjuvantchemotherapy in breast cancer? A prospective clinical study.** *BMC Cancer*. Aug 13; 4(1): 48, 2004

32. Mukherji A, Madholia V, Malhotra S, Singh P, Rekhi B, **Saxena S**, Aggarwal Y, Bhowmik K.T. **Multiple Myeloma Of The Breast –An Unusual Case Of Multiple Myeloma Of The Breast With Pathological Fractures Of Humerus and Femur**. *Jr. Of Clinical Radiotherapy and Oncology.4 (4):27-30, 2004.*

33. Mohanti. N.K, **Saxena . S**, Goyal .N.K, Singh. U.P, Arora .R.P **Delayed Cystectomy for T1G3 TCC of Urinary Bladder Managed**

initially by TURBT & Intravesical Immunotherapy (BCG+Interferon)—Rationale & our results. *Indian Journal of Urology, 2004, vol. 20: 2.*

34. Chintamani, Singhal V, Singh J P, Bansal A, **Saxena S**. **Half versus full vacuum suction drainage after modified radical mastectomy for breast cancer-A prospective randomized clinical trial** (ISRCT N24484328). *BMC Cancer*. 5:11, 2005**.**

35. Rekhi B, Bansal A, Bhatnagar D, Bhatnagar A, **Saxena S. Cytomorphological study of soft tissue neoplasms: role of fluorescent immunocytochemistry in diagnosis** *Cytopathology, 16(5) :219-26, 2005.*

36. Rekhi B, **Saxena S,** Chintamani**. Gastric outlet obstruction and cutaneous metastasis in Adenocarcinoid Tumor of Stomach- Unusual presentations with cytologic and Ultrastructural findings.** *Indian J Cancer. ,42(2):99-101,2005*

37. Ghai R, Rekhi B, **Saxena S**, Kapoor S. **An unusual presentation of Primary Lympoma of the Thyroid in a young male patient—A case report.** *I.J.P.M.,48(3): 385-387,2005*

38. Murthy NS, Chaudhary K, Saxena **S**. **Trends in Incidence of Cervical Cancer –Indian Scenario.** *Euro. J Can Prev. 2005 Dec; 14(6):513-8.*

39. A. Agarwal, S. Verma, U. Burra, NS Murthy, NK Mohanty and S. Saxena Ca"**Flow Cytometric analysis of Th1 and Th2 cytokines in PBMCs as a parameter of immunological dysfunction in patients of Superficial Transitional cell carcinoma of bladder".** *Cancer Immunology and Immunotherapy. 2006, 55(6), 734-743.*

40. N.K.Mohanty, **Sunita Saxena**, Uday Pratap Singh, Neeraj K.Goyal, R.P.Arora "Lycopene **as a chemoprevetvive agent in the treatment of High Grade Prostate Intraepithelial Neoplasia.**" *Urol Oncol. 2005 Nov-Dec; 23(6):383-5.*

41. Chintamani, Jai Prakash Singh, Mahesh K Mittal, **Sunita Saxena**, Anju Bansal, Ashima Bhatia, Pranjal Kulshreshtha **The role of P--glycoprotein expression in predicting response to neoadjuvant chemotherapy in breast cancer-a prospective clinical study.** *World Journal of Surgical Oncology* 2005 **3**:61

42. **Sunita Saxena** ; Bharat Rekhi ; Anju Bansal ; Ashok Bagga ; Chintamani C and N.S.Murthy: **Clinico-morphological patterns of Breast**

Cancers Including family history in a Delhi hospital, India- A Cross-sectional study *World Journal of Surgical Oncology 2005,* ***3:****67*

43. Burra UK, Singh A, **Saxena S**. **Eccrine porocarcinoma (malignant eccrine poroma): a case report.** *Dermatol Online J. 2005 Aug 1; 11(2):17.*

44. **Sunita Saxena**, Usha Agrawal, Abhilasha Agarwal*, Saurabh Verma, NS Murthy*, NK Mohanty **"Adjuvant Intravesical Therapy Based on In Vitro Cytotoxicity Assay In Management Of Superficial Transitional Cell Cancer of Urinary Bladder"** *BJU International 2006 Vol. 98 (5), 1012*

45. Bharat Rekhi, **Sunita Saxena "New Pot-pourri of Markers related to Invasive Breast Cancer"** *JIMSA 2006 19(1), Jan-Mar.*

46. Rekhi B, **Saxena S "Cytomorphology of Basal Cell Type Of Solid Ameloblastoma-A Case Report"** *J Of Cytology 2006;23:83-85*

47. **Sunita Saxena**, Anurupa Chakraborty, Mishi Kaushal Sanjeev Kotwal, Dinesh Bhatnagar, RS Mohil, Chintamani Chintamani, AK Aggarwal, Veena Sharma, PC Sharma, Gilbert Lenior and David Goldgar, Csilla Szabo **Contribution of germline BRCA1 and BRCA2 sequence alterations in to breast cancer in Northern India.** *BMC Medical Genetics 2006, 7:75.*

48. **Chatterjee I, Kapur S,** Mahanta J, Phukan RK, Barooah MN, Kataki AC, Purkayastha J**, Saxena S. Gene Expression profile in oesophageal cancer patients of Northeast region of India**. *J Cancer Res. Ther, 2(1): S 20, 2006.*

49. Singh A, **Saxena S**. **"Infiltrating Duct Carcinoma Of Breast, Metastatic to Axillary Lymph Nodes Harbouring Primary Tuberculous Lymphadinitis"** *Pathology Oncology Research,12(3),2006*

50. Singh Avninder, Amar Bhatnagar, Usha Agrawal and **Sunita Saxena**. **Isolated splenic metastasis from colorectal mucinous carcinoma: a case report** *International Journal of Gastrointestinal Cancer 2006;37(2-3):98-101*

51. N S Murthy, Usha K Burra, K Chaudhry, and **S Saxena" Trends in incidence of breast cancer-Indian Scenario".** *Europian Jr. Of Cancer Care. doi:10.1111/j.1365-2354.2006.*

52. Anurupa Chakraborty[1], N.S. Murthy[2], Chintamani[3], D Bhatnagar[3], R.S. Mohil[3], A. Bhatnagar[3], P.C. Sharma[4]**, Sunita Saxena[1] CYP 17 gene polymorphism and its association with high-risk North-Indian breast cancer patients"** *Journal of Human Genetics 52(2):159-165,2007*

53. IndranilChatterjee, SujalaKapur, JoydeepPurkayastha, Rupkumar Phukan, Amal Kataki, Jayanta Mahanta, **Sunita Saxena**. **Gene expression profile of esophageal cancer in North East India by cDNA microarray analysis.** *World Jr of Gastroentrology 2007; 13(9):1438-1444.*

54. Chintamani, Binita P Jha, Anju Bansal, **Sunita Saxena** and Dinesh Bhatnagar **The expression of mismatched repair genes and their correlation with clinicopathological parameters and response to neo-adjuvant chemotherapy in breast cancer** *International Seminars in Surgical Oncology.20074:5.*

55. Chintamani, Pranjal Kulshreshtha, Nidhi Sugandhi, Anju Bansal, Dinesh Bhatnagar and **Sunita Saxena. Is an aggressive approach justified in the management of an aggressive cancer-the squamous cell carcinoma of thyroid?** *International Seminars in Surgical Oncology* 2007, 4:8 doi:10.1186/1477-7800-4-8

56. Chintamani, Vinay Singhal, Anju Bansal, Dinesh Bhatnagar and **Sunita Saxena. Isolated colostomy site recurrence in rectal cancer -two cases with review of literature**. *World Journal of Surgical Oncology 2007* ***5****:52*

57. Tyagi I, Agarwal U, Amitabh V, Jain A K, **Saxena S. Thickness of Glomerular and Tubular basement membranes in preclinical and clinical stages of Diabetic Nephropathy.** *Indian Jr of Nephrology* 2008; 18(2):60-65.

58. Singh, A., Kapur, S. and Saxena, S. **Cytokeratins and gastrointestinal cancer: A brief review.** *Gastroenterol Today 12 (2008) 115.*

59. Sharma, M., Chintamani, Saxena, S. and Agarwal, U. **Squamous cell carcinoma arising in unilateral Warthin's tumor of parotid gland.** *J Oral Maxillo Facial Pathol 12 (2008).*

60. Anurupa Chakraborty[1], A.K Mishra[1], Abha Soni[1], Thodum Regina[1], D Bhatnagar,[2] A Bhatnagar[2], Chintamani,[2] Sunita Saxena[1] **VDR gene polymorphism(s) and breast cancer risk in North Indian Population.** *Cancer Detection and Prevention* **32** (2009) pp. 386-394

61. Chattopadhyay I, Phukan R, Vasudevan M, Singh A, Purkayastha J, Hewitt S, Kataki A, Mahanta J, Kapur S, Saxena S; **"Molecular profiling to identify molecular mechanism in esophageal cancer with familial clustering"** *Oncology Reports 21:1135-1146,2009*

62. Chintamani, T. Aeron, M. Mittal, D. Bhatnagar, U. Agarwal, S. Saxena **Are the structures preserved in functional neck dissections truly preserved**

functionally? – A prospective study of patients with head and neck cancer at a tertiary cancer care center. *Oral Oncology Supplement, Volume 3, Issue 1, July 2009, Page 175*

63. Murthy, N.S., Chaudhry, K., Nadayil, D., Agarwal, U.K. and Saxena, S. **Changing trends in incidence of breast cancer: Indian Scenario.** *Indian J Cancer 46 (2009) 73.*

64. Avninder Singh, ; Sujala Kapur,; Indranil Chattopadhyay,; Joydeep Purkayastha,; Jagannath Sharma,; Ashwani Mishra,; Stephen M. Hewitt; Sunita Saxena,. **Cytokeratin immunoexpression in esophageal squamous cell carcinoma of high-risk populationin Northeast India.** *Applied Immunohistochemistry & Molecular Morphology* , 17(5):419-424,Oct.2009

65. Mishi Kaushal, Indranil Chattopadhyay, Rupkumar Phukan, Joydeep Purkayastha, Jagadish Mahanta, Sujala Kapur, Sunita Saxena. **Contribution of germline *BRCA2* sequence alterations to risk of familial esophageal cancer in high-risk area of India.** *Disease of the Esophagus*. DOI:10.1111/j.1442-2050.2009.00975.x (published online),2010:23(1) 715.

66. Agrawal A, Agrawal U, Verma S, Mohanty N.K and SaxenaS. **Serum Th1 and Th2 cytokine balance in patients of superficial transitional cell carcinoma of bladder pre and post intravesical combination immunotherapy.** *Immunopharmacology and Immunotoxicology.2010:32(2)348-56*

67. Regina Devi T, Yadav DS, Mishra AK, Kaushal M, Ihsan R, Chattopadhayay I, Chauhan P, Sarma J, Zomawia E, Verma Y, Nandkumar A, Mahanta J, Phukan R, Kapur S, Saxena S: **Distribution of Glutathione S- transferase T1 and M1 genes polymorphisms in North East Indians. A potential report.** *Genetic Testing and Molecular Biomarkers.* 14(2);163-169,2010

68. Chintamani Chintamani, Rohan Khandelwal, Megha Tandon, Yashwant K, Pranjal Kulshreshtha, Tushar Aeron, Dinesh Bhatnagar, Anju Bansal, Sunita Saxena **Carcinoma developing in a fibroadenoma in a woman with a family history of breast cancer: a case report and review of literature** *Cases Journal 2009, 2:9348*

69. Chattopadhyay I, Singh A, Phukan R, Purkayastha J, Kataki A, Mahanta J,Saxena S, Kapur S .**Genome-wide analysis of chromosomal alterations in patients with esophagealsquamous cell carcinoma exposed to tobacco and betel quid from high-risk area in India** Mutation Research - Genetic Toxicology and Environmental Mutagenesis 696 (2010), pp. 130-138 DOI information: 10.1016/j.mrgentox.2010.01.001

70. Rakshan I, Chattopadhyay I, Phukan R, Mishra A K, Purkayastha J, Sharma J, Zomawia E, Verma Y, Nandkumar A, Mahanta J, Saxena S, Kapur S. **Role of EPHX1 gene polymorphisms in esophageal cancer of high-risk area in India.** Jr. of Gastroenterology and Hepatology 2010 Aug; 25(8):1456-62.

71. Usha Agrawal, Ashwani K Mishra, Payal Salgia, Saurabh Verma, Nayan K Mohanty, **Sunita Saxena. Role of Tumor Suppressor and Angiogenesis Markers in Prediction of Recurrence of Non Muscle Invasive Bladder Cancer.** Pathology and Oncology Research 17(1); 91-101,2011, DOI: 10.1007/s12253-010-9287-1

72. Dhirendra Singh. Yadav, Thoudam Regina Devi, Rakhshan Ihsan, AK Mishra, Mishi Kaushal, Indranil Chattopadhyay, Pradeep Singh Chauhan, Jagannath Sharma, Eric Zomawia, Yogesh Verma, A. Nandkumar, Jagadish Mahanta, Rupkumar Phukan, Sunita Saxena., Sujala Kapur. **Polymorphisms of Glutathione-S-transferase (GST) genes and the risk of aerodigestive cancers in Northeast Indian population.** Genetic Testing and Molecular Biomarker 14(5);1-9,2010

73. Pradeep Singh Chauhan1, Rakhshan Ihsan1, Dhirendra Singh Yadav1, Ashwani Kumar Mishra1, Bharat Bhushan 1, Abha Soni1, Mishi Kaushal1, Thoudam Regina Devi1,Sumita Saluja2, Dipendra Kumar Gupta2, Vishakha Mittal2, Sunita Saxena1,Sujala Kapur. **Association of *GST, EPHX* and *p53 codon 72* gene polymorphism with adult acute myeloid leukaemia.** DNA and Cell Biology 2011 Jan; 30(1): 39-46.

74. Chintamani**,** Pranjal Kulshreshtha, Anurupa Chakraborty, L C Singh, Ashwani K Mishra, Dinesh Bhatnagar, Sunita Saxena**. Androgen receptor status predicts response to chemotherapy, not risk of breast cancer in Indian women**. *World Journal of Surgical Oncology* 2010, **8**:64

75. Kaushal M, Mishra AK,Raju BS, Ihsan R, Chakraborty A, Sharma J, Zomawia E, Verma Y, Kataki A, Kapur S. Saxena S: **Betel quid chewing as an environmental risk factor for breast cancer.** Mutation Research - Genetic Toxicology and Environmental Mutagenesis 703(2010), 143-148

76. Rakhshan Ihsan, Thoudam Regina Devi, Dhirendra Singh Yadav, Ashwani Kumar Mishra, Jagannath Sharma, Eric Zomawia, Yogesh Verma, Rupkumar Phukan, Jagadish Mahanta, Amal Chandra Kataki, Sujala Kapur, Sunita Saxena. **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 and Cell Biology 2011 March; 30(3): 163-171

77. .Chintamani, Jp Singh, Megha Tandon, Rohan Khandelwal, Tushar Aeron, Sidharth Jain, Nikhil Narayan, Rahul Bamal, Yashwant Kumar, S Srinivas, Sunita Saxena **Vulval elephantiasis as a result of tubercular lymphadenitis: two case reports and a review of the literature** *Journal of Medical Case Reports* 2010 **4**:369

78. Abha Soni, Anju Bansal, L C Singh, Ashwani Kumar Mishra, Thoudam Regina, N K Mohanty, Sunita Saxena. **Gene expression profile and mutational analysis of DNA mismatch repair genes in carcinoma prostate in Indian population.** *OMICS: 2011 Feb 24* (Epub ahead of print)

79. Simmy Soni, Gayatri Rath, Chandra Prakash Prasad, Sudha Salhan, Arun Kumar Jain, Sunita Saxena. **Fas-Fas System in Molar Pregnancy.** *American Journal of Reproductive Immunology.* (In Press)

80. Simmy Soni, Gayatri Rath, Chandra Prakash Prasad, Sudha Salhan, Sunita Saxena, Arun Kumar Jain. **Apoptosis and Bcl-2 protein expression in Human Placenta over the course of Normal Pregnancy**. *Anatomia. Hisologia Embryologia.39(2010): 426-431*

81. Chintamani**,** Megha Tandon**,** Ashwini Mishra**,** Usha Agarwal and **Sunita Saxena. Sentinel lymph node biopsy using dye alone method is reliable and accurate even after neo-adjuvant chemotherapy in locally advanced breast cancer- a prospective study.** World Journal of Surgical Oncology 2011, 9:19

82. Agarwal S, Agrawal U, Mohanty NK, Saxena S. **Multilocular Cystic Renal Cell Carcinoma: A case report of rare entity.** Arch Pathol Lab Med—Vol135, March 2011

83. L C Singh, Anurupa Chakraborty, Ashwani K Mishra, Thoudam Regina Devi, Nidhi Sugandhi[2], Chintamani, Dinesh Bhatnagar, Sujala Kapur, **Sunita Saxena[*] "Study on predictive role of AR and EGFR family genes with response to Neo-adjuvant Chemotherapy in Locally Advanced Breast Cancer in Indian women"** Medical Oncology 2012, 29(2) : 539-546

84. Singh A, Mishra A K, Ylaya K, Hewitt S M, Sharma K C, Saxena S. **Wilms Tumor-1, Claudin-1 and Ezrin are useful Immunohistochemical markers that helps to distinguish Schwannoma from Fibroblastic Meningiomas.** Pathol. Oncol. Res. DOI 10.1007/s12253-011-9456-x

85. Anju Bansal, Abha Soni, Punita Rao, LC Singh, Ashwini Mishra, N K Mohanty, Sunita Saxena. **Implication of DNA repairs genes in prostate carcinogenesis in Indian men.** Indian Journal of Medical Research.(In Press)

86. Ihsan R, Chauhan PS, Mishra AK, Yadav DS, Kaushal M, Sharma JD, Zomawia E, Verma Y, Kapur S, Saxena S. **Multiple Analytic Approaches reveal distinct Gene- Environment interactions in Smokers in Lung Cancers. PLoS One**, 6(12): e29431, 2011.

87. Bansal A, Bhatnagar A, Saxena S. **Metastasizing granular cell ameloblastoma.** J Oral Maxillofac Pathol 2012;16:122-4.

88. Abha Sony, Anju Bansal, Aswini Kumar Mishra, Jyotsna Batra, L.C. Singh, Anurupa Chakraborty, Dhirendra Singh Yadav, N. K. Mohanty, Sunita Saxena **"Association of Androgen Receptor, Prostate Specific Antigen and CYP19 gene polymorphisms to Prostate Carcinoma and Benign Prostatic Hyperplasia in North Indian population"** Genetic Testing and Molecular Biomarker (In Press)

89. Mishra AK, Agrawal U, Negi S, Bansal A, Bhatnagar A, Bhatnagar D, Chintamani, Mohil R, Saxena S. **Study on expression of AR in Breast Cancer and its correlation with other steroid receptors and growth factors**. Indian J Med Res 135, June 2012, pp 843-852

90. Chauhan PS, Ihsan R, MishraAK,Yadav DS, Saluja S, Mittal V, Saxena S, and Kapur S: **High Order Interactions of Xenobiotic Metabolizing Genes and P53 Codon 72 Polymorphisms in Acute Leukemia.** Environmental and Molecular Mutagenesis (2012)

91. Mishi Kaushal, Ashwani. K. Mishra, Jaganath Sharma, Eric Zomawia , Amal Kataki, Sujala Kapur, Sunita Saxena **Genomic alterations in breast cancer patients in betel quid** 1 **and non betel quid chewers. PLoS One (Accepted)**

Books-

1. **Mishra AK, Chakraborty A, Saxena S: Significance of Vitamin D receptor polymorphisms in Breast Cancer—Multinomial Logistic Regression Analysis in Vitamin D,: Nutrition, Side effects and supplements. Editor: Stephanie R. Malone**
2. **Sunita Saxena: Molecular and Genetic Aspects of Lung Cancer. Chapter 13 pp 141-148 In Pathological and Occupational Lung Health. Editors: V.K.Vijyan, H.K.Tazelaar and Ritu Kulshrestha**