Genetic Polymorphism in the Phase II (*GSTM1* & *GSTT1* & *GSTP1*) Detoxification Genes with Risk for Occurrence Lung Cancer and Its Association With Clinico-pathological Features In North Indian Population

Dissertation

submitted in partial fulfillment of the requirement for the award of degree of

Master in Technology

in

Biotechnology

Submitted

by

Neha Sharma
(Regn. No. 601204016)

Under supervision of

Dr. Siddharth Sharma

Assistant Professor

DBTES, Thapar University

Department of Biotechnology

Thapar University

Patiala-147004, Punjab

July, 2014

Declaration

I, the undersigned, hereby declare that the research work presented in the M.Tech dissertation entitled **"Genetic Polymorphism in the Phase II (*GSTM1* & *GSTT1* & *GSTP1*) Detoxification Genes With Risk for Occurrence Lung Cancer and Its Association With Clinico-pathological Features In North Indian Population"** has been carried out by me under the supervision and guidance of Dr. Siddharth Sharma, Department of Biotechnology, Thapar University, Patiala.

Further, I declare that no part of this dissertation has been submitted for a degree or any other qualification of any other university or examining body in India/elsewhere.

Neha Sharma 4/7/14

Neha Sharma
(Reg. No. 601204016)

Certificate

This is to certify that dissertation entitled, **"Genetic Polymorphism in the Phase II (*GSTM1* & *GSTT1* & *GSTP1*) Detoxification Genes With Risk for Occurrence Lung Cancer and Its Association With Clinico-pathological Features In North Indian Population""** submitted by Ms. Neha Sharma in partial fulfillment of the requirements for the award of M. Tech in Biotechnology at Thapar University, Patiala is an authentic work carried out by her under our supervision and guidance.

To the best of our knowledge, the matter embodied in this dissertation has not been submitted to any other university/ institute for award of any Degree or Diploma.

Dr. Dinesh Goyal
Head & Professor
Department of Biotechnology

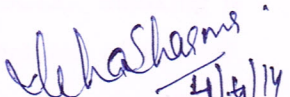
Dr. Siddharth Sharma
Assistant Professor
Department of Biotechnology

Dr. S. K. Mohapatra
Dean, Academic Affairs
Thapar University, Patiala

Acknowledgement

My sincere gratefulness goes to my supervisor Dr. Siddharth Sharma for his help, support and understanding, that made it possible for me to be successful in the pursuit and finalization of this project. I am thankful for their encouragement to keep going and to constantly learn new things.I would take this opportunity to thank, Dr. Siddharth Sharma for guiding me and correcting my mistakes along the way. I would also like to thank him for his guidance regarding the basic principles. I would like to extend my gratitude towards Dr. Dinesh Goyal, Head and Professor, Department of Biotechnology, Thapar University, for his support. I would also like to thank the members of the labs I visited and worked in for helping me with my experiments and for providing intelligent insights that made my work easier, thank you for being so friendly. I would like to thank Ms. Charu Bahl and MS Amrita Singh, Research Scholars, for their invaluable inputs regarding the finer parts of my work. I thank my friend Ms. Kushaldeep Kaur and my other lab mates, for helping me out with the smallest of things and for the support in performing all experiments and the staff members of the Department of Biotechnology, for providing a suitable environment to work in.

Finally, I would like to express my utmost gratitude to my parents, for their unconditional affection and support.



Neha Sharma

ABSTRACT

Glutathione transferases (GSTs), a multiple gene family of phase II enzymes, catalyses detoxification of endogenous compounds with glutathione and thus protects macromolecules from damage caused by carcinogenic agents. We carried out the study to investigate the potential role of *GSTT1*, *GSTM1* and *GSTP1* polymorphisms in likelihood of development of lung cancer, either separately or in combination. This case-control study consisted of 270 cases with lung cancer and 270 age and gender matched cancer free controls from Department of Pulmonary Medicine, Post Graduate Institute of Medical Education and Research (PGIMER) Chandigarh, India. The frequencies for the *GSTM1* null genotypes were 49.6 and 38.14 % among cases and controls, and for the *GSTT1* null genotype 18.52 and 13.70 % respectively. The distribution of the *GSTP1* Ile/Ile, Ile/Val and Val/Val genotypes was 83.33, 14.82 and 1.85 % in cases and 86.30, 11.85 and 1.85% in controls respectively.. In our studies *GSTM1* null genotype showed an elevated risk for lung cancer (OR= 1.65, 95%CI=1.16-2.3, *p*=0.005) but neither *GSTT1* null genotype nor the genotypes of *GSTP1* showed increased risk for overall lung cancer but when analysis was done for histological subtypes, *GSTT1 and GSTM1* null genotype showed a significant risk for developing ADCC (OR=2.32, 95%CI =1.27-4.25, *p*=0.006; OR=2, 95%CI =1.21-3.3, *p*=0.006). When analyzed in combinations, *GSTM1*, *GSTT1* and *GSTP1* genotypes did show an association with overall lung cancer and further these combinations had an elevated risk for ADCC and SCLC. Smoking is considered to be associated with lung cancer but in our case none of the genotypes of *GSTM1, GSTT1* and *GSTP1* suggested any relationship between smoking and lung cancer. Whereas, *GSTM1* null genotype showed an elevated risk for lung cancer in case of non smokers (OR=3.03, 95% CI- 1.48-6.39, *p*=0.002). The genotypes of three GSTs in triplet combinations did not show any significant relationship with overall lung cancer.

Contents

List of Tables

List of Figures

ABBREVIATIONS

GST	Glutathione S-Transferases
SCLC	Small Cell Lung Carcinoma
NSCLC	Non Small Cell Lung Carcinoma
ADCC	Adenocarcinoma
SQCC	Squamous Cell Carcinoma
C.I	Confidence Interval
CYP1A1	Cytochrome P450 1A1
BSA	Bovine Serum Albumin
DNA	Deoxyribonucleic acid
EDTA	Ethylene diamine tera cyclic acetic acid
OR	Odds Ratio
IARC	International Agency of Research on Cancer
PCR	Polymerase Chain Reaction
RFLP	Restriction Fragment Length Polymorphism
TBE	Tris Borate EDTA
TE	Tris Cl EDTA

CHAPTER 1

INTRODUCTION

Cancer is characterised by uncontrolled proliferation of cell. This uncontrolled growth leads to the accumulation of mass of the tissue *i.e.* "tumor". Every cell undergoes apoptosis and the cells which do not follow programmed cell death process become cancerous and one of the reasons may be mutation in DNA or damage to the genes involved in cell division or carcinogen metabolism.

Lung cancer is the leading cause of deaths in men and women worldwide. Over the past several years, it has been reckoned that tobacco smoking is related to lung cancer. But individual susceptibility to lung cancer could be different due to genetically determined variations in metabolism of tobacco derived carcinogens. Not that all smokers develop lung cancer but yes, people developing lung cancer do have smoking history. This indicates that identification of genetic variations can play a vital role in finding out the susceptibility factors for developing lung cancer.

Tobacco, asbestos and radon are the major risk factors of lung cancer. Human body has an inbuilt mechanism of detoxification for such carcinogens to prevent cancer. Xenobiotics metabolising enzymes are responsible for metabolism of many exogenous chemicals that are toxic, mutagenic or carcinogenic. Carcinogen detoxifying enzymes include the phase I enzymes and Phase II enzymes. Removal of undesirable or harmful compounds is done by phase I (transformation) and phase II (conjugation) enzymes. Both phase I and phase II enzymes work in unison for the removal of xenobiotic compounds.

Principally phase I detoxification system is mediated by cytochrome P450 multigene family. Various phase I metabolic reactions such as C, N and S oxidation and dealkylation are performed by cytochrome P450s (Androutsopoulos *et al.,* 2009). A number of genetic polymorphisms have now been established for several isoenzymes of the P450 enzyme system (i.e., in the *CYP1A1, CYP2A6, CYP2C9, CYP2C18, CYP2C19, CYP2D6,* and *CYP2E1* genes). Considering the fact that bioactivation by P450 enzymes may play an

important role in human drug toxicity and human cancer, these polymorphisms may result in large interindividual variations in the metabolism and toxicity of xenobiotics.

Phase II detoxification system mainly consists of Glutathion - S- Transferases which are soluble dimeric enzymes and are further classified as Alpha, Kappa, Mu, Pi and Theta. All act on different substrates converting them into soluble compound and facilitating them out of the body. GSTs are mainly involved in the detoxification of reactive compounds generated by phase I enzymes. Apart from detoxification of xenobiotics compounds, they have a potent role in chemical carcinogenesis. The genetic polymorphism so far detected in human population is in the genes encoding *GSTM1*, *GSTT1* and *GSTP1* (Wormhoudt *et al.*, 1999).

The polymorphism found in the different GSTs has been found to be playing a vital role in determining the metabolism capacity of an individual to detoxify carcinogens. Therefore, due to various differences in population ethnicity and inter individual variations at genetic levels, the study to find out role of these polymorphisms in GSTs is required. So, we are carrying out the study to find out association of *GSTP1*, *GSTT1* and *GSTM1* polymorphisms with risk of development of lung cancer.

CHAPTER 2

REVIEW OF LITERATURE

2.1 Lung Cancer

In both men and women, lung cancer is of two types based on histology, Small cell lung carcinoma and Non small cell lung carcinoma. Lung cancer is the second most common cancer apart from skin cancer. Lung cancer accounts for about 13% of all new cancers.

Risk factors likely to develop lung cancer are:

Tobacco smoke: Atleast 80% deaths are caused due to tobacco smoking and easy availability of manufactured cigarettes has increased this percent. More the number of packs you smoke greater is the risk for developing lung cancer.

Second hand smoke: If you don't smoke but inhalation of environment tobacco smoke also increases the risk of developing lung cancer. A non smoker who is exposed to tobacco smoke at his workplace or who lives with smokers has 20-30% greater chances of developing lung cancer.

Radon: It is a naturally occurring radioactive gas and in US it is the second leading cause of lung cancer according to US Environmental Protection Society.

Asbestos: It has been seen that both smokers and non smokers who are exposed to asbestos have a great tendency to develop mesothelioma (cancer that starts in the pleural lining of the lung). People who are exposed to asbestos and smoke also are at greater risk of developing lung cancer in comparison to these exposure factors separately. (http://www.cancerresearchuk.org)

2.2 Some Signs And Symptoms Of Lung Cancer

- A prolonged cough that worsens over time
- Hoarseness

- Problem in breathing and severe chest pain
- Loss of appetite and reduced weight
- Fatigue
- Shortness of breath
- Coughing up blood or phlegm

2.3 Diagnosis

Imaging tests like chest X-ray, CT scan, MRI scan, PET scan etc are used to diagnose lung cancer. Cytology tests and microscopic examination of the piece of abnormal mass of tissue (biopsy) are carried out. Then staging is done to decide the treatment and prognosis. TNM staging is done where T indicates the size of tumor and its spread; N indicates the spread of cancer to lymph nodes and M indicates whether the tumor has metastasised to other organs or not.

2.4 Histology

Lung cancer can be classified as follows:

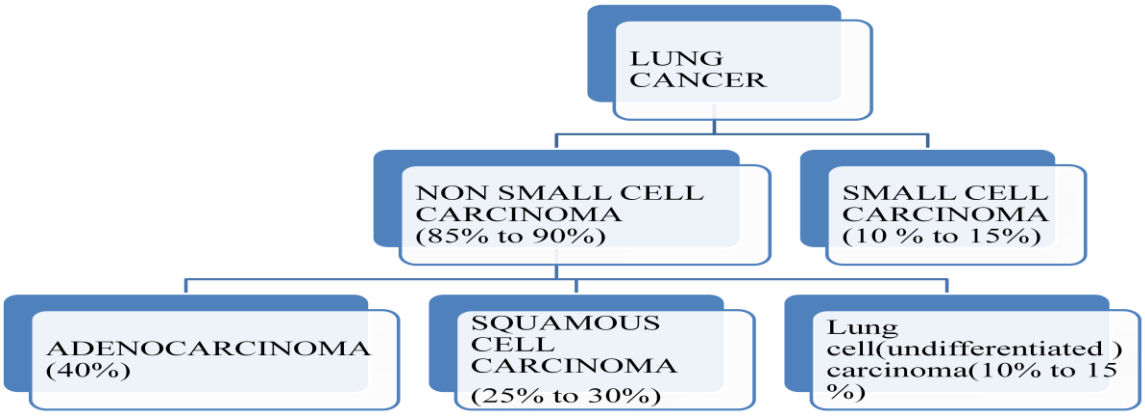


Figure: 2.1 Classification of lung cancer

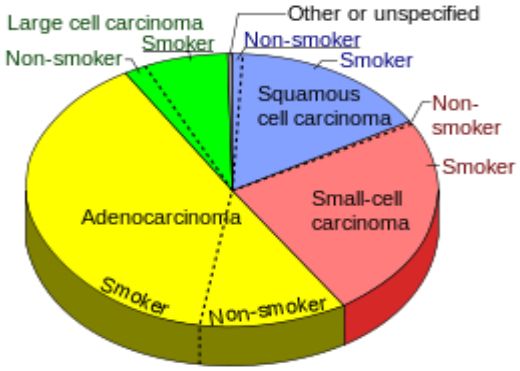


Figure2.2: Distribution of lung cancer subtypes among smokers and non smokers (Alberg *et al.*, 2010)

2.5 Incidence

Worldwide

Tobacco is the main risk factor causing lung cancer which is further responsible for causing 1 in 5 cancer deaths and almost three quarters of lung cancer deaths. In 2012, an estimated 14.1 million new cases of cancer occurred worldwide. Worldwide Lung cancer is the most common cancer in men. It has been estimated that more than 1 in 10 of all cancers diagnosed in men are lung cancers.

Table: 2.1 Estimated Incidence, Mortality and Prevalence Worldwide in 2012 (Globocan, 2012)

Estimated Numbers (thousands)	**Men**		**Women**	
	Cases	**Deaths**	**Cases**	**Deaths**
World	1242	1099	583	491
More developed regions	490	470	268	210
Less developed regions	751	682	315	281
WHO Africa region (AFRO)	12	11	6	6
WHO Americas region (PAHO)	178	149	146	113
WHO East Mediterranean region (EMRO)	26	23	7	6
WHO Europe region (EURO)	323	283	126	105
WHO South-East Asia region (SEARO)	116	104	46	42
WHO Western Pacific region (WPRO)	588	528	251	220
United States of America	112	92	102	76
China	459	422	193	175
India	54	49	17	15

India

Lung cancer scenario in India

Sex distribution of smokers (Khurana *et al.*, 1994)

Males: 33.4%

Females: 1.4%

Smoking prevalence in varied areas

Rural areas: 31.3%

Urban areas: 21.5%

Relative risk of developing lung cancer

Beedi smokers: 2.64

Cigarette smokers: 2.23

Overall: 2.45

Incidence rates (age-standardized) per 100,000 population (2008) (SIRO, 2011)

Males: 10.9

Females: 2.5

Increase in new cancer cases per 100,000 populations each year over 24 years (till 2005)

Chennai: 160%

Bangalore: 100%

Delhi: 40%

Mumbai: 60%

Most prevalent is NSCLC amongst lung cancers (Kenfield *et al.*, 2008)
Squamous cell carcinoma: 44.73%

Adenocarcinoma: 30.26%

2.6 Epidemiology

The risk factors of lung cancer are:

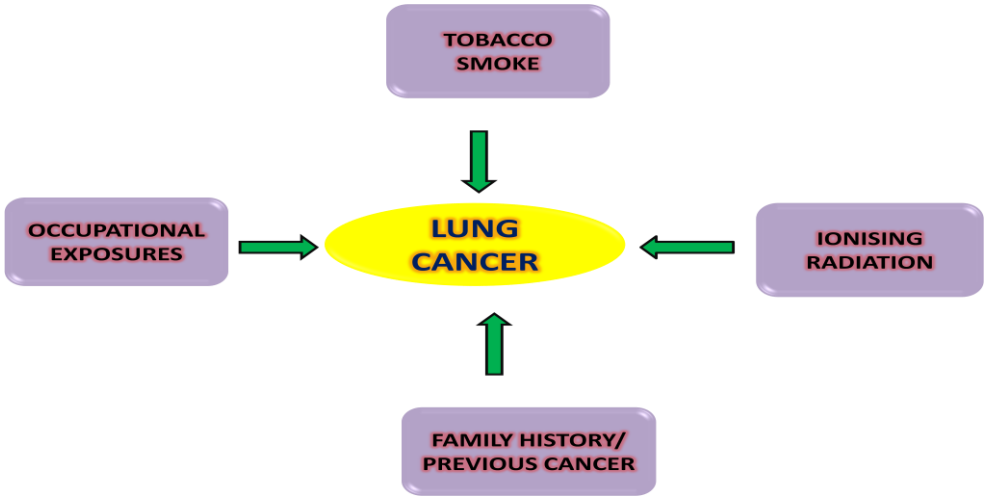


Figure 2.3: Various risk factors of lung cancer

Tobacco smoking: According to IARC 80% of the lung cancer cases are due to tobacco smoking. Cigarettes consist of several harmful chemicals. Nearly 4000 irritating, suffocating, dissolving, inflammable, toxic, poisonous, carcinogenic gases and substances and even radioactive compounds (nickel, polonium, plutonium, etc.) have been identified in tobacco smoke. Some of these are: Benzopyrene, dibenzopyrene, benzene, isoprene, toluene (hydorcarbons); naphthylamines; nickel, polonium, plutonium, arsenic, cadmium (metallic constituents); carbon dioxide, methane, ammonia, nitrogen dioxide, hydrogen sulphide (gases); methyl alcohol, éthanol, glycerol or glycerine, glycol (alcohols and esters); acetaldehyde, acrolein, acetone (aldehydes and ketones); cyanhydric or prussic acid, carboxyl derivatives (acids); chrysene, pyrrolidine, nicoteine, nicotinine, nicoteline, nornicotine, nitrosamines (alkaloids or bases); cresol (phenols), etc.(Parkin *et al.*,2011)

Genetic Factors

Almost 8 to 14% involvement of inherited factors has been estimated in causing lung cancer (Cogliano *et al*., 2011). The risk increases 2.4 times in the people who have family history of lung cancer. This is likely due to a combination of genes.

Inactivation of tumor suppressor genes and activation of oncogenes plays a vital role in lung cancer initiation. Cancer susceptibility increases due to oncogenes. Proto oncogenes when exposed to particular carcinogens are believed to turn into oncogenes (Salgia *et al.,* 2002)

Mutations in the *K-ras* proto-oncogene are responsible for 10–30% of lung adenocarcinomas (Herbst *et al.,* 2005). Proliferation, apoptosis, angiogenesis, and tumor invasion is being regulated by the epidermal growth factor receptor (EGFR). Mutations and amplification of EGFR are common in non-small-cell lung cancer and provide the basis for treatment with EGFR-inhibitor. Chromosomal damage can lead to loss of heterozygosity which can cause inactivation of tumor suppressor genes.

2.7 Detoxification System

In one's life time human body is exposed to various xenobiotics compounds and our body has developed complex enzymatic mechanisms to detoxify these compounds and render them out from the body. Environment, lifestyle, and genetic influences show a great deal of affect on these mechanisms and thus leading to individual variability. An impaired detoxification system has been associated to certain diseases including cancer, Parkinson's disease; immune dysfunction syndrome *etc* (Liska, 1988).Our detoxification system consists of phase I and phase II enzymes.

Phase I

Phase I (transformation) enzymes consists of Cytochrome P450 family which are mainly oxygenases and show first enzymatic defence against foreign compounds. The phase I enzymes by the process if biotransformation leads to the formation of reactive molecules which if not metabolised further by phase II enzymes may cause damage to protein, RNA and DNA.

Phase II

The main aim of phase II enzymes is to perform conjugating reactions and the conjugates formed are hydrophilic. The enzymes involved in phase II are mainly transferases *i.e.*

- UDP-GLUCURONOSYLTRANSFERASES (UGTs; EC 2.4.1.17)
- SULFOTRANSFERASES (SULTs; EC 2.8.2.1)

- N-ACETYLTRANSFERASES (NATs; EC 2.3.1.5)
- GLUTATHIONE S-TRANSFERASES (GSTs; EC 2.5.1.18)
- THIOPURINE S-METHYLTRANSFERASE (TPMT; EC 2.1.1.67)
- CATECHOL O-METHYL TRANSFERASE (COMT; EC 2.1.1.6)
  (Jancovaa *et al.*, 2010)

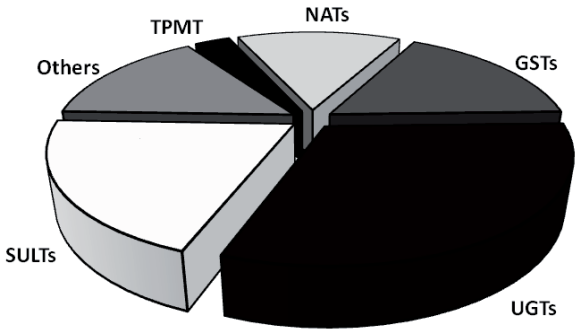


Figure 2.4: Participation of major phase II enzymes in the metabolism of clinically used drugs. (Jancovaa *et al.*, 2010)

Glutathione S-Transferases : The main contribution in the phase II metabolism of xenobiotic compounds is of GSTs. The epoxides derived from polycyclic aromatic hydrocarbons (PAHs) and alpha-beta unsaturated ketones are detoxified. GSTs are multigene family of enzymes involved in the detoxification and inactivation of various anticancer drugs and in a few instances, activation of wide variety of chemicals.

The nucleophilic attack of GSTs on their electrophilic substrates decreases their activity with cellular macromolecules. GSTs are virtually found in plant, animals and bacteria. Human GSTs are divided into three main families: cytosolic, mitochondrial and membrane-bound microsomal. The mammalian cytosolic family of GSTs exist as monomers and are catalytically active in a homo or heterodimeric state (Mannervik *et al.*, 2005). Mitochondrial GST referred to as GST kappa. In humans and rodents 4% of the total soluble proteins in the

liver are constituted by GSTs. The soluble dimeric proteins consist of subunits with molecular weight of 25kDa. Each subunit has as active site consisting of two functional regions; one that binds the physiological substrate glutathione and the other provides hydrophobic environment for the binding of electrophilic substrates.

2.8 Catalytic Function

The general reaction catalysed by GSTs is: GSH + R-X $\rightarrow$GSR + HX

The main function of the enzyme is:

1. The substrates are brought into close proximity with glutathione (GSH) by binding both, the GSH and the electrophilic substrate to the active site of the protein.
2. Nucleophilic attack of GSH is allowed on electrophilic substrate R-X after the activation of sulfhydryl group on GSH.

Other catalytic reactions identified are epoxide ring openings, nucleophilic aromatic substitution reactions, reversible additions to unsaturated aldehydes and ketones, isomerisation and for few GSTs peroxide reactions.

2.9 Substrates, Inhibitors And Inductors

Typical GST substrates

- Epoxides
- α,β-unsaturatedketes
- quinones
- sulfoxides
- esters
- peroxides
- ozonides
- Etacrynic acid-substrate for *GSTP1*
- *trans*-stilbene oxide is a diagnostic substrate for *GSTM1*
- Methylene chloride, ethylene dibromideor isoprene derivates have been shown to be conjugated by *GSTT1*
- 1-chloro-2,4-dinitrobenzene- ′universal′ GST substrate

GSTs Inhibitors

- synthetic and naturally-occurring phenols
- quinines
- derivatives of vitamin C
- retinoic acid as an inhibitor of human placental and liver glutathione transferases in the micromolar range
- inhibition of human GSTs by dopamine, α-methyldopaand glutathionyldopamine (Yokota *et al.*, 2010)

GSTs inductors

- Extracts of *Ginkgo biloba* induce *GSTP1*
- Extracts from cruciferous vegetables (e.g. broccoli, Brussels sprouts, cabbage) as well as grapefruit extract act as inducers of human GSTs (Williamson *et al*., 1997).

2.10 Tissue Specific Expression Of Human GSTs

Not all GST isoforms are expressed in all tissues because of difference in regulation of expression. Furthermore, some GSTs are polymorphic and deletion of *GSTT1* and *GSTM1* is prevalent in human population. Thus, individuals homozygous for *GSTT1* or *GSTM1* deletion will lack expression of the related enzymes in any tissue in the body. Due to such inter individual genetic differences and modulation by diet and xenobiotics compounds it is difficult to predict accurately the extent of expression of any GST gene in a given tissue. GST alpha are highly expressed in liver, kidneys and testis but not in lungs whereas *GSTP1* is highly expressed in brain and lung but not in liver. Mu, class of the GSTs, depending upon their subfamily is expressed to different extents in different tissues (Yang *et al.*, 2002).

2.11 Metabolism Of Xenobiotics And Endobiotics By GSTs

The ability of GSTs to inactivate xenobiotics and cytotoxic compounds within the cell has been extensively studied. Transferases play an important role in detoxifying noxious chemicals that may lead to mutagenic events. Example of such compounds includes benzo (a) pyrene and aflatoxin. Apart from detoxification of foreign compounds, there are numbers of by products of normal metabolism that are GST's substrate. Not all reactions catalysed by GSTs leads to the formation of less toxic products; there are instances where it leads to the

formation of more reactive and potentially carcinogenic compounds. Various enzymes are involved in the phase II of the detoxification system; they are as shown in the Figure 5.

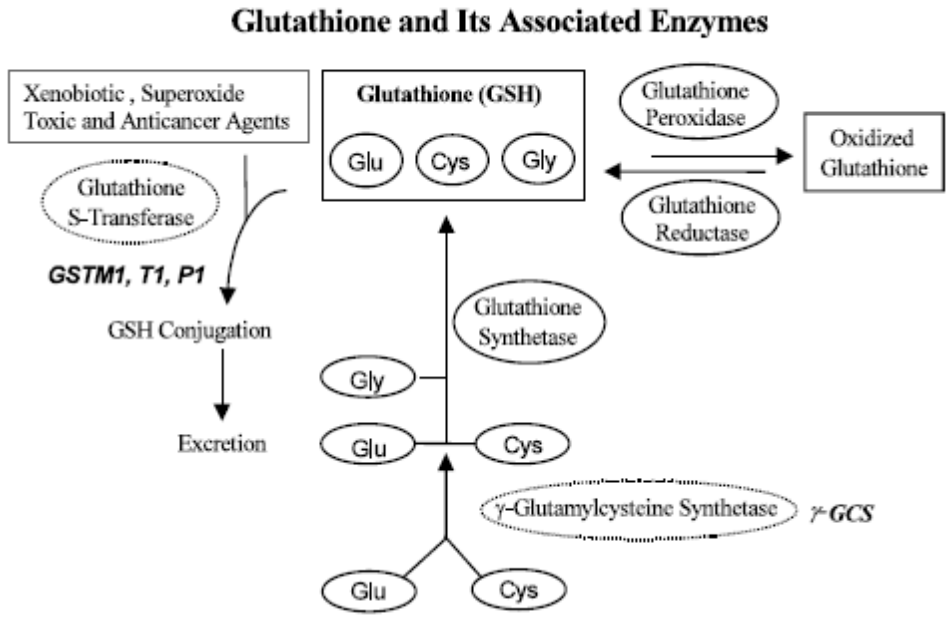


Figure 2.5: Glutathione and its associated enzymes. The three polymorphic genes in the study are *GSTM1*, *GSTT1* and *GSTP1* (Yang *et al.*, 2002)

The role of these enzymes has been studied in anticancer treatment sensitivity and responses too. Genetic polymorphisms with functional significance have been well characterised for GSH related enzymes. The polymorphism of the GST genes might alter the detoxification system and thus it is important to study the association of GST polymorphism with cancer risk (Yang *et al.*, 2002).

2.12 GST Polymorphism And Lung Cancer

GSTM1, *GSTT1* and *GSTP1* are found to be polymorphic in human population. As these enzymes are involved in the detoxification of various xenobiotics and carcinogenic substances, numerous epidemiological studies have examined their association with different types of cancers.

The major GST isoenzymes are *GSTM1*, *GSTT1* and *GSTP1* and the genes coding these enzymes are located on chromosome 1p13.3, 22q 11.2 and 11q 13 respectively and are shown

to be polymorphic in nature (Sharma *et al.,* 2014). Both the *GSTM1* and *GSTT1* polymorphism are represented by two active alleles and a non-functional null allele which is a homozygous deletion. *GSTT1* is also represented by functional allele and a non-functional (null) allele. For the *GSTP1* gene a point mutation exists in exon 5 which results in the substitution of the isoleucine to valine amino acid. This polymorphism results in the reduced catalytic activity of *GSTP1*. In case of *GSTP1* polymorphism three possible genotypes for the polymorphism at codon 105 are seen: Wild genotype (*Ile/Ile*), homozygous mutant (*Val/Val*) and heterozygous mutant (*Ile/Val*) (Zimniak *et al.,* 1994). The frequencies of *GSTM1* null genotype show variations between different populations and ethnicity. Both the Europeans (42–60%) & Asians (41–63%) exhibit a high frequency of the null genotype for *GSTM1* gene as compared with that of Africans (16–36%) (Cotton *et al.,* 2000; Hayes *et al.,* 2005) However, the frequency of *GSTT1* null genotypes is highly represented in case of African (14–57%) and Asian populations (35–48%) and somewhat less in Europeans (13.31%) (Sharma *et al.,* 2012). The frequency of mutant (*Val/Val*) *GSTP1* genotype is highly represented among African-Americans (19%), followed by Caucasians (6.5-11.7%) and least common among Japanese or Asian populations (0-3.1 %) (Gong *et al.,* 2012).

Homozygous deletion of *GSTM1* gene results in *GSTM1**0 allele and thus there is loss of enzyme activity. *GSTM1*-1 catalyzes the metabolism of a large number of potentially genotoxic compounds. The studies have shown that the polymorphism of *GSTM1* affects the metabolism of xenobiotics compounds and there could be DNA adducts formation which may lead to cancer. Various statistical studies related to *GSTM1* polymorphism has been carried out. The Meta analysis carried out by Raimondi *et al.,* 2006 has been summarised in Table 2.2.

Table 2.2: Meta analysis of lung cancer associated with *GSTM1* null genotype (Raimondi *et al.,* 2006)

	No. of studies	**Lung Cancer Patients**		**Controls**		**OR**
		Total	***GSTM1* Deficient (%)**	**Total**	***GSTM1* Deficient (%)**	
Total	43	7463	52.4	10789	50	1.17
Caucasians	20	4039	51.8	6000	50.2	1.10
Asians	12	1841	58.9	2787	53	1.33
African-Americans	3	383	27.4	503	24.3	1.19
Mexican-Americans	1	60	55	146	40.4	1.80
Mixed	4	785	58.1	894	56.2	1.10

In some studies deletion of *GSTT1* has been associated with lung cancer risk (Yang *et al.,* 2013) but in some populations no association has been reported (Raimondi *et al.,* 2006). Meta analysis carried out by Raimondi *et al.,* (2006) has been summarised in Table 2.3

Table 2.3: Meta- and Pooled Analysis of *GSTT1* and Lung Cancer (Raimondi *et al.,* 2006)

Authors	**No. Of Cases**	**No. Of Controls**	**Country**	**OR**
Reszka *et al.,* 2005	119	138	Poland	0.51
Alexandrie *et al.,* 2004	524	530	Sweden	0.92
Belogubova et. al, 2004	167	663	Russia	0.94
Harms *et al.,* 2004	110	119	US	1.55
Schneider *et al.,* 2004	446	622	Germany	0.89
Sobti *et al.,* 2004	100	76	India	1.30
Sørensen *et al.,* 2004	254	265	Denamrk	2.65
Dialyna *et al.,* 2003	122	178	Greece	1.64
Ruano-Ravina *et al.,* 2003	125	185	Spain	0.84

Lewis *et al.,* 2002	87	143	US	1.15
Stucker *et al.,*.2002	251	268	France	0.74
Hou *et al.,* 2001	184	162	Sweden	1.06
Liu *et al.,* 2001	1024	1176	US	0.89
Risch *et al.,* 2001	383	346	Germany	0.63
Chan-Yeung *et al.,* 2004	229	187	China	1.55
Wang *et al.,* 2003	112	119	Japan	1.08
Zhao *et al.,* 2001	223	187	Singapore	1.09
Gallegos-Arreola *et al.,* 2003–2004	52	178	Mexico	5.04
Malats *et al.,* 2000	122	121	Sweden, Germany, France, Italy, Russia, Rumania, Poland, and Brazil	0.62

Carcinogenic and toxic electrophiles are inactivated by *GSTP1* by conjugation PAH found in cigarette smoke are detoxified by GSTP. *GSTP1* plays an important role in protecting cells from cytotoxic and carcinogenic agents and is expressed in normal tissues. Exon 5 A to G transition (Ile 105 Val) results in significantly lower conjugation activity among individuals who carry one or more copies of the G allele (*Ile/Val* or *Val/Val*) compared with those who have (*Ile/Ile*). Although, *GSTP1* is the most abundant GST isoform in lung, the association between *GSTP1* polymorphisms and lung cancer are not well understood. A Meta analysis carried out by Cote *et al.,* (2008) has reported that there is no association between polymorphism of *GSTP1* and lung cancer risk. His reportings are summarized in Table 2.4.

Table 2.4: Meta analysis on association between at least one 105 Val at *GSTP1* exon5 and lung cancer (Cote *et al.,* 2008).

Asian Studies	**Case/Controls**	**Country**	**OR *Val/Val* & *Ile/Val* Vs *Ile/Ile***
Wang *et.al*,2003	112/119	China	1.61
Liang *et.al*,2005	227/227	China	1.22
Kiyohara *et.al*,2000	62/80	Japan	1.18
Lin *et.al*, 2003	198/332	Japan	1.27

Different studies conducted so far in various populations to examine the role of these genes towards individual susceptibility towards lung cancer risk has been contradictory. A few studies done in an Indian population have also had conflicting data. It might be plausible that gene-gene and gene-environment interactions might cause different results within the same population strata because of different lifestyles. Apart from that sample size and different clinico-pathological features such as histological subtypes, stage of tumor *etc.* might also differ between studies.

As we know that *GSTM1* and *GSTT1* do play a role in the process of detoxification it would be cogent that genetic polymorphism of these enzymes plays an important role towards host susceptibility towards lung cancer. On the other hand there has been no strong body of evidence to suggest that *GSTP1* polymorphisms alone might play a role towards pre-disposition towards lung cancer risk.

CHAPTER 3

AIM OF THE STUDY

1. To investigate the role of *GSTT1*, *GSTM1* and *GSTP1* polymorphism as a genetic modifier for risk towards lung cancer and its correlation with different clinico-pathological features within a North Indian population.

2. To investigate whether doublet or triplet combinations of polymorphisms for *GST* genes are a risk factor for lung cancer development.

3. To investigate the role of smoking as a modifier for lung cancer.

CHAPTER 4

MATERIAL AND METHODS

4.1 Study Subjects And Sample Collection

A total of 270 Lung cancer patients were recruited for this study from the Department of Pulmonary Medicine, Post Graduate Institute of Medical Education and Research (PGIMER) Chandigarh, India. This study had been reviewed and approved by the Institute ethics committee of PGIMER. Informed written consent was obtained from all participants or their representatives. In brief, eligible cases included all patients who were newly diagnosed with primary lung cancer. All the recruited patients were histopathologically diagnosed as having NSCLC and SCLC. There were no age, gender, smoking, histological, or TNM stage restrictions, but patients with a prior history of cancer were excluded from this study. During the same time, 270 unrelated individuals with no evidence of lung or other cancer who entered the hospital for health check-ups were enrolled as control group. Each control was pair-matched by sex, age (±10 years) and smoking parameters to a patient with lung cancer. These characteristics allowed us to obtain control population without any possible risk bias for lung cancer. A detailed questionnaire was completed for each case and control by a trained interviewer. The questionnaire included information on demographic and smoking characteristics. Smokers reported tobacco habits such as smoking of cigarette and/or beedi (a native cigarette like stick of coarse tobacco hand-rolled in a dry tembuhurni leaf). As an indication of cumulative smoking exposure, pack-years were calculated by the following formula: [(cigarettes or beedis per day / 20) X years smoked]. While medical information of cases, including Histology, TNM classification, clinical staging, primary tumor size, involvement of lymph node and metastasis were obtained from medical records of the hospital. Approximately 3-5ml of venous blood was collected from each participant.

4.2 Isolation Of DNA From Blood

Isolation of DNA from peripheral blood: Genomic DNA was isolated using standard Protein K digestion, phenol/chloroform extraction and ethanol precipitation method from whole blood samples of both cases and controls (Barlett and White's method).

Requirements

- Washing buffer
- Lysis buffer
- Phenol:Chloroform:Isoamylalcohol (25:24:1)
- Chloroform:Isoamylalcohol (24:1)
- Isopropanol
- TE buffer

Procedure

Preparation of Buffers

Washing buffer, Lysis buffer and TE buffer were prepared as shown in tables below.

Table 4.1: Preparation of Washing buffer

Stock Concentration	Working Concentration
1M sucrose	320 mM sucrose
100% Triton X-100	1% Triton X-100
100mM Magnesium Chloride	5mM Magnesium Chloride
100mM Tris-HCl pH (8.0)	10mM Tris-HCl pH (8.0)

Table4.2: Preparation of lysis Buffer

Stock Concentration	Working Concentration
1M Tris HCl pH (8)	400mM Tris HCl pH (8)
10% SDS	1% SDS
0.5M EDTA	60mM EDTA

5M NaCl	150mM NaCl
10mg/ml Proteinase-K	100 µg/ml Proteinase –K

Procedure of DNA Isolation

- Took 5ml of blood and added 5ml of washing buffer and mix it thoroughly. Centrifuged it at 3500 rpm for 5 minutes.
- Discarded the supernatant and added 5ml of washing buffer (1.6ml 1M Sucrose,0.5 ml Triton X-100, 0.25ml MgCl2, 0.5 ml 100mM Tris HCl and 0.26ml of water) to the pellet and resuspended the pellet in the buffer and centrifuged again (repeat this step thrice).
- Dissolved the pellet in 5ml of Lysis buffer (1 M Tris HCl 2ml,10% SDS 0.5ml ,0.5 M EDTA 0.6ml, 5M NaCl 0.15ml,10mg/ml Proteinase-K 0.05ml and water 1.7ml) and incubated at 44 °C overnight.
- Added an equal volume of Phenol: chloroform: Isoamyl alcohol (PCI) 25:24:1(2.5ml Phenol, 2.4 ml chloroform and 0.1ml isoamyl alcohol) and mixed the contents slowly.
- Centrifuged at 8000 rpm for 10 minutes at 4°C. Took the upper aqueous layer and again add PCI mix and centrifuged.
- Took the aqueous layer and added equal volume of Chloroform: Isoamyl alcohol (24:1).
- Centrifuged it at 6500 rpm for 5 minutes and took the upper layer.
- To the aqueous layer added equal volume of chilled isopropanol or 2.5 times volume of absolute ethanol and mixed it gently.
- Freeze it at -20°C for 1-2 hours.
- Centrifuged it at 12,000 rpm for 10 min at 4°C. Discarded the supernatant and washed the pellet of DNA with chilled 70% ethanol twice at 10,000 rpm for 5 minutes.
- Decanted ethanol and air dry the pellet.
- Dissolved the pellet in 50µl-150µl Tris-EDTA buffer depending on the size of DNA pellet (Bartlett & White, 2003).

4.3 DNA Quantification

The nanodrop spectrophotometer holds 1µl of sample without the need of traditional containment devices such as cuvettes and capillaries. Using fibre optic technology and surface tension, the sample is held in place between two optical surfaces that define the path length in vertical orientation. Removal of fixed containment devices from the system allows the pathlenghth to change in real time for a sample. This essentially eliminates the need to perform dilutions.

Procedure

- Pipetted 1µl of deionised water onto the lower optical surface of Nanodrop (Thermo Scientific) to clean it
- Opened the nanodrop software and selected nucleic acid module
- Performed a blank measurement by loading 1µl of TE and selecting "blank" from the screen
- Measured the nucleic acid sample by loading 1µl of DNA sample and selecting "measure"
- Concentration and purity of DNA samples were calculated automatically

DNA concentration otherwise can be calculated as:

DNA concentration (µg/ml) = O.D at 260nm X 50 X Dilution factor

Where 50µg/ml of DNA is equal to 1 O.D

Purity of DNA = O.D at 260nm/O.D at 280nm

NOTE: A ratio of ~1.8 indicates ―pure for DNA; a ratio of ~2.0 is generally accepted as pure for RNA. If the ratio is appreciably lower in either case, it may indicate the presence of protein, phenol or other contaminants.

4.4 Resolution Of DNA Fragments On Agarose Gels

Requirements

- Electrophoresis buffer (TAE or TBE)

- Ethidium bromide solution
- Electrophoresis-grade agarose
- 6X loading dye
- DNA molecular weight markers
- Horizontal gel electrophoresis apparatus
- Gel casting platform
- Gel combs (slot formers)
- DC power supply

Procedure

Preparing 5X TBE (1000ml)

Tris base - 54 g

Boric Acid - 27.5g

EDTA (0.5M) - 20ml

Make up final volume with water.

Preparing 6X Loading Dye (20ml)

0.25% Bromophenol blue - 0.05gm

0.25% Xylene Cyanol - 0.05gm

40% Sucrose - 8gm

Make up final volume with TE buffer.

Preparing the gel

- Prepared an adequate volume of electrophoresis buffer.
- Added the desired amount of electrophoresis-grade agarose to a volume of electropho- resis buffer sufficient for constructing the gel. For example for genomic

DNA 0.7% gel (0.7g agarose in 100ml 0.5X TBE) was prepared while for the PCR products 1.7% gel (1.7g agarose in 100ml 0.5X TBE buffer) was prepared.

- Melted agarose was cooled to 55°C in a water bath before pouring onto the gel platform. to prevent warping of the gel apparatus.
- Before pouring ethidium bromide solution was added to the electrophoresis buffer to a final concentration of 0.3µg/ml to facilitate visualization of DNA when seen under UV transilluminator.
- Poured the melted agarose onto gel casting apparatus between 0.5 and 1 cm thick and inserted the gel comb, making sure that no bubbles are trapped underneath the combs and all bubbles on the surface of the agarose were removed before the setting of the gel.

Loading and running the gel

- After the gel got hardened, gel comb was withdrawn taking care not to tear the sample wells.
- Placed the gel casting platform containing the set gel in the electrophoresis tank. Added sufficient electrophoresis buffer to cover the gel until the tops of the wells are submerged. Made sure no air pockets were trapped within the wells.
- DNA samples were prepared by mixing 5µl DNA with 2µl of 6X loading dye and 2ul water in case of genomic DNA or by mixing 5µl DNA with 2µl of 6X loading dye in case of PCR product.
- Samples were typically loaded into the wells with micropipette. Care was taken to prevent mixing of the samples between wells.
- Appropriate DNA molecular weight marker was also loaded in case of PCR products.
- Connected the electrodes to a power pack, turned on the power, and allowed the electrophoresis run at 60 V until the marker dyes migrated the desired distance.
- Turned off the electric power, disconnected the leads, and discarded the electrophoresis buffer from the reservoirs
- DNA was visualized by placing the gel on a UV transilluminator (Bio Imaging Systems) and then photographed by using Gel Doc (BIORAD).

4.5 Multiplex Polymerase Chain Reaction (PCR) Amplification Of *GSTM1* And *GSTT1*

Considerable time and effort can be saved by simultaneously amplifying multiple sequences in a single reaction, a process referred to as multiplex polymerase chain reaction (PCR). Multiplex PCR requires that primers lead to amplification of unique regions of DNA, both in individual pairs and in combinations of many primers, under a single set of reaction conditions. In addition, methods must be available for the analysis of each individual amplification product from the mixture of all the products. For a successful multiplex PCR assay, the relative concentration of the primers, concentration of the PCR buffer, balance between the magnesium chloride and deoxynucleotide concentrations, cycling temperatures, and amount of template DNA and Taq DNA polymerase are important. An optimal combination of annealing temperature and buffer concentration is essential in multiplex PCR to obtain highly specific amplification products. Magnesium chloride concentration needs only to be proportional to the amount of dNTP, while adjusting primer concentration for each target sequence is also essential (Markoulatos *et al.*, 2002).

Requirements

- 10X PCR buffer
- BSA
- Forward Primer
- Reverse Primer
- dNTP‘s
- Taq DNA polymerase
- Water
- DNA sample

Note: Primer sequences need to be chosen to uniquely select for a region of DNA, avoiding the possibility of mishybridization to a similar sequence nearby. Primers should not easily anneal with other primers in the mixture (either other copies of same or the reverse direction primer); this phenomenon can lead to the production of 'primer dimer' products

contaminating the mixture. Primers should also not anneal strongly to themselves, as internal hairpins and loops could hinder the annealing with the template DNA.

Given below are the primers used.

GSTM1 primers:

Forward primer	5’ GAA CTC CCT GAA AAG CTA AAG C 3’
Reverse primer	5’ GTT GGG CTC AAA TAT ACG GTG G 3’
Band size	480bp

GSTT1 primers:

Forward primer	5’-TTC CTT ACT GGT CCT CAC ATC TC 3’
Reverse primer	5’-TCA CCG GAT CAT GGC CAG CA 3’
Band size	215bp

Albumin primers:

Forward primer	5’-GCCCTCTGCT AACAAGTCCT AC-3’
Reverse primer	5’-GCC CTA AAA AGA AAA TCG CCA ATC-3’
Band size	312bp

Table 4.3: Reaction mixture of multiplex PCR carried out for *GSTM1* and *GSTT1* (20 reactions)

Reagent	**Stock Concentration**	**Final Reaction Concentration**	**Quantity Used**
Additive 1 BSA	1000 µg/ml	100 µg/ml	33 µl
PCR Buffer (Mg Conc.)	10 X 25mM	1 X 1.5mM total	33 µl
GSTM1 Primer (Forward)	10µM	0.5 µM	16.50 µl
GSTM1 Primer (Reverse)	10µM	0.5 µM	16.50 µl
GSTT1 Primer (Forward)	10µM	0.5 µM	16.50 µl

GSTT1 Primer (Reverse)	10µM	0.5 µM	16.50 µl
Albumin Primer (Forward)	10µM	0.3 µM	9.90 µl
Albumin Primer (Reverse)	10µM	0.3 µM	9.90 µl
Taq Polymerase	2.0U/µl	1.5mM	16.50 µl
dNTP	10mM each	0.2mM each	6.60 µl
PCR Grade Water			87.1 µl
DNA Template	100ng/µl	400ng	4 µl

Table 4.4: Cycling profile of PCR for *GSTT1* and *GSTM1*

Steps	Temperature	Time
1. Initial Denaturation	95°C	5min
2. Denaturation	94°C	1 min
3. Annealing	59°C	1 min
4. Polymerization	72°C	1 min
5. Final Extension	72°C	5 min

Step 2, 3 and 4 carried out for 30 cycles.

4.6 PCR Amplification Of *GSTP1*

Polymerase Chain Reaction (PCR) is a very sensitive assay in which a single DNA molecule can be amplified, and single-copy genes can be extracted out of complex mixtures of genomic sequences. PCR can also be utilized for rapid screening and/or sequencing of inserts directly from aliquots of individual phage plaques or bacterial colonies.

Requirements

- 10X PCR buffer
- BSA
- Forward Primer
- Reverse Primer
- dNTP‘s
- Taq DNA polymerase
- Water
- DNA sample

The primers used are:

GSTP1 primers

Forward primer 5’-ACC CCA GGG CTC TAT GGG AA-3’

Reverse primer 5’-TGA GGG CAC AAG AAG CCC CT-3’

Band size 176bp

Table 4.5: Reaction mixture of PCR carried out for *GSTP1* (10 reactions)

Reagent	Stock Concentration	Final Reaction Concentration	Quantity Used
Additive 1 BSA	1000 µg/ml	100 µg/ml	16.5 µl
PCR Buffer (Mg Conc.)	10 X 25mM	1 X 1.5mM total	16.5µl
GSTP1 Primer (Forward)	10µM	0.5 µM	8.25 µl

GSTP1 Primer (Reverse)	10µM	0.5 µM	8.25 µl
Taq Polymerase	2.0U/µl	1.5mM	5.5 µl
dNTP	10mM each	0.2mM each	3.3 µl
PCR Grade Water			84.70 µl
DNA Template	100ng/µl	400ng	2µl

Table 4.6: Cycling profile of PCR for *GSTP1*

Steps	Temperature	Time
1. Initial Denaturation	95°C	5min
2. Denaturation	94°C	30 sec
3. Annealing	58°C	45 sec
4. Polymerization	72°C	45 sec
5. Final Extension	72°C	5 min

Step 2, 3 and 4 carried out for 30 cycles

4.7 Restriction Digetsion Of *GSTP1*

This enzymatic technique can be used for cleaving DNA molecules at specific sites, ensuring that all DNA fragments that contain a particular sequence have the same size; furthermore, each fragment that contains the desired sequence has the sequence located at exactly the same position within the fragment. The enzyme used by us was:

***BsmA1*:** The source microorganism from which it has been islolated is *Bacillus stearothermophilus* A664. *BsmAI* is an isoschizomer of *Alw26I* and *BcoDl*

Restriction site for the enzyme is

5'….GTCTC (N)⇩……3'
3'….CAGAG(N)……⇧5'

Procedure

The total reaction mixture of 20µl consisted of 2.2 µl 10X NEB 4 buffer, 0.25µl (3U) of 50U/ml *BsmA1* enzyme (NEB), and 10µl of PCR amplified product and 7.55µl water. The buffer used for the process was provided with the enzyme by NEB for increased activity of enzyme. All the samples were incubated at 37°C overnight. The enzyme reaction was stopped by keeping the samples at -20°C and the samples were loaded in 2.5% polyacrylamide gel and the results were visualised by staining them with ethidium bromide.

4.8 Statistical Analysis

Differences in the distributions of demographic characteristics between the cases and controls were evaluated using the Chi-square tests (χ^2 test) for the categorical data and student *t* test for continuous variables. The Hardy–Weinberg equilibrium theory ($p^2+2pq+q^2=1$; where p is the frequency of the wild-type allele and q is the frequency of the variant allele) was used both in cases and controls to calculate the genotype frequencies of *GSTP1* gene polymorphism using χ^2 test. Pearson's χ^2 test was used to determine whether there was any significant difference in allele and genotype frequencies between cases and controls. To assess the risk for lung cancer and GST polymorphisms adjusted Odds Ratio (ORs) along with 95% Confidence Intervals (CI) were calculated using logistic regression analysis with adjustment for possible confounders (age and pack-years of smoking as continuous variables; and gender as a nominal variable). The homozygous wild genotype AA (*Ile/Ile*) of the *GSTP1* gene was used as the reference in calculating the ORs and 95%CI. All p values were two sided, and a p value of <0.05 was considered statistically significant. All the statistical analyses were performed with Medcalc version 9.3.6.0 (Medcalc Software, Ostend, Belgium) and SPSS Version 20.0. (Chicago, IL, USA).

CHAPTER 5

RESULTS AND DISCUSSION

5.1 DNA Isolation

DNA was isolated from blood (as discussed in section 4.2) using a simple and efficient procedure and the samples were run on 0.7% gel. This total DNA from blood samples were used as template in PCR.

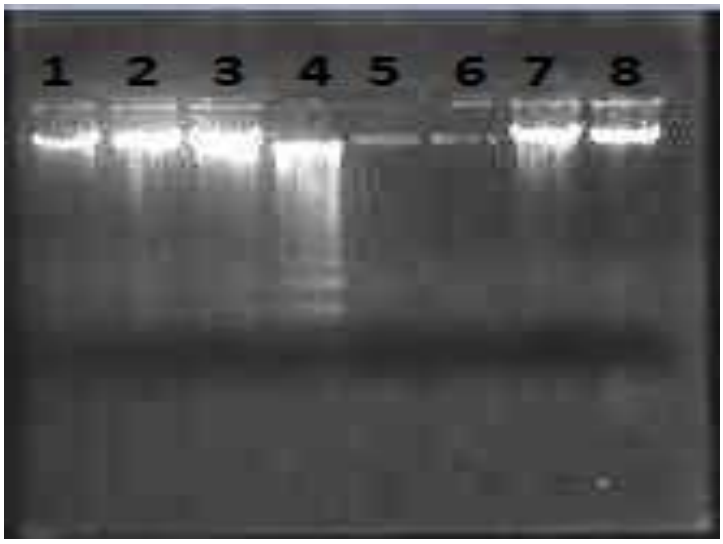


Figure 5.1: Genomic DNA isolated from peripheral blood

5.2 PCR (Multiplex) Amplification Of *GSTM1* And *GSTT1*

In order to amplify *GSTM1*, *GSTT1* and albumin gene suitable sets of primers were used. The temperature cycling parameters as employed for *GSTM1* and *GSTT1* during PCR are given in Table 4.4. The PCR amplified products were separated on 1.7% agarose gel containing ethidium bromide. The DNA bands were clearly visible and distinct which indicated that the primer combinations worked well for all the genes. Figure 5.2 shows the PCR amplified DNA products obtained using set of primer pair specific for *GSTM1*, *GSTT1* and Albumin. DNA bands of 215bp 480bp, and 312bp were obtained for *GSTT1*, *GSTM1* and Albumin respectively.

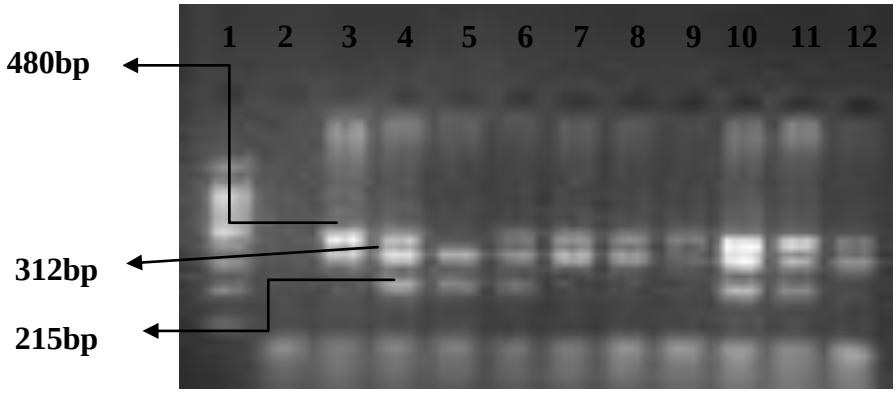


Figure 5.2: PCR amplified DNA products of *GSTM1* and *GSTT1* gene for LC samples
Lane 1: 100 bp ladder; Lane 2 Negative control; Lane 3, 7, 8 9 12: *GSTT1*; Lane 5:*GSTM1*

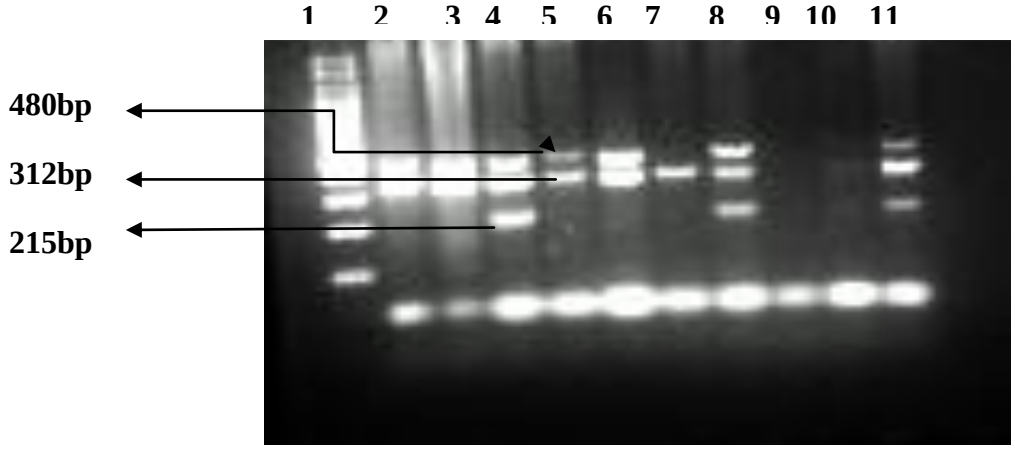


Figure 5.3: PCR amplified DNA products of *GSTM1* and *GSTT1* gene for LCC samples
Lane 1: 100 bp ladder; Lane 2, 3, 5, 6: *GSTT1*; Lane 3: *GSTM1*, *GSTT1*; Lane 7: Albumin; Lane 9: Negative control

5.3 PCR Amplification Of *GSTP1*

In order to amplify *GSTP1 gene* suitable sets of primers were used. The temperature cycling parameters for *GSTP1* as employed during PCR are given in Table 4.6. The PCR amplified products were separated on 1.7% agarose gel containing ethidium bromide. The DNA bands

were clearly visible and distinct which indicated that the primer combinations worked well for the gene. Figure 5.3 shows the PCR amplified DNA products obtained using set of primer pair specific for *GSTP1*.A clear and distinct band of 176bp was obtained.

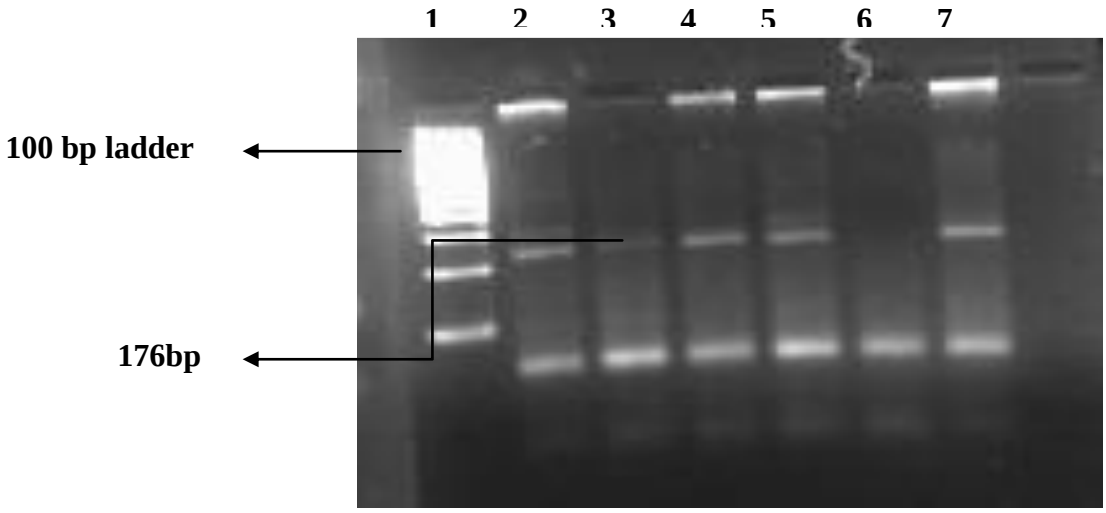


Figure 5.4: PCR amplified DNA product of *GSTP1* gene
Lane 1: Ladder; Lane 2-7 (176bp, *GSTP1* PCR product)

5.4 Restriction Digestion Of *GSTP1*

PCR product of *GSTP1* was digested with the restriction enzyme *BsmA1*. Then the digested samples were allowed to run on 2 % polyacrylamide gel containing ethidium bromide. Genotypic analysis was done as shown in Figure 5.4 and Figure 5.5. PCR product of 176bp when not digested by *BsmA1* digested was indicative of wild type genotype ($A^{Ile/Ile}A$), whereas the presence of two bands of 91bp and 85bp represented the mutant ($G^{ValVal}G$) genotype and the presence of three bands of 176-bp, 91bp and 85bp band represented the heterozygous genotype ($A^{Ile/Val}G$)

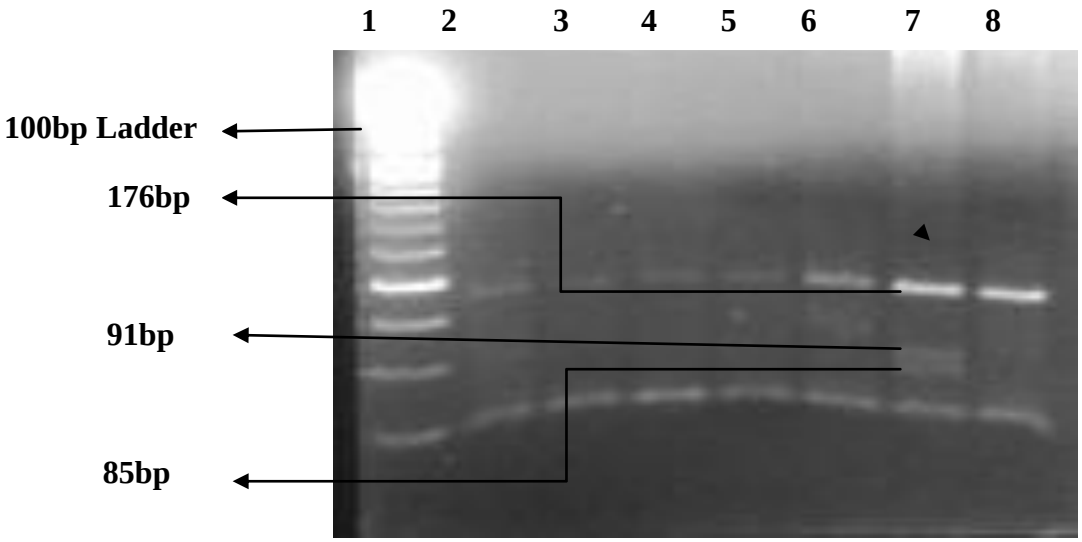


Figure 5.5: PCR RFLP genotype analysis of *GSTP1*
Lane 1 100bp Ladder; Lane 2-6, 8: Wild Type; Lane 7 Heterozygote

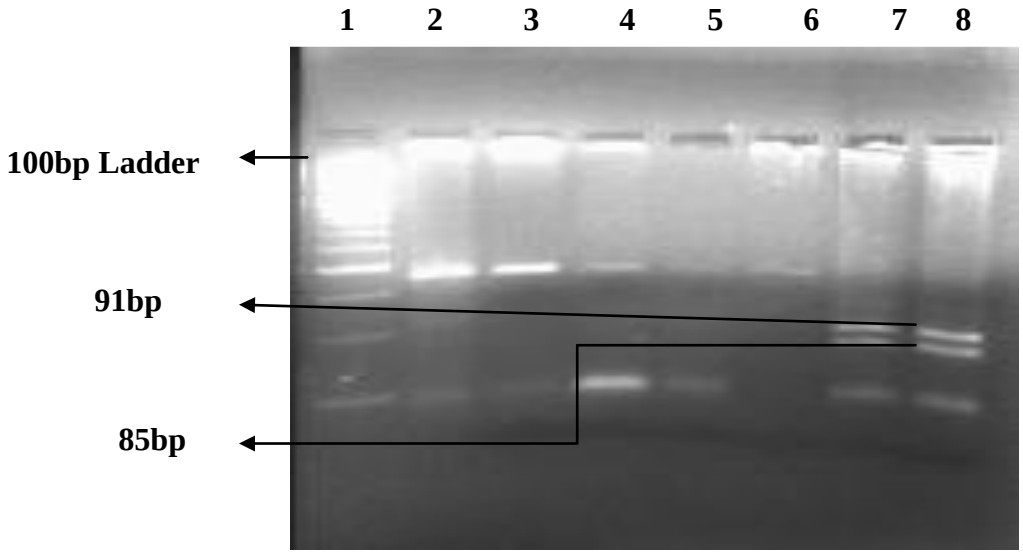


Figure 5.6: PCR RFLP genotype analysis of *GSTP1* gene
Lane 1: Ladder Lane 2-6: Wild Type Lane 7-8: Mutant

5.5 Epidemiology

Table 5.1: Distribution of demographic characteristics of LC cases and controls

Variable	**Cases, n (%) N = 270**	**Controls, n (%) N = 270**	***p* – value**
Age (years) Mean ± SD Range	 57.39 ± 10.60 (29 – 86)	 53.23 ± 10.49 (22 – 83)	 < 0.0001
Gender Male Female	 235 (87.04) 35 (12.96)	 233 (86.30) 37 (13.70)	 0.899
Smoking Status Smokers Non – Smokers	 212 (78.51) 58 (21.49)	 193 (71.48) 77 (28.51)	 0.06
Pack Years Mean ± SD	 34.079 ± 32.13	 25.42 ± 20.38	 0.001
Histologic Types SQCC ADCC SCLC	 115 (42.59) 86 (31.85) 69 (25.56)		
TNM Staging I II III IV Unclassified	 2 (0.75) 12 (4.44) 120 (44.44) 111 (41.11) 25 (9.26)		
Tumor Size T1 T2 T3 T4 Unknown	 11(4.07) 32(11.86) 66(24.44) 120(44.44) 41(15.19)		
Lymph Node Involvement N0 N1 N2 N3 N4 Unknown	 45 (16.68) 32 (11.85) 93 (34.44) 60 (22.22) 6 (2.22) 34 (12.59)		
Metastasis M0 M1 Unknown	 131 (48.52) 105 (38.89) 34(12.59)		

SD = Standard Deviation

$p < 0.05$ was considered statistically significant

p-values were derived from pearson Chi – square test except age; Student t-test was used for age. All p- values are two – sided.

Characteristics of Subjects

Demographic characteristics of the study groups including age, gender, smoking status, pack years, histological subtypes, TNM staging and other clinical parameters are shown in Table 5.1 The case-control study conducted consisted of 270 lung cancer cases and 270 controls. The mean age of cases was 57.39±10.60 (range 29-86), whereas the mean age of all control subjects were 53.23± 10.49 (range 22-83). The study comprised of 235 (87.04%) males and 35 (12.96%) of females in the case group, and 233 (86.30%) males and 37 (13.70%) females in the control group. There was no significant difference in distribution of males and females between controls and patients suggesting adequate matching (p=0.899). In the present study, 78.51% of the cases were smokers and 21.49% were non-smokers, where as in the control group 71.48% were smokers and 28.51% were non-smokers. Smoking came out to be a significant risk factor for lung cancer suggesting an association between smoking and lung cancer risk ($p < 0.06$). Furthermore, among the smokers numbers of pack years were significantly higher in cases as compared to controls (34.079 ± 32.13 vs 25.42 ± 20.38): $p < 0.001$). Variables such as age, gender, smoking exposure were further adjusted for any residual confounding effect using multivariate logistic regression analyses. Of the 270 Lung cancer cases 115 (42.59%) were squamous cell carcinoma, 86 (31.85%) were adenocarcinoma, 69 (25.56%) small cell carcinoma. TNM stage data was available in 245 (90.74 %) of 270 patients (Stage I: 2 (0.75%), II: 12(4.44%), III: 120 (44.44%), IV: 111 (41.11%). Tumor size T1 and T2 had vey less frequency of 4.07 and 11.86% whereas T3 and T4 had a frequency of 24.44 and 44.44 %. While examining lymph node involvement, N0 had a frequency of 16.68% whereas N1, N2, N3 and N4 had frequency of 11.85, 34.44, 22.22 and 2.22% respectively. Among all the cases for which TNM data was available 48.52% of the cases had M0= 48.52% and M1 had a frequency of 38.89%.

5.6 Relationship Of *GSTM1*, *GSTT1* AND *GSTP1* Polymorphism With Lung Cancer

Table 5.2a: Genotype distribution and risk of lung cancer associated with the *GSTM1* and *GSTT1* genotypes

		GSTM1				***GSTT1***			
		Positive **n (%)**	**Null** **n (%)**	**OR[1] (95% CI)**	***p***	**Positive** **n (%)**	**Null** **n (%)**	**OR[1] (95% CI)**	***P***
Control	N	167 (61.86)	103 (38.14)	1[1]		233 (86.30)	37 (13.70)	1[1]	
Cases	270	136 (50.4)	134 (49.6)	1.65 (1.16-2.3)	**0.005**	220 (81.48)	50 (18.52)	1.4 (0.87-2.2)	0.15
SQCC	115	63 (54.78)	52 (45.22)	1.34 (0.84-2.1)	0.2	102 (88.70)	13 (11.30)	0.8 (0.4-1.61)	0.5
ADCC	86	39 (45.35)	47 (54.65)	2.0 (1.21-3.3)	**0.006**	62 (72.10)	24 (27.90)	2.32 (1.2-4.25)	**0.006**
SCLC	69	34 (49.27)	35 (50.73)	1.59 (0.91-2.78)	0.1	55 (79.71)	14 (20.29)	1.46 (0.7-3.05)	0.3

[1]OR odds ratio and (95% CI) confidence interval were calculated by logistic regression and adjusted for age, gender and smoking

Table 5.2b: Genotype distribution and risk of lung cancer associated with the *GSTP1*

		GSTP1					
	N	***Ile/Ile*** **n (%)**	***Ile/Val*** **n (%)**	***Val/Val*** **n (%)**	***Ile/Val Val/Val*** **n (%)**	**OR[1] (95% CI)**	***p***
Control	270	233 (86.30)	32 (11.85)	5 (1.85)	37 (13.70)	1[1]	
Cases	270	225 (83.33)	40 (14.82)	5 (1.85)	45 (16.67)	1.30 (0.80-2.12)	0.2
SQCC	115	100 (86.96)	14 (12.17)	1 (0.87)	15 (13.04)	0.98 (0.5-1.92)	0.96
ADCC	86	74 (86.05)	11 (12.79)	1 (1.16)	12 (13.95)	1.09 (0.50-2.34	0.8
SCLC	69	51 (73.91)	15 (21.74)	3 (4.35)	18 (26.47)	2.67 (1.34-5.32)	**0.004**

[1]OR odds ratio and (95% CI) confidence interval were calculated by logistic regression and adjusted for age, gender and smoking

Table 5.2a shows the genotype distribution of *GSTM1* and *GSTT1* genotypes and adjusted ORs and 95% CIs for lung cancer and histological subtypes. The prevalence of *GSTM1* null genotype was represented in cases (49.6%) than in controls (8.14%) whereas that of the *GSTT1* null genotype in the cases and control was 18.52% and 13.70% respectively. The frequency of ADCC and SCLC in case of *GSTM1* null genotype was 54.65 and 50.73%. Similarly in case of *GSTT1* null genotype, the frequency for ADCC was 27.90% and for SCLC it was 20.28%. Table 5.2b shows the distribution of *GSTP1* genotypes and adjusted ORs and 95% CIs for lung cancer as well as histological subtypes. Due to the limited number of heterozygous and mutant alleles, we combined these genotypes for the statistical analysis throughout the study. Previous studies have reported that *GSTM1* null genotype is associated with lung cancer but *GSTT1* and *GSTP1* genotypes are not related with lung cancer risk (Houlsten *et al.*, 1988; Harris *et al*., 1988). In our studies *GSTM1* null genotype also showed an elevated risk for lung cancer (OR= 1.65, 95%CI=1.16-2.3, p=0.005) but neither *GSTT1* null genotype (OR=1.4, 95% CI=0.87-2.26, p=0.15) nor the genotypes of *GSTP1* showed increased risk for overall lung cancer (data not shown for *GSTP1*) but when analysis was done for histological subtypes, *GSTT1 and GSTM1* null genotypes showed a significant risk for developing ADCC when compared with the controls (OR=2.32, 95%CI =1.27-4.25, p=0.006) and *GSTP1* (*Ile/Val + Val/Val*) genotype showed a significant risk for SCLC (OR=2.67, 95%CI =1.34-5.32, p=0.004). *GSTM1* null genotype was also found to be statistical significant for ADCC (OR=2, 95%CI =1.21-3.3, p=0.006).

5.7: Relationship Between Genotypic Combinations Of *GSTT1*, *GSTM1* and *GSTP1* With Lung Cancer

Table 5.3: Distribution of genotype combinations of *GSTM1* and *GSTT1* and risk of occurrence of lung cancer associated with the genotypic combinations

Genotype combination		**Control n (%) N=270**	**Cases n (%) N=270**	**OR (95% CI)**	***p***
GSTM1	***GSTT1***				
Positive	Positive	145 (53.70)	109 (40.37)	1[1]	
Null	Null	15 (5.55)	109 (40.37)	2.26 (1.1-4.6)	**0.02**
Positive	Null	22 (8.15)	27 (10)	1.51 (0.79-2.87)	0.20
Null	Positive	88 (32.60)	111 (41.11)	1.70 (0.16-2.50)	**0.005**

OR odds ratio and (95% CI) confidence interval were calculated by logistic regression and adjusted for age, gender and smoking

[1]Comparison between *GSTM1* (+)/*GSTT1* (+) and the other combined genotypes of *GSTM1/GSTT1* in relation to lung cancer.

Table 5.4: Distribution of genotype combinations of *GSTM1*, *GSTP1* and risk of occurrence of lung cancer associated with the genotypic combinations

Genotype combination		**Control n (%) N=270**	**Cases n (%) N=270**	**OR (95% CI)**	***p***
GSTM1	***GSTP1***				
Positive	*Ile/Ile*	148 (54.81)	109 (40.37)	1[1]	
Null	*Ile/Val+Val/Val*	18 (6.67)	18 (6.67)	1.54 (0.74-3.21)	0.24
Null	*Ile/Ile*	85 (31.48)	116 (42.96	1.97 (1.34-2.90)	**0.0004**
Positive	*Ile/Val+Val/Val*	19 (7.04)	27 (10)	2.14 (1.1-4.19)	**0.02**

OR odds ratio and (95% CI) confidence interval were calculated by logistic regression and adjusted for age, gender and smoking

[1]Comparison between *GSTM1* (+)/*GSTP1 (Ile/Ile)* and the other combined genotypes of *GSTM1/GSTP1* in relation to lung cancer.

Table 5.5: Distribution of genotype combinations of *GSTM1 GSTP1*, *GSTT1* and risk of occurrence of lung cancer associated with these genotypic combinations

Genotype combination			**Control n (%) N=270**	**Cases n (%) N=270**	**OR (95% CI)**	***p***
GSTT1	***GSTP1***					
Positive	*Ile/Ile*		201 (74.44)	184 (68.2)	1[1]	
Null	*Ile/Val+Val/Val*		5 (1.9)	11 (4.08)	2.3 (0.76-6.92)	0.13
Positive	*Ile/Val+Val/Val*		32 (11.85)	36 (13.33)	1.18 (0.69-2.02)	0.52
Null	*Ile/Ile*		32 (11.85)	39 (14.44)	4.14 (2.01-8.51)	**0.0001**
GSTM1	***GSTT1***	***GSTP1***				
Positive	Positive	*Ile/Ile*	128 (47.41)	87 (32.22)	1[2]	
Null	Null	*Ile/Val Val/Val*	3 (1.11)	4 (1.48)	1.73 (0.35-8.4)	0.49

OR odds ratio and (95% CI) confidence interval were calculated by logistic regression and adjusted for age, gender and smoking

[1]Comparison between *GSTT1* (+)/*GSTP1* (+) and the other combined genotypes of *GSTT1/GSTP1* in relation to lung cancer.

[2]Comparison between *GSTM1* (+)/*GSTT1* (+) *GSTP1* (+) and the other combined genotypes of *GSTM1/GSTT1/GSTP1* in relation to lung cancer

Examining the frequencies of *GSTM1*/*GSTT1* genotypic combinations, it was seen that he occurrence of genotypic combinations of *GSTM1*/*GSTT1* was more in cases as compared to controls (*GSTM1*null/*GSTT1*null, *GSTM1*positive/*GSTT1*null and *GSTM1*null/*GSTT1* positive; 40.3%, 10% and 41.11% respectively and for controls 5.55%, 8.15% and 32.60% respectively). Whereas *GSTM1* positive/*GSTT1* positive genotypic combination was more in controls (53.70%) as compared to cases (40.37%). The reference for all the combinations of *GSTM1*/*GSTT1* was *GSTM1* (+)/*GSTT1* (+).The null genotypes of *GSTM1* and *GSTT1* in combination represented an increased risk towards lung cancer as shown in Table 5.3.

We had taken three different combinations of *GSTM1*/*GSTP1* genotypes along with the reference *GSTM1* (+)/*GSTP1* (*Ile/Ile*) for further studies, Table 5.4. A twofold increased risk for lung cancer was seen in combination of *GSTM1* null/*GSTT1* null (OR=2.26, CI=1.1-4.6, *p*=0.02) and further another combination *i.e. GSTM1* positive*/GSTP1 (Ile/Val+Val/Val)* showed an elevated risk for lung cancer with OR=2.14, CI=1.1-4.19 and statistically significant *p* value of 0.02.

The relevance of different combinations of *GSTT1* and *GSTP1* genotypes with lung cancer has been shown in Table 5.5.Taking *GSTT1* (+)/*GSTP1* (*Ile/Ile*) as reference for all the combinations of *GSTT1*/*GSTP1*. When all the combinations were analysed statistically after adjusting age, sex and smoking, only the combination of *GSTT1* (-)/*GSTP1*(*Ile/Ile*) was found to be significantly associated with overall lung cancer with an enhanced risk of four fold with the p value of 0.0001 which is statistically significant. The combinations of three genotypes (*GSTM1*, *GSTT1* and *GSTP1*) altogether did not show any relevance with lung cancer as is evident from the Table 5.5

5.8:Association Of Lung Cancer Related To Combination Of GST Genotypes According To Histological Subtype

Table 5.6: Genotype distribution of *GSTM1*/*GSTT1* combinations and risk of lung cancer subtypes associated with it

Genotype combination		**Control N=270**	**ADCC**	**OR (95% CI)**	***p***	**SCLC**	**OR (95% CI)**	***P***
GSTM1	***GSTT1***							
Positive	Positive	145	26	1[1]		27 (39.13)	1[1]	
Null	Null	15	11	4.09 (1.10-10.2)	**0.002**	6 (8.7)	2.87 (0.9-8.8)	**0.06**
Positive	Null	22	13	3.4 (1.47-7.81)	**0.003**	7 (10.14)	1.71 (0.63-4.62)	0.2
Null	Positive	88	36	2.4 (1.34-4.32)	**0.003**	29 (42.03)	1.67 (0.91-3.08)	**0.09**

OR odds ratio and (95% CI) confidence interval were calculated by logistic regression and adjusted for age, gender and smoking
[1]Comparison between *GSTM1* (+)/*GSTT1* (+) and the other combined genotypes of *GSTM1/GSTT1* in relation to lung cancer types

Table 5.7 Genotype distribution of *GSTM1*/*GSTP1* combinations and risk of lung cancer subtypes associated with it

Genotype combination		**Control N=270**	**ADCC**	**OR (95% CI)**	***p***	**SCLC**	**OR (95% CI)**	***P***
GSTM1	***GSTP1***							
Positive	*Ile/Ile*	148	33	1[1]		26		
Null	*Ile/Val Val/Val*	18	5	1.4 (0.46-4.25)	0.55	10	3.77 (1.43-9.94)	**0.007**
Null	*Ile/Ile*	85	42	2.4 (1.41-4.21)	**0.001**	25	1.54 (0.80-2.95)	0.10
Positive	*Ile/Val Val/Val*	19	6	1.27 (0.43-3.7)	0.65	8	3.47 (1.26-9.58)	**0.01**

OR odds ratio and (95% CI) confidence interval were calculated by logistic regression and adjusted for age, gender and smoking
[1]Comparison between *GSTM1* (+)/*GSTP1* (+) and the other combined genotypes of *GSTM1/GSTP1* in relation to lung cancer types

Table 5.8: Genotype distribution of *GSTM1*/*GSTT1*/ *GSTP1* combination and risk of lung cancer subtypes associated with it

Genotypic combinations			**Control N=270**	**ADCC**	**OR (95% CI)**	***p***
GSTM1	***GSTT1***	***GSTP1***				
Positive	Positive	*Ile/Ile*	128	22	1[1]	
Null	Null	*Ile/Ile* *Val/Val*	3	3	5.0 (0.86-29.3)	**0.07**

OR odds ratio and (95% CI) confidence interval were calculated by logistic regression and adjusted for age, gender and smoking
[1]Comparison between *GSTM1* (+)/*GSTT1* (+) *GSTP1* (*Ile/Ile*) and the other combined genotypes of *GSTM1/GSTT1* in relation to lung cancer types

When subgroups were examined in combinations according to various histological subtypes of lung cancer we found association for adenocarcinoma and SCLC. Statistical significant association was observed when *GSTM1* null and *GSTT1* null genotypes were analysed separately with the risk of developing ADCC (Table 5.6) but more pronounced effect of four times was seen when these genotypes were conjointly analysed for ADCC risk. A significant increased risk for ADCC was also evident for the genotypic combination of *GSTM1* positive and *GSTT1* null (OR=3.4, 95% CI=1.47-7.81, *p*=0.003), *GSTM1* null and *GSTT1* positive (OR=2.4, 95% CI=1.34-4.32, *p*=0.03), *GSTM1* null and *GSTP1 Ile/Ile* (OR=2.4, 95% CI=1.41-4.21, *p*=0.001), *GSTM1* null, *GSTT1* null and *GSTP1 Ile/Val+Val/Val* (OR=5.0, 95% CI=0.86-29.3, *p*=0.07). The adjusted OR for the ADCC was found to be most significant for *GSTM1* null/GSTT null genotypic combination. Although no statistically significant associations were found between the separate *GSTM1* and *GSTT1* polymorphisms and overall SCLC risk (Table 5.6), the combination of the *GSTM1* null genotype and *GSTT1* null genotype was significantly related to SCLC, risk being elevated 2.87 times as compared to the combined *GSTM1* present genotype and *GSTT1* present genotype. On further examination the other genotypic combinations that were showing increased risk for SCLC were *GSTM1* null and *GSTP1 Ile/Ile* (OR=1.67, 95% CI=0.91-3.08, *p*=0.09), *GSTM1* null and *GSTP1 Ile/Val+ Val/Val* (OR=3.77, 95% CI=1.43-9.94, *p*=0.007) as well as *GSTM1* positive and *GSTP1 Ile/Val+Val/Val* (OR=3.47, 95% CI=1.26-9.58**,** *p*=0.01).

5.9 Effects Of Smoking Status On The Association Between *GST* Gene Polymorphisms And Lung Cancer Risk

Table 5.9: Interaction of *GSTM1* genotype and tobacco smoking on the overall risk of lung cancer

Smoking	Genotype	Cases (%)	N Cases	Controls (%)	N Control	OR (95%CI)	*P*
	GSTM1						
Non smokers	Positive	24 (41.37)	58	54 (70.12)	77	1[1]	
Non smokers	Null	34 (58.62)		23 (29.87)		3.03 (1.48-6.39)	**0.002**
Smokers	Positive	112 (52.83)	212	11 (58.54)	193	1[2]	
Smokers	Null	100 (47.16)		80 (41.45)		1.27 (0.851.91)	0.2355
Light smokers PY<25	Positive	44 (45.36)	97	61 (55.45)	110	1[3]	
Light smokers PY<25	Null	53 (54.63)		49 (44.54)		1.43 (0.81-2.5)	0.2
Heavy Smokers PY>=25	Positive	68 (59.13)	115	52 (62.65)	83	1[4]	
Heavy Smokers PY>=25	Null	47 (40.87)		31 (37.34)		1.17 (0.64-2.1)	0.5

[1]ORs and 95% CIs were calculated by logistic regression analysis with the *GSTM1 (+)* genotype as the reference group for non smokers and adjusted for age and gender.
[2]ORs and 95% CIs were calculated by logistic regression analysis with the *GSTM1 (+)* genotype as the reference group for smokers and adjusted for age and gender.
[1]ORs and 95% CIs were calculated by logistic regression analysis with the *GSTM1 (+)* genotype as the reference group for Light smokers and adjusted for age and gender.
[1]ORs and 95% CIs were calculated by logistic regression analysis with the *GSTM1 (+)* genotype as the reference group for Heavy smokers and adjusted for age and gender.

Table 5.10: Interaction of *GSTT1* genotype and tobacco smoking on the overall risk of lung cancer

Smoking	Genotype	Cases (%)	N Cases	Controls (%)	N Control	OR (95%CI)	*P*
	GSTT1						
Non smokers	Positive	47 (81.03)	58	64 (83.12)	77	1[1]	
Non smokers	Null	11 (18.96)		13 (16.88)		2.18 (0.84-5.66)	0.1
Smokers	Positive	173 (81.60)	212	169 (87.56)	193	1[2]	
Smokers	Null	39 (18.39)		24 (12.44)		1.46 (0.83-2.11)	0.18
Light smokers PY<25	Positive	78 (80.41)	97	97 (88.18)	110	1[3]	
Light smokers PY<25	Null	19 (19.58)		13 (11.82)		1.65 (0.75-3.62)	0.2
Heavy Smokers PY>=25	Positive	95 (82.60)	115	72 (86.75)	83	1[4]	
Heavy Smokers PY>=25	Null	20 (17.39)		11 (13.25)		1.3 (0.56-2.98)	0.5

[1]ORs and 95% CIs were calculated by logistic regression analysis with the *GSTT1 (+)* genotype as the reference group for non smokers and adjusted for age and gender.

[2]ORs and 95% CIs were calculated by logistic regression analysis with the *GSTT1 (+)* genotype as the reference group for smokers and adjusted for age and gender.

[1]ORs and 95% CIs were calculated by logistic regression analysis with the *GSTT1 (+)* genotype as the reference group for Light smokers and adjusted for age and gender.

[1]ORs and 95% CIs were calculated by logistic regression analysis with the *GSTT1 (+)* genotype as the reference group for Heavy smokers and adjusted for age and gender.

Table 5.11: Interaction of *GSTP1* genotypes and tobacco smoking on the overall risk of lung cancer

GSTP1	**Non Smokers**		**OR (95%CI)**	***p***	**Smokers**		**OR (95%CI)**	***p***
	Cases %	Controls %			Cases %	Controls %		
Ile/Ile	46 (79.31)	64 (83.12)	1[1]		179 (84.43)	169 (87.56)	1[2]	
Ile/Val, Val/Val	12 (20.69)	13 (16.88)	1.5 (0.56-4.0)	0.4	33 (15.57)	24 (12.44)	1.23 (0.69-2.21	0.4

[1]ORs and 95% CIs were calculated by logistic regression analysis with the *GSTP1 Ile/Ile* genotype as the reference group for non smokers and adjusted for age and gender.
[2]ORs and 95% CIs were calculated by logistic regression analysis with the *GSTP1 Ile/Ile* genotype as the reference group for smokers and adjusted for age and gender.

Table 5.12: Odds ratio for risk of lung cancer and *GSTP1* genotype by smoking status, adjusted for age, sex, pack years of smoking.

GSTP1	**Light PY<25 Smokers**		**OR (95%CI)**	***P***	**Heavy PY>=25 Smokers**		**OR (95%CI)**	***p***
	Cases %	Controls %			Cases %	Controls %		
Ile/Ile	82 (84.54)	94 (85.45)	1[1]		95 (82.61)	75 (90.36)	1[1]	
Ile/Val, Val/Val	15 (15.46)	16 (14.55)	0.93 (0.4-2.0)	0.86	20 (17.39)	8 (9.64)	1.66 (0.67-4.1)	0.26

[1]ORs and 95% CIs were calculated by logistic regression analysis with the *GSTP1 Ile/Ile* genotype as the reference group for Light smokers and adjusted for age and gender.
[1]ORs and 95% CIs were calculated by logistic regression analysis with the *GSTP1 Ile/Ile* genotype as the reference group for Heavy smokers and adjusted for age and gender.

Results of assessment of interactions between GST polymorphism and smoking status are summarized in Table 5.9, 5.10, 5.11 and 5.12. In this case control study of 540 subjects, there were 58 non smokers in cases and 77 in controls. 58.62% prevalence of *GSTM1* null genotype was found in cases. It was determined that the non smokers having *GSTM1* null genotype were more prone to lung cancer (OR=3.03, 95% CI=1.48-6.39, *p*=0.002). It is evident from Table 5.2a the proportion of *GSTT1* null genotype is very less in cases. The *GSTT1* null genotype showed two times increased risk of lung cancer but this was not statistically significant. Smokers were further stratified into light and heavy smokers. Heavy smokers having packyears>=25 and light smokers having packyears<25. Both the heavy smokers and light smokers having genotype *GSTM1* null and *GSTT1* null were identified to be statistically insignificant for the risk of lung cancer. However a decreasing trend was observed in *GSTT1* null genotype where non smokers having OR=2.18, 95% CI=0.84-5.66, *p*=0.1, light smokers having OR=1.65, 95% CI= 0.75-3.62, *p*=0.2 and heavy smokers having OR=1.3, 95% CI=0.56-2.98, *p*=0.18. In case of *GSTP1*, no association was observed between smoking and lung cancer in any of the *GSTP1* genotypes.

Though tobacco smoking is considered to be the major risk factor in the etiology of lung cancer but epidemiologic studies have shown that genetic variability can play a vital role in inducing different responses to carcinogens and thus can alter the susceptibility for developing lung cancer. Phase II drug metabolising enzymes are considered to be partly accountable for the individual differences in genetic susceptibility. Polymorphism of GSTs has been widely studied for their correlation with lung cancer in various ethnic populations (Stephanie *et al.,* 2000; Kiyohara *et al.,* 2000; Yeunga *et al.,* 2004) but there have been inconsistency in the results and that may be due to the fact that the studies are being carried out in different populations with different heredity backgrounds and lifestyle or habits (e.g. smoking). In a decade lung cancer cases have increased remarkably in the North Indian population and this might be due to changed lifestyle, habits and increase in environmental pollution (air pollution). So, our study was based on finding out the correlation between the GST polymorphism and the lung cancer susceptibility in the North Indian population with greater sample size and compared the data with the results of studies performed on different populations.

Studies conducted earlier demonstrated that the null genotype of *GSTM1* and *GSTT1* had no statistical significant association with lung cancer (Sobti *et al.,* 2004) whereas we had contradictory results for *GSTM1* showing 1.65 fold increased risk for lung cancer but *GSTT1*

showed no statistical significance towards lung cancer which is consistent with the earlier studies carried out (Schneider *et al.,* 2004; Stewart *et al.,* 2003). Our findings related to *GSTM1* studies is consistent with the Meta analysis carried out by Benhamou *et al.,* 2002.The distribution of allelic frequencies of *GSTP1* in the North Indian population was contradictory to the studies carried out by Bid *et al.,* 2010. There is not much information on the role of *GSTP1* in lung cancer. It is found to be abundant in lungs (Kiyohara *et al.,* 2000) and also detoxify inhaled carcinogens. The data here reported for *GSTP1* genotypic frequencies did not show any significant association with lung cancer. It was suggested that the exon 5 polymorphism of *GSTP1* did not increase the risk of lung cancer; a Meta analysis by Xu CH *et al.,* 2014. In India, lung adenocarcinoma has been increasing over the last few decades. Our data suggests that the patients with ADCC showed an increased frequency of *GSTM1* null genotype. Moreover *GSTM1* null genotype had two fold increased risk for ADCC and this result is similar to the one reported by Sunaga *et al.,* 2002 and Stewart *et al.,* 2003.We have found a significant association of ADCC with *GSTT1* null genotype although Wang *et al.,* 2003 did not find any association; Wang *et al.,* 2003 and Sunaga *et al.,* 2002, found a significant increase in lung cancer risk with *GSTT1* null genotype. Thus our results are in line with Wang *et al.,* 2003 and Sunaga *et al.,* 2002. Our study was unable to identify a statistical significant association between *GSTM1* null, *GSTT1* null and SQCC risk. But Wang *et al.,* 2003, did show an association between *GSTT1* null genotype and SQCC in Han population in China. Also for *GSTM1* our results were not in line with the one reported by Sobti *et al.,* 2004. We found statistically significant increase in developing SCLC with *GSTP1* exon 5 polymorphism. There is another report (Stucker *et al.,* 2004) which shows remarkable risk of 3.6 fold for SCLC in relation to *GSTP1*.

Various investigators have emphasised on finding out the link between genotypic combinations and lung cancer susceptibility. Few studies have suggested that individuals with more than one defective genotype in GST genes are at higher risk of lung cancer than the one with single gene defect (el-Zein *et al.,* 1997; Saarikoski *et al.,* 1998). It may be possible to correlate lung cancer susceptibility and polymorphisms in a better way when simultaneous analysis of polymorphisms would be done (Ada *et al.,* 2010). So, we included combined investigation of *GSTM1, GSTT1* and *GSTP1*genotypes. Our results reported that the subjects who were having lacking both genes *GSTM1* and *GSTT1* were at higher risk of lung cancer which is in concordance with the studies carried out by Saarikoski *et al.,* 1998 . We also found that the combination of *GSTP1* variant with *GSTM1* null or *GSTM1* positive, and the

combination of *GSTP1* variant with *GSTT1* null genotypes had a remarkable increased risk of lung cancer overall. Although in our studies *GSTT1* alone did not show any significant association but conjointly with *GSTP1* variant had an increased risk of 4 times to lung cancer. There are similar studies which have found same results as ours like that carried out by Sorensen *et al.,* 2007 and Sreeja *et al.,* 2008 whereas Ahmet *et al.,* 2012, Sobti *et al.,* 2008 and Yang *et al.,* 2007 did not find such association. Different substrates are present in tobacco smoke and other environmental inhalations and to detoxify them different GSTs are involved. Thus we can say that analysing GST polymorphism in combination is beneficial as it is evident from our studies too.

The combinations of *GSTM1*/*GSTP1* and *GSTT1*/*GSTP1* showed increased risk for SCLC as well as adenocarcinoma. ORs for adenocarcinoma were particularly high when one of the genotype was null in *GSTM1* and *GSTT1* combination. None of the combinations revealed any association with SQCC. The same results have been obtained by Isabelle Stucker *et al*., 2002. And our results are in complete contradiction to the one obtained by Sobti *et al.,* 2008 and Saarikoski *et al.,* 1988.Though *GSTM1, GSTT1* and *GSTP1* separately did not show an increased risk to overall lung cancer but deficit genotypes of *GSTM1, GSTT1* in combination with exon 5 polymorphism (*GSTP1 Ile/Val, Val/Val*) showed an elevated risk of 5 fold to adenocarcinoma. There is less information available till date associated with triple combinations of GSTs and their relation with histological subtypes of lung cancer.

It has been noted that gene-environment interactions are also playing a major role in altering lung cancer susceptibility. The effect was seen in the subjects who were tobacco smokers and in our case prevalence of bidi smokers was seen. Bidis are more carcinogenic than cigarettes. We collected the detailed information related to smoking habits and further stratified smokers into heavy (pack years >=25) and light smokers (pack year<25). Risk associated with level of exposure to smoke and GST polymorphism is yet unclear; some reported association between higher level of exposure (Cote *et al.,* 2005) and some showed no association between smoking and lung cancer (Wang *et al.,* 2003). Our findings did not suggest any strong relationship between smoking and *GSTT1* null, *GSTM1* null genotypes. Infact *GSTP1* exon 5 polymorphism also did not have any relation between smokers and lung cancer. When the *GSTM1*null genotype was considered for non smokers we had 3 fold increased risk to lung cancer.

There is a lot of ambiguity in the studies done in relation to find out the effect of GST genes on lung cancer and as a solution we need to have ample sample size in order to find out the small effects. The major strength of our study is that we studied the three major GST enzymes in a large population reporting their frequencies separately as well as in combinations. Further, we examined the relationship between genotype combinations and lung cancer and also specifically studied the link between the genotypic combinations of all the three major GSTs and the histological subtypes of lung cancer. We have also analysed the effect of smoking on lung cancer.

CHAPTER 6

CONCLUSION

The present case/control study pertains to patients visiting the Post Graduate Institute of Medical Education and Research, which is a referral center for patients from states like Haryana, Himachal Pradesh, Punjab, Uttar Pradesh, Jammu & Kashmir and Chandigarh. The following points are evident from the present study

- *GSTM1* null genotype was significantly associated with the overall lung cancer whereas genotypes of *GSTT1* and *GSTP1* did not reveal any association when analysed separately
- However in doublet combinations of *GSTM1*, *GSTT1* and *GSTP1*, an association was seen with overall lung cancer
- The combinations of GSTs showed an elevated risk for ADCC and SCLC
- The null genotype of *GSTM1* and *GSTT1* showed an elevated risk for ADCC
- Significant association was found between non-smokers and null genotype of *GSTM1* indicating elevated risk of lung cancer among non-smokers
- There was no relationship between triplet combinations of GSTs and overall lung cancer
- However, *GSTM1* null/*GSTT1* null/*GSTP1*(heterozygous) combination had an increased risk for ADCC

In conclusion, this case-control study provides evidence that *GSTM1* polymorphism is associated with a significantly increased risk of lung cancer. Moreover, doublet and triplet combinations of GST genes showed remarkable risk for lung cancer. The combination of these variants can be considered as biomarkers for lung cancer thus our present work despite the limitation of small sample size will provide guidelines for future work.

REFERENCES

1. Markoulatos P, Siafakas N, Moncany M, (2002). Multiplex polymerase chain reaction: a practical approach. *J Clin Lab Anal,* **16**(1):47-51.

2. Androutsopoulos VP**,** Tsatsakis AM and Spandidos DA, (2009). Cytochrome P450 CYP1A1: wider roles in cancer**.** *BMC Cancer,* **9**:187**.**

3. Wormhoudt LW, Commandeur JNM, and Vermeulen NPE, (1999). Genetic Polymorphisms of Human N-Acetyltransferase, Cytochrome P450,Glutathione-S-Transferase, and Epoxide Hydrolase Enzymes: Relevance to Xenobiotic Metabolism and Toxicity. *Crit Rev Toxicol,* **29**:59–124.

4. Wander GS, Khurana SB, Gulati R, Sachar RK, Gupta RK, Khurana S, Anand, (1994). Epidemiology of coronary heart disease in a rural Punjab population prevalence and and correlation with various risk factors. *Indian Heart J,* **46***:319–*23.

5. Kenfield SA, Wei EK, Stampfer MJ, Rosner BA, Colditz GA, (2008). Comparison of Aspects of Smoking Among Four Histologic Types of Lung Cancer. *Tob Control*, **17**:198-204.

6. Parkin DM, Boyd L, Walker LC, (2011). The fraction of cancer attributable to lifestyle and environmental factors in the UK in 2010. *Br J Cancer,* **105**:S77-S81.

7. Cogliano VJ, Baan R, Straif K, *et al*., (2011). Preventable Exposures Associated With Human Cancers. *J Natl Cancer I,* **103**:1827-39.

8. Salgia R, Maulik G, Shrikhande A, Kijima T, Patrick CM, Morrison PT, (2002). Role of the hepatocyte growth factor receptor, c-Met, in oncogenesis and potential for therapeutic inhibition. *Cytokine Growth F R,* **13**:41–59.

9. Herbst RS, Bajorin DF, Bleiberg H, Blum D, Hao D, Johnson BE, Ozols RF, Demetri GD, Ganz PA, Kris MG, *et al*.,(2005).Clinical Cancer Advances 2005: major research advances in cancer treatment, prevention, and screening--a report from the American Society of Clinical Oncology. *J Clin Oncol,* **24:190**–205.

10. Liska DJ, (1988). The detoxification enzyme systems. *Altern Med Rev,* **3**:187-98.

11. Jancovaa P, Anzenbacherb P, Anzenbacherova E, (2010). Phase II Drug metabolizing enzymes. *Biomed Pap Med Fac Univ Palacky Olomouc Czech Repub.* **154**:103–116.

12. Mannervik B, Board PG, Hayes JD, Listowsky I, Pearson WR, (2005). Nomenclature for mammalian soluble glutathione transferases. *Methods Enzymol,* **401**:1-8.

13. Yokota J, Shiraishi K, and Kohno T, (2010).Genetic Basis for Susceptibility to Lung Cancer: Recent Progress and Future Directions. *Cancer Res,* **109:51**-72.

14. Yang P , Yokomizo A , Tazelaar HD, Marks RD, Lesnick TG, Miller DL, Sloan JA, Edell ES, Meyer RL , Jett J, Liu W, (2002). Genetic determinants of lung cancer short-term survival: the role of glutathione-related genes. *Lung Cancer J-IASLC,* **35**: 221–229.

15. Sharma A, Pandey A, Sharma S, Chatterjee I, Mehrotra R, Sehgal A, Sharma JK, (2014). Genetic polymorphism of glutathione S-transferase P1 (*GSTP1*) in Delhi population and comparison with other global populations. *Meta Gene,* **2:**134–142.

16. Zimniak P, Nanduri B, Pikula S, *et.al,,* (1994). Naturally occurring human glutathione S-transferase GST-PI-1isoforms with isoleucine and valine in position 105 differ in enzymatic properties. *Eur J Biochem,* **15:893**–9.

17. Cotton SC, Sharp L, Little J, Brockton N, (2000). Glutathione S-transferase polymorphisms and colorectal cancer: A Huge review. Am. *J. Epidemiol,* **151**:7–32.

18. Hayes J, Flanagan J, Jowsey I, Hayes J, Flanagan J, Jowsey I, (2005).Glutathione transferases. Annu. Rev. Pharmacol. *Toxicol,* **45**: 51–88.

19. Sharma A, Pandey A, Sardana S, Sehgal S, Sharma JK, (2012). Genetic polymorphisms of *GSTM1* and *GSTT1* genes in Delhi and comparison with other Indian and Global populations. *Asian Pac. J. Cancer Prev*, **13**:5647–5652.

20. Yang X, Qiu M-T, Hu J-W, Wang X-x, Jiang F, *et al.,* (2013) *GSTT1* Null Genotype Contributes to Lung Cancer Risk in Asian Populations: A Meta-Analysis of 23 Studies. *PLoS ONE* **8**:e62181.

21. Houlsten RS, (1998).Glutathione S-transferase M1 status and lungcancer risk: a meta-analysis, Cancer Epidemiol. *Biomarkers Prev*, **8**:675–682.

22. Harris MJ, Coggan M, Langton L, Wilson SR, Board PG, (1998). Polymorphism of the Pi class glutathione S-transferase in normal and cancer patients. *Pharmacogenetics* **8** :7–31.

23. Stephanie J, Yuan JM, Chung FL, Gao Y, Coetzee GA, Ross RK, Mimi CY, (2000). Isothiocyanates, glutathione S-transferase M1 and T1 polymorphisms, and lung-cancer risk: a prospective study of men in Shanghai, China. *Clin Nutr*, **356**: 724–729.

24. Kiyohara C, Yamamura K, Nakanishi Y, Takayama K, Hara N, (2000). Polymorphism in *GSTM1*, *GSTT1*, and *GSTP1* and Susceptibility to lung cancer in a Japanese Population. *Asian Pacific J Cancer*, **1**:293-298.

25. Chan-Yeunga M, Tan-Un KC, Man Ip MS ,Tsang KWT, Ho SP, Man Ho JC, Chan H, Lam WK, (2004). Lung cancer susceptibility and polymorphisms of glutathione-S-transferase genes in Hong Kong. *Lung Cancer J-IASLC*, **45:155**—160.

26. Sobti RC, Sharma S, Joshi A, Jindal SK and Janmeja A, (2004).Genetic polymorphism of the *CYP1A1*, *CYP2E1*, *GSTM1* and *GSTT1* genes and lung cancer susceptibility in a north Indian population. *Mol Cell Biochem*, **266**:1–9.

27. Schneider J, Bernges1 U, Philipp M, Woitowitz H, (2004). *GSTM1*, *GSTT1*, and *GSTP1* polymorphism and lung cancer risk in relation to tobacco smoking. *Cancer Lett*, **208**:65–74.

28. Nazar-Stewart V, Vaughan T, Stapleton P, Van Loo J, Blades B, Eaton DL, (2003). A population based study of glutathione S-transferase M1, T1 and P1 genotypes and risk for lung cancer. *Lung Cancer J-IASLC*, **40**:247-258.

29. Benhamou S *et al.*, (2002). Meta and pooled analysis of the effects of glutathione S-transferase M1 polymorphisms and smoking on lung cancer. *Carcinogenesis*, **8**:1343-1350.

30. Bid HK, Konwar R, Saxena M, Chaudhari P, Agarwal CG, Banerjee M, (2010). Association of glutathione S-transferase (*GSTM1,*T1, P1) gene polymorphisms with type 2 diabetes mellitus in north Indian population. *J Postgrad Med,* **56**:176–181.

31. Xu CH, Wang Q, Zhan P, Qian Q, Yu LK, (2014). *GSTP1* Ile105Val polymorphism is associated with lung cancer risk among Asian population and smokers: An updated meta-analysis. *Mol Bio Rep*, **41**:4199-212.

32. Wang Y Yang H, Li L, Wang H, (2010). Glutathione S-transferase T1 gene deletion polymorphism and lung cancer risk in Chinese population: A meta-analysis. *Cancer Epidemiol,* **34**:593-597.

33. Sobti RC, Kaur P, Kaur S, Janmeja AK, Jindal SK, Kishan J, Raimondi S., (2008). Combined effect of *GSTM1*, *GSTT1* and *GSTP1* polymorphisms on histological subtypes of lung cancer. *Biomarkers*, **13**:282-95

34. Wang J, Deng Y, Cheng J, Ding J, Tokudome S, (2003).GST genetic polymorphisms and lung adenocarcinoma susceptibility in a Chinese population. *Cancer Lett,* **201**:185-193.

35. Sunaga N, Kohno T, Yanagitani N, Sugimura H, Kunitoh H, Tamura T, Takei Y, Tsuchiya S, Saito R, Yokota J, (2002). Contribution of the NQO1 and *GSTT1* polymorphisms to lung adenocarcinoma susceptibility. *Cancer Epidemio Biomarkers Prev*, **11**:730-8.

36. Stucker I, Hirvonen A, deWaziers I, Cabelguenne A, Mitrunen K, Cenee S, *et al.,* (2002). Genetic polymorphisms of glutathione S-transferases as modulators of lung cancer susceptibility, *Carcinogenesis,* **23**:1475–1481.

37. el-Zein R, Zwischenberger JB, Wood TG, Abdel-Rahman SZ, Brekelbaum C, Au WW, (1997) Combined genetic polymorphism and risk for development of lung cancer. *Mutat Res*, **381**:189–200.

38. Williamson G, DuPont MS, Wanigatunga S, Heaney RK, Musk SRR, Fenwick GR, Rhodes MJC, (1997). Induction of glutathione S'-transferase activity in hepG2 cells by extracts from fruits and vegetables. *Food Chem*, **60**:157–160.

39. Ada AO, Kunak SC, Hancer F, Bilgen S, Suzen SH, Alpar S, Gulhan M, Kurt B, Iscan M, (2010). CYP and GST polymorphisms and survival in advanced non-small cell lung cancer patients. *Neoplasma,* **57**:512–521.

40. Sorensen M, Raaschou-Nielsen O, Brasch-Andersen C, Tjønneland A, Overvad K, Autrup H, (2007) Interactions between *GSTM1*, *GSTT1* and *GSTP1* polymorphisms and smoking and intake of fruit and vegetables in relation to lung cancer. *Lung Cancer J-IASLC,* **55**:137–144.

41. Yang M, Choi Y, Hwangbo B, Lee JS, (2007) Combined effects of genetic polymorphisms in six selected genes on lung cancer susceptibility. *Lung Cancer J-IASLC***, 57**:135–142.

42. Ahmet O, Ada, Semih C. Kunak, Figen Hancer, Emre Soydas, Sibel Alpar, Meral Gulhan, Mumtaz Iscan, (2012). Association between *GSTM1*, *GSTT1*, and *GSTP1* polymorphisms and lung cancer risk in a Turkish population. *Mol Biol Rep*, **39**:5985–5993.

43. Sreeja L, Syamala V *et al.*, (2008). *GSTM1*, T1 and P1 polymorphisms susceptibility and outcome in lung cancer patients. *J Exp Ther Oncol*, **7**:73-85.

44. Saarikoski ST, Voho A, Reinikainen M, Anttila S, Karjalainen A, Malaveille C, Vainio H, Husgafvel-Pursiainen K, Hirvonen A, (1988). Combined effect of polymorphic GST genes on individual susceptibility to lung cancer. *Int J Cancer,* **77**:516-21.

45. Cote ML, Kardia SLR, Wenzlaff AS, Land SJ and Schwartz AG, (2005). Combinations of glutathione S-transferase genotypes and risk of early-onset lung cancer in Caucasians and African Americans: a population study. *Carcinogenesis,* **26:**811-819.

46. Raimondi S *et al.,* (2006).Meta- and Pooled Analysis of *GSTT1* and Lung Cancer: a HuGE-GSEC Review. *Am J Epidemiol.* **164:**1027-42.

47. Cote ML *et al.,* (2008). Meta- and Pooled Analysis of *GSTP1* Polymorphism and Lung Cancer: A HuGE Review. *Am J Epidemiol,* **169:**802-814.

48. http://www.cancerresearchuk.org/cancerinfo/cancerstats/types/lung/riskfactors/lung-cancer-risk-factors.

49. Gong M, Dong W, Shi Z, Xu Y, Ni W, *et al.,* (2012). Genetic polymorphisms of *GSTM1*, *GSTT1* and *GSTP1* with prostate cancer risk: A Meta- analysis of 57 studies. *PLoS ONE,* **7**:e50587.

50. J. Alberg AJ, Malcolm V. Brock, Samet JM, (2010). Epidemiology of lung cancer: looking to the future. *J Clin Oncol,* **23**:3175-3185.

APPENDIX

1. 0.5M EDTA: Dissolved 9.306g of disodium salt of EDTA in 20ml of deionised water, and then adjusted the pH to 8.0 by 1 M sodium hydroxide. Sterilized the solution by autoclaving.
2. 10% SDS: Dissolved 1g of SDS in 10ml of deionised water.
3. 100mM Tris-Cl (pH 8.0): Dissolved 0.32g of Tris-Cl in 10 ml of deionised water, then adjusted the pH to 8.0 by 1M sodium hydroxide. Sterilized the solution by autoclaving.
4. 10mg/ml Proteinase K: Dissolved 10mg Proteinase K in 1ml of double distilled water. Sterilized the solution by autoclaving.
5. 1mg/ml BSA: Dissolved 100mg of BSA in 100ml of deionised sterile water and kept at 4 C overnight.
6. 5M Sodium chloride (NaCl): Dissolved 5.85g of sodium chloride in 20ml of deionised water. Sterilized the solution by autoclaving.
7. 5X TBE buffer: Dissolved 54g of Tris base and 27.5g of boric acid in 980ml of double distilled water and then added 20ml of 0.5 EDTA. Sterilized the solution by autoclaving.
8. Ethidium Bromide (10mg/ml): Dissolved 1g of ethidium bromide in 100ml of water. Mixed the solution properly.
9. Magnesium chloride (MgCl2) (100mM): Dissolved 0.41gms of MgCl2 in 20ml of deionised water and sterilized by autoclaving.
10. Sucrose (1M): Dissolved 3.41 g of sucrose in 10 ml of deionised water and sterilized by autoclaving.
11. TE buffer (pH 8.0): Added 1ml of 100mM Tris-Cl (pH 8.0) and 200 µl of 0.5M EDTA solution to 8.8 ml of deioinsed water. Sterilized the solution by autoclaving.
12. Triton X- 100 (10%): Took 100 µl of TritonX-100 and mixed with 900 µl of deionised water and mixed properly.