CONTRIBUTION OF DRUG METABOLIZING ENZYMES IN GENE-GENE AND GENE-ENVIRONMENT INTERACTIONS IN LUNG CARCINOGENESIS

by

Sutira Lerdtragool

B.S., Chiang Mai University, 1991

M.S., Mahidol University, 1994

Submitted to the Graduate Faculty of

the Graduate School of Public Health in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2002

University of Pittsburgh

Graduate School of Public Health

This dissertation was presented

by

Sutira Lerdtragool

It was defended on

September 16, 2002

and approved by

Billy W. Day, Ph.D., Associate Professor, Department of Pharmaceutical Sciences

Jill M. Siegfried, Ph.D., Professor, Department of Pharmacology

John W. Wilson, Ph.D., Assistant Professor, Department of Biostatistics

Phouthone Keohavong, Ph.D., Associate Professor, Department of Environmental and Occupational Health

Dissertation Director: Marjorie Romkes, Ph.D., Associate Professor, Center for Clinical Pharmacology

ACKNOWLEDGEMENTS

I would like to thank my great major advisor, Dr. Romkes, who inspired me to do this research since I took her class in the first year of my Ph.D. study. I am very grateful for her patience, help, support and advice during this learning process. Her contribution has been invaluable.

I am also grateful to Dr. Siegfried who helped me to acquire all the samples from Pittsburgh area and also for her comments and suggestions on my research. I appreciate Dr. Keohavong, who sometimes spoke Thai to me, for his kindness, advice and guidance not only on my research but also on other scientific projects. I also would like to acknowledge my appreciation of, Dr. Day, who always knew the answer to questions in the seminar or journal club that I attended during my Ph.D. training. He exemplified to me of an example of a competent scientist. Without his advice and input, I would not have seen some of the limitations of my research. I would like to thank Dr. Wilson for his helpful suggestions for the statistical analyses.

I also would like to thank Dr. Yang and other Mayo Clinic staff members for helping me to obtain the lung cancer samples and demographic data from the Mayo Clinic.

I am so grateful for the Thai government giving me an opportunity to pursue my Ph.D. degree. Also, I would like to acknowledge the Lung SPORE (NIH IP60 DE 13059-01) for the funding support in the final year of my study.

I would like to thank friends, colleagues and others, who have helped me in my work and personal life, for their contribution to my success. I am so grateful for my family and Marut Buranarach for their endless support.

Without them, I would not have made this achievement.

CONTRIBUTION OF DRUG METABOLIZING ENZYMES IN GENE-GENE AND GENE-ENVIRONMENT INTERACTIONS IN LUNG CARCINOGENESIS

Sutira Lerdtragool, Ph.D.

University of Pittsburgh, 2002

A panel of metabolic enzyme genetic polymorphisms, which are involved in cigarette carcinogen metabolism, also a DNA repair gene involved in the nucleotide excision repair pathway were evaluated for associations with lung cancer risk in 203 lung cancer cases and 205 controls and in a case-only analysis of 177 lung cancer patients. Significant relationships between predicted high CYP1A2 activity, *CYP1B1* (**1/*3* or **3/*3*), *GSTM3*A/*A*, and *XPD* (Lys/Gln or Gln/Gln) genotypes and lung cancer risk were observed (adjusted ORs, 2.05; 95%CI, 1.13-3.73, 2.6; 95%CI, 1.19-5.69, 1.84; 95%CI, 1.03-3.31, 2.56; 95%CI, 1.45-4.51, respectively). The predicted mEPHX intermediate or high activity genotype also increased risk approximately 3-fold among females. The combined effects of carrying multiple genetic polymorphisms (gene-gene interaction) or of gene-environment interactions, for example, between the *CYP1B1* (**1/*3* or **3/*3*) genotype and packyear also resulted in significant levels of increased risk and early age onset of lung cancer.

CYP mRNA expression levels were measured in 20 pairs of lung tumor and histologically normal tissues to evaluate the potential for local metabolic activation and as a smoking exposure biomarker. The detection of CYP1B1 and CYP2E1 mRNA expression in lung tissue suggests that local bioactivation of procarcinogens may occur. A relationship between lung CYP2E1 mRNA expression levels and cigarette smoke and/or other environmental carcinogens exposure was observed. The significant increase in the levels of CYP2E1 mRNA expression in lung

tumors compared to their corresponding histological normal adjacent tissues among current and former smokers and nonsmokers who were exposed to petroleum and/or other environmental exposures further suggest a mechanistic link between environmental carcinogens exposures and lung cancer development. The results indicate that individual susceptibility to lung cancer determined by endogenous host factors such as genetic polymorphisms in metabolic and DNA repair genes, family history of lung and other cancers, early age onset, and interindividual differences in capacity of local procarcinogens bioactivation could interact with each other and/or interact with individual exposures or other exogenous factors such as cigarette smoke, environmental carcinogens and occupational exposures in modifying lung cancer risk.

TABLE OF CONTENTS

TABLE OF TABLES

TABLE OF FIGURES

LIST OF ABBREVIATIONS

DNA	2’-deoxy-5’-ribonucleic acid
CYP	Cytochrome P450
OR	Odds ratio
CI	Confidence interval
mEPHX	Microsomal epoxide hydrolase
B[*a*]P	Benzo[*a*]pyrene
NNK	4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone
NNN	N’-nitrosonornicotine
4-ABP	4-aminobiphenyl
PM	Poor metabolizer
EM	Extensive metabolizer
GST	Glutathione S-transferase
NAT	N-acetyltransferase
SNPs	Single nucleotide polymorphisms
MPO	Myeloperoxidase
PMNs	Polymorphonuclear granulocytes
A-T	Ataxia-telangiectasia
XP	Xeroderma pigmentosum
NER	Nucleotide excision repair

UV	Ultraviolet
AHH	Aromatic hydrocarbon hydroxylase
DMBA	Dimethylbenz[*a*]anthracene
DB[*a*,*l*]P	Dibenzo[*a*,*l*]pyrene
UM	Ultra-rapid metabolizer
IM	Intermediate metabolizer
GSH	Glutathione
BPDE	Benzo[*a*]pyrene diolepoxide
CDNB	1-chloro-2,4-dinitrobenzene
Ile	Isoleucine
ATP	Adenosine triphosphate
TFIIH	Transcription factor IIH
NNAL	4-(methylnitrosamino)-1-(3-pyridyl)-1-butanol
PhIP	2-amino-1-methyl-6-phenylimidazo[4,5-*b*]pyridine
PAHs	Polycyclic aromatic hydrocarbons
PCR	Polymerase chain reaction
RFLP	Restriction fragment length polymorphism
AhR	Aryl hydrocarbon receptor
Hsp90	Heat-shock protein 90
XRE	Xenobiotic-response elements
BEC	Bronchial epithelial cells
RNA	Ribonucleic acid
mRNA	Messenger ribonucleic acid

RT	Reverse transcription
ELISA	Enzyme-linked immunosorbent assay
FAF-ELOSA	Fluorescein-antifluorescein-based enzyme-linked oligonucleotide sorbent assay
TMB	Trimethylbenzidine
FAM	6-carboxyfluorescein
TAMRA	6-carboxytetramethylrhodamine
FRET	Forster resonance energy transfer
CCD	Charged-coupled device
C_T	Threshold cycle
GPAS	General patient access system
PFH	Patient and family history
IRB	Institutional review board
FH+	Positive family history
BUN	Blood urea nitrogen
pH	Logarithmic measure of hydrogen ion concentration
RBC	Red blood cell
xg	Times g-force
dH_2O	Deionized water
DMSO	Dimethylsulfoxide
dNTP	Mixed solution of 2'-deoxyadenosine 5'-triphosphate, 2'-deoxycytidine 5'-triphosphate, 2'-deoxyguanosine 5'-triphosphate, and 2'-deoxythymidine 5'-triphosphate

DEPC	Diethyl pyrocarbonate
U	Unit
SSC	Standard sodium citrate
BSA	Bovine serum albumin
HRP	Horseradish peroxidase
PBS	Phosphate buffered saline
FCS	Fetal calf serum
MMLV	Moloney Murine Leukemia Virus
GUS	Glucuronidase
PY	Packyear
SD	Standard deviation
N	Number
M	Male
F	Female
P	P-value
CV	Coefficient of variance
NADPH	Nicotinamide adenine dinucleotide phosphate

INTRODUCTION

1.1. Lung cancer

Lung cancer has become a leading cause of cancer deaths for both men and women. In the US, the incidence of lung cancer is increasing and is second only to prostate cancer in men and breast and colorectal cancer in women (1). Lung cancer has become a major public heath problem. Due to relatively poor prognoses associated with current treatment regimens, prevention is clearly optimal for reduction of lung cancer mortality. Significant effort continues towards the development of effective methods for cancer risk assessment, early detection and new treatment strategies. A thorough understanding of the mechanisms of lung carcinogenesis is also necessary to achieve these goals. Lung cancer risk has a multifactorial basis, including both genetic and environmental components. Tobacco cigarette smoking continues to be the major cause of lung cancer. The risk of lung cancer from passive exposure and other carcinogenic exposures such as asbestos, radon, arsenic, chromium, nickel and carcinogens in diet have been noted (2-10). There is also evidence of genetic predisposition for lung cancer, including data showing increased lung cancer incidence in relatives of lung cancer patients (11-15).

Genetic host factors can interact with environmental carcinogens to place individuals at increased risk of lung cancer. The gene-gene and gene-environment interactions paradigms have been used to improve models of cancer risk assessment. Further characterization of gene-gene and gene-environment interactions would be very useful in public health intervention strategies for lung cancer. For example, an understanding of these interactions will aid in the identification

of individuals who would most benefit from modification of hazardous lifestyles, targeted chemoprevention protocols and reduction of involuntary exposure to carcinogens by regulation. Although many epidemiological studies have identified causes of lung cancer that have provided data for preventive strategies in general populations, individual determinants of susceptibility, for example susceptibility to cigarette smoking and family history of lung cancer, have not been adequately described. Ninety percent of lung cases in the US are attributed to cigarette smoking but only 10% of all smokers develop lung cancer. This observation implies that host factors and other environmental agents also influence individual susceptibility to tobacco smoke. The variability in host susceptibility to lung cancer can in part be explained by differences in metabolism, DNA repair, genetic instability and altered oncogene or tumor suppressor gene expression. Conventional epidemiological methods may not be sufficiently sensitive for the detection of genetic predisposition. Therefore, molecular epidemiological approaches are being explored by a number of investigators. The molecular epidemiology approach seeks to identify human cancer risk based on individual exposures and susceptibilities to cancer. The overall goal is to develop, apply and validate biomarkers of human risk in order to enhance cancer risk assessment. The combination of information from carcinogenicity tests in laboratory animals with classical and molecular epidemiology can help to identify human cancer risk, elucidate mechanisms of carcinogenesis and lead to cancer prevention strategies.

1.1.1. Epidemiology of lung cancer

Lung cancer epidemiological studies have facilitated the discovery of causes of this disease. Unfortunately, epidemiological results may not be sufficient for the elucidation of the mechanism(s) by which lung cancer is produced. Nevertheless, epidemiology has the advantage that it focuses on the disease and the deaths that actually occur from that disease.

Epidemiological data has provided sufficient evidence that lung cancer is the most common fatal cancer throughout the world. Case-control and cohort studies in the 1950s and 1960s demonstrated that cigarette smoking was the single greatest risk factor for lung cancer (16-19). There is extensive literature regarding specific agents that are associated causally with the epidemic rise of lung cancer (20-32). Since 1980, considerable progress has been made in understanding the multifactor etiology of lung cancer (33-39). It has been shown that the dynamics of the mortality rate change over time. Lung cancer mortality rates have differed markedly between the sexes. The age-adjusted lung cancer mortality rate in the US increased from 4 to 74 cases per 10,000 in men and from 4 to 27 per 10,000 in women between 1930 and 1987 (40). During the early 1990s, lung cancer incidence and mortality rates in men have begun to level off. The increase in lung cancer in men has ceased at younger ages and is slowing at older ages, in agreement with the differences between rates of decline of smoking prevalence between sexes in which 37% vs. 18% of men to women have stopped smoking since the mid 1960s. For women, lung cancer incidence has increased dramatically since the beginning of the 1990s. Also among women, lung cancer death rates have climbed faster than for any other cancer (41,42). Furthermore, women are starting to smoke at a younger age (43). Information on demographic factors, for example, socioeconomic status, and education could help in understanding the differences in rate of incidence and mortality of lung cancer between men and women.

A strong inverse association of lung cancer incidence with income and education is apparent among male. Among females, however, positive associations between income, education and lung cancer rates have been observed (44). In the US, the occurrence of lung cancer also varies among racial and ethnic groups. Mortality rates have been higher in African American men than

in Caucasian men, whereas the rates have been comparable among African American and Caucasian women. Lung cancer mortality has been lower in Hispanic men than in Caucasian men (41). Before 1970, regional differences in mortality rates for lung cancer were observed only among men. In the late 1980s, the geographic patterns also clearly emerged among women (45). The rates of adenocarcinoma and oat cell carcinoma increased in Caucasian men from 1969 through 1988, while the rate of squamous cell carcinoma decreased. In contrast, these rates have increased for all major types in Caucasian women (40). Adenocarcinoma is the predominant histological cell type among women whereas squamous cell carcinoma predominates among men. The proportion of adenocarcinoma is greater in non-smokers than in smokers, and in particular in women (40).

1.1.2. Risk factors for lung cancer

I. Smoking

In the past, tobacco was consumed largely in the form of snuff, chewing tobacco, pipes, and cigars. More recently, tobacco has been consumed primarily as cigarettes. The explosive growth of cigarette smoking is because of the development of cigarette manufacturing equipment, the invention of safety matches, and the introduction of mass marketing techniques. In addition, changes in the characteristics of the tobacco in cigarettes, which deliver a more pleasant sensation to smokers, has led to the dramatic rise in cigarette consumption (46). However, the use of cigarettes is significantly more hazardous than pipes or cigars. Pipe or cigar smokers are less likely to inhale when they smoke than cigarette smokers because the smoke produced by pipes and cigars is harsher and more alkaline than that produced by cigarettes (47). Also, the lower pH of cigarette smoke decreases the absorption of nicotine, the addictive agent in tobacco

smoke, across the oral mucosa. The inhalation of smoke into the lung is required to facilitate the absorption (48). As more cigarette smoke is inhaled, more toxic and carcinogenic compounds in the cigarette smoke are deposited in the airway and alveoli, and absorbed into the lung and the body.

Cigarette smoke contains several constituents in both the particulate and gas phases of mainsteam smoke. Tar is the total particulate substance of the smoke after nicotine and water have been removed. Compounds remaining in the gas phase of smoke such as carbon dioxide, carbon monoxide, nitrogen oxides, acetone, acrolein, acetonitrile, pyridine, *N*-nitrosodimethylamine and *N*-nitrosoethylmethylamine. Table 1 shows some of carcinogenic agents in cigarette smoke that are contained in the particulate phase (46).

Table 1. Carcinogenic agents in cigarette smoke

Compounds	Amount per cigarette
Polycyclic aromatic hydrocarbon (PAH)	
Benzo[*a*]pyrene (B[*a*]P)	20-40 ng
Aza-arenes	
Quinoline	1- 2 μg
N-nitrosamines	
4-(methylnitrosamino)-1-(3-pyridyl)-1- butanone (NNK)	0.08-0.77 μg
N′-nitrosonornicotine (NNN)	0.12-3.7 μg
Aromatic amines	
4-aminobiphenyl (4-ABP)	2-5 ng
Aldehydes	
Acetaldehyde	18-1,400 mg

Source: reference (46)

I.I. Lung cancer and tobacco smoking

Studies in the 1950s suggested an association between smoking and lung cancer, which have been further supported by evidence from animal carcinogenicity studies of tobacco smoke. The accumulating evidence during the 1960s to the 1980s firmly established this causal relationship and also showed that differences in various aspects of smoking behavior affect lung cancer risk (49). The relative risk increases with duration of smoking and earlier age of starting to smoke. Dose-response relationships between the amount smoked and the relative risk of lung cancer have been consistently found in epidemiological studies. A dose-response relationship is evidence of a causal association of smoking with lung cancer (40). In the 1950s, manufacturers introduced the lower-tar filter cigarettes into the market. The tar yield has dropped from approximately 35 mg in 1950s to approximate 13 mg currently (41). Studies during the 1960s and 1970s addressed the consequences of switching from nonfiltered, high-tar cigarettes to the lower-tar filtered cigarettes. These studies reported that filter smokers compared with nonfilter smokers showed a lower risk of lung cancer (50-53). There was 20-30% reduction in the odds ratio for lung cancer associated with smoking reduced tar yield cigarettes (54). However, reduction of risk from smoking filter or lower-tar cigarettes may be insignificant if smokers compensate for the lower tar intake per cigarette by increasing the number of cigarettes smoked per day or by changing the manner of smoking, e.g., puff frequency or duration (49).

Inhalation is the major route by which lung tissue is exposed to carcinogens found in tobacco smoke. Depth of inhalation is considered to be a factor influencing the level of risk in individual smokers. There are studies that have shown an association between smokers who reported inhaling deeply or frequently and higher risk of lung cancer (49). The more recent studies carried out with larger sample sizes documented a thorough investigation of risk of different histologic

types of lung cancer that may be influenced by different parameters of tobacco smoke exposure. These studies showed higher relative risks for squamous and small or oat cell cancers than for adenocarcinomas, especially for different parameters of tobacco smoke exposure including numbers of cigarettes smoked, years of smoking, and amount and duration of tobacco smoke among men and women (54-56). A decline in relative risk with cessation of smoking or with smoking filter compared with nonfilter cigarettes was more apparent for squamous and small or oat cell cancers than for adenocarcinoma of the lung. However, when frequency and depth of inhalation were considered, the relative risk for adenocarcinoma of the lung equaled or exceeded the relative risk for squamous cell cancers (56). Although it is certain that cigarette smoking increases the risk of the major histologic types of lung cancer, interpretation of these findings in a biologically based framework is presently difficult (49).

The different levels of cigarette smoking can help to explain the differences in rates of lung cancer among several ethnic groups. In the US, the age-adjusted lung cancer incidence rates are highest in African Americans, intermediate in Caucasians, and lowest in Hispanics (49). The incidence rates of lung cancer in African-Americans are 47% higher than in Caucasians and may be explained by differences in smoking habits. African American male smokers smoke cigarettes with greater tar and nicotine yields than Caucasians. Also, the former prefer to smoke menthol cigarettes, which may promote deeper inhalation of cigarette smoke (57). The rates of lung cancer in Hispanics are 58% lower than in Caucasians. Hispanic men tend to be lighter smokers as they smoke about half the number of cigarettes per day compared to Caucasian men. Also, they start to smoke at a later age (49). The high rates of lung cancer among Chinese women have received significant attention recently, as the prevalence of smoking among Chinese women is low. Results from several studies showed that Chinese female smokers have a two- to four-fold

increased risk of lung cancer compared to never smokers. They smoke fewer cigarettes and start smoking at a later age. Chinese females have lower attributable risk associated with smoking compared with Caucasian females due to lower percent of Chinese females who have ever smoked. In addition to smoking, risk factors such as personal and family history of lung diseases and exposure to various sources of indoor air pollution, including exposure to oil volatiles from cooking and coal-burning devices, may in part, explain the high rates of lung cancer among Chinese women (49).

Studies evaluating smoking cessation and lung cancer incidence show a 20-90% reduction in risk among former smokers compared with current smokers. The relative risk of lung cancer among ex-smokers decreases sharply with increasing years of abstinence and reaches a plateau when adjusted for duration of smoking. In addition, the magnitude of risk reduction among ex-smokers is considerably less among heavy smokers or ex-smokers who had smoked for longer periods than for light smokers or those who had smoked for shorter periods (58). The proportion of lung cancer in developed countries attributable to cigarette smoking was estimated as 83% to 94% in men and 57% to 80% in women. The proportion of lung cancers that are due to smoking would be decreased if smoking were totally eliminated. The encouragement of smoking cessation among current smokers and prevention of smoking initiation among school-age children seems to be the best prospect for reducing lung cancer incidence. Although the proportion of current smokers in the US declined from 50.2% to 31.7% in men and from 31.9% to 26.8% in women between 1965 and 1987, the reduction in smoking prevalence has been greatest among more educated men and as a result, smoking is increasingly becoming a habit associated with lower socioeconomic status (40). The plateau of age-specific rates for lung cancer among men is consistent with the reported success of smoking cessation programs among

men. With the continued rise of incidence rates of lung cancer among women, extensive health promotion research has focused on designing effective strategies to delay the smoking start age among young women, as well as facilitate their smoking cessation (40,45).

II. Environmental tobacco smoke

Environmental tobacco smoke is produced from the burning of the cigarette between puffs called sidestream smoke and from mainstream smoke exhaled by the smoker. There are significant differences in the composition of mainstream and side stream smoke. Most chemical components are emitted from cigarettes in about 2-5 times greater amounts in sidestream smoke than in the mainstream smoke, however, some nitrosamines are emitted at levels 50 times higher. Environmental tobacco smoke is mainly an indoor pollution problem, and exposure levels depend on the intensity of smoking, room size, and air exchange. More than 30 epidemiological studies have been published on environmental tobacco smoke and lung cancer. Some of these studies investigated the increased risk of lung cancer in nonsmoking females in relation to smoking habits of husbands. The results were inconclusive, with studies reported either positive or negative associations. However, after combining the results of these studies (total n = 272,387), a pooled relative risk analysis of lung cancer in nonsmoking women living with a smoker was 1.23 (95% confidence interval 1.11-1.36) (2). Many of the studies also investigated the effect of environmental tobacco smoke exposure during childhood and adulthood in relation to an increasing risk of lung cancer. Although some reports have suggested an increase in lung cancer risk among men with occupational exposure to environmental tobacco smoke and among children if the mother smoked, no consistent association was evident from other studies (2).

Epidemiological studies of passive smoking are confronted by a number of challenges. For example, the rarity of lung cancer in lifetime nonsmokers, the greater dilution of environmental

tobacco smoke compared to smoke inhaled by the active smokers, the difficulty of obtaining accurate exposure information by self-reports, the lack of a biological marker for long-term exposure assessment, misclassification of smokers as nonsmokers, and the possibility of confounding by other risk factors, including diet or exposure to cooking fumes (for example women in China) (40). However, taking into consideration the presence of carcinogens in environmental tobacco smoke, exposure estimates in passive smokers, exposure-response relationships in smokers, and epidemiological evidence of lung cancer in passive smokers, environmental tobacco smoke is of significant carcinogenic importance.

III. Occupational exposure

i) Asbestos

Numerous studies have demonstrated increased lung cancer risk among asbestos workers. The overall lung cancer risk was elevated 23% for workers who have been exposed to asbestos and the risk in a dose dependent fashion (3). Occupational asbestos exposure interacts with smoking to synergistically enhance lung cancer risk (59). Individuals who are exposed to both cigarette smoke and asbestos are at a 50-90 fold higher risk than unexposed individuals. The precise mechanism of the multiplicative interaction between cigarette smoke and asbestos remains unclear. Two mechanisms that have received considerable attention involve an asbestos-mediated enhanced delivery of mutagenic PAH compounds in tobacco smoke to the respiratory epithelium (60,61) and asbestos-induced chromosomal mutation and aneuploidy (62-64). In the first hypothesized model, the carcinogenic (genotoxic) potential of B[*a*]P may be enhanced several-fold by the promoter (epigenetic) effect of asbestos particles. This promoting effect could be related to the fact that when B[*a*]P is adsorbed onto the particles, there is a resulting enhanced transport and uptake of the carcinogen into epithelial cell membranes. Therefore, it is more likely

that individuals who are exposed to both cigarette smoke and asbestos have higher levels of B[*a*]P in the lung epithelium, resulting in a higher proportion of DNA damage compared to those in asbestos unexposed smokers. For the second hypothesized model, it is still controversial whether an increased susceptibility of asbestos-exposed individuals to B[*a*]P induced lung cancer results from an enhanced sensitivity to the induction of genetic damage by B[*a*]P or the direct physical interaction of the asbestos fibers with the chromosome or structural proteins of the spindle apparatus that causes missegregation of chromosomes during mitosis resulting in aneuploidy. One or both of these models may explain the mechanisms that underlie the enhancement of lung cancer risk resulting from an interaction of asbestos and smoking.

ii) Arsenic

There is substantial epidemiological evidence that inhalation of inorganic arsenic is a cause of lung cancer in the workplace. The mechanisms by which inorganic arsenic species induce lung cancer are unknown. Arsenic has generally failed to either initiate or promote cancer in experimental animals. A few studies have reported that arsenic can cause cell transformation, chromosomal aberrations, and sister chromatid exchanges, and can inhibit DNA repair in animal systems (4,5,65). Arsenic has recently been shown to induce oncogene amplification in cell culture, suggesting a role in the progression of initiated cells in the lung (5). Limited data are available on the interaction between occupational arsenic exposure and tobacco smoking. However, most of the epidemiological studies that provided data on the interaction between occupational exposure and tobacco smoking point toward a less than multiplicative interaction (66).

iii) Radon

Radon (radon-222) is chemically inert radioactive gas at typical environmental temperatures. Radon forms from the decay of the radium-226 that is in most soils and rocks. During the natural decay, some atoms leave the soil or rock and enter the surrounding air or water resulting in release radon to indoor and outdoor air and also to the air of underground passages and mines (67). Radon is classified as a group I carcinogen based on evidence that shows radon is carcinogenic in humans and a clear dose-response relationship from animal studies. Radon decays with a half-life of 3.82 days into radon decay products, polonium-218 and polonium-214. These two decay products emit alpha particles that are high energy and high mass particles consisting of two protons and two neutrons and highly effective in damaging tissues. Therefore, when people inhale radon, emissions of alpha particles take place within the lung, and genetic material in cells lining the airways may be damaged, possibly resulting in lung cancer. Lung cancer risk has been associated with exposure to radon decay products in different populations of underground miners in the US, Canada, Sweden, Czechoslovakia, etc. The interaction between tobacco smoking and radon exposure has been investigated in a number of epidemiological studies. These study results indicate that radon and smoking act synergistically in the etiology of lung cancer (67).

Based on studies of underground miners, it is firmly established that exposure to relatively high levels of radon and its decay products causes an increased risk of lung cancer in humans (40). The possible effects of radon in homes has led to public health concerns when radon and its progeny were detected in the air of buildings and there was an understanding of the mechanism by which radon reaches indoor air. The concentration of radium in soil and rock varies over several orders and accounts for most of the variation in radon concentration between dwellings.

The entry of radon into a building is determined by the structural characteristics of, and the flow of radon-containing air into, the building. Routes of radon entry include cracks and holes in basements, and cracks in concrete floors (67). Conditions of radon exposure in homes differ from that in mines. It is possible that long-term exposure to the relatively low radon levels in indoor air may cause a greater risk than expected based on linear extrapolation from the levels typical of mines (6). Inconsistent results among epidemiological studies of residential radon exposure in relation to lung cancer may be due to inadequate sample size, misclassification of exposure, and confounding by cigarette smoking. It is to be hoped that the large number of epidemiological studies on the effects of environmental exposure to radon that are currently in progress will provide useful evidence for defining the level of radon exposure at which an increase in risk is observed, and the interaction between radon exposure and cigarette smoking (67).

iv) Silica

During the past decade, there has been interest in a possible link between silica exposure and the risk of lung cancer. Epidemiological study findings of pulmonary cancer risks among silica-exposed workers show relative risk levels, which are consistently elevated across studies within the range of 1.3-1.7. After adjusting for age and smoking, the risk of lung cancer remains elevated. Evidence for a dose-response relationship has been found. A temporal sequence between exposure and response is observed and animal studies are confirmatory (7). The results of epidemiological studies that assessed the interaction between exposure to silica and tobacco smoking are however inconsistent. Although most of these studies point toward a less than multiplicative interaction, no conclusion can be drawn at present on the model of interaction between tobacco smoking and silica exposure (66). One complicating factor is the fact that in some of these studies, silicosis is used as a surrogate marker for exposure to silica. Furthermore,

none of these investigations included nonsilicotic subjects into the interpretation, which may provide clear evidence for the carcinogenicity of silica or for the role that tobacco smoke may play in it.

v) Chromium and Nickel

Exposure to chromium and chromium compounds is also associated with an increased risk of developing lung cancer. Little information is available on smoking habits for chromium-exposed workers and a possible interaction between smoking and chromium exposure has not been examined to date. Epidemiological investigations of the association of nickel exposure with lung cancer have been conducted. A risk of lung cancer has been found among mining, smelting, and refining operations workers (8). Again, there is little data regarding the effect of smoking and nickel exposure on lung cancer risk.

IV. Air pollution

It has long been suspected that pollution of outdoor air by industry, vehicles, power plants, and residential fire might contribute to excess lung cancer incidence. Outdoor air pollution is a complex mixture that is variable over time and region. Moreover, biologically significant interactions among the mixtures' components may occur. Studies of air pollution and lung cancer is further complicated by the overwhelming effect of cigarette smoking, occupational, and lifestyle factors (68). However, urban residence is still associated with increased lung cancer risk after adjustment for smoking and occupation (28). These may explain the 1.5-2.0 fold greater lung cancer incidence in cities compared to rural areas. Outdoor air pollution has worsened and lung cancer rates increased, it is plausible to postulate a causal link. Thus, while the overall contribution of air pollution is difficult to determine, exposure to polluted air is likely to contribute to a modest percentage of lung cancer incidence.

V. Endocrine factors

The higher proportion of nonsmokers and of lung adenocarcinoma among females compared to male lung cancer cases suggests a possible role of endocrine factors such as menstrual history, reproductive history, and use of exogenous hormones (40). One study suggested that the increased risk of adenocarcinoma of the lung after adjustment for smoking in Chinese women was associated with short menstrual cycles (69). There are other studies which have reported observations that support the role of the endocrine system in lung cancer among women. For example, the detection of steroid receptors in some lung cancers, a higher-than-expected rate of adenocarcinoma in lung among more than 10 year survivors of endometrial cancer, and an apparent increase in the risk of lung cancer in women receiving exogenous estrogens, including hormone replacement therapy (70-72). However, more investigation is needed before any conclusions regarding the relationship of endocrine factors to lung cancer risk among women can be made.

VI. Diet

The dietary factors most intensively studied with respect to lung cancer are retinoids in the form of vitamin A or its precursor carotenoids. Experimental studies including animal models have provided evidence in support of a protective effect of retinoids against lung cancer. For example, retinoid protect the integrity of the epithelium in the presence of chemical carcinogens (73). Many epidemiological studies have investigated the relationship between diet and lung cancer in humans. Consistent evidence of a substantial protective effect has been observed. Carotenoids are antioxidants under certain conditions. Their ability to neutralize free radicals is an accepted mechanism underlying the apparent protective effect on lung cancer. Some studies have found a protective effect of vegetable consumption, suggesting that other antioxidants such

as vitamin C and vitamin E may also be protective against lung cancer. (9,74). Based on the consistency of the epidemiological observations that smokers with lower intakes of carotene appear to be at higher risk after adjusting for cigarette smoking, chemoprevention trials have investigated the effect of retinoids on inhibition and reversion of lung cancer carcinogenesis with premalignant end-points such as bronchial metaplasia and dysplasia and sputum atypia. Even though retinoids have showed suppressing effect on lung carcinogenesis in animal models, they have failed to inhibit carcinogenesis in humans. Moreover, beta-carotene has been shown to be associated with greater lung cancer incidence and mortality (75). A study of a combination of beta-carotene and retinyl palmitate in 18,000 high-risk smokers in US found a 28% increase in the incidence of lung cancer. A similar finding was found in a study in Finland for the group taking beta-carotene (76). Therefore, these agents should not be regarded as harmless, but as having potential toxicities. Fat intake may also play a key role in lung carcinogenesis. In animal models, it has been shown that high intake of dietary fat can promote chemically induced pulmonary tumors (77,78). Several case-control studies provided some support for the association between diets high in fat and increased lung cancer risk. These studies showed greater risk among people with higher levels of fat or cholesterol intake in their diets (9,10,79). There are only a few reports studying the interaction of diets high in fat, which are also low in fruits and vegetables.

1.2. Genetic Susceptibility

It is therefore well established that smoking and other environmental exposures are major risk factors for lung cancer. However, not all smokers and individuals who are exposed to environmental carcinogens develop lung cancer. As a result, it has been hypothesized that other factors influence interindividual differences in susceptibility to smoking and environmental

exposures and may play a major role in lung carcinogenesis. It is now becoming increasingly evident that lung cancer has a genetic basis in addition to the well-known contribution of environmental factors. This evidence includes results of recent genetic epidemiological studies, metabolic phenotyping, cytogenetic studies, and oncogene and growth factor expression analyses. A brief review of the evidence on familial aggregation that may modify lung cancer risk follows. The contribution of genetic predisposition to lung cancer that results from individual differences in metabolism and DNA repair is then extensively described.

1.2.1. Familial aggregation

Several epidemiological studies have investigated the contribution of familial aggregation to lung cancer by examining lung cancer aggregation in family members of lung cancer patients in order to identify gene(s) responsible for lung cancer. Some evidence of familial susceptibility to lung cancer development has been observed. For example, there is a 2- to 4-fold greater risk of lung cancer among smoking and nonsmoking proband relatives after controlling for cigarette smoking (12,15). Furthermore, some studies found the greatest risk of lung cancer among subjects with first-degree relatives with lung cancer and a modest increase in lung cancer risk with other tobacco-related cancers reported in first-degree relatives (36,80-82). Sellers *et al.* (83) reported that the pattern of lung cancer in families ascertained through a proband with lung cancer is consistent with Mendelian segregation of an autosomal allele and predicted that a rare codominant gene is responsible for the age at onset distribution. According to this model, at the mean level of tobacco consumption, individuals who inherited two copies of the high risk allele have a mean age at onset of lung cancer of 38 years, roughly 24 years earlier than individuals who inherit only one high risk allele, which show a mean onset of 62 years. For individuals who are presumed to carry no high-risk alleles, the mean age at onset is 87 years (83). However, both

the differential exposure to tobacco across generation, and sex differences in tobacco exposure could have a profound effect on the modeling of lung cancer susceptibility.

Since most individuals in familial aggregation studies of lung cancer had some history of cigarette smoking, conclusions regarding the genetic contribution to the etiology of lung cancer cannot be made. Detailed information on cigarette exposure will be required to separate the contribution of genetics and the environment. Investigation into the genetics of lung cancer has been evaluated in individuals with lung cancer and little exposure to cigarettes. There are several genetic epidemiological studies that show familial aggregation of lung cancer among nonsmoking lung cancer cases with an earlier age at onset (84,85). These studies investigated the role of genetic factors and the effects of cigarette smoking, and passive smoking in lung cancer risk among relatives of nonsmoking probands. The first study observed that first-degree relatives (ages 40-59 years) of nonsmoking lung cancer cases have an over 6-fold increased risk of lung cancer compared to relatives of controls (84). No increased risk was seen in relatives of probands age 60 years or more. Similar results were found by another study in which no evidence of a major genetic effect was detected among older probands' families (≥60 years). However, in families of older probands, eliminating exposure to cigarette smoking not only can reduce lung cancer risk by over 85% among smokers, but also reduced the risk by over 60% among passive-smokers. In younger probands' families, a Mendelian codominant model with significant modifying effects of smoking best explained a genetic effect (85). These results suggest the presence of a high-risk gene, which contributes to early-onset lung cancer in a population where the probands are nonsmokers. Also, the results demonstrate that the attributable risk from the putative high-risk allele declines with age while the attributable risk from tobacco smoking increases with age.

Nevertheless, results of epidemiologic studies on familial aggregation require careful interpretation. Although the confounding effect of cohort differences in exposure may be controlled by stratification of the age of probands, a relatively small proportion of families have been studied among the young proband group. With family members sharing lifestyle and other environmental factors, it is difficult to provide conclusive evidence that accumulation of a disease has a genetic origin. Another limitation in familial aggregation studies is an inaccurate estimation of cigarette exposure, such as family history recall bias, potentially results in differential misclassification. Taken together, the findings from familial aggregation of lung cancer studies suggest that a genetic component might act as both independent risk factor and an effect modifier of exogenous risk factors, with smoking being the most important in lung cancer etiology.

1.2.2. Drug metabolizing enzyme polymorphisms

A potentially important host factor influencing individual susceptibility to tobacco smoke is interindividual differences in the metabolism of carcinogens in tobacco smoke, and of agents from environmental and occupational exposure. Most environmental carcinogens, including those in tobacco smoke, do not exert their biological effects per se, but require metabolic activation before they interact with cellular macromolecules. Many compounds are converted to reactive electrophilic metabolites by the oxidative (phase I) enzymes for example, cytochrome P450s (CYPs). These intermediates may then be conjugated with endogenous compounds and detoxified by phase II enzymes to become water soluble and thus readily excreted. Depending on the structure of the parent compound, adequate solubility in water may be obtained following after phase I reactions. Conversely, some compounds may be metabolized by phase II reactions without phase I metabolism. Differences in the metabolic activation and detoxification pathways

of environmental agents including tobacco smoke are likely to be a major source of interindividual variation in levels and types of DNA mutation and susceptibility to cancer.

Cytochrome P450s (CYPs) are unique hemoproteins that bind carbon monoxide to give a characteristic absorption spectrum at 450 nm. Hepatic CYPs are the major enzyme system involved in the phase I metabolism and toxicity of many drugs and xenobiotics (86). To date, many human CYPs have been characterized, and there are differences in their patterns of substrate specificity, regulation and expression. There are large interethnic variations in the CYP genes that have been investigated (87-91). Interindividual variation in drug or xenobiotic metabolism can occur as a result of specific CYP genetic polymorphisms in which individuals are either poor metabolizers (PMs) or extensive metabolizers (EMs). For example, human CYP genetic polymorphisms occur in a number of the CYP genes including *CYP1A1* (92-94), *CYP1A2* (95,96), *CYP1B1* (97-99), *CYP2D6* (100-102) and *CYP2E1* (103,104). CYP genetic polymorphisms can result in variability of enzyme expression and activity levels.

Phase II enzymes such as glutathione S-transferase (GST) and N-acetyltransferase (NAT) are generally involved in phase II detoxification pathways. To date, human GST isozymes have been identified which are referred to as Alpha, Mu, Pi, Theta, and Zeta (105). GST isozymes have been shown to be polymorphic, resulting in reduced enzyme activity (106-111). Not only have large interindividual variations in enzyme activities been demonstrated for several GSTs, but also ethnic differences in the frequency distribution of some GST allelic variants have been shown (112,113). The human N-acetyltransferase enzymes are involved in transformation of drugs and chemicals, including the activation of arylamines in tobacco smoke. Genetic polymorphisms that result in a slow acetylator phenotype affects about 50% of Caucasians (114). The previous understanding has been that the ultimate origin of acetylation polymorphism is

ascribable to the *NAT2** gene locus. However, recent evidence shows that *NAT1** may in fact also participate in the acetylation polymorphism (115). A wide range of values for NAT activity in different population groups has been reported. The *NAT1* fast acetylator alleles are found in frequencies ranging from 15% to 50%, whereas the *NAT2* slow acetylator alleles range from 5% to 90% among different ethnic groups (116).

Recently, polymorphisms in genes encoding for some other enzymes involved in the metabolism of tobacco smoke and environmental carcinogens, such as microsomal epoxide hydrolase and myeloperoxidase, have also been investigated as host factors that can modulate interindividual susceptibility. Microsomal epoxide hydrolase (mEPHX) encoded by the *EPHX1* gene, is a critical biotransformation enzyme that catalyzes the hydrolysis of a large number of epoxide intermediates, which arise frequently from the oxidation of xenobitics and environmental compounds by CYPs. Single nucleotide polymorphisms (SNPs) in the *EPHX1* gene may result in either increased or decreased enzyme activity. Myeloperoxidase (MPO) is a heme-containing enzyme, which is present in azurophilic granules of human polymorphonuclear granulocytes (PMNs) (117). Animal models have demonstrated that inhalation of cigarette smoke results in an accumulation of neutrophils in the lung. In humans, cigarette smoke may attract neutrophils to the lung by stimulating alveolar macrophages to release a potent chemotactic factor for neutrophils (118). Myeloperoxidase activates carcinogens including B[*a*]P as well as aromatic amines in tobacco smoke. A single base substitution (G to A) within the Alu element preceding the *MPO* gene results a lack of a strong general transcription factor SP1 binding site. The allele which has an A residue instead of G confers severalfold less transcriptional activity in transient transfection assays (119).

1.2.3. DNA repair gene polymorphisms

DNA repair systems are essential for the maintenance of genomic integrity. Both endogenous and exogenous exposures to carcinogens or genotoxic agents cause cell-cycle delays that allow cells to repair DNA damage. Another important source of interindividual variability in relation to the development of cancer is variability in DNA repair capacity. Several inherited genetic instability syndromes including Ataxia-telangiectasia (A-T), Fanconi anemia and Bloom's syndrome are characterized by both chromosomal instability and high risk of cancer (120). Xeroderma pigmentosum (XP) is caused by a deficiency in nucleotide excision repair (NER) and characterized by extreme susceptibility to ultraviolet (UV) light associated with skin cancer is another example of an association between deficiency in DNA damage repair and increased cancer risk (120). In addition to these rare syndromes, individuals differ widely in their capacity to repair DNA damage from both exogenous agents such as tobacco smoke and endogenous reactions such as oxidations. Some of these interindividual differences are likely to have a genetic origin. Recently, forty different amino acid substitution variants have been identified in nucleotide excision repair genes (*XPD,* and *XPF*), a gene involved in double-strand break repair/recombination genes (*XRCC3*), and a gene functioning in base excision repair and the repair of radiation-induced damage (*XRCC1*) with variant allele frequencies ranging from 0.04-0.50 (121). Five additional single-nucleotide polymorphisms in the coding regions of the *XPF* gene have been reported (122). Khan *et al.* (123) reported a new intronic, biallelic poly (AT) insertion/deletion polymorphism (*XPC-PAT*) of the DNA repair gene *XPC* that is in linkage disequilibrium with a single-nucleotide polymorphism in *XPC* exon 15. The biological effects of these polymorphisms have not been elucidated, but some of these variants may be associated with a reduced repair capacity and increased cancer susceptibility. Therefore, polymorphisms in

DNA repair genes may be associated with differences in the repair efficiency of DNA damage and may also influence susceptibility to environmental and occupational carcinogens and predisposition to cancer.

1.3. Genetic susceptibility biomarkers

Genetic polymorphisms in the enzymes responsible for activation (phase I) and detoxification (phase II) of tobacco carcinogens which may modulate the dose of these carcinogens that lung tissue is exposed and in the DNA repair genes affecting the repair efficiency of DNA damage from tobacco carcinogens in those lung tissues, respectively, are therefore likely to be an important source of interindividual differences in lung cancer susceptibility. In order to assess the impact of genetically polymorphic biotransformation and DNA repair on interindividual differences in lung cancer susceptibility, genetic susceptibility biomarkers of polymorphisms in metabolic and DNA repair genes involved in tobacco carcinogens metabolism that reflect the mechanism of action of those carcinogens have been developed and validated. It is assumed that these validated genetic susceptibility biomarkers will be strong predictors of an individual's risk of lung cancer. Several studies that have been demonstrated the usefulness of genetic susceptibility biomarkers for polymorphisms associated with metabolism and repair genes in increasing the strength of an association between exposure and lung cancer. The study results have provided considerable support for a causative association between exposure and onset of cancer that most notably by increasing the magnitude of observed relative risks and by providing mechanistic explanation for the development of lung cancer.

1.3.1. CYP1A1

The *CYP1A1* gene product, aromatic hydrocarbon hydroxylase (AHH), is primarily a extrahepatic enzyme that catalyzes the first step in the metabolism of PAHs, such as B[*a*]P in cigarette smoke, to the ultimate DNA-binding, carcinogenic agent. The *CYP1A1* gene is induced by exposure to exogenous agents such as B[*a*]P, other PAHs and dioxin (124). There are 10 allelic variants within the *CYP1A1* gene, which have been reported and are listed in Table 2.

Table 2. *CYP1A1* allele nomenclature

Allele	Nucleotide changes	Amino acid substitution	Enzyme activity		References
			In vivo	In vitro	
CYP1A1*1A	**None**	None	**Normal**	**Normal**	**(125–127)**
*CYP1A1*1B*	C3219T	None			(128)
*CYP1A1*1C*	G3229A	None			(128)
CYP1A1*2A	**T3801C**	**None**			**(129)**
*CYP1A1*2B*	A2455G, T3801C	I462V			(93)
CYP1A1*2C	**A2455G**	**I462V**		**Normal**	**(93,130,131)**
CYP1A1*3	**T3205C**	None			**(94)**
CYP1A1*4	**C2453A**	**T461N**			**(132)**
*CYP1A1*5*	**C2461A**	R464S			(133)
*CYP1A1*6*	G1636T	M331I			(133)

Modified from http://www.imm.ki.se/CYPalleles/

The relationship between *CYP1A1* genotypes and phenotypic expression has not been extensively studied. AHH inducibility is correlated with the *CYP1A1*2A* but not the *CYP1A1*2B* allele (134,135). The *CYP1A1*2A* variant has been associated with increased lung cancer susceptibility in Japanese populations (92) but this finding was not confirmed in Caucasians (136,137), possibly due to a lower allele frequency of 12% in Caucasians compared

to 33.2% in Japanese (138). The functional significance of African-American specific polymorphism (*CYP1A1*3*) is unclear. Also, the relevance of this variant allele to lung cancer susceptibility is controversial (139,140). Cascorbi *et al.* (132) reported the allele frequency of 2.87% for *CYP1A1*4* among both lung cancer patients and controls. There are several studies which have evaluated the association between the currently known *CYP1A1* polymorphisms and lung cancer risk and have suggested that *CYP1A1* polymorphisms may explain the interindividual variation in lung cancer susceptibility (141-144).

1.3.2. CYP1A2

CYP1A2 is involved in the metabolic activation of endogenous estrogens and several carcinogens such as aromatic and heterocyclic amines, and nitroaromatic compounds. In humans, CYP1A2 protein has only been detected in the liver. Genetic polymorphisms in the *CYP1A2* gene account in part for the wide interindividual differences in CYP1A2 activity that observed in humans. The functional significance of two allelic variants of *CYP1A2* is known. *CYP1A2*1C,* identified by a point mutation from guanine to adenine at position -2964 in the 5′-flanking region of the *CYP1A2* gene (95), results in a significant decrease of CYP1A2 transcription and activity. The *CYP1A2*1F* variant allele (C-734A) results in increased inducibility. Other allelic variants of *CYP1A2* are shown in Table 3.

Table 3. *CYP1A2* allele nomenclature

Allele	Nucleotide changes	Amino acid substitution	Enzyme activity		References
			In vivo	In vitro	
CYP1A2*1A	**None**	**None**	**Normal**	**Normal**	**(145,146)**
*CYP1A2*1B*	T1545C	None			(95,147)
CYP1A2*1C	**G-3858A**	**None**	**Decreased**		**(148)**
*CYP1A2*1D*	-2464delT	None			(149)
*CYP1A2*1E*	T-740G	None			(149)
CYP1A2*1F	**C-164A**	**None**	**Higher inducibility**		**(149,150)**
*CYP1A2*1G*	T-740G, T1545C	None			(151)
*CYP1A2*1H*	A951C, T1545C	None			(151)
*CYP1A2*2*	C63G	F21L			(152)
*CYP1A2*3*	G1042A, T1545C	D348N			(151)
*CYP1A2*4*	A1156T	I386F			(151)
*CYP1A2*5*	G1217A	C406Y			(151)
*CYP1A2*6*	C1291T	C431Y			(151)

Modified from http://www.imm.ki.se/CYPalleles/

CYP1A2 enzyme activity is also induced by environmental and dietary compounds such as those from tobacco smoke and charbroiled or high-temperature cooked meat. Several epidemiological studies have been conducted to investigate the association of CYP1A2 activity, alone or in combination with other CYPs, and cancer risk. Most studies have focused on the relationship between CYP1A2 phenotype, environmental exposure and risk of colorectal or urinary bladder cancer. It has been shown that elevated CYP1A2 activity is a risk factor for lung cancer, especially among individuals who are smokers and regularly consume charbroiled or high-temperature cooked meat. These individuals are exposed to high amounts of PAHs,

aromatic and heterocyclic amines, and nitroaromatic compounds that can induce the CYP1A2 activity in liver and are metabolized by CYP1A2. This results high levels of reactive intermediates that can enter the circulation, and be absorbed across the lung epithelium where DNA adduct formation may occur as been observed in *in vivo* animal feeding studies (153,154) as well as identification of arylamine-DNA adducts in human peripheral lung tissue (155). Another study investigated the association between CYP1A2 activity and lung cancer risk among non-smoking Chinese women. It was found that individuals with higher CYP1A2 activity might have a higher risk of lung adenocarcinoma. Furthermore, a significant increased risk of lung adenocarcinoma (OR 6.9, 95% CI, 1.3-37.6) was observed in subjects with slow NAT2/high CYP1A2 activity. Although this finding is limited by the small numbers of subjects in the study, further data on the contribution of dietary heterocyclic amines, endogenous estrogens and CYP1A2/NAT2 activities to risk of lung cancer especially adenocarcinoma among non-smoking women may explain a mechanism involved in lung cancer etiology.

1.3.3. CYP1B1

CYP1B1 has recently been identified and cloned. CYP1B1 is constitutively expressed in extrahepatic tissues including lung and mammary tissues, and catalyzes the activation of both PAHs and aryl amines (156). CYP1B1 is very active in catalyzing the activation of lung carcinogens such as (+)- and (-)-B[*a*]P-7,8-diol, 7,12-dimethylbenz[*a*]anthracene (DMBA)-3,4-diol, and dibenzo[*a,l*]pyrene (DB[*a,l*]P-11,12-diol. There are nineteen allelic variants of human *CYP1B1*, including those leading to amino acid substitutions (97,98) as shown in Table 4.

Table 4. *CYP1B1* allele nomenclature

Allele	Nucleotide changes	Amino acid substitution	Enzyme activity		References
			In vivo	In vitro	
CYP1B1*1	**None**	**None**	**Normal**	**Normal**	**(157)**
*CYP1B1*2*	C142G, G355T	R48G, A119S			(98,158)
CYP1B1*3	**C4326G**	**L432V**			**(97,98)**
*CYP1B1*4*	A4390G	N453S			(97,98)
*CYP1B1*11*	G171C	W57C			(98)
*CYP1B1*12*	G182A	G61E			(98)
*CYP1B1*13*	501insT	frameshift			(98)
*CYP1B1*14*	G841T	E281X			(98)
*CYP1B1*15*	863insC	frameshift			(159)
*CYP1B1*16*	Deletion	splicing error			(159)
*CYP1B1*17*	4096del13	frameshift			(159)
*CYP1B1*18*	G4125T	G365W			(98)
*CYP1B1*19*	C4168T	P379L			(98)
*CYP1B1*20*	G4191A	E387K			(98)
*CYP1B1*21*	G4201A	R390H			(98)
*CYP1B1*22*	4232dup10	frameshift			(98)
*CYP1B1*23*	C4342T	P437L			(98)
*CYP1B1*24*	G4377del	frameshift			(98)
*CYP1B1*25*	C4437T	R469W			(98)
*CYP1B1*26*	4435dup27	frameshift			(98)

Modified from http://www.imm.ki.se/CYPalleles/

An association between *CYP1B1* polymorphisms and increased risk of lung cancer, in particular squamous cell carcinoma, has been observed only with the *CYP1B1*2* variant (99). However, it has also been reported that the *CYP1B1*2*, *CYP1B1*3* variants are associated with increased activation of some of the dihydrodiol metabolites of PAHs (1.2 to 1.5-fold) (160). Although the effect of *CYP1B1* single nucleotide polymorphisms on catalytic activity has not been completely evaluated, interindividual differences in the metabolism of procarcinogens found in tobacco smoke as a result of genetic polymorphisms of *CYP1B1* gene may contribute to increased susceptibility to lung cancer.

1.3.4. CYP2D6

CYP2D6 plays a role in the metabolism of the tobacco specific nitrosamine, 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK). Genetic polymorphisms in the *CYP2D6* gene result in four major phenotypes, ultra-rapid metabolizers (UM), extensive metabolizers (EM), intermediate metabolizers (IM), and poor metabolizers (PM). In Caucasians, 3-10% of individuals are poor metabolizers as a result of either absent or inactive *CYP2D6* gene (161). There are fifty-three variant alleles of *CYP2D6* gene have been characterized (90). The most common enzyme-inactivating allelic variants of *CYP2D6* are *CYP2D6*3A*, which has a single nucleotide deletion in exon 5 resulting in frameshift and *CYP2D6*4A*, which contains several silent single nucleotide substitutions and a point mutation in the consensus sequence for the splice site at the intron 3/exon 4 splice junction. The *CYP2D6* gene deletion is termed as *CYP2D6*5*. Five allelic variants (the wild-type *CYP2D6*1A, *2, *2B, *4* and **5*) account for about 87% of all alleles. The remaining alleles occur with a frequency of 0.1%-2.7%.

The possible association between the *CYP2D6* polymorphisms and lung cancer risk has been investigated in several studies. It was first reported that extensive metabolizers are at markedly

increased risk of lung cancer (35). However this association between *CYP2D6* and lung cancer susceptibility was not confirmed in several other studies (162-164). More recently, a weak association between the *CYP2D6* genetic polymorphism and lung cancer risk was reported following a meta-analysis of the data from all published studies (165) and confirmed by a recent population-based case control study that co-analyzed a number of variables such as smoking history and occupational exposure to environmental pollutants (166). Although the molecular basis of impaired activity of CYP2D6 is now well understood, varying *CYP2D6* allele frequencies in different populations, methods used to assign CYP2D6 phenotype or genotype, selection of the control population, smoking history etc., makes the significance of the data obtained from several studies difficult to determine.

1.3.5. CYP2E1

Cytochrome P450 2E1 activates several known carcinogens including *N*-nitrosamines, benzene, styrene, and vinyl chloride. *CYP2E1* expression is dramatically enhanced by ethanol, benzene and tobacco smoke exposures. The *CYP2E1* gene is also reported to have genetic variation with significant differences of allele frequencies among different racial groups.

Table 5. *CYP2E1* allele nomenclature

Allele	Nucleotide changes	Amino acid substitution	Enzyme activity		References
			In vivo	In vitro	
CYP2E1*1A	**None**	**None**	**Normal**	**Normal**	**(167)**
*CYP2E1*1B*	C9893G	None			(168,169)
*CYP2E1*1C*	6 repeats in the 5' flanking region	None			(170)
*CYP2E1*1D*	8 repeats in the 5' flanking region	None	Increase in activity after alcohol exposure and in obese subjects		(170,171)
*CYP2E1*2*	G1132A	R76H		Reduced	(172)
CYP2E1*3	G10023A	V389I		Normal	(172)
*CYP2E1*4*	G4768A	V179I		Normal	(173)
*CYP2E1*5A*	G-1293C, C-1053T, T7632A	None			(103,174,175)
CYP2E1*5B	**G-1293C, C-1053T**	**None**			**(103,174)**
*CYP2E1*6*	T7632A	None			(175)
*CYP2E1*7A*	T-333A	None			(173)
*CYP2E1*7B*	G-71T, T-333A	None			(173)
*CYP2E1*7C*	T-333A, A-352G	None			(173)

Modified from http://www.imm.ki.se/CYPalleles/

The *CYP2E1*6* allele is present at an allele frequency of 8-11% in Caucasians and African-Americans but at a frequency of 24-29% in Oriental populations. The allelic frequencies of *CYP2E1*5B* variant allele were 2-5% in Caucasians and African-Americans but 19-28% in Oriental populations. Since CYP2E1 is important in the metabolic activation of various *N*-nitrosamines, including the potent tobacco-specific procarcinogen 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK), differences in CYP2E1 activity as a result of polymorphisms in the *CYP2E1* gene may be responsible for variations in host susceptibility to lung cancer. The *CYP2E1*6* variant was reported to be associated with lung cancer among Japanese (104) but not among Caucasians and African-Americans (176). There have been also contradictory and inconclusive reports associating the *Rsa*I polymorphism with susceptibility to lung cancer and smoking exposure (175,177,178). A major reason for this discrepancy might be ethnic differences in allele frequencies. According to a meta-analysis of the seven published case-control studies on the *CYP2E1* 5′-flanking region polymorphism and lung cancer, a possible higher risk for the homozygous wild type genotype was indicated. However, stratification by ethnicity did not reveal significant differences in lung cancer risk associated with the 5′-flanking region polymorphism among Caucasians, African-Americans, and Asians (179). There are few data on the relationship between *CYP2E1* genotypes and phenotypes in humans. Furthermore, the inducibility of *CYP2E1* by cigarette smoking for individuals with different genotypes remains unknown. The association between the *CYP2E1* polymorphisms, their phenotypes and lung cancer risk needs to be further explored.

1.3.6. GSTs

The glutathione S-transferases (GSTs) include a supergene family of phase II detoxifying enzymes that catalyze reactions between glutathione (GSH) and electrophilic compounds,

including carcinogens present in the diet and tobacco smoke such as PAHs. There is evidence for polymorphisms in the GST genes among the Mu, Theta, and Pi gene families. Three alleles for *GSTM1* polymorphisms have been reported. The *GSTM1*0* allele (*GSTM1* null genotype) is a result of a homozygous gene deletion of the *GSTM1* gene locus resulting in no GSTM1 activity (107). The other two variant alleles (*GSTM1*A* and *GSTM1*B*) differ by only a single amino acid for a cytosine to guanine at position 534, which results in the substitution of amino acid lysine to asparagines. These two alleles encode GSTM1 enzyme that shows similar catalytic activity and are classified together as the positive conjugator phenotype. The *GSTM1*0* allele has a high prevalence and approximately 40-60% of Caucasians have the null genotype (180). Recently, two alleles have been identified for the *GSTM3* gene (*GSTM3*A* and *GSTM3*B*). The *GSTM3*B* allele differs from the *GSTM3*A* allele by containing a 3 base-pairs deletion in intron 6 of the *GSTM3* gene which forms a recognition motif for the YY1 transcription factor and has been postulated to increase GSTM3 transcription. *GSTM3*B* is in a linkage disequilibrium with *GSTM1*A*. Thus, individuals with *GSTM1*A*/*GSTM3*B* may express more GSTM3 than those with *GSTM1*0*/*GSTM3*A* or *GSTM1*B*/*GSTM3*A* because *GSTM3*A* is not inducible by the YY1 transcription factor (105). *GSTM3* genotype frequencies were reported by Inskip *et al.* (181) at a frequency of 84.2% for *GSTM3*A* and of 15.8% for *GSTM3*B*.

GSTM1 is involved in the detoxification of carcinogens found in tobacco smoke such as the diol epoxide of B[*a*]P. The possible association between *GSTM1* null genotype and lung cancer risk has been investigated in several studies. The earliest studies reported that individuals lacking enzyme activity were at increased risk of lung cancer. Since then many studies that have classified subjects according to the presence or homozygous deletion of the *GSTM1* gene have yielded conflicting data and the influence of the GSTM1 on lung cancer risk remains unclear. A

meta-analysis of 51 published or unpublished studies involving about 20,000 individuals indicated that *GSTM1* null conferred a weak but significant increased risk of lung cancer (OR 1.20, 95% CI, 1.10-1.29) (182). GSTM3 is one of the most abundant expressed GSTs in human lung and also plays an important role in the detoxification of PAHs including B[*a*]P (183). There are a few studies that assess the *GSTM3* polymorphisms in association with lung cancer susceptibility. These studies found no significant contribution of *GSTM3* genotypes to the risk of lung cancer (184,185).

A gene deletion polymorphism is also common for the *GSTT1* (108). About 20-40% of Caucasians are homozygotes for the null allele, *GSTT1*0* (112). GSTT1 mainly plays a role in the detoxification of reactive hydrocarbons such as ethylene oxide, alkyl halides and diepoxybutane (108). Furthermore, GSTT1 has activity towards epoxides, for example, reactive epoxide metabolites of butadiene, suggesting that individuals with the null genotype at *GSTT1* may be at high risk of smoking-related cancers such as lung cancer (180). However, the role of *GSTT1*0* as a susceptibility factor for lung cancer has not been well established. Two association studies evaluating the *GSTT1* null genotype and lung cancer risk found no association of the *GSTT1* deletion polymorphism and lung cancer (186,187).

GSTP1 metabolizes many carcinogenic compounds including B[*a*]P diolepoxide (BPDE). Four *GSTP1* alleles have been identified as shown in Table 6.

Table 6. *GSTP1* allele nomenclature

Allele	Nucleotide changes	Amino acid substitution	Enzyme activity		References
			In vivo	In vitro	
GSTP1*A	**None**	**None**	**Normal**	**Normal**	**(110)**
GSTP1*B	**A313G**	**I105V**	**Reduced**	**Reduced**	**(110,113,188)**
GSTP1*C	**A313G, C341T**	**A114V**			**(110)**
*GSTP1*D*	C341T	A114V			(113)

The *GSTP1*B* variant allele has been shown to reduces enzyme activity as measured *in vitro* by 1-chloro-2,4-dinitrobenzene (CDNB) conjugation (113,188). However, enzymes with Val^{105} have a 7-fold higher efficiency for PAH diol epoxides than the enzymes with Ile^{105} (105). It was also suggested by Hu *et al.* (189) that subjects homozygous for *GSTP1*B* alleles would less susceptible to the carcinogenic effects of B[*a*]P than heterozygotes or wild-type homozygotes, due to more efficient detoxification capacity for the ultimate carcinogen of benzo[a]pyrene, (+)-anti-BPDE. Little is known about the role of *GSTP1* polymorphisms in individual response to environmental carcinogen exposures. Since GSTP1 is the most abundant GST isoform in human lungs (183), it is a particularly interesting candidate modifier of individual lung cancer susceptibility. However, the association between *GSTP1* variant alleles and lung cancer risk is still unknown (184,190,191).

1.3.7. NAT

N-Acetyltransferases metabolize a number of aromatic and heterocyclic carcinogens such as arylamines, heterocyclic amines, and 4-aminobiphenyl. In humans, two genes, *NAT1* and *NAT2*, are responsible for *N*-acetyltransferase activity. Aromatic and heterocyclic amine carcinogens are inactivated (*N*-acetylation) or activated (*O*-acetylation) by NAT1 and/or NAT2. NAT2 activity is

highest in the liver and gastrointestinal tract, whereas NAT1 activity is expressed in extrahepatic tissues. However, the relative contribution of hepatic versus extrahepatic activation and/or inactivation of aromatic and heterocyclic amine carcinogens is not fully understood (192). Both *NAT1* and *NAT2* are polymorphic resulting in a fast and slow acetylator phenotype. To date, there are 27 and 24 allelic variants of human *NAT2* and *NAT1* as shown in Table 7 and 8, respectively. *NAT2* alleles containing the G191A, T341C, A434C, G590A, and/or G857A substitutions are associated with the slow acetylator phenotype (192). Five of the *NAT2* variant alleles (*NAT2*5A*, *NAT2*5B*, *NAT2*5C*, *NAT2*6A*, and *NAT2*7B*) have been shown to account for more than 95% of the slow *NAT2* acetylator genotypes in Caucasians. For the *NAT1** gene locus; *NAT1*14*, *NAT1*15*, *NAT1*17*, *NAT1*19*, and *NAT1*22* have been clearly shown to reduce NAT1 activity (193). The *NAT1*10* allele has been reported to be associated with both higher NAT1 activity in bladder and colon tissues and higher DNA adduct levels in the colon tissues (193).

Table 7. *NAT2* allele nomenclature

Allele	Nucleotide changes	Amino acid substitution	Enzyme activity	References
NAT2*4	**None**	**None**	**Normal**	**(194,195)**
NAT2*5A	**T341C, C481T**	**I114T**	**Reduced**	**(196)**
NAT2*5B	**T341C, C481T, A803G**	**I114T, K268R**	**Reduced**	**(195,196)**
*NAT2*5C*	T341C, A803G	I114T, K268R	Reduced	(196)
*NAT2*5D*	T341C	I114T		(197)
*NAT2*5E*	T341C, G590A	I114T, R197Q		(197)
*NAT2*5F*	T341C, C481T, C759T, A803G	I114T, K268R		(198)
*NAT2*6A*	C282T, G590A	R197Q		(196)
NAT2*6B	**G590A**	**R197Q**	**Reduced**	**(196,199)**
*NAT2*6C*	C282T, G590A, A803G	R197Q, K268R		(199)
*NAT2*6D*	T111C, C282T, G590A	R197Q		(200)
*NAT2*7A*	G857A	G286Q		(196)
NAT2*7B	**C282T, G857A**	**G286Q**	**Reduced**	**(194)**
*NAT2*12A*	A803G	K268R		(196)
*NAT2*12B*	C282T, A803G	K268R		(196)
*NAT2*12C*	C481T, A803G	K268R		(199)
*NAT2*13*	C282T			(196,197)
NAT2*14A	**G191A**	**R64Q**	**Reduced**	**(201,202)**
*NAT2*14B*	G191A, C282T	R64Q		(202)
*NAT2*14C*	G191A, T341C, C481T, A803G	R64Q, I114T, K268R		(197,203)
*NAT2*14D*	G191A, C282T, G590A	R64Q, R197Q		(197,203)
*NAT2*14E*	G191A, A803G	R64Q, K268R		(197)
*NAT2*14F*	G191A, T341C, A803G	R64Q, I114T, K268R		(203)

Table 7. (cont'd)

Allele	Nucleotide changes	Amino acid substitution	Enzyme activity	References
*NAT2*14G*	G191A, C282T, A803G	R64Q, K268R		(200)
*NAT2*17*	A434C	Q145P		(204)
*NAT2*18*	A845C	K282T		(204)
*NAT2*19*	C190T	R64W		(205)

Modified from http://www.louisville.edu/medschool/pharmacology/NAT.html

Table 8. *NAT1* allele nomenclature

Allele	Nucleotide changes	Amino acid substitution	Enzyme activity	References
*NAT1*3*	C1095A	None		(206)
*NAT1*4*	None	None	Normal	(115)
*NAT1*5*	G350C, G351C, G497C, G498C, G499C, A884G, 976del1, 1105del1	R117T, R166T, E167Q		(207)
*NAT1*10*	T1088A, C1095A	None		(115)
*NAT1*11*	C-344T, A-40T, G445A, G459A, T640G, 1065-1090del9, C1095A	V149I, S214A		(208)
*NAT1*14A*	G560A, T1088A, C1095A	R187Q	Reduced	(209)
*NAT1*14B*	G560A	R187Q	Reduced	(210)
*NAT1*15*	C559T	R187Stop	Reduced	(209,210)
*NAT1*16*	1091insAAA, C1095A	None		(211)
*NAT1*17*	C190T	R64W	Reduced	(212,213)
*NAT1*18A*	1064-1087del3, T1088A, C1095A	None		(214)
*NAT1*18B*	1064-1091del3	None		(215)
*NAT1*19*	C97T	R33Stop	Reduced	(213)
*NAT1*20*	T402C	None		(213)
*NAT1*21*	A613G	M205V		(213)
*NAT1*22*	A752T	D251V	Reduced	(213)
*NAT1*23*	T777C	None		(213)
*NAT1*24*	G781A	E261K		(213)
*NAT1*25*	A787G	I263V		(213)
*NAT1*26A*	1066-1091insTAA, C1095A	None		(216)

Table 8. (cont'd)

Allele	Nucleotide changes	Amino acid substitution	Enzyme activity	References
*NAT1*26B*	1066-1091insTAA	None		(217)
*NAT1*27*	T21G, T777C	None		(218)
*NAT1*28*	1085-1090delTAATAA	None		(219)
*NAT1*29*	T1088A, C1095A, 1025del	None		(220)

Modified from http://www.louisville.edu/medschool/pharmacology/NAT.html

Tobacco smoke contains the aromatic amines 4-aminobiphenyl, 2-napthylamine, and o-toluidine which are activated and inactivated by NAT1 and/or NAT2. There are studies that have been published with regard to NAT2 polymorphisms and lung cancer risk, however results are inconclusive, with either negative or an increased risk for lung cancer in smokers with homozygous rapid acetylator (*NAT2*4/*4*) genotype (197,221-224). The role of *NAT1* genotype in lung cancer risk has also been investigated with conflicting results (192,224). The role of NAT as a susceptibility biomarker to lung cancer requires further investigation.

1.3.8. Microsomal epoxide hydrolase

Microsomal epoxide hydrolase (mEPHX) is involved in phase I metabolism and is capable of catalyzing the hydrolysis of highly reactive epoxide intermediates to less reactive and more water soluble trans-dihydrodiol derivatives (225). However, some trans-dihydrodiols generated from PAHs are substrates for further metabolism to more reactive carcinogenic polycyclic hydrocarbon diol epoxides such as B[*a*]P diolepoxide (BPDE). Thus, mEPHX plays a dual role in the detoxification and activation of procarcinogens, depending on the substrate. Two allelic variants of *mEPHX* have been reported as shown in Table 9. These polymorphisms are located within the coding region of the *mEPHX* gene at exons 3 and 4.

Table 9. *mEPHX* allelic variants

Allele	Nucleotide changes	Amino acid substitution	Enzyme activity		References
			In vivo	In vitro	
mEPHX	**None**		**Normal**	**Normal**	**(226)**
mEPHX3	**T17673C**	**Y113H**		**Decreased**	**(226)**
mEPHX4	**A24448G**	**H139R**		**Increased**	**(226)**

In vitro cDNA expression studies indicate that the *mEPHX3* allele is involved with decreasing mEPHX activity by 40% whereas the *mEPHX4* allele results in a 25% increase in mEPHX activity (226). As a result of these polymorphisms, there are four metabolic phenotypes, which are fast, normal, slow and very slow hydrolase activities in the general population.

The relationship between *mEPHX* polymorphisms and lung cancer risk has been investigated in several studies. Several studies found an association between high activity *mEPHX* genotypes and increased risk of lung cancer (227,228). However, one study reported a slight increase in the frequency of the very slow phenotype in lung cancer patients (229). Zhou *et al.* (225) observed the effect of very low activity of *mEPHX* genotype on increasing risk of lung cancer among nonsmokers and light smokers. However, among heavy smokers, the very low activity of *mEPHX* genotype was associated with decreased risk of lung cancer. Moreover, an interaction between *mEPHX* genotype and cumulative smoking exposure in lung cancer risk among squamous cell carcinoma subgroup was found in this study. Polymorphisms identified in the *mEPHX* gene may therefore also have an important role in determining the susceptibility of an individual to lung cancer.

1.3.9. Myeloperoxidase

Myeloperoxidase (MPO) is an enzyme located in granules of polymorphonuclear neutrophils and monocytes. Exposure to environmental insults, including tobacco smoke, stimulates the recruitment of neutrophils into human lung tissue with the local release of MPO. MPO activates various carcinogens associated with tobacco smoke including B[*a*]P and aromatic amines. Moreover, MPO is involved in the transformation of B[*a*]P-7,8-diol to an ultimate carcinogenic metabolite (230). As a result of a single base substitution (G to A) within the Alu element 463 base pair preceding the *MPO* gene, the A allele confers less transcriptional activity compared to the allele that contains a G residue. The effect of this polymorphism on susceptibility to lung cancer has been studied. London *et al.* (231) first reported a significant association between the myeloperoxidase *MPO* polymorphism (A/A genotype) and decreased risk of lung cancer. A second study also demonstrated the protective effects of the *MPO* variant allele and reduced lung cancer risk (232).

1.3.10. XPD

The *XPD* gene product is a member of the DNA repair subfamily and functions as single-strand DNA-ATP-dependent 5′-3′ DNA helicase within the basal transcription factor IIH (TFIIH) complex which participates in DNA unwinding during nucleotide excision repair (NER) (121). The NER pathway repairs DNA damage from both UV radiation and bulky DNA adducts. NER is a complex DNA repair process that consists of approximately 30 proteins involved in sequential damage recognition, DNA unwinding, incision of the damaged DNA strand on both sides of the lesion, excision of the oligonucleotide containing the damage and gap-filling DNA synthesis followed by strand ligation. The DNA bulky adducts generated during tobacco smoke

carcinogen metabolism including B[*a*]P or other PAHs and arylamines are removed effectively by the NER pathway. Recently, seventeen variant alleles of *XPD* gene were reported (121).

Table 10. *XPD* variant alleles

Gene	Nucleotide changes	Amino acid substitution	Enzyme activity		Allele frequency
			In vivo	In vitro	
XPD	**None**		**Normal**	**Normal**	
XPD	A18814G				0.21
XPD	C18980T				0.04
XPD	C22541A				0.25
XPD	A22559C				0.04
XPD	C22812T				0.33
XPD	C23047G	I199M			0.04
XPD	C23051T	H201Y			0.04
XPD	G23591A	D312N			0.42
XPD	C32983T				0.04
XPD	G34382C				0.29
XPD	C34706T				0.25
XPD	C34750T				0.04
XPD	C34770T				0.04
XPD	C35326T				0.25
XPD	35788insG				0.04
XPD	G35790C				0.29
XPD	**A35932C**	**K751Q**			**0.29**

Modified from reference (121)

The impact of these *XPD* polymorphisms on DNA repair is not known. A study of the effect *XPD* polymorphisms and proficiency for repair of X-ray-induced chromatid breaks and gaps by Lunn *et al.* (233) found that individuals carrying the *XPD*-751Lys/Lys genotype had significantly more ionizing radiation-induced chromatid aberrations than individuals homozygous for the *XPD*-751Gln allele. Therefore, the Lys751 allele may alter the XPD protein product resulting in suboptimal repair of X-ray induced DNA damage. It has been hypothesized that polymorphisms in DNA repair *XPD* gene may be associated with differences in the repair efficiency of DNA damage and genetic variants in *XPD* may also be markers for lung cancer susceptibility. One recent study investigated the hypothesis that the polymorphisms in DNA repair genes may influence an individual's risk of lung cancer and found that the *XPD* codon 312 Asp/Asp genotype have almost twice the risk of lung cancer when compared to the Asp/Asn + Asn/Asn combined genotype (234). This finding is consistent with that observed by Lunn *et al*. (233) in which individuals with the *XPD*-312Asp/Asp genotype showed more chromatid aberrations than the *XPD*-312Asn carriers, although this was not statistically significant. For *XPD*-751 polymorphisms, the result did not show any significant association with lung cancer risk. However, it was found that the *XPD*-751 polymorphism appeared to be in linkage disequilibrium with the *XPD*-312 polymorphism which individuals who carrying the *XPD*-312 Asn allele had the *XPD*-751Gln allele, while most *XPD*-312Asp allele carriers had the *XPD*-751Lys allele. This observation agrees with the higher chromatid aberrations observed among individuals who had *XPD*-312Asp/Asp or *XPD*-751Lys/Lys genotypes in the report of Lunn *et al.* (233). Although the functional significance of the *XPD* polymorphisms has not been fully elucidated, based on these preliminary results, additional molecular epidemiological studies of lung cancer risk in relation to the *XPD* polymorphisms should be pursued.

1.4. Gene-gene and gene-environment interactions

All of the enzymes described in the previous section have been hypothesized to be genetic susceptibility biomarkers for lung cancer. However, the underlying complexity of carcinogen metabolism, including overlapping substrate specificities for the CYP enzymes, and the involvement of multiple genes activation, detoxification and DNA repair suggests that gene-gene interactions are likely to further influence cancer risk. For example, the tobacco smoke carcinogen, B[*a*]P, can be activated by CYP1A1, CYP1A2, CYP1B1, CYP2C8, CYP2C9, CYP2C18, and CYP3A4. These B[*a*]P reactive metabolites can be further detoxified by both GSTM1 and GSTT1 as indicated in Table 11. Moreover, the bulky adducts generated by B[*a*]P-7,8-diol 9,10-epoxide are repaired by the NER pathway consisting of multiple genes. As shown in Fig 1, B[*a*]P is activated mainly by CYP1A1 to B[*a*]P-7,8-epoxide which then can be detoxified by glutathione conjugation and excreted. B[*a*]P-7,8-epoxide may also be detoxified by microsomal epoxide hydrolase to B[*a*]P-7,8-dihydrodiol which then can be activated by CYP3A4 to a highly reactive metabolite, B[*a*]P-7,8-diol 9,10-epoxide, which can be detoxified by glutathione conjugations and excreted or can interact with DNA to give rise to DNA adducts which may or may not be repaired. Thus, variation in activity in any combination of these enzymes may influence cumulative risk, for example, individuals with high activity variants of CYP1A1, microsomal epoxide hydrolase, and CYP3A4 together with less activity of GSTM1, GSTT1 and nucleotide excision repair enzymes would hypothesized to be at higher risk than any one of the risk genotypes alone. Therefore, the effects on one gene may compensate for changes in others or enhance the effect of another.

Table 11. Metabolizing enzymes responsible for metabolism of carcinogens in tobacco and cigarette smoke

Compounds	Metabolizing enzymes
PAHs e.g. B[*a*]P	CYP1A1, 1A2, 1B1, 2C8, 2C9, 2C18, 3A4, GSTM1, GSTT1
Aza-arenes e.g. Quinoline	CYP1A1, 1A2, 2E1, 3A, NAT2, GST, mEPHX
N-nitrosamine e.g. NNN, NNK, 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanol (NNAL) e.g. 4-ABP	CYP1A2, 2A6, 2B1, 2D6, 2E1, 3A4 CYP1A1, 1A2, NAT
N-heterocyclic amines eg. 2-amino-1-methyl-6-phenylimidazo [4,5-*b*] pyridine (PhIP)	GSTA1, GSTP1, NAT
Aldehydes e.g. Acetaldehyde	CYP1A2, 2E1, 4A2, ALDH2, GST
Miscellaneous compounds e.g. 1,3-butadiene	CYP2A6, 2E1, GSTT1

Modified from reference (235)

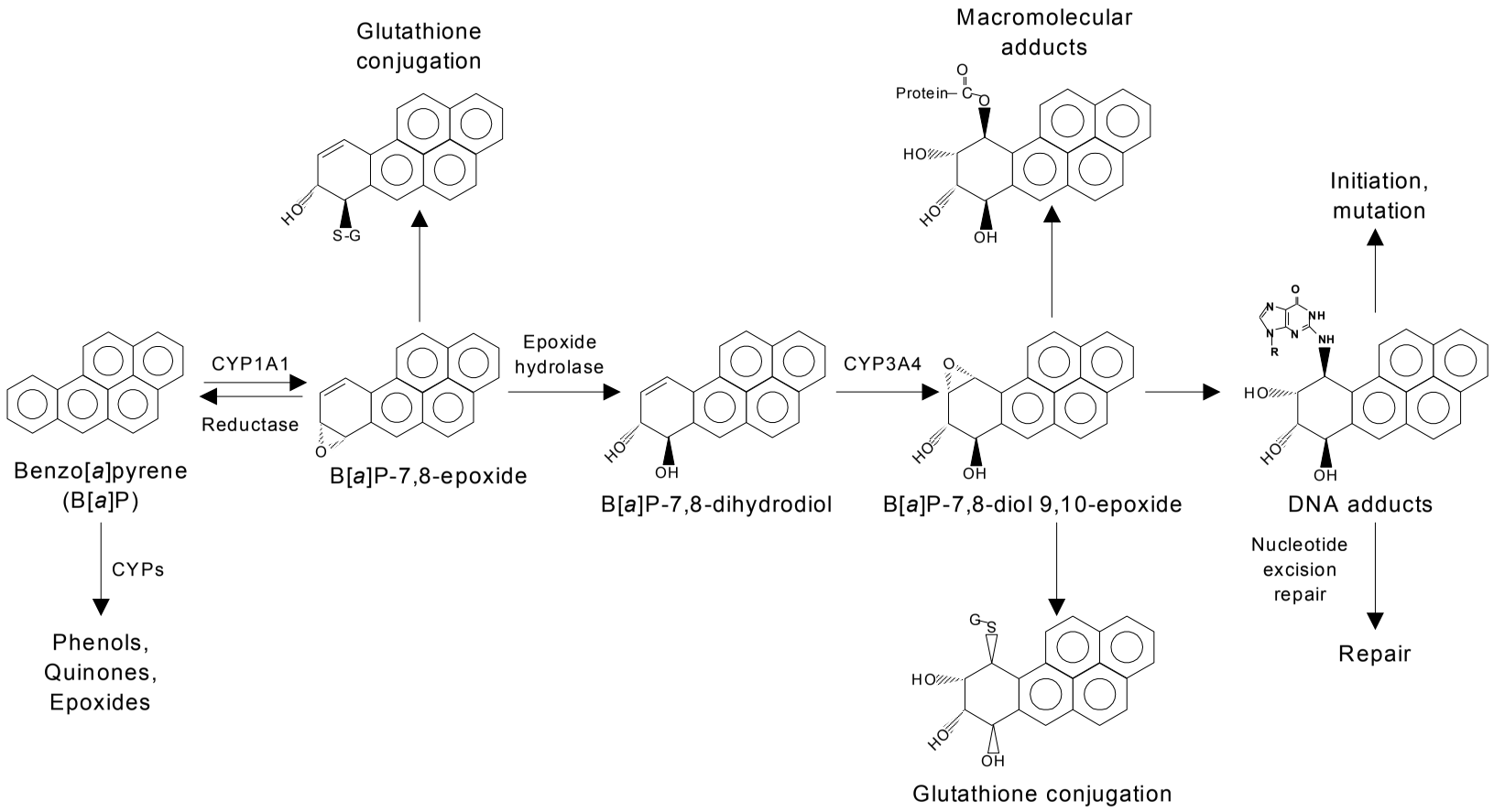


Figure 1 Metabolic pathway of benzo[*a*]pyrene

1.4.1. Gene-gene interaction

The study of interactions between at risk genotypes and lung cancer risk can include interactions between phase I, II and DNA repair enzymes. The interactions between phase I and phase II enzyme genetic polymorphisms have been investigated to some extent. For example, Japanese case control studies showed that the combination of *CYP1A1*2B* and GSTM1 null genotypes conferred a remarkably high risk of lung cancer (OR 5.8, P<0.001) (236) and a 41-fold increased risk of lung cancer particularly in low-dose cigarette smokers (237). Moreover, a dose-dependent increase in risk of lung cancer up to twenty-twofold was observed among heavy Japanese smokers with a combined *CYP1A1*2A/2A* and *GSTM1* null genotype (238). This finding is difficult to reproduce in Caucasians due to the significantly lower frequency of *CYP1A1* homozygous variant allele (*CYP1A1*2A/2A* frequency 12%) compared to Japanese populations (*CYP1A1*2A/2A* frequency 33%). However, Alexandrie *et al.* (239) reported that lung cancer cases who diagnosed before 66 years of age and carrying both *CYP1A1*2A/2A* and *GSTM1* null alleles has higher risk of squamous cell carcinoma. Thus, the combination between *CYP1A1* homozygous variant allele and *GSTM1* null genotype is likely to influence the risk of lung cancer in Caucasians in particular histological and age of diagnosis subgroups.

The interactions between phase II enzymes have been studied mainly within the GST family because some isoforms demonstrate overlap in substrate specificities, for example, GSTP1 detoxifies epoxides of PAHs similarly to GSTM1 and GSTM3. Moreover, Anttila *et al.* (240) observed that lung expression of GSTM3 was significantly higher in subjects with a *GSTM1*-positive allele than in subjects with a homozygous *GSTM1* null genotype. The mechanism for this observation is unknown but may reflect the linkage of *GSTM3*B* with *GSTM1*A* as suggested by Inskip *et al.* (181) and consequent association of *GSTM3* allele with the other

GSTM1 alleles, especially *GSTM1*0*. It is conceivable that combinations of at risk *GST* genotypes could result in higher risk of lung cancer.

Jourenkova *et al.* (241) investigated the effect of concurrent *GSTM1* and *GSTT1* null genotypes on lung cancer risk and found a significantly increasing risk of lung cancer among individuals who smoked more than 30 pack-years with both *GSTM1* and *GSTT1* null genotypes (OR 3.5, 95% CI, 1.2-10.7). The potential interaction of *GSTM1* and *GSTP1* genotypes in relation to lung cancer risk was examined by Ryberg *et al.* (190) and Kihara *et al.* (242). In the first study, the highest DNA adduct levels were observed in patients with a combination of *GSTM1* null and *GSTP1*A/*B* or *GSTP1*B/*B* genotype compared to all other genotype combinations (P=0.011). The risk of the *GSTM1* null genotype for lung cancer is enhanced in the presence of the *GSTP1*A/*B* or *GSTP1*B/*B* genotype among Japanese male smokers was reported in a second study with a smoking adjusted OR 2.58, 95% CI, 1.26-5.30 compared to that in the *GSTP1*A* or wild-type allele group (smoking adjusted OR 1.17, 95% CI, 0.77-1.79). Jourenkova-Mironova *et al.* (185) studied the potential role of the combination of at risk *GST* gene polymorphisms in susceptibility to lung cancer. The significant interactions between pack-years of smoking (≥ 35 pack-years) and the combined *GSTM3*A/*A* and *GSTP1*A/*B* or *GSTP1*B/*B* genotypes, and *GSTM3*A/*A, GSTP1*A/*B* or *GSTP1*B/*B*, and *GSTM1* null genotypes on increased risk of lung cancer among smokers (sex and age adjusted OR 2.1, 95% CI, 1.0-4.2 and OR 2.7, 95% CI 1.2-6.0, respectively) were observed.

There are relatively few data on interactions between phase I enzymes and lung cancer risk. Furthermore, no significant effects involving in the interactions between phase I enzymes, and phase I or II enzymes and DNA repair. Even though CYP1A1, 1A2, 2D6 and 2E1 show overlap

in substrate specificities of compounds found in tobacco smoke, low frequencies of the rare alleles for these loci may contribute to the difficulty in evaluating possible interactions.

Therefore, these significant interactions arising from the combinations of at risk genotypes among phase I and II enzymes and within phase II enzymes on the risk of lung cancer support the rationale for studying interactions between different metabolic enzyme polymorphisms that can affect the relative balance between activation and detoxification of tobacco related carcinogens in each individual, and increasing susceptibility to lung cancer.

1.4.2. Gene-environment interaction

Many genes and environmental exposures contribute to the carcinogenic process. The effects can be additive or multiplicative, and are modifiable by interindividual variation in genetic function and environmental exposures. In this dissertation research project, a hypothesis test is that **cancer risk may be influenced not only by variability in phase I and II metabolic and DNA repair enzymes, but environmental influences such as active or passive smoking, air pollution, workplace pollution, and pesticides may further modify cancer risk.** As shown in Fig. 2, the environmental factors contribute into the complex interactions between host susceptibility factors and environmental factors, gene-environment interaction, in the multistage process of carcinogenesis, which results in interindividual variation in cancer risk. Thus, lung cancer risk from exposure to tobacco smoke varies widely from person to person; depending in part on the status of metabolic and DNA repair genes that determine how cells activate, detoxify tobacco related carcinogens, and repair DNA damages. Activated carcinogen metabolites may bind to DNA and form DNA adducts, many of which can induce genetic mutations. Therefore, if individuals who are exposed to tobacco carcinogens have an increased capacity to activate

carcinogens, reduced capacity to detoxify reactive metabolites and repair the damage, they might form more carcinogen-DNA adducts and subsequently have an increased risk of lung cancer.

Several gene-environment interactions for lung cancer risk have been demonstrated (139,144,161,178,184,185,190,237,238,241,243-249)(Table 12). These study results indicate that genetic polymorphisms in carcinogen metabolizing genes have been shown to modify the effects of carcinogen exposure on lung cancer risk and DNA adducts.

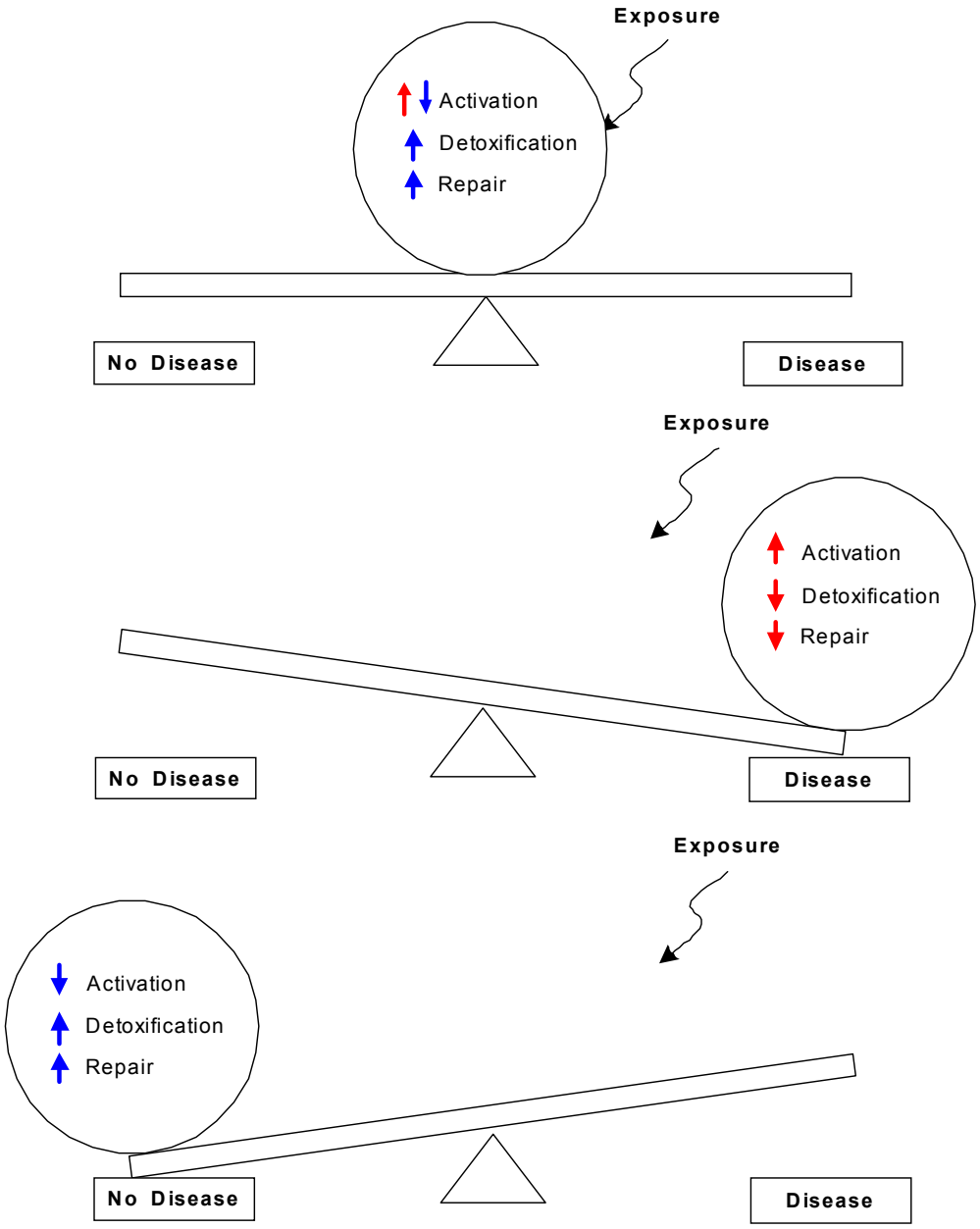


Figure 2. Gene-environment interaction

Table 12. Effects of gene-environment interaction between drug metabolizing enzymes and smoking, occupational, and environmental exposures to lung cancer

Drug metabolizing enzyme allele	Degree of cigarette smoking exposure	Relative Risk	References
*CYP1A1*2A/*2A*	Cumulative cigarette dose $<3 \times 10^5$	Increase (squamous cell carcinoma) OR 7.31, 95% CI, 2.13-25.12	(243)
*CYP1A1*2A/*2A*	Cumulative cigarette dose of $29.1 \pm 15.5 \times 10^4$	Increase (adenocarcinoma) OR 3.25, 95% CI, 1.40-7.56	(244)
*CYP1A1*1A/*2A & CYP1A1*2A/*2A*	≤ 30 pack-years	Increase OR 2.03, 95% CI, 1.03-4.01	(144)
*CYP1A1*1A/*3 & CYP1A1*3/*3*	Mean pack-years = 5	Increase (adenocarcinoma) OR 2.6, 95% CI, 1.1-6.3	(139)
*2D6*1A/*2* (EM)	< 30 pack-years	Increase (small cell carcinoma) OR 3.6, 95% CI, 1.1-11.9	(161)
*CYP2D6*3A/*3A or CYP2D6*4B/*4B or CYP2E1*1A/*6*	None	Increase (7-methyl-deoxyguanosine adduct)	(245)
*CYP2E1*1A/*6*	Low	Increase ($P<0.005$)	(178)
*GSTM1*0*	< 40 pack-years	Increase OR 1.77, 95% CI, 1.1-2.82	(246)
*GSTM1*0*	Levels of smoky coal exposure ≥ 130 tons	Increase OR 5.2, 95% CI, 2.1-12.6	(247)
*GSTM1*0*	Mean cigarette exposure = 25 cigarettes per day	Increase (DNA adducts in mononuclear blood cells, $P=0.05$)	(248)
*GSTM1*0*	Low of occupational PAH exposure <1.5 μg B[a]P/m^3 or High cigarette smoke exposure	Increase (B[*a*]P diolepoxide-DNA adducts in mononuclear blood cells, $P<0.0001$)	(249)

Table 12. (cont'd)

Drug metabolizing enzyme allele	Degree of cigarette smoking exposure	Relative Risk	References
*CYP1A1*1A/*2A* & *GSTM1*0*	Low of occupational PAH exposure <1.5 μg B[a]P/m^3 or High cigarette smoke exposure	Increase (B[*a*]P diolepoxide-DNA adducts in mononuclear blood cells, P<0.015)	(249)
*CYP1A1*1A/*2A* & *GSTM1*0*	Σof cigarettes smoked per day×years of smoking ≥ 800	Increase (squamous and small cell carcinoma) OR 16.4, 95% CI, 4.78-60.2	(238)
*CYP1A1*2A/*2A* & *GSTM1*0*	Cumulative cigarette dose < 32.1×10^4	Increase OR 16, 95% CI, 3.76-68.62	(237)
*GSTM1*0* & *GSTT1*0*	> 30 pack-years	Increase OR 3.5, 95% CI, 1.2-10.7	(241)
*GSTM1*0* & *GSTT1*0*	≤ 40 pack-years	Increase OR 2.9, 95% CI, 1.1-7.7	(184)
*GSTP1*A/*B* or *GSTP1*B/*B*, and *GSTM1*0*	≥ 35 pack-years	Increase OR 2.4, 95% CI, 1.1-5.1	(185)
*GSTP1*A/*B* or *GSTP1*B/*B*, and *GSTM1*0*	Mean = 35 pack-years	Increase (Lung tissue DNA adducts, P=0.011)	(190)
NAT1 slow	Mean = 25 cigarettes per day	Increase (DNA adducts in mononuclear blood cells, P=0.05)	(248)

Interestingly, these gene-environment interaction studies showed the genetic effects on the susceptibility to lung cancer at low levels of cigarette exposure in subjects with *CYP1A1*2A, CYP1A1*3*, *CYP2D6*2, CYP2D6*3A*, *CYP2D6*4B*, *CYP2E1*6, GSTM1*0,* or combined *CYP1A1*2A* & *GSTM1*0* alleles. These results indicate genetic host susceptibility might be important in individuals who smoke less and are among the population that show a high percent of the rare allelic frequencies such as *CYP1A1*2A* and *GSTM1* null genotypes. For example, Nakachi *et al*. (237) reported the risk of lung cancer was 16 times among light smokers who had *CYP1A1*2A/*2A* & *GSTM1*0* genotypes compared to smokers without these genetic polymorphisms. However, the risk was increased only 1.25 times among heavier smokers compared to light smokers who had the same 'at risk' genotypes. This low risk difference for cigarette dose levels among individuals with 'at risk' genotypes may be reflected by an overwhelming environmental factor (heavy smoking) on the genetic effects. Therefore, it is important to include genotyping of susceptible genes, which are involved in carcinogen metabolism, and the exposure levels for identification of very susceptible individuals to environmental carcinogens. Moreover, it is likely that people who smoked less or are nonsmokers might be exposed to other environmental carcinogens that further interacted with genetic factors, resulting in increasing risk in those individuals. Kato *et al.* (245) observed higher levels of 7-methyldeoxyguanosine adduct in lung tissues of nonsmokers who had *CYP2D6*3A/*3A* or *CYP2D6*4B*4B* or *CYP2E1*1A/*6* genotypes. It was suggested that the effect of the genotypes on higher adduct levels was mostly in nonsmokers who exposed to either passive tobacco smoke or to *N*-nitrosamine exposures other than tobacco smoke.

A dose-response relationship between the amount of cigarette smoking exposure and lung cancer risk have been observed. There are significant associations between high or moderate

levels of cigarette exposure and increased risk of lung cancer in individuals who had combined *CYP1A1*2A* & *GSTM1*0*, *GSTM1*0* & *GSTT1*0*, or *GSTM1*0* & *GSTP1*A*/or**B* genotypes (184,185,237,241). Moreover, high DNA adduct levels were detected in subjects who were exposed to occupational or environmental carcinogens and had *GSTM1* null, *NAT1* slow genotypes, *CYPA1*1A/*2A* & *GSTM1*0*, and *GSTM1* null & *GSTP1*B*. The results of these studies revealed that genetic polymorphisms of carcinogen metabolizing enzymes affect DNA adducts in human leukocytes, and lung tissues and risk of lung cancer from occupational and/or environmental carcinogen exposures (190,248,249).

Thus, recent knowledge of the genetic basis for individual metabolic variation and increased individual susceptibility to environmentally induced cancer, especially with reference to smoking-induced lung cancer has led to important paradigm, "gene-environment interaction". Lung cancer susceptibility due to cigarette smoking and/or other chemical carcinogen exposure is likely to be modulated by an individual's phenotype for a number of enzymes, including both activating and detoxifying enzymes involved in the metabolism of a single carcinogen or mixtures of carcinogens. Given the number and variability in expression of carcinogen metabolizing and DNA repair enzymes and the complexity of chemical exposures, the assessment of a panel of polymorphic enzyme combined with other biomarkers, such as biomarkers of exposure may be helpful in demonstrating gene-environment interaction in lung cancer susceptibility. The screening of a panel of candidate genes to assess lung cancer susceptibility was utilized in this research project.

1.5. Pharmacogenetics

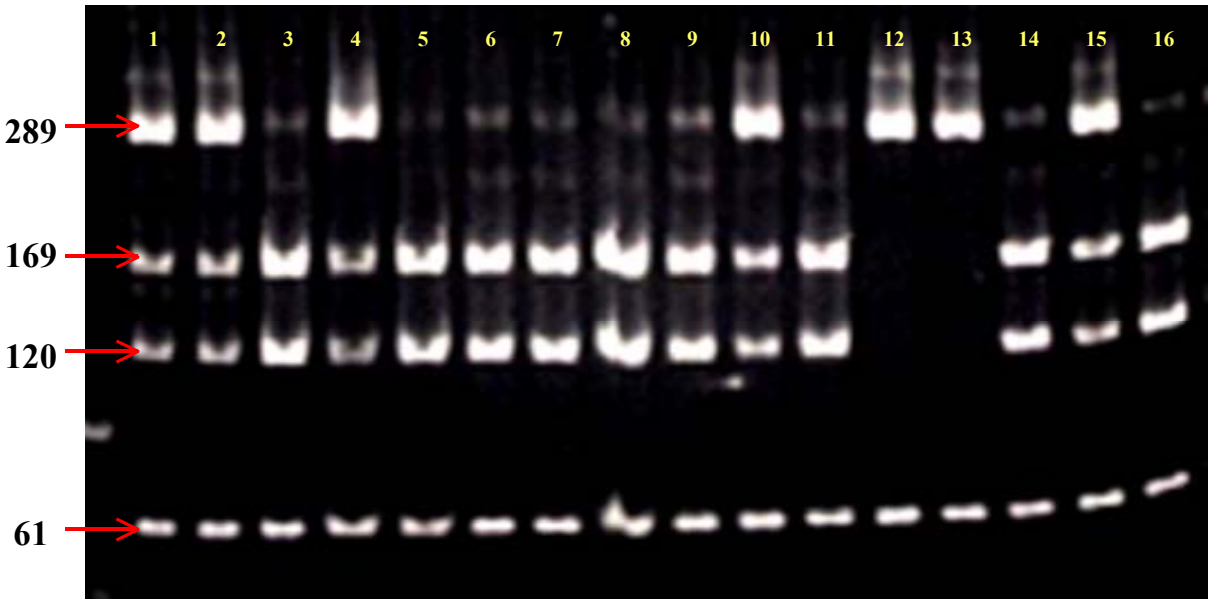
In order to accomplish the ultimate goal of human cancer risk assessment, the study of metabolic polymorphisms in cancer epidemiology requires the analytical tools to determine

genetic predispositions that result in increased sensitivity or resistance to exogenous exposures. Pharmacogenetics is the study of variability in individual responses to drugs and chemicals as a consequence of differences in individual genetic background. Through pharmacogenetic analyses, questions relating to interindividual differences in drug response, drug-drug interactions and cancer etiology and biology, such as the interaction between environmental and genetic factors in the development of common tumors, can be answered. Pharmacogenetics mainly uses two technologically different approaches, which are not mutually exclusive but rather are complementary. One approach involves phenotyping or assaying for individual metabolic capacity for certain drugs and the other, genotyping, involves direct screening of gene(s) encoding for enzymes involved in metabolism.

The assessment of in vivo enzyme activity level(s) can be determined following the administration of a therapeutic dose of a probe substrate, which is metabolized by a specific enzyme. A number of probe drugs have been employed for the determination of phenotype by measuring the relative amount of parent compound and metabolite excreted in the urine, or the metabolic ratio. Although phenotyping analysis is a well-established experimental approach that can give an accurate overall assessment of an individual's drug-metabolizing capacity, the limitation includes the lack of many validated probes, complicated procedures, invasiveness, high costs, and the results are confounded by a number of host and environmental factors. The drug metabolism phenotype confounders include age, hormonal status, disease, drug-drug interactions and dietary habits. Basically, the use of probe drug should give straightforward information about the genotype if the probe drug is exclusively metabolized by one enzyme that encoded by one allele, and the polymorphism is a differentiation between an active and an inactive allele. However, for example, an individual may inherit one copy of *CYP2D6*2A* allele

(normal allele) with two copies of the gene arranged in tandem and another copy of *CYP2D6*4B* allele which is a gene-inactivating allele resulting in an intermediate metabolizer phenotype. In this case, phenotype and genotype do not directly corresponding each other, and the prediction of CYP2D6 activity becomes dependent on the phenotyping assay.

Genotyping analysis is a method that predicts metabolic capacity on the basis of the exact allelic inheritance of each individual, and can be performed using a variety of complementary polymerase chain reaction (PCR)-based techniques. The most common of these methods is RFLP-PCR (RFLP, restriction fragment length polymorphism). It is an approach to detect genetic mutations including base pair changes, small deletions and insertions, which are located within a restriction enzyme recognition sequence. These genetic changes either create or destroy the recognition site of an endonuclease that resulting in different DNA size fragments after cleavage by the corresponding endonuclease between individuals who have genetic changes and who have not. Then, the different fragment size pattern can be detected by gel electrophoresis with ethidium bromide staining. Figure 3 is an ethidium bromide gel picture showing three possible genotypes for the polymorphism at position –463 of the *MPO* gene after the 350 bp size of PCR products were digested with an *Aci* I enzyme. A single base substitution (G to A) results an absence in one of the two *Aci* I recognition sites in a 350 bp size PCR product of the *MPO* gene. Individuals who have homozygous A/A allele show two bands, at 289, and 61 bp (lanes 12 and 13) whereas those homozygous G/G individuals show three bands, at 169, 120, and 61 bp (lanes 3, 5, 6, 7, 8, 9, 11, 14, and 16). Heterozygotes show all four bands (lanes 1, 2, 4, 10, and 15).



Lanes 3, 5, 6, 7, 8, 9, 11, 14, and 16 are G/G genotype

Lanes 12, and 13 are A/A genotype

Lanes 1, 2, 4, 10, and 15 are G/A genotype

Figure 3. Three possible genotypes from single nucleotide polymorphism (G or A) at 463 position of *MPO* gene detected by RFLP-PCR method

The genotyping approach is less laborious, less expensive, not invasive, not confounded by environmental factors, and gives an unequivocal genetically based prediction of individual drug metabolism compared to phenotyping analysis, although the functional status can not be determined in relevance of the polymorphism and cancer susceptibility. There is also wide variation of enzyme activity within a genotype, and of distribution of the genotype in different populations. However, genotyping assays provide an assessment of risk that reflects exogenous chemical carcinogen metabolism from the recent exposure as well as one's lifetime ability to activate and detoxify carcinogens. Therefore, genotyping is a method that is reliable, inexpensive, simple, safe, and rapid, and can be applied by large-scale screening to determine individual relative risk of developing cancer is being focused.

1.6. mRNA expression

In addition to inherited factors for the detection of sensitive populations, the potential for local metabolic activation of tobacco related procarcinogens in lung tissues is also another useful marker in determining the relative risk for developing lung cancer. *CYP1A1* and *CYP2E1* have been known to be expressed in extrahepatic tissues such as lung and lymphocytes (250-253). Recently, a new member of the human *CYP1* family, *CYP1B1*, was discovered (157,254). It is expressed in normal lung, kidney, uterus, prostrate and mammary. CYP1B1 is involved in the activation of B[*a*]P-7,8-diol and 2-aminoanthracene. (156). The levels of CYP1A1, CYP1B1 and CYP2E1 expression in lung tissues are known to be induced by carcinogens found in cigarette smoke (255-257). Lung tissues that are exposed to cigarette smoke are more likely to be exposed to more potent carcinogenic metabolites by CYPs induction. The expression of CYPs in lung tissue is therefore likely to have important metabolic and lung cancer susceptibility consequences.

In addition to the hypothesis that lung tissue expression of metabolic enzymes are important factors in modulating lung cancer risk, CYP mRNA expression levels (CYP1A1, CYP1B1 and CYP2E1) in lung tumors and normal tissues are likely to represent a smoking exposure biomarker. The differences in the levels of each CYP mRNA expression between lung tumor and histologically normal adjacent tissues would suggest a mechanistic link between cigarette smoking exposure and lung carcinogenesis. Moreover, the potential for CYP smoking induction can be determined if there is differences in the expression of CYPs between lung smoking and nonsmoking tumors. Thus, the investigation of CYP mRNA expression in lung tumor and histologically normal tissue will be helpful in determining the potential for local metabolic activation of procarcinogen agents in the lung.

1.7. Cigarette exposure biomarkers

While it is accepted widely that inhalation of chemical procarcinogens in cigarette smoke increase risk for lung cancer, this hypothesis is based on the assumption that procarcinogens are activated in lung tissues through some mechanism. This is because lung tissues are the main targets of cigarette related carcinogens exposure. Moreover, activated procarcinogens are highly reactive with proteins as well as nucleic acids, so as a consequence the carcinogenic effects are generally thought to be mainly involved in the cells in which the chemical are activated. There is considerable interindividual variation in lung cancer risk that has been attributed to an interaction between exogenous and endogenous factors. Both active and passive cigarette exposure, and environmental pollution exposures are considered to be a potentially important exogenous determination of risk. Polymorphisms in metabolic enzyme genes responsible for activation and inactivation of procarcinogens are important endogenous determinants. Therefore, measuring

interindividual variation in levels of gene expression in lung tissues is likely to be an approach to determine both exogenous and endogenous factors of lung cancer risk.

In order to evaluate a gene that represents a cigarette smoke and environmental pollution exposure biomarker, it must be expressed in the tissue of interest. Also, it must be inducible to a degree that is detectable by available methodology and over range that allows determination of exposure and response relationships. In contrast, a biomarker of susceptibility needs not to be inducible as long as it is expressed at different constitutive levels in different individuals. For example, CYP1A1, CYP1B1, and CYP2E1 are expressed in extrahepatic tissues including bronchial epithelial cells and also are inducible by PAHs. Furthermore, interindividual variation in quantitative levels of expression of AHH activity was detected and was associated with interindividual variation in the amount of B[*a*]P DNA adducts observed in bronchial epithelial cells (257). Thus, CYP1A1, CYP1B1, and CYP2E1 may potentially serve as biomarkers both for exposure to procarcinogens and for susceptibility of lung cancer (local bioactivation biomarker).

1.7.1. CYP1A1

The CYP1A1 gene product, AHH, catalyzes the first step in the metabolism of PAHs found in cigarette smoke to potent carcinogenic metabolites. Several investigators demonstrated that cigarette smoke induces AHH activity and proposed a relationship between AHH activity and lung cancer. McLemore *et al.* (256) found 89% (17 of 19) of normal lung tissues from active cigarette smokers and 0% (0 of 5) from nonsmokers expressed CYP1A1 mRNA. Moreover, a time-dependent decrease in expression of the *CYP1A1* gene occurs following cessation of cigarette smoking. The effect of CYP1A1 expression induced by cigarette smoke in lung tissues can last up to 60 days after cessation of smoking (255). Elevated levels of AHH activity in lung tissue from recent smokers corresponded with the extent of conversion of B[*a*]P-7,8-diol to

tetraols and BPDE-DNA adduct levels, and were also greater than those in lung tissues from nonsmokers or exsmokers (258,259). These results show that cigarette smoking induces CYP1A1 expression, and also suggest that there is a cigarette smoke-inducible pathway leading to BPDE-DNA adducts in smokers' lungs. The AHH activity mediated mainly by CYP1A1 in human lung was mainly involved in the metabolism of B[*a*]P to 3-hydroxyB[*a*]P (260,261). This suggests a role of CYP1A1 in local bioactivation of PAHs found in cigarette smoke in human lung tissue.

CYP1A1 mRNA expression induced by cigarette smoke is due to the presence of PAHs in cigarette smoking. Induction of CYP1A1 enzyme occurs at the level of transcription and via a specific receptor protein, the aryl hydrocarbon receptor (AhR) (262). The Ah receptor (AhR) is a ubiquitous protein that reacts as a potent transcription factor following activation by binding of ligands such as polycyclic hydrocarbon compounds. The latent AhR is a cytosolic protein held in a complex with the molecular chaperone heat-shock protein 90 (Hsp90). The binding of ligand to the AhR induces a conformational change in the receptor and subsequent nuclear translocation. Once AhR has been transported to the nucleus, a dimerization process with a nuclear protein, Arnt, occurs and Hsp90 is released from the AhR. The AhR/Arnt is able to bind specific DNA target sequences or xenobiotic-response elements (XRE), which function as enhancer regions. *CYP1A1* has six XRE sequences in the promotor region approximately 1 kilobase upstream. The binding of the AhR/Arnt complex to XRE sequences in the *CYP1A1* gene initiates transcription of the gene.

1.7.2. CYP1B1

CYP1B1 is able to catalyze the metabolism of both PAHs and aryl amines to active metabolites (156). CYP1B1 is expressed in normal human lung tissue and is induced by cigarette

smoke. Willey *et al.* (257) reported that *CYP1B1* was expressed at higher levels in bronchial epithelial cells (BEC) of smokers compared with nonsmokers. However, large interindividual variation in the levels of cigarette smoke induced *CYP1B1* expression was observed which might result from hereditary variation in constitutive level of expression and/or inducibility, differing exposures to occupational or environmental pollutants or tobacco smoke, how soon the induction starts after smoking and how long the effect lasts in BEC. Not only was *CYP1B1* gene expression detected in human lung and induced by cigarette smoking, but also the expression of CYP1B1 mRNA was significantly higher in lung tumors than nontumors (263,264). These findings together with the involvement of CYP1B1 in activation of B[*a*]P and aryl amines found in cigarette smoke suggests that CYP1B1 expression in lung tissue contributes to the lung carcinogenesis process.

CYP1B1 gene expression induced by cigarette smoke in lung tissue is associated with the presence of PAHs. Regulation of *CYP1B1* gene expression is thought to proceed through a similar mechanism as *CYP1A1*. Analysis of the 5′ flanking region of the *CYP1B1* gene has revealed a number of similarities to the upstream region of the *CYP1A1* gene. The human *CYP1B1* gene has nine core XRE motifs within a 2.5 kilobase region 5′ of the transcription start site and at least three of these XRE seem to be functional in mediating AhR ligand-induced transcriptional of the *CYP1B1* gene (157).

1.7.3. CYP2E1

The *CYP2E1* gene product is involved in the metabolism of low-molecular-weight compounds found in cigarette smoke, such as 1,3-butadiene, and *N*-nitrosamines such as *N*-nitrosodimethylamine, *N*-nitrosodiethylamine and NNK. CYP2E1 is expressed in human extrahepatic tissue such as lung (265). CYP2E1 expression is induced by cigarette smoke. The

inductive effect may be due to the presence of PAHs, pyridine, acetone, and benzene in cigarette smoke. The mechanism of CYP2E1 induction is thought to be through protein and/or mRNA stabilization (266).

As a result of the presence of CYP2E1 in the human lung and the main involvement of CYP2E1 in α-hydroxylation of cigarette-specific *N*-nitrosamines, it is likely that CYP2E1 also contributes to the activation of pulmonary procarcinogens. Several studies have detected major DNA adducts derived from the NNK metabolic activation pathway which are 7-methylguanine, O^6-methylguanine, and pyridyloxobutyl adducts in human lung (245,267-269). The levels of bronchial 7-methylguanine-DNA adducts were statistically higher in smokers than those in nonsmokers (268,269). This was confirmed in the report of Foiles *et al.* (267) showing that higher tobacco-specific nitrosamine DNA adducts were observed in lung smokers than those in nonsmokers. The detection of methyl and pyridyloxobutyl adducts in DNA from smokers' lungs is consistent with the ability of human lung tissue to metabolically activate NNK, and the contribution of CYP2E1 in local bioactivation of cigarette derived procarcinogens in lung tissue.

1.8. mRNA quantitation

Variability in the expression of enzymes metabolizing carcinogens and procarcinogens derived from cigarette smoke may therefore contribute to individual susceptibility to lung carcinogenesis. This variation could be a result of enzyme induction, interindividual hereditary constitutive variation in expression, and environmental exposure. The induction of these enzymes may be a measure of environmental exposure to chemical carcinogens. For example, induced CYP1A1, CYP1B1, and CYP2E1 mRNA expression in lung tissue may represent exposure biomarkers of cigarette smoke, and environmental chemical carcinogens exposures. Furthermore, quantitation of gene expression including CYP1A1, CYP1B1, and CYP2E1 in

specific lung cells may be important in understanding the molecular mechanism of lung carcinogenesis.

An accurate human risk assessment requires sensitive methods to evaluate dose-response relationships especially following low levels occupational and environmental exposures. The quantitation of mRNA expression can be determined by several techniques including Northern blotting analysis, RNase protection assay, and quantitative reverse transcription (RT)-PCR. Northern blotting analysis, and quantitative RT-PCR followed by Southern blotting technique have been used to quantitate lung mRNA expression of CYP1A1, CYP1B1, and CYP2E1 in several studies (256,257,263). However, these techniques have some limitations. For example, Northern blotting has low sensitivity and requires large amounts of starting material to detect low level of mRNA. For example, one study used ^{32}P-labeled RNA target gene hybridized with target mRNA of samples after Northern blotting, the target mRNA expressed as a percent of the control mRNA. This determination is considered as semi-quantitative because it depends on estimation of the intensity of images of the bands by densitometry. Post-PCR product detection steps such as gel electrophoresis and Southern blotting are time consuming, tedious, and generally involve radioactive material, for example, ^{32}P, which raise safety and time concerns associated with ^{32}P-labelling RNA or PCR products.

1.8.1. FAF-ELOSA

In 1990, the new enzyme-linked immunosorbent assay (ELISA) format based technique called fluorescein-antifluorescein-based enzyme-linked oligonucleotide sorbent assay (FAF-ELOSA), was developed (270). It is an immunological based technique for the quantitative analysis of polymerase chain reaction (PCR) products. FAF-ELOSA is a variation of the sandwich-type ELISA. The PCR products are labeled with both biotin and fluorescein (Fig. 4).

The biotin label is used in place of primary antibody. The secondary antibody (FAF antibody) detects fluorescein-labeled PCR products. The fluorescein label is used for its antigenic rather than its fluorescent properties. Then, biotinylated DNA are immobilized on streptavidin-coated microtiter plates and can be quantitated by an ELISA specific for the antigenic group. The FAF antibody is linked to horseradish peroxidase, which allows quantification of PCR product by enzymatic production of a colored compound from trimethylbenzidine blue (TMB). The substrate development can be read at an absorbance of 450-630 nm. A standard curve is generated for each PCR reaction by using a serial dilution of a synthetized single-strand DNA oligonucleotide identical in sequence to the target DNA. Then, the concentration of the target DNA can be determined by comparing the absorbance obtained from sample with that obtained from a standard. The sensitivity of the immunological detection system that employs horseradish peroxidase linked to anti-fluorescein antibodies is high: 1 microliter of the PCR mixture obtained after approximately 25 cycles of amplification of 1 ng/microliter genomic template DNA is sufficient for the detection of human single-copy genes (270). The procedure does not require electrophoretic separation and/or hybridization with radioactive probes. This technique has overcome many of difficulties of those by classic quantitative PCR techniques such as the determination of relative differences in target mRNA amounts, safety, and time concerns.

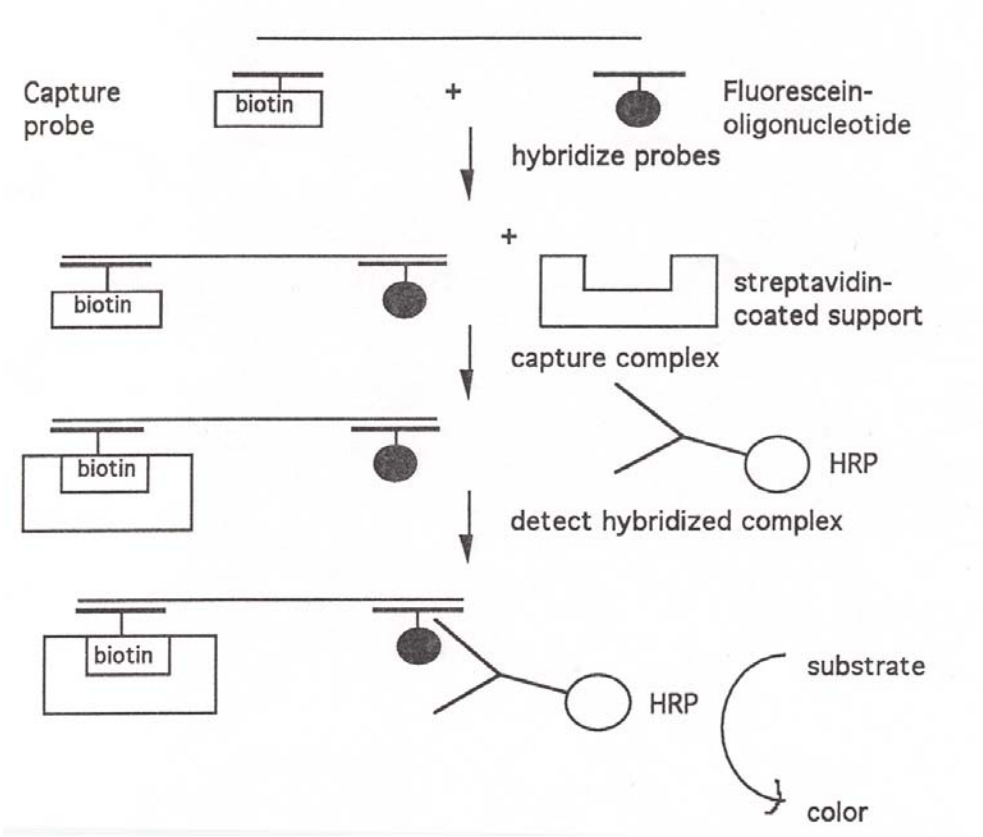


Figure 4. Schematic of the fluorescein-antifluorescein-based enzyme-linked oligonucleotide sorbent assay (FAF-ELOSA)

Carcillo *et al.* (271) was the first study that applied the FAF-ELOSA technique to quantitate CYP mRNA expression in human tissue. Also, the relative sensitivity of FAF-ELOSA versus that of the ^{32}P-labeled PCR/Southern blotting technique in quantitation of PCR products used to measure specific mRNA was compared in this study. Compared to the ^{32}P-labeled PCR/Southern blotting technique, FAF-ELOSA is a sensitive assay for quantitative of mRNA through PCR amplification of a known mRNA sequence. It can be performed in a relatively short time, is nonradiometric, and can be quantified to a known standard. Moreover, the strategy of labeling biotin and fluorescein to PCR products was modified in this study. A biotin-labeled PCR product was done by designing a biotin-labeled oligonucleotide termed capture probe and hybridized to the PCR product (Fig. 5). A fluorescein-labeled oligonucleotide designed to hybridize to the PCR product by incorporating as a primer in the complementary strand. The capture probe was designed within 20 basepairs (bp) of the 3′ fluorescein primer incorporated into the PCR product. This strategy provides two advantages. First, cost and time is minimized as no primer is required. Second, another level of specificity is obtained as any nonspecific PCR product that may have homology with two 20 bp that are about 200 bp apart is less likely to have homology with a third 20 bp that is only 20 bp from the 3′ fluorescein-labeled primer. Subsequently, the FAF-ELOSA is another technique that can be applied to the quantitation analysis of specific mRNAs of interest in massive samples for cancer risk assessment.

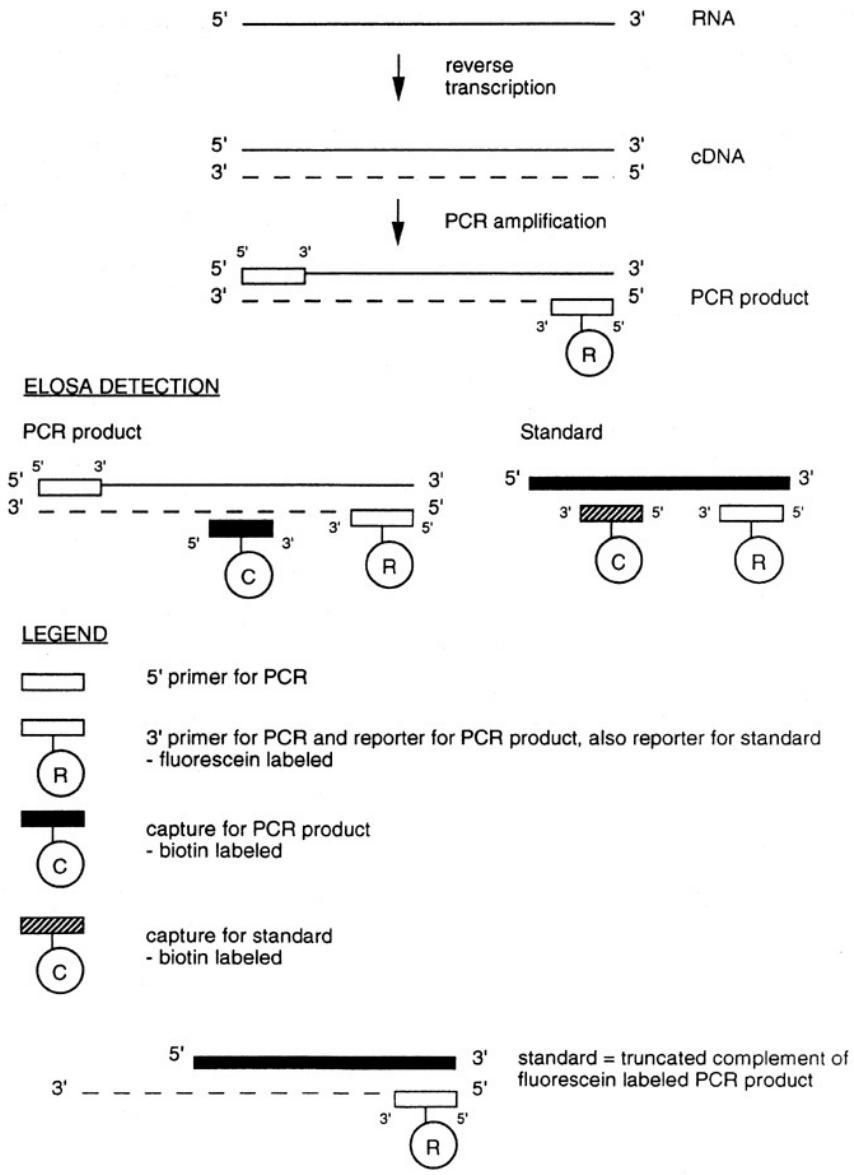


Figure 5. FAF-ELOSA using fluorescein primers and biotin labeled internal capture probes incorporated into both target and standard PCR products (271)

1.8.2. TaqMan® assay

Recently, the real-time quantitative RT-PCR technique (TaqMan® assay) has been developed as an approach to quantitate mRNA expression. Quantitation and sequence detection by TaqMan® technology and the ABI Prism® 7700 Sequence Detection system (Perkin-Elmer Applied Biosystems, Foster City, CA) has provided significant advances in terms of performance, rapid, sensitivity, accuracy, and high-through capability over other quantitative PCR methods (272-274). Real-time detection eliminates using post-PCR processing such as gel electrophoresis, since the detection occurs during each PCR cycle. Reactions are characterized by the point during cycling when amplification of the PCR product is first detected, rather than by the amount of PCR product accumulated after a fixed number of cycles. The 5′ nuclease assay provides a real-time method for detecting specific amplification products based on the use of a fluorogenic probe designed to hybridize within the target sequence and to generate a signal that accumulates during PCR cycling in a manner proportional to the concentration of amplification products (275). The probe is a nonextendable oligonucleotide with both a reporter fluorescent dye and a quencher dye attached. The reporter fluorescent dye FAM (6-carboxy-fluorescein) is covalently linked to the 5′ end. The quencher fluorescent dye TAMRA (6-carboxy-tetramethyl-rhodamide) is linked to the 3′ end. Because of the proximity of the quencher, the fluorescence emitted by the reporter dye is greatly reduced by Forster resonance energy transfer (FRET) through space (276). During PCR cycling, the probe specifically anneals downstream from one of the PCR primers of the corresponding template as shown in Fig. 6. Then, it is cleaved by the 5′ to 3′ exonuclease activity of Taq DNA polymerase as this primer is extended. This cleavage of the probe separates the reporter dye from quencher dye resulting in an increase of fluorescence emission of the FAM reporter without affecting the emission of quencher dye. Cleavage removes

the probe from the target strand, allowing primer extension to continue to the end of the template strand. This cleavage event occurs in each PCR cycle without interference in the enzymatic reaction since the exonuclease of Taq polymerase acts only if the fluorogenic probe is annealed to the target and the enzyme cannot hydrolyze the probe when it is free in solution. Thus, the increase of fluorescence intensity is proportional to the amount of produced PCR product.

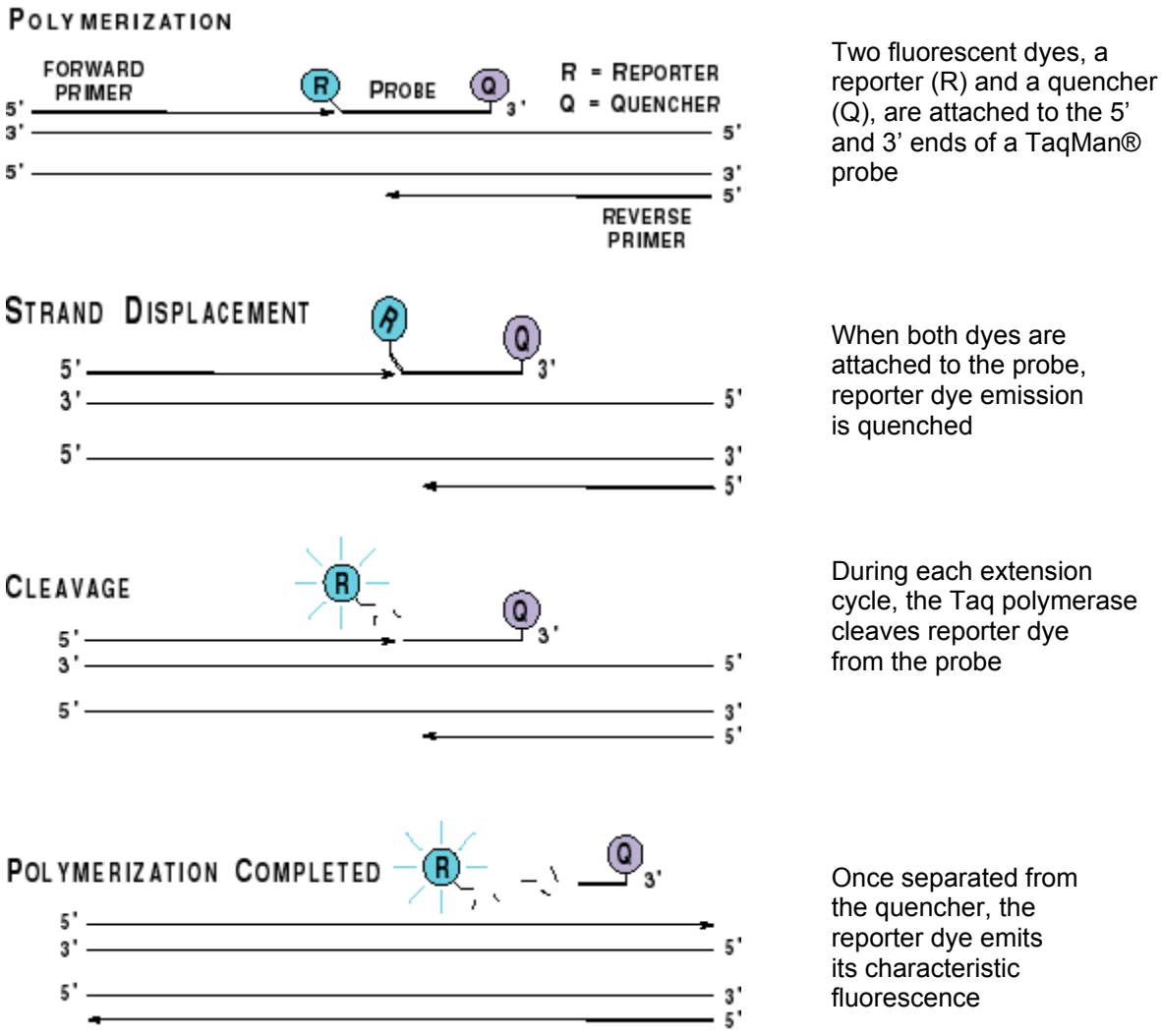
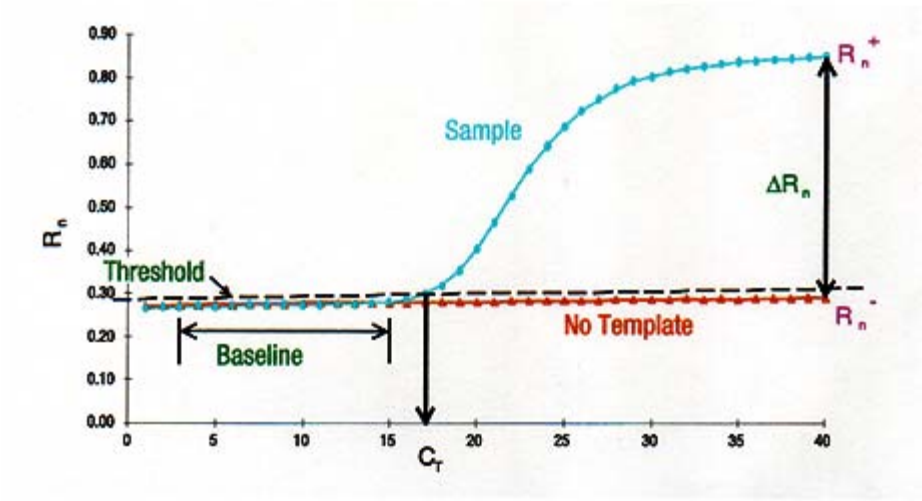


Figure 6. TaqMan® assay, polymerization associated 5' to 3' nuclease activity of Taq DNA polymerase acting on a fluorogenic probe during one extension phase of PCR (277)

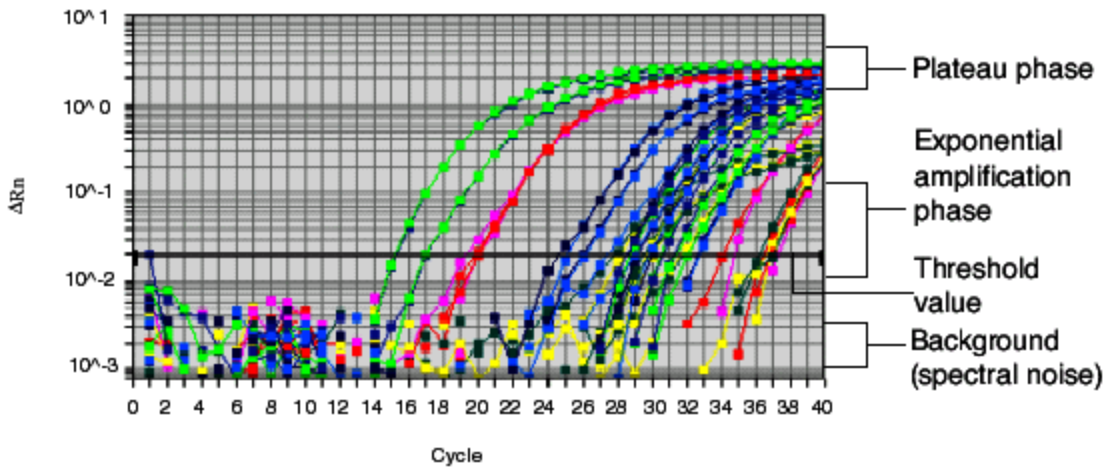
The fluorescence emitted from reporter dye is detected by the Sequence Detection System (Perkin-Elmer Applied Biosystems, Foster City, CA). This system includes a built-in 96-well thermal cycler, a laser directed by fiber optic cables to each of the 96 sample wells, a charged-coupled device (CCD) detector, and real-time sequence detection software. A laser is used to evoke photons from the fluorescein reporter molecules. The fluorescence emission travels back through the cables to a CCD camera detector. For each sample, the CCD camera collects the emission data between 520 and 660 nm once every few seconds. The software analyzes the data by first calculating the contribution of each component dye to the experimental spectrum. Each reporter signal is then divided by the fluorescence of an internal reference dye (ROX) in order to normalize for non-PCR related fluorescence fluctuations occurring well-to-well or over time. The emission intensity of the quencher dye that remains relatively constant during amplification is used as an internal control to normalize fluorescence emission and calculate the value term ΔR_n. ΔR_n is calculated from the equation $\Delta R_n = (R_n+) - (R_n-)$. (R_n+) is the emission intensity of the reporter divided by the emission intensity of the quencher during a specific amplification cycle. (R_n-) is the emission intensity of the reporter divided by the emission intensity of the quencher prior to amplification. Therefore, ΔR_n represents the amount of annealed probe cleaved by the 5′ to 3′ exonuclease activity of Taq DNA polymerase during amplification (273). An average ΔR_n for each cycle is calculated and is plotted versus cycle number that generates an amplification plot (Fig. 7).



Source: http://www.appliedbiosystems.com/

Figure 7. A single amplification plot of real-time quantitative PCR by TaqMan® assay

In the initial cycles of PCR, there is little change in fluorescence signal. This defines the baseline for the amplification plot. An increase in fluorescence above the baseline indicates the detection of accumulated PCR product. A fixed fluorescence threshold can be set above the baseline. The default threshold value is the average standard deviation of R_n within the defined baseline region, multiplied by an adjustable factor. The sequence detection software calculates the threshold value as ten standard deviations from the baseline. The parameter C_T (threshold cycle) is defined as the fractional cycle number at which the fluorescence passes the fixed threshold. The higher the starting copy number of the target, the sooner a significant increase in fluorescence, and a lower C_T value is observed (Fig. 8). A standard curve is plotted between a log of initial of target copy number for a set of standards versus C_T. Quantitation of the amount of target in unknown samples is accomplished by measuring C_T and using the standard curve to determine starting copy number.



Source: http://docs.appliedbiosystems.com/pebiodocs/04310255.pdf

Figure 8. Amplification plot showing terms used in TaqMan® assay and the plot shift to higher CT value that reveal the lesser in starting copy number of the target

The many advantages of the TaqMan® technology over other quantitative PCR methods has resulted in the acceptance of this approach for widespread measuring gene expression (274,278-281).

OBJECTIVES

This investigation involved a molecular epidemiological study and the analysis of metabolic enzyme genotype and CYP mRNA expression levels in order to investigate the relative contribution of genetic risk factors and environmental interactions to susceptibility of lung cancer. The following hypotheses were evaluated:

1. Genetic polymorphisms in drug metabolizing enzymes, specifically *CYP1A1, CYP1A2, CYP1B1, CYP2D6, CYP2E1, MPO, mEPHX, GSTM1, GSTM3, GSTP1*, *GSTT1,* and *NAT2**, which are involved in the metabolism of tobacco smoke component and in a DNA repair gene (*XPD*), are associated with risk of lung cancer.

2. An increasing risk of lung cancer is associated with gene-gene and gene-environment interactions, for example, a combination of drug metabolizing enzyme polymorphisms, history of environmental exposure, family history of lung or other cancers.

3. Lung tumor and normal tissues for mRNA expression of the CYPs that are involved in tobacco smoke carcinogen metabolism (CYP1A1, CYP1B1 and CYP2E1) represent smoking exposure biomarkers.

The specific aims of this study were to:

1) conduct a molecular epidemiological study to determine if genetic polymorphisms in the drug metabolizing enzymes *CYP1A1, CYP1A2, CYP1B1, CYP2D6, CYP2E1, MPO, mEPHX, GSTM1, GSTM3, GSTT1* and *NAT2**, and in the DNA repair gene (*XPD*) represent lung cancer susceptibility biomarkers using polymerase chain reaction (PCR) based techniques. Blood samples from lung cancer cases and healthy controls were genotyped.

2) evaluate the hypothesis that individuals with drug metabolizing enzyme polymorphisms combined with an environmental exposure(s) of family history of lung or other cancers and younger age-at-diagnosis are at greater risk for lung cancer development.

3) determine if CYP mRNA expression levels (CYP1A1, CYP1B1 and CYP2E1) in lung tumors and normal tissues are a smoking exposure biomarker.

4) compare the relative sensitivity of the RT-PCR FAF-ELOSA assay with the sensitivity of the TaqMan® assay for quantitation of CYP1A1, CYP1B1, CYP2E1 mRNA in lung tumors and normal tissues.

MATERIAL AND METHODS

3.1. Study Design

3.1.1. Type of study

This study was a population based case/control study. All subjects included in the study were recruited from either the Pittsburgh, Pennsylvania area or Olmsted County, Minnesota. Those subjects who were diagnosed with primary lung cancer were categorized as cases and those who were without any malignant disease were defined as controls.

3.1.2. Cases and controls selection

i). Case ascertainment in Pittsburgh area

Cases were ascertained from a retrospective case identification. The core tissue bank facility for the Early Detection Research Network: Patients at risk for head, neck and lung cancers was utilized for the identification of all patients diagnosed with primary lung cancer. Subjects were diagnosed during May 1992 through June 1994 at the University of Pittsburgh Cancer Institute. Potential study participants were provided with an IRB approved consent form (IRB#960116) to read, and the study protocol was explained to them. Any questions they had were answered, after which they were required to sign the consent form in the presence of a witness. A copy of the signed consent form was maintained in the subject hospital charts and a copy retained by the investigator. Subjects were interviewed about their health history and habits, and their family's

history of cancer. In addition, blood samples were collected, frozen and stored in the University of Pittsburgh Cancer Institute Tissue Bank for further experimental analyses.

ii). Case ascertainment in Olmsted County

In collaboration with Dr. Ping Yang at the Mayo Research Foundation, two methods of case identification were used for prospective case identification for patients diagnosed during June 1, 1997 through December 31, 1998, and retrospective case identification for patients diagnosed during January 1, 1995 through May 31, 1997. Criteria for study eligibility were cases diagnosed with primary lung cancer and with a positive family history of lung cancer (at least one first-degree or two second-degree relatives), a positive family history of other cancers (at least two first-degree or three second-degree relatives), diagnosed under the age of 50, or were lifetime non-smokers (less than 100 cigarettes/lifetime). Cases who had primary lung cancer but did not meet other criteria as described above were still recruited and categorized as controls for a case-only analysis.

Prospective case identification (Fig. 9)

To identify and track all patients with a new diagnosis of primary lung cancer, the study coordinator obtained a daily list of all lung cancer patients through the Department of Pathology and/or CoPath, a software package designed to generate reports of all pathologic diagnoses of specimens examined at the Mayo Clinic. Upon identification of a patient diagnosed with lung cancer, the study coordinator accessed the General Patient Access System (GPAS) to locate the patient's medical record. A board-certified genetic counselor reviewed the patient's medical record to determine if the patient met study eligibility criteria. When the information was deemed incomplete to determine study eligibility, the physician was notified and an alert flag was inserted into the medical record. This flag was an alert for the clinicians who were treating

or evaluating the patient to review the patient's history to determine if the patients met the study criteria, and then to notify the study coordinator.

If the patient met study criteria, or was systematically selected as a control (10% of the patients not meeting study inclusion criteria), the study coordinator then arranged a convenient time for the genetic counselor to review the Patient and Family History (PFH) form with the patient. The genetic counselor then assisted the patient in its completion and necessary expansion. As genetic counselors are trained to record medical, social, and family histories, complete and accurate information on the PFH form would be of benefit to the clinicians treating the patient. Additionally, in assessing the patient's family history of cancer, the genetic counselor expanded upon the family history by including information on second-degree relatives, types of cancer, ages of diagnosis, and relevant exposure (active and passive tobacco smoke) and occupation histories. If the patient was no longer being seen at Mayo for treatment, or the study coordinator was unable to arrange a convenient time to meet with the patient, the genetic counselor would call the patient to review and complete the PFH and expand the family history. During the visit, the genetic counselor also explained the research study, provided the patient with a brochure describing the study, and provided the patient with an IRB approved consent form (IRB#76-97 at Mayo Clinic, and IRB#960123 at Pittsburgh).

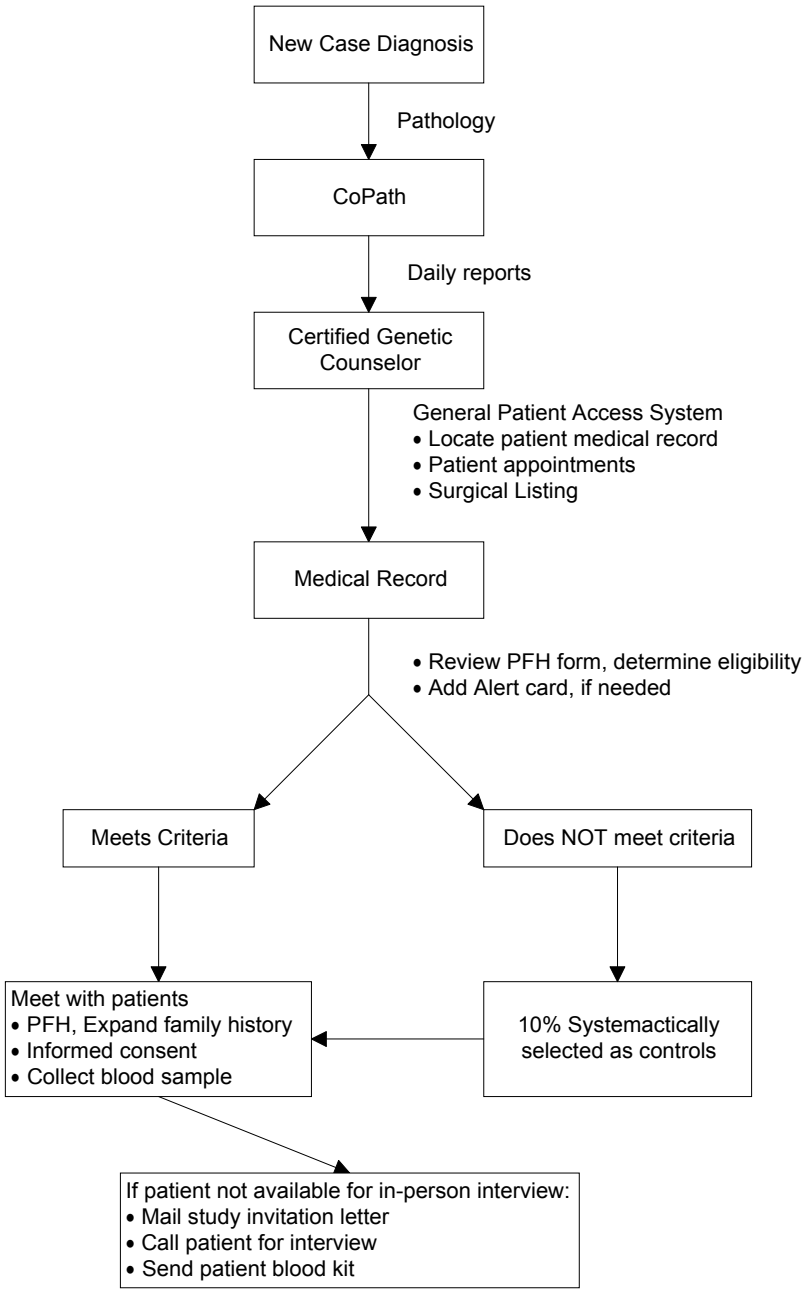


Figure 9. Prospective case identification at the Mayo Clinic

Retrospective case identification (Fig. 10)

Case sources utilized for identifying all patients diagnosed with lung cancer include Medical Diagnostic Index, Surgical Index, and Tumor Registry. Considering the very high case-fatality rate and unstable nature of the clinical patient population, the patients who were either Olmsted County residents at diagnosis or had a positive family history of lung cancer were ascertained.

Upon receipt of the patient's medical record, the study coordinator determined study eligibility status as described above. Patients who met study eligibility criteria were invited to participate in the study as follows:

a) If the patient was actively being seen at the Mayo Clinic, the genetic counselor arranged to meet with the patient at a scheduled appointment. During the visit, the genetic counselor reviewed and expanded the family history, described the study, and obtained consent (IRB#76-97) to participate.

b) If the patient was not actively being seen at the Mayo Clinic, the patient was sent an invitation letter and was called by the genetic counselor two weeks later. If the patient gave the written or verbal consent (IRB#76-97) to participate in the study, the family history was reviewed and the PFH form was completed and expanded the same way as for the newly diagnosed cases. Meanwhile, the patient was mailed a blood kit, instructions, and a written consent form. In the event that the patient was deceased, paraffin-embedded tissue blocks or other forms of preserved tissues were obtained from the Mayo Tumor Registry.

The study coordinator maintained a database that included clinic number, patient name, physician, date of lung cancer diagnosis, histological diagnosis, vital status, date of death, study eligibility status, date patient was visited or called by the genetic counselor, patient consent,

tumor/tissue sample request date, blood collection date, date blood kit sent, date blood kit returned, and laboratory number of all biosamples.

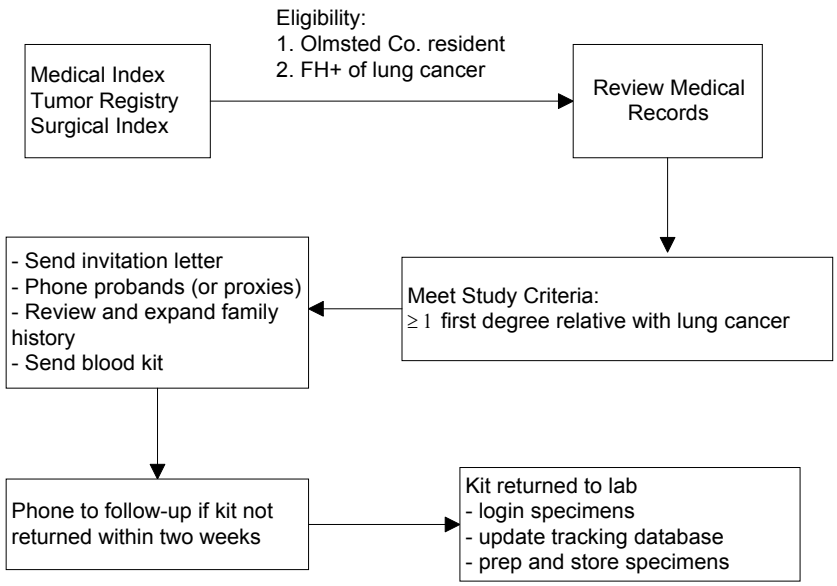


Figure 10. Retrospective case identification at the Mayo Clinic

iii). Control subject ascertainment at the Center for Clinical Pharmacology, University of Pittsburgh (Fig. 11)

Normal healthy controls were ascertained by prospective control identification. Healthy subjects were at least eighteen years, included both genders, smokers and non-smokers, and both Caucasians and African Americans. Subjects were recruited from the Center for Clinical Pharmacology's pool of healthy volunteers. Each subject passed a screening evaluation based on medical history, physical examination and the following biochemical and urinalysis tests: BUN, creatinine, electrolytes, liver function tests, total protein, albumin, complete blood count, urine pH, hemoglobin, protein, glucose, ketones, specific gravity and microscopic examination of sediment. Female subjects of childbearing potential were tested to exclude pregnancy. Subjects enrolled in this study were provided with an IRB approved consent form (IRB#990562) to read, and the study protocol was explained to them. Any questions they have were answered after which they were required to sign the consent form in the presence of a witness. A copy was maintained in the subject hospital charts and a copy retained by the investigator. After the initial evaluation, subject blood samples were collected for further experimental analyses.

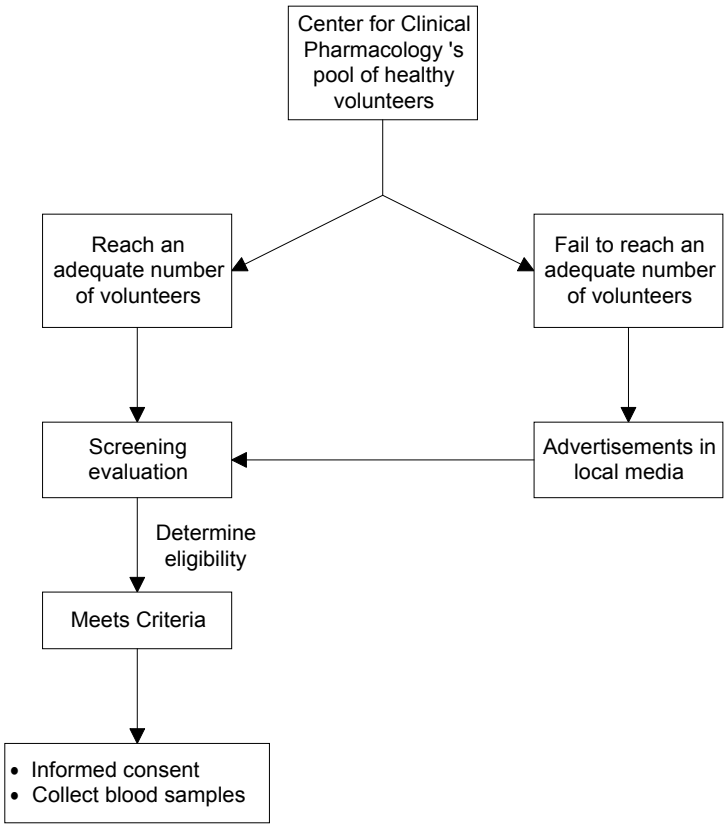


Figure 11. Prospective control identification at the Center for Clinical Pharmacology, University of Pittsburgh

iv). Control subject ascertainment at the University of Pittsburgh Medical Center

Controls were ascertained from a retrospective identification. The National Tissue and Serum Bank: Patients at Risk for Colorectal Carcinoma was utilized for the identification of patients who were screened for the development of colon cancer but resulted in normal colons and polyps. Subjects were diagnosed during May 1992 through June 1994 at the University of Pittsburgh Medical Center. All study participants were provided with an IRB approved consent form (IRB#0105103) to read and the study protocol was explained to them. Any questions they had were answered, after which they signed the consent form in the presence of a witness. A copy of the signed consent form was maintained in the subject hospital charts and a copy retained by the investigator. Subjects completed the questionnaires about their health, family's history of cancer, and risk factors for cancer. In addition, blood samples were collected, frozen and stored in the University of Pittsburgh Cancer Institute Tissue Bank for further experimental analyses.

3.2. Genotyping

3.2.1. DNA isolation

Genomic DNA was isolated from whole blood, buffy-coat, or peripheral blood lymphocyte samples from cases and controls according to standard protocols as described below.

i). Genomic DNA isolation from whole blood with the PureGene Kit (Gentra Systems)

Whole blood (150 μl) was mixed thoroughly with 450 μl of RBC lysis solution and incubated at room temperature for ten minutes. This solution was centrifuged at 15,000 x *g* for 30 seconds, then the supernatant was removed and discarded. The pellet was resuspended by vortexing the tube vigorously and mixed with 150 μl of cell lysis solution. The solution was incubated at 37^{o}C for 15 minutes with 0.75 μl of RNase A. This was followed by the addition of 50 μl of protein precipitation solution, the tube was inverted gently 10 times, and centrifuged at 15,000 x *g* for 5 minutes. The supernatant was transferred into a new tube, mixed with 150 μl of isopropanol, inverted about 50 times, and centrifuged at 15,000 x *g* for 5 minutes. The supernatant was discarded. The pellet was washed with 150 μl of 70% aqueous ethanol, and centrifuged at 15,000 x *g* for 2 minutes. The supernatant was removed and the pellet was dried in a vacuum centrifuge for 5 minutes. dH_2O (50 μl) was added to the pellet. DNA was rehydrated at 65^{o}C for 1 hour and at room temperature overnight.

ii). DNA isolation from the buffy coat fraction prepared from 3 ml whole blood with PureGene Kit (Gentra Systems)

The buffy coat fraction (150 μl) sample was mixed with 450 μl of RBC lysis solution and incubated for 10 minutes at room temperature. The tube was inverted once during the incubation. After centrifugation at 2,000 x *g* for 10 minutes, supernatant was removed, leaving behind the

white cell pellet in about 100-200 µl of the residual liquid. The tube was vortexed vigorously to resuspend the cells in the residual liquid. Cell lysis solution (300 µl) was added and mixed with buffy coat solution by pipetting up and down to lyse the cells. The cell lysate was mixed with 15 µl of RNase A solution by inverting the tube 25 times and incubated at 37 ^{o}C for 15-60 minutes. The sample was cooled to room temperature and mixed with 100 µl of protein precipitation solution. The mixture was vortexed vigorously at high speed for 20 seconds and centrifuged at 2,000 x *g* for 10 minutes. The supernatant was removed into a clean 1.5 ml centrifuge tube containing 300 µl of isopropanol and mixed by inverting tube gently 50 times. The DNA was centrifuged at 2,000 x *g* for 3 minutes and the supernatant was discarded. The DNA pellet was washed with 300 µl of 70% aqueous ethanol and centrifuged at 2,000 x *g* for 1 minute. The ethanol layer was poured off. The DNA was air dried for 10-15 minutes and then rehydrated with 250 µl of DNA hydration solution at 65^{o}C for 1 hour and overnight at room temperature.

iii). Isolation of DNA from 1-2 Million Cultured Cells (PureGene Kit Method – Gentra Systems)

One to two million cells were centrifuged at 14,000 x *g* for 5 seconds to pellet the cells. The supernatant was discarded, leaving behind 10-20 µl of residual liquild. The tube was vortexed vigorously to resuspend the cells. Cell lysis solution (300 µl) was added to the tube and the mixture was pipeted up and down to lyse the cells. The mixture was mixed with 1.5 µl of RNase A Solution and the tube was inverted 25 times and incubated at 37^{o}C for 5 minutes. The sample was placed on ice for 5 min, mixed with 100 µl of protein precipitation solution by vortexing tube vigorously at high speed for 20 seconds, then centrifuged at 14,000 x *g* for 1 minute. The supernatant was poured into a clean 1.5 ml centrifuge tube containing 300 µl of isopropanol. The tube was inverted gently 50 times and centrifuged at 14,000 x *g* for 1 minute. Aqueous ethanol

(70%, 300 μl) was added after the supernatant was discarded. The sample was centrifuged at 14,000 x *g* for 1 minute and the ethanol layer was poured off. The DNA was rehydrated with 50 μl of DNA hydration solution and mixed by vortexing at medium speed for 5 seconds. Then, DNA was incubated at 65^{o}C for 5 minutes and vortexed at medium speed for 5 seconds.

3.2.2. Genotyping assays

A panel of metabolic enzyme and DNA repair genetic polymorphisms was screened by using methods as indicated in Table 13 with some modifications to the published procedure. Each variant allele listed in Table 2-7, 9, 10 was selected for analysis based on the allele frequency, predicted enzyme phenotype and previously reported associations with lung cancer risk. Each PCR reaction was optimized by varying the volume used of 10X PCR buffer II (100mM Tris-HCl, pH 8.3, 500mM KCl), $MgCl_2$, dimethylsulfoxide (DMSO), dNTP mix, and gelatin. Also, the final volume of PCR reaction and the type of DNA polymerase used were important for obtaining the optimal condition for each PCR reaction. The assay validation was accomplished by sequencing each PCR product and comparison with published sequences. Negative control and positive controls with known genotypes including homozygous wild type, heterozygous, and homozygous variant allele were included in each PCR reaction.

Table 13. PCR assays used in amplification DNA fragments for detection genetic polymorphisms in variant alleles

Allele	Reference	Modification
*CYP1A1*2A* and *CYP1A1*3*	(282)	The PCR reaction was performed in a final volume 50 µl using 5 µl of DMSO, 5 µl of 10X PCR Buffer II, 2 µl of $MgCl_2$, 1 µl of 10 mM mix dNTP, 1 µl of each forward and reverse primer (400 ng/µl), 0.25 µl of gelatin, and 0.35 µl of AmpliTaq DNA polymerase at 94^{o}C 1 minute, 55^{o}C 1 minute, and 72^{o}C 1 minute for 40 cycles
*CYP1A1*2C* and *CYP1A1*4*	(132)	The PCR reaction was performed in a 100µl final volume using 10 µl of 10 X PCR Buffer II, 6 µl of $MgCl_2$, 2 µl of 10 mM mix dNTP, 1 µl of each forward and reverse primer (400 ng/µl), and 0.25 ul AmpliTaq Gold DNA polymerase at 94°C 30 second, 62°C 30 second, and 72°C 30 second for 40 cycles
*CYP1A2*1C*	(148)	The PCR reaction was performed in a final volume 50 µl using 5 µl of DMSO, 5 µl of 10 X PCR Buffer II, 3 µl of $MgCl_2$, 1 µl of 10 mM mix dNTP, 1 µl of each forward and reverse primer (400 ng/µl), 0.25 µl of gelatin, and 0.25 µl AmpliTaq Gold DNA polymerase at 94°C 1 minute, 56°C 1 minute, and 72°C 1 minute for 40 cycles
*CYP1A2*1F*	(150)	The PCR reaction was performed in a final volume 50 µl using 5 µl of DMSO, 5 µl of 10 X PCR Buffer II, 2 µl of $MgCl_2$, 1 µl of 10 mM mix dNTP, 1 µl of each forward and reverse primer (400 ng/µl), 0.25 µl of gelatin, and 0.25 µl AmpliTaq Gold DNA polymerase at 94°C 1 minute, 64°C 1 minute, and 72°C 1 minute for 40 cycles
*CYP1B1*3*	(159)	The PCR reaction was performed in a final volume 50 µl using 5 µl of DMSO, 5 µl of 10X PCR Buffer II, 2.5 µl of $MgCl_2$, 1 µl of 10 mM mix dNTP, 1 µl of each forward and reverse primer (400 ng/µl), 0.25 µl of gelatin, and 0.25 µl of AmpliTaq Gold DNA polymerase at 94°C 30 second, 52°C 30 second, and 72°C 30 second for 40 cycles

Table 13. (cont'd)

Allele	Reference	Modification
*CYP2D6*4*	(164)	The PCR reaction was performed in a final volume 50 µl using 5 µl of DMSO, 5 µl of 10X PCR Buffer II, 2 µl of $MgCl_2$, 1 µl of 10 mM mix dNTP, 1 µl of each forward and reverse primer (400 ng/µl), 0.25 µl of gelatin, and 0.25 µl of AmpliTaq Gold DNA polymerase at 94°C 30 second, 60°C 30 second, and 72°C 30 second for 40 cycles
*CYP2E1*5B*	(177)	The PCR reaction was performed in a final volume 50 µl using 5 µl of DMSO, 5 µl of 10X PCR Buffer II, 2 µl of $MgCl_2$, 1 µl of 10 mM mix dNTP, 1 µl of each forward and reverse primer (400 ng/µl), and 0.25 µl of AmpliTaq Gold DNA polymerase at 94°C 30 second, 60°C 30 second, and 72°C 30 second for 40 cycles
GSTM1	(283)	The PCR reaction was performed in a final volume 50 µl using 5 µl of DMSO, 5 µl of 10X PCR Buffer II, 3 µl of $MgCl_2$, 1 µl of 10 mM mix dNTP, 1 µl of each *GSTM1* and beta-globin primer, 0.25 µl of gelatin, and 0.25 µl of AmpliTaq Gold DNA polymerase at 94°C 30 second, 55°C 30 second, and 72°C 30 second for 40 cycles
*GSTM3*B*	(181)	The PCR reaction was performed in a final volume 50 µl using 5 µl of DMSO, 5 µl of 10X PCR Buffer II, 2.5 µl of $MgCl_2$, 1 µl of 10 mM mix dNTP, 1 µl of each forward and reverse primer (400 ng/µl), and 0.25 µl of AmpliTaq Gold DNA polymerase at 94°C 30 second, 58°C 30 second, and 72°C 30 second for 40 cycles
*GSTP1*B*	(284)	The PCR reaction was performed in a final volume 50 µl using 5 µl of DMSO, 5 µl of 10X PCR Buffer II, 2.5 µl of $MgCl_2$, 1 µl of 10 mM mix dNTP, 1 µl of each forward and reverse primer (400 ng/µl), 0.25 µl of gelatin, and 0.25 µl of AmpliTaq Gold DNA polymerase at 94°C 30 second, 55°C 30 second, and 72°C 30 second for 40 cycles

Table 13. (cont'd)

Allele	Reference	Modification
*GSTP1*C*	(284)	The PCR reaction was performed in a final volume 50 µl using 5 µl of DMSO, 5 µl of 10X PCR Buffer II, 2.5 µl of $MgCl_2$, 1 µl of 10 mM mix dNTP, 1 µl of each forward and reverse primer (400 ng/µl), 0.25 µl of gelatin, and 0.25 µl of AmpliTaq Gold DNA polymerase at 94°C 30 second, 55°C 30 second, and 72°C 30 second for 40 cycles
GSTT1	(108)	The PCR reaction was performed in a final volume 50 µl using 8.5 µl of Buffer B, and 1 µl of 10 mM mix dNTP, 1 µl of each *GSTT1* and beta-globin primer, and 0.35 µl of AmpliTaq DNA polymerase at 94^{o}C 1 minute, 55^{o}C 1 minute, and 72^{o}C 1 minute for 40 cycles
*NAT2**	(202)	The PCR reaction was performed in a final volume 100 µl using 20 µl of Buffer F, 1 µl of 10 mM mix dNTP, and 1 µl of forward and reverse primer (400 ug/ul, and 0.35 µl AmpliTaq DNA polymerase at 94^{o}C 1 minute, 57^{o}C 1 minute, and 72^{o}C 3 minutes for 40 cycles. Additional PCR products digested with *Ava*II for *NAT2*6A* allele
mEPHX3	(285)	The PCR reaction was performed in a final volume 50 µl using 2 µl of DMSO, 5 µl of 10X PCR Buffer II, 2 µl of $MgCl_2$, 1 µl of 10 mM mix dNTP, 1 µl of each forward and reverse primer (400 ng/µl), 0.25 µl of gelatin, and 0.25 µl of Ampli Taq Gold DNA polymerase at 94°C 10 second, 56°C 30 second, and 72°C 45 second for 40 cycles
mEPHX4	(226)	The PCR reaction was performed in a final volume 50 µl using 5 µl of DMSO, 5 µl of 10X PCR Buffer II, 2 µl of $MgCl_2$, 1 µl of 10 mM mix dNTP, 1 µl of each forward and reverse primer (400 ng/µl), 0.25 µl of gelatin, and 0.25 µl of AmpliTaq Gold DNA polymerase at 94°C 30 second, 61°C 45 second, and 72°C 1 minute for 40 cycles

Table 13. (cont'd)

Allele	Reference	Modification
MPO	(231)	The PCR reaction was performed in a final volume 50 µl using 5 µl of DMSO, 5 µl of 10X PCR Buffer II, 3 µl of $MgCl_2$, 1 µl of 10 mM mix dNTP, 1 µl of each forward and reverse primer (400 ng/µl), 0.25 µl of gelatin, and 0.25 µl of AmpliTaq Gold DNA polymerase at 94°C 30 second, 61°C 30 second, and 72°C 30 second for 40 cycles
XPD exon 23		The PCR reaction was performed in a final volume 50 µl using 5 µl of DMSO, 5 µl of 10X PCR Buffer II, 2 µl of $MgCl_2$, 1 µl of 10 mM mix dNTP, 1 µl of each forward and reverse primer (400 ng/µl), 0.25 µl of gelatin, and 0.25 µl of AmpliTaq Gold DNA polymerase at 94°C 1 minute, 54°C 1 minute, and 72°C 1 minute for 40 cycles

All PCR reactions were performed using DMSO from Sigma, 10X PCR Buffer II from Applied Biosystems, $MgCl_2$ from Applied Biosystems, 10 mM mix dNTP from Pharmacia, forward and reverse primers from Midland, gelatin from Sigma, AmpliTaq or AmpliTaq Gold DNA polymerase from Applied Biosystems, Buffer F and B from Invitrogen.

Table 14. Expected RFLP-PCR gel fragment sizes of each variant allele

Allele	Restriction endonuclease	Predicted pattern of gel bands
*CYP1A1*2A*	*Msp*I	*CYP1A1*1/*1* wide type - 1032 bp *CYP1A1*1/*2A* heterozygote - 1032, 826, 206 bp *CYP1A1*2A/*3* heterozygote -556, 260, 206 bp *CYP1A1*2A/*2A* homozygous variant - 826, 206 bp
*CYP1A1*3*	*Msp*I	*CYP1A1*1/*1* wide type - 1032 bp *CYP1A1*1/*3* heterozygote - 1032, 772, 260 bp *CYP1A1*2A/*3* heterozygote -556, 260, 206 bp *CYP1A1*3/*3* homozygous variant - 772, 260 bp
*CYP1A1*2C*	*BsrD*I	*CYP1A1*1/*1* wild type - 149, 55 bp *CYP1A1*1/*2C* heterozygote - 204, 149, 55 bp *CYP1A1*2C/*2C* homozygous variant - 204 bp
*CYP1A1*4*	*Bsa*I	*CYP1A1*1/*1* wild type - 139, 65 bp *CYP1A1*1/*4* heterozygote - 204, 139, 65 bp *CYP1A1*4/*4* homozygous variant - 204 bp
*CYP1A2*1C*	*Dde*I	*CYP1A2*1/*1* wild type - 596 bp *CYP1A2*1/*1C* heterozygote - 596, 464, 132 bp *CYP1A2*1C/*1C* homozygous variant - 464, 132 bp
*CYP1A2*1F*	*Bsp*120I	*CYP1A2*1/*1* wild type - 709, 211 bp *CYP1A2*1/*1F* heterozygote - 920, 709, 211 bp *CYP1A2*1F/*1F* homozygous variant - 920 bp
*CYP1B1*3*	*Eco*571	*CYP1B1*1/*1* wild type - 182, 89 bp *CYP1B1*1/*3* heterozygote - 271, 182, 89 bp *CYP1B1*3/*3* homozygous variant - 271 bp
*CYP2D6*4*	*Pal*I	*CYP2D6*1A/*1A* wild type - 83, 44, 36 bp *CYP2D6*1A/*4* heterozygote - 83, 44, 40, 39, 36 bp *CYP2D6*4/*4* homozygous variant - 44, 40, 39, 36 bp
*CYP2E1*5B*	*Pst*I	*CYP2E1*1A/*1A* wild type - 412 bp *CYP2E1*1A/*5B* heterozygote - 412, 290, 122 bp *CYP2E1*5B/*5B* homozygous variant - 290, 122 bp

Table 14. (cont'd)

Allele	Restriction endonuclease	Predicted pattern of gel bands
GSTM1	-	*GSTM1* wild type - equal intensity between 268 and 215 bp *GSTM1* heterozygote - half intensity of 215 bp compared to 268 bp *GSTM1* null - 268 bp
*GSTM3*B*	*Mnl*I	*GSTM3*A/*A* wild type - 125, 86, 51, 11 bp *GSTM3*A/*B* heterozygote - 134, 125, 86, 51, 11 bp *GSTM3*B/*B* homozygous variant - 134, 125, 11 bp
*GSTP1*B*	*BsmA*I	*GSTP1*A/*A* wild type - 176 bp *GSTP1*A/*B* heterozygote - 176, 93, 83 bp *GSTP1*B/*B* homozygous variant - 93, 83 bp
*GSTP1*C*	*Cac*81	*GSTP1*A/*A* wild type - 110, 34, 26 bp *GSTP1*A/*C* heterozygote - 136, 110, 34, 26 bp *GSTP1*C/*C* homozygous variant - 136, 34 bp
GSTT1	-	*GSTT1* wild type - equal intensity between 480 and 268 bp *GSTT1* heterozygote - half intensity of 480 bp compared to 268 bp *GSTT1* null - 268 bp
*NAT2*5A* or *NAT2*5B*	*Kpn*I	*NAT2*4/*4* wild type - 660, 433 bp *NAT2*4/*5A* or *NAT2*4/*5B* heterozygote - 1093, 660, 433 bp *NAT2*5A/*5A* or *NAT2*5B/*5B* homozygous variant - 1093 bp
*NAT2*6A*	*Taq*I & *Ava*II	*NAT2*4/*4* wild type - 380, 317, 170, 122, 104 bp *NAT2*4/*6A* heterozygote - 380, 317, 274, 170, 122, 104 bp *NAT2*6A/*6A* homozygous variant - 380, 317, 274, 122 bp
*NAT2*7B*	*BamH*I	*NAT2*4/*4* wild type - 811, 282 bp *NAT2*4/*7B* heterozygote - 1093, 811, 282 bp *NAT2*7B/*7B* homozygous variant - 1093 bp
*NAT2*14A*	*Alu*I & *Msp*I	*NAT2*4/*4* wild type - 759, 189, 91, 53 bp *NAT2*4/*14A* heterozygote - 759, 280, 189, 91, 53 bp *NAT2*14A/*14A* homozygous variant - 759, 280, 53 bp
mEPHX3	*Asp*I	*mEPHX3* (Tyr/Tyr) wild type - 231 bp *mEPHX3* (Tyr/His) heterozygote - 231, 209, 22 bp *mEPHX3* (His/His) homozygous variant - 209, 22 bp

Table 14. (cont'd)

Allele	Restriction endonuclease	Predicted pattern of gel bands
mEPHX4	*Rsa*I	*mEPHX4* (His/His) wild type - 295, 62 bp *mEPHX4* (His/Arg) heterozygote - 295, 174, 121, 62 bp *mEPHX4* (Arg/Arg) homozygous variant - 174, 121, 62 bp
MPO	*Aci*I	*MPO* (G/G) wild type - 169, 120, 61 bp *MPO* (G/A) heterozygote - 289, 169, 120, 61 bp *MPO* (A/A) homozygous variant - 289, 61 bp
XPD23	*Pst*I	*XPD23* (Lys/Lys) wild type - 575, 128 bp *XPD23* (Lys/Gln) heterozygote - 575, 128, 65, 63 bp *XPD23* (Gln/Gln) homozygous variant - 575, 65, 63 bp

The restriction endonuclease *Msp*I, *BsrD*I, *Bsa*I, *Dde*I, *Pst*I, *Mnl*I, *BsmA*I, *Cac*81, *Kpn*I, *Taq*I & *Ava*II, *BamH*I, *Alu*I, *Rsa*I, and *Aci*I were purchased from New England Biolabs, Inc.; *Bsp*120I, and *Eco*571 were purchased from MBI Fermentas, Inc.; *Pal*I and *Asp*I were purchased from Stratagene and Roche, respectively.

3.3. Quantitation of mRNA expression levels

Lung tumor and normal adjacent tissues were obtained from cases recruited at Mayo Clinic, Minnesota. Normal liver tissue was obtained from donor liver transplant tissue at the University of Pittsburgh Medical Center. The lung and liver tissue samples were kept at -80 ^{o}C until RNA was isolated. Total RNA isolation from snap frozen tissues was performed using the method described below.

3.3.1. Isolation of total RNA from tissue (Purescript Kit Method)(Gentra Systems)

Frozen tissue 250 mg was mixed with 300 μl cell lysis solution in a microgrinder on ice, and then transferred to a clean centrifuge tube. This mixture was then combined with 100 μl of protein-DNA precipitation solution by inverting the tube gently 10 times, and placed on ice for 10 minutes. After the mixture was centrifuged at 15,000 x *g* for 5 minutes, the supernatant was pipetted into a clean tube containing 300 μl of isopropanol and mixed by inverting tube about 50

times. The sample was centrifuged at 15,000 x *g* for 5 minutes and supernatant was poured off. The pellet was washed with 300 µl of 70% aqueous ethanol and centrifuged at 15,000 x *g* for 2 minutes. The supernatant was discarded and the pellet was dried in vacuum centrifuge for 5 minutes. Diethyl pyrocarbonate treated deionized H_2O (25 µl)(DEPC dH_2O) was added to the RNA and 5 µl was taken to 995 µl of DEPC dH_2O. Optical density at 260 nm and 280 nm was read on UV-2101PC, UV-VIS scanning spectrophotometer (Shimadzu) to determine the total RNA concentrations.

Detection of Contamination of total RNA with genomic DNA

The *CYP2D6* genotyping PCR reaction was performed using 1 µg of RNA in order to detect genomic DNA contamination in the total RNA preparation. The PCR reaction was performed using the method described in Table 13. The PCR product was eletrophoresed through a 8 % acrylamide gel. Then, the gel was stained with ethidium bromide (Sigma) for 20 minutes, exposed to UV light, and visualized with the FOTODYNE® gel photograph system. Expected results were no bands for RNA samples, no band with blank, and a band in the lane of the DNA positive control. If a band existed in the RNA sample lane, the RNA was contaminated with genomic DNA and was treated as follows.

DNase treatment for contaminated RNA

RNA was incubated with 25 µl of DNase buffer (80mM Tris-HCl buffer, 16mM $MgCl_2$, pH 7.6) and 1 µl of DNase/10 µg nucleic acids (Stratagene) at 37°C for 1 hour. The solution was brought up to 250 µl total volume with DEPC dH_2O (Sigma). Tri-reagent (Molecular Research Center) (750 µl) was added and the RNA isolation protocol was then repeated. Absence of genomic DNA was confirmed as described above.

3.4. Methods for RNA quantitation

One of two methods was used for the quantitation of mRNA expression.

3.4.1. FAF-ELOSA

3.4.1.1. Reverse Transcription (RT) of Total RNA to cDNA

Total RNA was reverse transcribed and the cDNA product amplified. Briefly, 10 μg of total RNA was mixed with 0.5 μl of RNasin RNase inhibitor (40U/μl) (Promega) and 1 μl of random hexamers (2.5 μM) in a total volume of 9.5 μl. The mixture was heated to 94^{o}C for 2 minutes and cooled on ice for 5 minutes. To this mixture, 0.5 μl of RNasin RNase inhibitor (40U/μl)(Promega), 4 μl of 5X Moloney murine leukemia virus (MMLV) buffer (20mM Tris-HCl (pH 7.5), 100 mM NaCl, 0.1mM EDTA, 1mM dithiothreitol, 0.01% (v/v) nonidet P40, 50% (v/v) glycerol) (Gibco BRL), 2 μl of 0.1 M dithiothreitol (Gibco BRL), and 2 μl of dNTP mix (10 mM)(Pharmacia) was added and incubated at 41^{o}C for 15 minutes, and 2 μl of Moloney-murine leukemia virus reverse transcriptase (Gibco BRL) was added and incubated for an additional 60 minutes at 41^{o}C. The mixture was heated to 99^{o}C for 5 minutes to inactivate the RT reaction and cooled to 4^{o}C.

3.4.1.2. PCR reaction for amplification of cDNA

The PCR reaction for amplification of human β-actin cDNA was performed by the method as described by Carcillo *et al.* (271). In a final volume of 100 μl, 10 μl of DMSO (Sigma), 10 μl of 10X PCR Buffer II (Perkin Elmer), 4 μl of $MgCl_2$ (Perkin Elmer), 2 μl of 10 mM mix dNTP (Pharmacia), 1 μl of fluorescein-labeled 3′ primer and unlabeled 5′ primer (400 μg/μl) (Midland), and 0.5 μl gelatin (Sigma) were added to 2 μl of cDNA (0.5 μg/μl). Hot Start PCR was run at 94^{o}C 5 minutes, 86^{o}C 1.5 minutes, and 94^{o}C 1.5 minutes and 0.5 μl AmpliTaq DNA

polymerase (Perkin Elmer) was added during the 86°C step. Then, the reaction was continued at 94°C for 1 minute, 51°C for 1 minute, and 72°C for 3 minutes for 40 cycles followed by a 10 minutes at 72°C extension and a 4°C soak. The CYP1A1, CYP1B1, and CYP2E1 cDNA were amplified under similar conditions, except the annealing step was performed at 59°C for CYP1A1 and CYP2E1. The PCR product was electrophoresed through a 8 % acrylamide gel. Then, the gel was stained with ethidium bromide (Sigma) for 20 minutes, exposed to UV light, and visualized with the FOTODYNE® system. The expected band sizes of CYP1A1, CYP1B1, CYP2E1, and β-actin PCR products were 202, 165, 122, and 139 bp, rescpectively.

3.4.1.3. Quantitative analysis of the fluorescein-labeled PCR product using FAF-ELOSA

Four PCR reactions from a single cDNA preparation were performed. Aliquots (10 μl) were taken at 20, 22, 24, 26, and 28 cycles to ensure quantitation during the exponential phase of the PCR. Quantification of PCR products was carried out using the procedure of Carcillo *et al.* (271). Briefly, 5 μl of the aliquots were diluted with deionized water to a total volume of 20 μl and heat denatured at 100°C for 5 minutes and then cooled to 4°C. The standard solutions (20 μl) or denatured sample (20 μl) was pipetted into the streptavidin-coated microtiter plate wells (DuPont). Hybridization buffer (100 μl of 10% formamide, 6x SSC, 1.2% Triton X-100, 0.12% BSA) containing 10 nM bitotin-labeled captured probe and 5 nM fluorescein-labeled reporter probe for the standard only was added to each well and incubated at 37°C for 1 hour. The plate was washed 6 times with 1x DuPont plate wash buffer (phosphate buffered saline, 10x with 0.5% Tween 20/phosphat-gepufferte)(DuPont). Anti-fluorescein horseradish peroxidase (HRP) conjugate (100 μl, 1:200)(DuPont) in antifluorescein diluent buffer (PBS, 2% Tween, 5% FCS, 0.2% casein) was added to the wells and incubated at room temperature in dark for 30 minutes. The wells were washed six more times. Tetramethylbenzidene blue (TMB) substrate (100 μl)

(ScyTek) was added and incubated in the dark for 1 hour at room temperature. The reaction was stopped with TMB stop solution (ScyTek). Product color development was read on a MicroQuant plate reader (Bio Tek) as the difference in absorbances at 450 nm and 630 nm. A standard curve was generated for each PCR reaction by using a serial dilution of a synthesized single-stranded DNA oligonucleotide identical in sequence to the target DNA. The concentration of the synthetic DNA was determined spectrophotometrically. The concentration of target DNA in sample was determined by comparing the different values in absorbances at 450 nm and 630 nm read on a MicroQuant plate reader (Bio Tek) obtained from a sample with that obtained from a standard.

3.4.2. TaqMan® assay

3.4.2.1. Reverse Transcription of Total RNA to cDNA

The RT was carried out in a 100 μl volume consisting of 10 μl of 10X PCR buffer II (Perkin Elmer), 30 μl of $MgCl_2$ (25mM)(Perkin Elmer), 4 μl of dNTP mix (25mM)(Pharmacia), 5 μl of random hexamers (100μM), 1 μl of RNase Inhibitor (40U/μl)(Promega), 1.25 μl of M-MLV reverse transcriptase (200U/μl)(Gibco BRL), and 45 to 180ng total RNA. Reactions were incubated at 25^{o}C for 10 minutes, 48^{o}C for 30 minutes, and 95^{o}C for 5 minutes. "No RT" controls were carried out in all cases using the same RT reaction mix but substituting DEPC dH_2O for MMLV reverse transcriptase.

3.4.2.2. Real time quantitative analysis of target PCR product

The following gene expression patterns were studied: β-actin, β-glucuronidase (β-GUS), CYP1A1, CYP1B1, and CYP2E1. The PCR primers and fluorogenic probes were designed using Primer Express software (PE Biosystems). TaqMan primers were designed to span introns in the

genomic DNA in order to avoid signal from contaminating genomic DNA. The β-GUS, CYP1A1, CYP1B1, and CYP2E1 PCR reactions were performed in duplicate in a final 25 μl reaction volume using 2.5 μl of 10x TaqMan buffer A (Perkin Elmer), 5.5 μl of $MgCl_2$ (25mM)(Perkin Elmer), 3 μl of dNTP mix (25mM)(Pharmacia), 0.25 μl of AmpliTaq Gold DNA polymerase (Perkin Elmer), 0.25 μl of primers and probe (10μM)(Midland), and 5 μl of cDNA. Two-step PCR cycling was carried out at 95°C for 12 minutes, and followed by 95°C for 20 seconds and 60°C for 1 minute for 40 cycles.

The β-actin PCR reaction was performed in duplicate in a final 50 μl reaction volume consisting 1x TaqMan PCR buffer A (Perkin Elmer), 3.5 mM $MgCl_2$ (25mM)(Perkin Elmer), 200 μM each dNTP (Perkin Elmer), 300 nM β-actin primers (Perkin Elmer), 200 nM β-actin probe (Perkin Elmer), 0.025 U/μl AmpliTaq Gold DNA polymerase (Perkin Elmer), 0.01 U/μl AmpErase UNG (Perkin Elmer), and 5 μl of cDNA. Two-step PCR cycling was carried out at 95°C for 10 minutes, and followed by 95°C for 15 seconds and 60°C for 1 minute for 40 cycles. At the end of the PCR, baseline and threshold values were set in the ABI 7700 Prism software and the amplification plot and calculated C_T values were determined.

3.4.2.3. Relative expression calculation

Relative mRNA expression was calculated using the comparative method (PE Biosystems). All data were controlled for quantity of RNA input by performing measurements on an endogenous reference gene β-GUS. In addition, results on RNA from lung tissues were normalized to results obtained on RNA from liver tissues. Briefly, the analysis was performed by calculating the difference in C_T values (ΔC_T) for each RNA sample by;

$\Delta C_T = \Delta C_T$ (target gene) – ΔC_T (β-GUS)

Then $\Delta\Delta C_T$ was generated by:

$\Delta\Delta C_T = \Delta C_T$ (lung RNA) – ΔC_T (liver RNA)

Since all RNAs were reverse transcribed at three different RNA concentrations, three values of $\Delta\Delta C_T$ were calculated for each gene on each RNA sample: $\Delta\Delta C_T$ (180ng), $\Delta\Delta C_T$ (90ng) and $\Delta\Delta C_T$ (45ng). The mean of these $\Delta\Delta C_T$ measurements was used to calculate the amount of target, normalized to an endogenous reference and relative to liver tissue sample.

Relative expression = $2^{-\Delta\Delta CT}$

3.5. Statistical analysis of results

The genotyping results and complete information on age, gender, smoking status, packyear, and occupational exposures obtained from both case and control populations were recorded in a Microsoft Excel file and then imported and analyzed with the SAS statistical computer program (Version 6: SAS Institute Inc., Cary, NC). The statistical analysis was restricted to Caucasians in order to minimize the confounding effect of allele frequency variation by ethnicity. The log likelihood ratio test was used to test the linearity of continuous variables such as age and packyear. If the test result showed non-linearity of any continuous variables, the transformed covariates were created and further tested for linearity. The log likelihood tests showed a linear association between lung cancer risk and age but not for the packyear variable. Then, the transformed packyear value of packyear+packyear-square was used instead of packyear as suggested by an examination of the linear relationship between the transformed packyear and the log odds of lung cancer risk. Unconditional multivariate logistic regression analysis was performed to compare the relative contribution of risk by different factors, for example, single gene loci polymorphisms and combined at risk genotypes, and control for confounding factors and variable interactions. The model used a stepwise variable selection to obtain ORs and 95% confidence intervals allowing for inclusion of both continuous and categorical independent

variables and interaction terms among confounding factors such as age, race and smoking status and between confounding factors and at risk genotype(s). Odds ratios and their 95% confidence interval for risk of lung cancer in the entire lung cancer cases, in different histological subtypes (adenocarcinoma and squamous cell carcinoma), males, females, and cases with no history of occupational subgroups by single or combined at risk genotypes, were calculated from these models. The interaction terms between at risk genotype(s) and smoking status and/or transformed packyear and/or occupational exposure were fit into logistic regression models to examine the relationship between log odds of lung cancer and the gene-environment interactions.

A case-only analysis was used to determine gene-gene and gene-environment interactions for a combination of drug metabolizing enzyme and DNA repair gene polymorphisms, history of environmental exposure, family history of lung or other cancers and age-at-diagnosis on increasing risk of lung cancer. The contribution of both metabolic enzyme and DNA repair gene polymorphisms on greater risk of lung cancer development among at risk lung cancer cases was assessed by comparing the differences in numbers of at risk lung cancer cases (cases who had family history of lung and other cancers and early age onset) who had one at risk genotype or combined genotypes compared to sporadic cases (cases who did not have family history of lung and other cancers and late age at diagnosis) using the χ^2-test (Minitab Release 11.1, Minitab Inc.). The relative risk was also estimated using SAS software by using logistic regression models to obtain odd ratios and their 95% confidence intervals.

Analysis tools of the student t-test for two sample assuming unequal variances and paired two sample for means were used to determine the differences in mean levels of CYPs mRNA expression between lung tumors and histologically normal adjacent tissues.

RESULTS

4.1. Case control study

4.1.1. Characteristics of study population

There were 203 lung cancer cases and 205 controls in the present study. All controls and 61 lung cancer cases were recruited from the Pittsburgh, Pennsylvania area. One hundred forty two lung cancer cases were from Olmsted County, Minnesota. The distribution of age, ethnicity, smoking status, packyear and history of occupational exposures was not statistically different between Pittsburgh and Minnesota lung cancer case groups. Therefore, lung cancer cases from both areas were pooled to give the study population of lung cancer cases in the case control and case only analyses. The demographic characteristics of the study populations are given in Table 15. The percentage of males and ever smokers were higher among cases compared to controls. Ninety-six percent of cases and seventy-seven percent of controls were Caucasians. The remainder were African-American or of other ethnicity. The frequency distribution of occupational exposure among cases and controls was similar. Although the mean age and packyear of cases and controls were different, both age and packyear variables were adjusted in the logistic regression models when the relative risk values were calculated for predicting the association of risk factors and disease outcome. Cases were categorized into five subgroups based on histopathological cell types, gender and history of occupational exposure. Characteristics of cases in each subgroup are shown in Table 16.

Table 15. Distribution of characteristics of the controls and lung cancer cases

Variable	Number of individuals (%)	
	Controls (N=205)	Cases (N=203)
Gender		
Male	108 (52.7)	130 (64.4)
Female	97 (47.3)	72 (35.6)
Ethnicity		
Caucasians	157 (77)	196 (96.5)
African Americans	42 (20.5)	4 (2)
Others	5 (2.5)	3 (1.5)
Smoking Status[a]		
Ever smokers	107 (52.5)	176 (91.7)
Nonsmokers	97 (47.5)	16 (8.3)
Age (mean ± SD)	50 ± 16	64 ± 12.1
Packyear		
0 < PY ≤ 30	63 (60.6)	47 (31.1)
30 < PY ≤ 60	28 (26.9)	59 (39.1)
60 < PY	13 (12.5)	45 (29.8)
mean ± SD	31 ± 30	51.8 ± 34
Occupational Exposures		
Asbestos exposure	9 (4.6)	17 (8.5)
Others[b]	45 (22.8)	45 (22.5)
None	143 (72.6)	138 (69)
Histological cell types[c]		
Adenocarcinoma		110 (58.5)
Squamous cell carcinoma		62 (33.0)
Others[d]		16 (8.50)

[a] Ever smokers (ex-/current smokers who smoke ≥ 100 cigarettes lifetime); nonsmokers (< 100 cigarettes lifetime)
[b] Includes silicon, petroleum, radon, nickel chromium and others
[c] Includes small call carcinoma, large cell carcinoma, and other
[d] Missing data on 15 cases

Table 16. Characteristic of lung cancer cases among subgroups

Variable	Number of individuals (%)				
	Adenocarcinoma (N = 108[a])	Squamous cell carcinoma (N = 54[b])	Male (N = 130[c])	Female (N = 72[d])	Non occupational exposure (N = 141[e])
Gender					
Male	64 (59.2)	39 (72)	130 (100)	0	74 (52.9)
Female	44 (40.8)	15 (28)	0	72 (100)	66 (47.1)
Ethnicity					
Caucasians	108 (100)	54 (100)	126 (97.7)	70 (98.6)	135 (97.8)
African Americans	0	0	3 (2.30)	1 (1.40)	3 (2.20)
Others	0	0	0	0	0
Smoking Status					
Ever smokers	90 (86.5)	54 (100)	117 (94.4)	59 (86.8)	118 (90.8)
Nonsmokers	14 (13.5)	0	7 (5.60)	9 (13.2)	12 (9.20)
Age (mean ± SD)	65 ± 12	65 ± 11	63 ± 13	66 ± 11	65 ± 12
Packyear					
0 < PY ≤ 30	34 (38.6)	14 (28.0)	35 (33.0)	25 (43.1)	44 (38.6)
30 < PY ≤ 60	29 (33.0)	20 (40.0)	40 (37.7)	19 (32.8)	41 (36.0)
60 < PY	25 (28.4)	15 (32.0)	31 (29.3)	14 (24.1)	29 (25.4)
mean ± SD	51 ± 29	53 ± 30	51 ± 30	53 ± 41	53 ± 37
Occupational Exposures					
Asbestos	9 (8.30)	4 (7.40)	17 (13.2)	0	0
Others[f]	25 (23.6)	10 (18.5)	39 (30.2)	6 (8.60)	0
None	72 (31.9)	40 (74.1)	73 (56.6)	64 (91.4)	0

[a] Data shown only Caucasian cases; missing data 4, 20, and 2 cases on smoking, packyear, and occupational exposures, respectively
[b] Data shown only Caucasian cases; missing data 5 cases on packyear
[c] Missing data 1, 6, 24, and 1 cases on ethnicity, smoking, packyear, and occupational exposures, respectively
[d] Missing data 1, 4, 14, and 2 cases on ethnicity, smoking, packyear, and occupational exposures, respectively
[e] Missing data 1, 3, 11, and 27 cases on gender, ethnicity, smoking, and packyear, respectively
[f] Includes silicon, petroleum, radon, nickel chromium and others

4.1.2. Single gene loci polymorphisms and lung cancer risk

Analyses of each 'at risk' genotype association with lung cancer risk were evaluated by a priori hypotheses based on the relevant function of each single gene to cigarette carcinogen metabolism and the impact of the single nucleotide polymorphism on predicted enzyme phenotype. Logistic regression models were fit by using a step-wise procedure to examine the relationship between the log odds of lung cancer and each 'at risk' genotype after adjusting for known risk factors such as age, race, smoking status, and interaction term between these factors where appropriate. Then, the odds ratio and their 95%CIs for the risk of lung cancer were calculated from these models.

The frequencies of 18 genetic polymorphisms screened were similar between cases and controls as shown in Table 17. However, significantly higher genotype frequencies of predicted high CYP1A2 activity, *GSTM3*A/*A*, and *XPD* exon 23 (Lys/Gln and Gln/Gln) genotypes among cases were observed with crude ORs of 3.5 (95% CI, 2.22-5.52), 1.93 (95% CI, 1.27-2.93), and 1.80 (95% CI, 1.19-2.72), respectively. After adjustment for age, race, smoking status and/or packyear and/or occupational exposures in logistic regression models, the adjusted ORs were 2.05 (95% CI, 1.13-3.73), 1.84 (95% CI, 1.03-3.31), and 2.56 (95% CI, 1.45-4.51) (Table 18-20), respectively.

Table 17. Genotype frequencies among controls and lung cancer cases

Genotype	Controls N (%)	Cases N (%)
CYP1A1		
**1A/*1A*	106 (58.6)	144 (75.4)
**1A/*2A*	43 (23.8)	17 (8.9)
**1A/*2C*	6 (3.3)	4 (2.1)
**1A/*4*	6 (3.3)	11 (5.8)
**2A/*2A*	3 (1.7)	5 (2.6)
**2A/*2C*	12 (6.6)	8 (4.2)
**2A/*4*	5 (2.7)	2 (1)
CYP1A2		
**1A/*1A*	15 (7.8)	18 (9.4)
**1A/*1F*	75 (38.9)	66 (34.6)
**1C/*1C*	1 (0.5)	0 (0)
**1C/*1F*	22 (11.4)	4 (2.1)
**1F/*1F*	80 (41.4)	103 (53.9)
Predicted 1A2 activity		
Low	141 (78.3)	97 (50.8)
High	39 (21.7)	94 (49.2)
CYP1B1		
**1/*1*	56 (27.8)	60 (31.3)
**1/*3*	90 (44.8)	84 (43.7)
**3/*3*	55 (27.4)	48 (25)
CYP2D6		
**1A/*1A*	56 (61.5)	78 (39.2)
**1A/*4B*	33 (36.3)	115 (57.8)
**4B/*4B*	2 (2.2)	6 (3)
CYP2E1		
**1A/*1A*	179 (90)	184 (92.9)
**1A/*5B*	20 (10)	14 (7.1)
**5B/*5B*	0 (0)	0 (0)
GSTM1		
Positive	88 (42)	86 (42.4)
Null	119 (58)	117 (57.6)
GSTM3		
**A/*A*	106 (54.1)	132 (69.5)
**A/*B*	68 (34.7)	51 (26.8)
**B/*B*	22 (11.2)	7 (3.7)
GSTT1		
Positive	133 (65.8)	157 (78.1)
Null	69 (34.2)	44 (21.9)
GSTP1		
**A/*A*	94 (48.5)	90 (46.6)
**A/*B*	66 (34.0)	70 (36.4)
**A/*C*	0 (0)	1 (0.5)
**B/*B*	12 (6.2)	7 (3.6)
**B/*C*	22 (11.3)	24 (12.4)
**C/*C*	0 (0)	1 (0.5)

Table 17. (cont'd)

Genotype	Controls N (%)	Cases N (%)
*NAT2**		
Fast acetylator	39 (43.3)	91 (45.3)
Slow acetylator	51 (56.7)	110 (54.7)
mEPHX		
Tyr113/Tyr113	89 (52)	91 (48.9)
Tyr113/His113	54 (31.6)	61 (32.8)
His113/His113	28 (16.4)	34 (18.3)
mEPHX		
His139/His139	105 (60.7)	118 (63.4)
His139/Arg139	55 (31.8)	56 (30.1)
Arg139/Arg139	13 (7.5)	12 (6.5)
Predicted mEPHX activity		
Very low	19 (11.3)	21 (11.3)
Low	46 (27.4)	58 (31.2)
Intermediate	60 (35.7)	66 (35.5)
High	43 (25.6)	41 (22)
MPO		
G463/G463	100 (50.3)	113 (58.9)
G463/A463	82 (41.2)	67 (34.8)
A463/A463	17 (8.5)	12 (6.3)
XPD		
Lys571/Lys571	97 (54.8)	77 (40.3)
Lys571/Gln571	59 (33.3)	87 (45.6)
Gln571/Gln571	21 (11.9)	27 (14.1)

When lung cancer cases were stratified into subgroups according to gender, histological cell types, and history of occupational exposures, a significant association between predicted mEPHX high and intermediate activity or *GSTP1* (**A/*B* and **B/*B*) genotypes and lung cancer risk were observed among females and males, with adjusted ORs 2.77 (95% CI, 1.02-7.51), and 2.09 (95% CI, 1.03-4.25) (Table 21, 22), respectively. Not only was there a relationship between lung cancer risk in all cases by *GSTM3*A/*A* genotype, but also among females with adjusted OR 3.03 (95% CI, 1.09-8.38) (Table 19). Furthermore, relative risk of lung cancer in all cases and in each subgroup for individuals who had reduced DNA repair capacity predicted by *XPD* exon 23 (Lys/Gln and Gln/Gln) genotypes was 2- to 3-fold higher than individuals who had *XPD* exon 23 (Lys/Lys) genotype (Table 20).

Table 18. Number of cases/controls and odds ratios of lung cancer in relation to *CYP1A2* genotype

Genotype	Cases/controls		OR[a,b] (95% CI)
	Predicted CYP1A2 high activity	Predicted CYP1A2 low activity	
*CYP1A2*1C* and *CYP1A2*1F*			
All cases	59/33	75/100	**2.05[c] (1.13-3.73)**
Males	45/21	45/19	1.17 (0.54-2.51)
Females	23/14	29/47	2.19 (0.82-5.89)
Adenocarcinoma	33/35	47/104	1.79 (0.92-3.48)
Squamous cell carcinoma	25/35	23/28	0.92 (0.39-2.12)
Non occupational exposures	46/28	56/77	0.90 (0.43-1.88)

[a] Adjusted odds ratios (OR) for age, race, smoking status (ever smokers, nonsmokers), and packyear (packyear+packyear square); CI, confidence interval

[b] Odds ratio calculated comparing prevalences of predicted high CYP1A2 versus low and normal CYP1A2 activity

[c] Adjusted odds ratios (OR) for age, race, smoking status (ever smokers, nonsmokers), packyear (packyear+packyear square), and occupational exposures

Table 19. Number of cases/controls and odds ratio of lung cancer in relation to *GSTM3* genotype

Genotype	Cases/controls		OR[a,b] (95% CI)
	***GSTM3*A/*A* genotype**	***GSTM3* (*A/*B and *B/*B) genotype**	
GSTM3			
All cases	96/86	38/57	**1.84[c] (1.03-3.31)**
Males	61/33	29/17	1.35 (0.68-2.65)
Females	40/37	12/26	**3.03 (1.09-8.38)**
Adenocarcinoma	53/91	27/58	1.17 (0.62-2.20)
Squamous cell carcinoma	37/50	11/28	2.15 (0.84-5.48)
Non occupational exposures	73/63	29/41	1.53 (0.81-2.87)

[a] Adjusted odds ratios (OR) for age, race, smoking status (ever smokers, nonsmokers), and packyear (packyear+packyear square); CI, confidence interval

[b] Odds ratio calculated comparing prevalences of *GSTM3*A/*A* versus *GSTM3* (**A/*B* and **B/*B*) genotypes

[c] Adjusted odds ratios (OR) for age, race, smoking status (ever smokers, nonsmokers), packyear (packyear+packyear square), and occupational exposures

Table 20. Number of cases/controls and odds ratio of lung cancer in relation to *XPD* exon 23 genotype

Genotype	Cases/controls		OR[a,b] (95% CI)
	***XPD* (Lys/Gln and Gln/Gln) genotype**	***XPD* (Lys/Lys) genotype**	
XPD exon 23			
All cases	89/62	45/69	**2.56[c] (1.45-4.51)**
Males	63/20	42/26	**2.20[d] (1.11-4.39)**
Females	33/28	19/32	**2.72 (1.10-6.74)**
Adenocarcinoma	54/65	27/71	**2.75 (1.43-5.28)**
Squamous cell carcinoma	30/30	18/43	**2.66[d] (1.15-6.12)**
Non occupational exposures	64/47	37/52	**2.33 (1.25-4.32)**

[a] Adjusted odds ratios (OR) for age, race, smoking status (ever smokers, nonsmokers), and packyear (packyear+packyear square); CI, confidence interval

[b] Odds ratio calculated comparing prevalences of *XPD* exon 23 (Lys/Gln and G/n/Gln) versus *XPD* exon 23 (Lys/Lys) genotypes

[c] Adjusted odds ratios (OR) for age, race, smoking status (ever smokers, nonsmoker), packyear (packyear+packyear square), and occupational exposures

[d] Adjusted odds ratios (OR) for age, race, and smoking status (ever smokers, nonsmokers)

Table 21. Number of cases/controls and odds ratio of lung cancer in relation to combined *mEPHX* exon 3 and 4 genotypes

Genotype	Cases/controls		OR[a,b] (95% CI)
	Predicted mEPHX high and intermediate activity	**Predicted mEPHX low and very low activity**	
mEPHX exon 3 and 4			
All cases	70/69	59/56	1.02 (0.60-1.73)
Males	44/30	43/15	0.65 (0.33-1.27)
Females	28/27	22/31	**2.77 (1.02-7.51)**
Adenocarcinoma	40/72	37/59	1.14 (0.61-2.14)
Squamous cell carcinoma	27/40	25/31	0.87 (0.37-2.04)
Non occupational exposures	51/52	47/46	1.01 (0.55-1.86)

[a] Adjusted odds ratios (OR) for age, race, smoking status (ever smokers, nonsmokers), and packyear (packyear+packyear square); CI, confidence interval

[b] Odds ratio calculated comparing prevalences of predicted high and intermediate versus low or very low mEPHX activity

Table 22. Number of cases/controls and odds ratio of lung cancer in relation to *GSTP1* genotype

Genotype	Cases/controls		OR[a,b] (95% CI)
	***GSTP1* (**A/*B* and **B/*B*) genotype**	***GSTP1 *A/*A* genotype**	
GSTP1			
All cases	90/99	90/94	1.51 (0.88-2.61)
Males	58/47	58/56	**2.09 (1.03-4.25)**
Females	32/52	32/38	0.88 (0.35-2.18)
Adenocarcinoma	44/72	53/76	1.24 (0.67-2.32)
Squamous cell carcinoma	27/36	25/41	1.66 (0.73-3.79)
Non occupational exposures	55/75	66/67	1.23 (0.65-2.33)

[a] Adjusted odds ratios (OR) for age, race, and smoking status (ever smokers, nonsmokers); CI, confidence interval

[b] Odds ratio calculated comparing prevalences of *GSTP1*(**A/*B* and **B/*B*) versus *GSTP1*A/*A* genotypes

4.1.3. Gene-gene interactions and lung cancer risk

For the gene-gene interaction analyses, the possible interactions among 'at risk' genotypes based on biological mechanisms of cigarette smoke carcinogen metabolism were identified prior to inclusion in the logistic regression models. The associations between gene-gene interactions and lung cancer risk were found among all cases who both had and did not have a history of occupational exposures for combinations of the *XPD* exon 23 (Lys/Gln and Gln/Gln) genotype and others at risk genotypes such as *CYP1B1* (**1/*3* and **3/*3*) genotype, predicted mEPHX high and intermediate activity, *GSTM3*A/*A* genotype, and *GSTP1* (**A/*B* and **B/*B*) genotype after adjusting for age, race, smoking status, and packyear (Table 23, 24). A significant gene-gene-gene interaction was also found to be associated with a 2-fold increased risk of lung only among cases who had *XPD* exon 23 (Lys/Gln and Gln/Gln), *GSTM1* null and *GSTM3*A/*A* genotypes, and history of occupational exposures compared to reference individuals (Table 23).

The interactions between *XPD* exon 23 (Lys/Gln and Gln/Gln) genotype and *GSTM3*A/*A* genotype, or *GSTP1* (**A/*B* and **B/*B*) genotype also showed approximately a 3-fold significant increasing risk of lung cancer among males who had a combination of these two at risk genotypes compared to reference individuals who carried wild type alleles for the *XPD* exon 23 and *GSTP1* genes but the variant allele of *GSTM3* gene after adjusting for age, race, and smoking status (Table 25). Among females, not only gene-gene interactions between between *XPD* exon 23 (Lys/Gln and Gln/Gln) genotype, and other at risk genotypes (Table 26) were found to be associated with increased risk of lung cancer but also interactions between predicted mEPHX high and intermediate activity, and other at risk genotypes such as *CYP1B1* (**1/*3* and **1/*3*), *MPO* (G/G), and *GSTM3*A/*A* (Table 27). The lung cancer risk increased from 2.5- to 4- fold in combinations of the *XPD* exon 23 (Lys/Gln and Gln/Gln) and *MPO* (G/G), or *GSTM3*A/*A*

genotypes, and of predicted mEPHX high and intermediate activity and *GSTM3*A/*A* genotype, but the rate of increase was greater to be 5- to 6-fold in lung cancer risk if female cases had combinations of *XPD* exon 23 (Lys/Gln and Gln/Gln) and *CYP1B1* (**1/*3* and **3/*3*) genotypes, or predicted mEPHX high and intermediate activity, and of predicted mEPHX high and intermediate activity and *MPO* (G/G) genotype. In addition, a gene-gene-gene interaction was found to significantly increase risk of lung cancer among females who had predicted mEPHX high and intermediate activity and *GSTM1* null and *GSTM3*A/*A* genotypes after adjusting for age, race, smoking status, and packyear (Table 27).

The interactions between two combined at risk genotypes were in the same direction when the different histological cell types (adenocarcinoma and squamous cell carcinoma) were compared separately to controls (Table 28, 29). However, the interactions between two at risk genotypes of phase II enzymes (*GSTM3*A/*A* and *GSTP1*A/*B* and **B/*B*), and between *XPD* exon 23 (Lys/Gln and Gln/Gln) and *MPO* (G/G) genotypes were statistically significant (P=0.03 and 0.04, respectively) in squamous cell carcinoma subgroup (Table 29) but not in the adenocarcinoma subgroup. In contrast, among adenocarcinoma cases only, combinations of *XPD* exon 23 (Lys/Gln and Gln/Gln) genotype and predicted mEPHX high and intermediate activity, or predicted mEPHX intermediate activity showed the relative risk of lung cancer 2.3- and 6.3-fold, respectively.

4.1.4. Gene-environment interactions and lung cancer risk

For the gene-environment interaction analyses, all of the possible interactions among the genotype, smoking status, transformed packyear (packyear+packyear square), and occupational exposure variables were included in the logistic regression models. The only first order interaction between genotype and the transformed packyear that was found to be significant in

the models was an interaction term between *CYP1B1* (**1/*3* and **3/*3*) genotype and the transformed packyear variable. The final model in all cases that predicted relative risk of lung cancer by *CYP1B1* (**1/*3* and **3/*3*) genotype included the *CYP1B1* (**1/*3* and **3/*3*) genotype variable and the interaction term between *CYP1B1* (**1/*3* and **3/*3*) and the transformed packyear (P=0.008) and was adjusted for age, race, smoking status, and occupational exposures (Table 30). In addition, the interaction term between *CYP1B1* (**1/*3* and **3/*3*) and the transformed packyear was significant (P<0.05) in final models for predicting risk of lung cancer by combined *XPD* exon 23 (Lys/Gln and Gln/Gln) and *CYP1B1* (**1/*3* and **3/*3*) genotypes in all cases, females, and adenocarcinoma subgroups, and by combined predicted mEPHX high and intermediate activity and *CYP1B1* (**1/*3* and **3/*3*) genotype among females after adjusting for age, race, smoking status, and/or occupational exposures (Table 23, 26, 28, and 27; respectively).

The second order interactions between genotype and the transformed packyear of predicted mEPHX high and intermediate activity and *CYP1B1* (**1/*3* and **3/*3*) and the transformed packyear variables, and of the *XPD* exon 23 (Lys/Gln and Gln/Gln) genotype and predicted mEPHX intermediate activity and the transformed packyear variables were found to be significant at P-value less than 0.05 in final models of relative risk of lung cancer prediction by a combination of predicted mEPHX high and intermediate activity and *CYP1B1* (**1/*3* and **3/*3*) genotype among females and by combined the *XPD* exon 23 (Lys/Gln and Gln/Gln) genotype and predicted mEPHX intermediate activity in adenocarcinoma subgroup after adjusting age, race, smoking status, and the transformed packyear (Table 27 and 28; respectively).

Table 23. Odds ratios for the association of combined *XPD* exon 23, *CYP1B1*, *mEPHX*, *GSTM1*, *GSTM3*, and *GSTP1* genotypes with lung cancer in all cases

Genotype	Cases/controls	OR[a] (95% CI)
XPD exon 23 and *CYP1B1* *XPD* (Lys/Gln and Gln/Gln) and *CYP1B1* (**1/*3* and **3/*3)*	63/39	**3.09[b,c] (1.70-5.63)**
XPD exon23 and *mEPHX* *XPD* (Lys/Gln and Gln/Gln) and mEPHX high and intermediate activity	50/30	**2.11[d] (1.13-3.94)**
XPD exon 23 and *GSTM3* *XPD* (Lys/Gln and Gln/Gln) and *GSTM3* (**A/*A*)	62/33	**3.18[e] (1.73-5.85)**
XPD exon 23 and *GSTM1* and *GSTM3* *XPD* (Lys/Gln and Gln/Gln) and *GSTM1* null and *GSTM3* (**A/*A*)	40/25	**2.02[f] (1.05-3.89)**
XPD exon 23 and *GSTP1* *XPD* (Lys/Gln and Gln/Gln) and *GSTP1* (**A/*B* and **B/*B*)	46/28	**2.46[g] (1.29-4.71)**

[a] Adjusted odds ratios (OR) for age, race, smoking status (ever smokers, nonsmokers), packyear (packyear+packyear square), and occupational exposures; CI, confidence interval

[b] Adjusted odds ratios (OR) for age, race, smoking status (ever smokers, nonsmokers), packyear (packyear+packyear square), occupational exposures, and interaction term between *CYP1B1* (**1/*3* and **3/*3)* genotype and packyear (packyear+packyear square); CI, confidence interval

[c] Cases/controls (73/129) with *XPD* (Lys/Lys) and *CYP1B1*1/*1* genotype serve as the reference category

[d] Cases/controls (82/121) with *XPD* (Lys/Lys) genotype and predicted mEPHX low or very low activity serve as the reference category

[e] Cases/controls (73/130) with *XPD* (Lys/Lys) and *GSTM3* (**A/*B* and **B/*B*) genotype serve as the reference category

[f] Cases/controls (95/138) with *XPD* (Lys/Lys) and *GSTM1* null and *GSTM3* (**A/*B* and **B/*B*) genotype serve as the reference category

[g] Cases/controls (91/135) with *XPD* (Lys/Lys) and *GSTP1* (**A/*A*) genotype serve as the reference category

Table 24. Odds ratios for the association of combined *XPD* exon 23, *CYP1B1*, *MPO*, *GSTM3*, and *GSTP1* genotypes with lung cancer among individuals without occupational exposures

Genotype	Cases/controls	OR[a] (95% CI)
XPD exon23 and *mEPHX*		
XPD (Lys/Gln and Gln/Gln) and mEPHX high and intermediate activity	33/22	**2.40[b] (1.18-4.85)**
XPD exon23 and *MPO*		
XPD (Lys/Gln and Gln/Gln) and *MPO* (G/G)	39/29	**1.97[c] (1.03-3.80)**
XPD exon23 and *GSTM3*		
XPD (Lys/Gln and Gln/Gln) and *GSTM3* (**A/*A*)	45/25	**2.26[d] (1.17-4.38)**
XPD exon23 and *GSTP1*		
XPD (Lys/Gln and Gln/Gln) and *GSTP1* (**A/*B* and **B/*B*)	30/22	**2.07[e] (1.02-4.19)**

[a] Adjusted odds ratios (OR) for age, race, and smoking status (ever smokers, nonsmokers), and packyear (packyear+packyear square); CI, confidence interval

[b] Cases/controls (68/98) with *XPD* (Lys/Lys) genotype and predicted mEPHX low and very low activity serve as the reference category

[c] Cases/controls (63/99) with *XPD* (Lys/Lys) genotype and *MPO* (G/A and A/A) serve as the reference category

[d] Cases/controls (58/100) with *XPD* (Lys/Lys) and *GSTM3* (**A/*B* and **B/*B*) genotype serve as the reference category

[e] Cases/controls (73/104) with *XPD* (Lys/Lys) and *GSTP1* (**A/*A*) genotype serve as the reference category

Table 25. Odds ratios for the association of combined *XPD* exon 23, *GSTM3*, and *GSTP1* genotypes with lung cancer among males

Genotype	Cases/controls	OR[a] (95% CI)
XPD exon23 and *GSTM3*		
XPD (Lys/Gln and Gln/Gln) and *GSTM3* (**A/*A*)	39/11	**2.79[b] (1.27-6.13)**
XPD exon23 and *GSTP1*		
XPD (Lys/Gln and Gln/Gln) and *GSTP1* (**A/*B* and **A/*B*)	32/8	**3.17[c] (1.35-7.43)**

[a] Adjusted odds ratios (OR) for age, race, and smoking status (ever smokers, nonsmokers); CI, confidence interval

[b] Cases/controls (74/78) with *XPD* (Lys/Lys) and *GSTM3* (**A/*B* and **B/*B*) genotype serve as the reference category

[c] Cases/controls (83/82) with *XPD* (Lys/Lys) and *GSTP1* (**A/*A*) genotype serve as the reference category

Table 26. Odds ratios for the association of combined *XPD* exon 23, *mEPHX*, *MPO*, and *GSTM3* genotypes with lung cancer among females

Genotype	Cases/controls	OR[a] (95% CI)
XPD exon23 and *CYP1B1* *XPD* (Lys/Gln and Gln/Gln) and *CYP1B1* (**1/*3* and **3/*3)*	27/17	**5.69[b,c] (2.01-16.07)**
XPD exon23 and *mEPHX* *XPD* (Lys/Gln and Gln/Gln) and mEPHX high and intermediate activity	17/9	**5.81[d] (1.78-18.93)**
XPD exon23 and *MPO* *XPD* (Lys/Gln and Gln/Gln) and *MPO* (G/G)	25/20	**2.66[e] (1.05-6.74)**
XPD exon23 and *GSTM3* *XPD* (Lys/Gln and Gln/Gln) and *GSTM3* (**A/*A*)	25/16	**3.14[f] (1.22-8.13)**

[a] Adjusted odds ratios (OR) for age, race, smoking status (ever smokers, nonsmokers), and packyear (packyear+packyear square); CI, confidence interval

[b] Adjusted odds ratios (OR) for age, race, smoking status (ever smokers, nonsmokers), packyear (packyear+packyear square), and interaction term between *CYP1B1* (**1/*3* and **3/*3)* genotype and packyear (packyear+packyear square); CI, confidence interval

[c] Cases/controls (26/64) with *XPD* (Lys/Lys) and *CYP1B1*1/*1* genotype serve as the reference category

[d] Cases/controls (34/66) with *XPD* (Lys/Lys) genotype and predicted mEPHX low or very low activity serve as the reference category

[e] Cases/controls (28/59) with *XPD* (Lys/Lys) and *MPO* (G/A and A/A) genotype serve as the reference category

[f] Cases/controls (27/63) with *XPD* (Lys/Lys) and *GSTM3* (**A/*B* and **B/*B*) genotype serve as the reference category

Table 27. Odds ratios for the association of combined *mEPHX*, *CYP1B1*, *MPO*, *GSTM1*, and *GSTM3* genotypes with lung cancer among females

Genotype	Cases/controls	OR[a] (95% CI)
mEPHX and *CYP1B1* mEPHX high and intermediate activity and *CYP1B1* (**1/*3* or **3/*3)*	21/15	**9.10[b,c] (2.43-34.13)**
mEPHX and *MPO* mEPHX high and intermediate activity and *MPO* (G/G)	24/13	**5.47[d] (1.90-15.81)**
mEPHX and *GSTM3* mEPHX high and intermediate activity and *GSTM3* (**A/*A*)	24/15	**3.74[e] (1.41-9.86)**
mEPHX and *GSTM1*and *GSTM3* mEPHX high and intermediate activity and *GSTM1* null and *GSTM3* (**A/*A*)	20/11	**3.04[f] (1.12-8.23)**

[a] Adjusted odds ratios (OR) for age, race, smoking status (ever smokers, nonsmokers), and packyear (packyear+packyear square); CI, confidence interval

[b] Adjusted odds ratios (OR) for age, race, smoking status (ever smokers, nonsmokers), packyear (packyear+packyear square), and interaction term between combined genotypes and packyear (packyear+packyear square); CI, confidence interval

[c] Cases/controls (28/61) with predicted mEPHX low and very low activity and *CYP1B1*1/*1* genotype serve as the reference category

[d] Cases/controls (27/63) with predicted mEPHX low and very low activity and *MPO* (G/A and A/A) genotype serve as the reference category

[e] Cases/controls (26/61) with predicted mEPHX low and very low activity and *GSTM3* (**A/*B* and **B/*B*) genotype serve as the reference category

[f] Cases/controls (30/65) with predicted mEPHX low and very low activity, *GSTM1* positive, and *GSTM3* (**A/*B* and **B/*B*) genotype serve as the reference category

Table 28. Odds ratios for the association of combined *XPD* exon 23, *CYP1B1*, *mEPHX*, *GSTM3*, and *GSTP1* genotypes with lung cancer among adenocarcinoma subgroup

Genotype	Cases/controls	OR[a] (95% CI)
XPD exon23 and *CYP1B1* *XPD* (Lys/Gln and Gln/Gln) and *CYP1B1* (**1/*3* and **3/*3)*	37/42	**2.84[b,c] (1.44-5.61)**
XPD exon23 and *mEPHX* *XPD* (Lys/Gln and Gln/Gln) and mEPHX intermediate activity	16/17	**6.35[d,e] (1.99-20.24)**
XPD exon23 and *mEPHX* *XPD* (Lys/Gln and Gln/Gln) and mEPHX high and intermediate activity	29/31	**2.31[f,g] (1.12-4.76)**
XPD exon23 and *GSTM3* *XPD* (Lys/Gln and Gln/Gln) and *GSTM3* (**A/*A*)	37/35	**2.89[h] (1.45-5.77)**
XPD exon23 and *GSTP1* *XPD* (Lys/Gln and Gln/Gln) and *GSTP1* (**A/*B* and **B/*B*)	24/28	**2.43[i,j] (1.13-5.19)**

[a] Adjusted odds ratios (OR) for age, race, smoking status (ever smokers, nonsmokers), and packyear (packyear+packyear square); CI, confidence interval

[b] Adjusted odds ratios (OR) for age, race, smoking status (ever smokers, nonsmokers), packyear (packyear+packyear square), and interaction term between *CYP1B1* (**1/*3* and **3/*3)* genotype and packyear (packyear+packyear square); CI, confidence interval

[c] Cases/controls (43/93) with *XPD* (Lys/Lys) and *CYP1B1*1/*1* genotype serve as the reference category

[d] Adjusted odds ratios (OR) for age, race, smoking status (ever smokers, nonsmokers), packyear (packyear+packyear square), and interaction term between combined genotypes and packyear (packyear+packyear square); CI, confidence interval

[e] Cases/controls (41/73) with *XPD* (Lys/Lys) genotype and predicted mEPHX low or very low activity serve as the reference category

[f] Adjusted odds ratio (OR) for age, race and smoking status (ever smokers, nonsmokers)

[g] Cases/controls (64/94) with *XPD* (Lys/Lys) genotype and predicted mEPHX low or very low activity serve as the reference category

[h] Cases/controls (41/93) with *XPD* (Lys/Lys) and *GSTM3* (**A/*B* and **B/*B*) genotype serve as the reference category

[i] Adjusted odds ratio (OR) for age, race, smoking status (ever smokers, nonsmokers), and occupational exposures

[j] Cases/controls (54/92) with *XPD* (Lys/Lys) and *GSTP1* (**A/*A*) genotype serve as the reference category

Table 29. Odds ratios for the association of combined *XPD* exon 23, *CYP1B1*, *MPO*, *GSTM3*, and *GSTP1* genotypes with lung cancer among squamous cell carcinoma subgroup

Genotype	Cases/controls	OR[a] (95% CI)
XPD exon23 and *CYP1B1* *XPD* (Lys/Gln and Gln/Gln) and *CYP1B1* (**1/*3* and **3/*3)*	26/18	**2.95[b] (1.26-6.93)**
XPD exon23 and *MPO* *XPD* (Lys/Gln and Gln/Gln) and *MPO* (G/G)	21/17	**2.51[c] (1.01-6.15)**
XPD exon23 and *GSTM3* *XPD* (Lys/Gln and Gln/Gln) and *GSTM3* (**A/*A*)	24/18	**3.23[d] (1.32-7.88)**
XPD exon23 and *GSTP1* *XPD* (Lys/Gln and Gln/Gln) and *GSTP1* (**A/*B* and**B/*B*)	17/14	**3.15[e] (1.13-8.77)**
GSTM3 and *GSTP1* *GSTM3* (**A/*A*) and *GSTP1* (**A/*B* and **B/*B*)	24/23	**2.75[f] (1.14-6.64)**

[a] Adjusted odds ratios (OR) for age, race, and smoking status (ever smokers, nonsmokers); CI, confidence interval
[b] Cases/controls (26/56) with *XPD* (Lys/Lys) and *CYP1B1*1/*1* genotype serve as the reference category
[c] Cases/controls (31/57) with *XPD* (Lys/Lys) genotype and *MPO* (G/A and A/A) serve as the reference category
[d] Cases/controls (28/56) with *XPD* (Lys/Lys) and *GSTM3* (**A/*B* and **B/*B*) genotype serve as the reference category
[e] Cases/controls (35/58) with *XPD* (Lys/Lys) and *GSTP1* (**A/*A*) genotype serve as the reference category
[f] Cases/controls (28/54) with *GSTM3* (**A/*B* and **B/*B*) and *GSTP1* (**A/*A*) genotype serve as the reference category

Table 30. Number of cases/controls and odds ratios of lung cancer in relation to *CYP1B1* genotype

Genotype	Cases/controls		OR[a,b] (95% CI)
	***CYP1B1* (**1/*3* and **3/*3)* genotype**	***CYP1B1*1/*1* genotype**	
CYP1B1			
All cases	94/95	41/51	**2.60[c] (1.19-5.69)**
Males	65/76	33/30	0.94 (0.47-1.89)
Females	38/40	15/24	3.69[d] (0.94-14.54)
Adenocarcinoma	58/100	23/52	2.34[d] (0.99-5.53)
Squamous cell carcinoma	35/51	13/28	1.55 (0.61-3.96)
Non occupational exposures	72/110	34/35	0.70 (0.36-1.35)

[a] Adjusted odds ratios (OR) for age, race, smoking status (ever smokers, nonsmokers), and packyear (packyear+packyear square); CI, confidence interval
[b] Odds ratio calculated comparing prevalences of *CYP1B1* (**1/*3* and **3/*3*) versus *CYP1B1*1/*1* genotypes
[c] Adjusted odds ratios (OR) for age, race, smoking status (ever smokers, nonsmokers), packyear (packyear+packyear square), occupational exposures, and interaction term between *CYP1B1* (**1/*3* and **3/*3)* genotype and packyear (packyear+packyear square); CI, confidence interval
[d] Adjusted odds ratios (OR) for age, race, smoking status (ever smokers, nonsmokers), packyear (packyear+packyear square), and interaction term between *CYP1B1* (**1/*3* and **3/*3)* genotype and packyear (packyear+packyear square); CI, confidence interval

4.2. Case only analysis

4.2.1. Characteristics of case only study population

There were a total of 177 lung cancer cases evaluated in this analysis. All cases were Caucasians. The demographic characteristics of these cases including age, smoking status, packyear and occupational exposures are summarized in Tables 31 and 32. Cases were categorized into 6 subgroups by family history of lung and other cancer(s), age at diagnosis, smoking history, and histological cell types (Table 33). Cases who did not meet the above criteria were called sporadic and used as the referent group. There were 23 cases who fell in more than one group. Overall mean age, mean packyear, and gender distribution of cases in each subgroup and in the sporadic group were similar (Table 34). However, there was a difference in mean age for cases who at least one first-degree relative(s) with lung cancer compared to all sporadic cases. Mean packyear of younger age onset and uncommon histological cell type subgroups were statistically lower than that of all sporadic cases. In addition, the ratio of lifetime nonsmoker females to males was higher than that of all sporadic cases. After sporadic cases with history of occupational exposures were excluded, the distribution of gender in lifetime nonsmokers was not different from that of sporadic cases (Table 35). All cases were combined and classified by history of occupational exposures into two groups in order to evaluate the overall effect of combined risk factors such as metabolic and DNA repair gene polymorphisms and family history of lung and other cancer(s) on increasing risk of lung cancer related to the history of occupational exposures.

4.2.2. Metabolic and DNA repair gene polymorphisms and lung cancer risk among at risk lung cancer cases

The contribution of both metabolic enzyme and DNA repair gene polymorphisms on greater risk of lung cancer development among at risk lung cancer cases was assessed by using the χ^2-test. The relative risk were estimated by comparing the differences in numbers of individuals who had one at risk genotype or combined genotypes in each subgroup and all case groups compared to sporadic cases.

There was a gene-gene interaction between *CYP1A2* and *NAT2** at risk genotypes among lung cancer cases with at least one first-degree relative(s) with lung cancer as indicated by an overrepresentation of percentages of cases with combined predicted high CYP1A2 activity and slow *NAT2** genotype compared to sporadic cases (Fig. 12 and Table 36). In group 2 or a group of cases with at least two first-degree relatives with other cancer(s), *GSTM1* null genotype alone was not statistically significant for an increased risk of lung cancer compared to cases also with this 'at risk' genotype but no family history of cancer(s). However, the risk was significant when *GSTM1* null genotype was combined with the *XPD* exon 23 (Lys/Gln and Gln/Gln) genotype. (Fig. 13 and Table 37). A contribution of DNA repair gene and phase II enzyme polymorphisms (*XPD* exon 23 (Lys/Gln and Gln/Gln) and *GSTP1*A/*B* or **B/*B* genotypes) for early age onset of lung cancer development was found as shown in Fig.14 and Table 38).

Associations between single loci polymorphisms or gene-gene interactions and lung cancer risk were found among lifetime nonsmokers as shown in Fig 15. There was an overrepresentation of the number of lifetime nonsmoker cases who had predicted high CYP1A2 activity compared to ever smoker sporadic cases (Table 39). Furthermore, the combination of predicted CYP1A2 activity and *NAT2** slow genotype, and predicted reduced DNA repair

capacity and *GSTM1* null genotype were associated with increased risk of lung cancer among these lifetime nonsmokers.

When all lung cancer cases were combined together and categorized into two subgroups according to history of occupational exposures, the contribution of interactions of metabolic enzymes and DNA repair gene polymorphisms in combination with other risk factors such as family history of cancer(s), early age onset, smoking history, and/or occupational exposures in lung cancer development were evaluated. It was found the associations between two different combined at risk genotypes (predicted high CYP1A2 activity and slow acetylator, and *MPO* (G/G) and *GSTM1* null genotype) and increasing risk of lung cancer among ever smoker cases with family history of lung cancer and other cancer(s), and age at diagnosis less than or at 50 years old compared to ever smokers with no other risk factor(s) (Fig. 16 and Table 40). Furthermore, cases who had the above other 'at risk' factor(s), also history of occupational exposures, and interactions between *XPD* exon 23 (Lys/Gln and Gln/Gln), or predicted mEPHX intermediate and high activity and *GSTP1* (**A/*B* or **B/*B)* genotypes were at greater risk of lung cancer compared to lung cancer cases who were ever smokers but no other risk factor(s) (Fig 17 and Table 41).

Table 31. Age and exposure histories of lung cancer cases (N=177)

	Mean age (±SD)	Non-cigarette Smokers (%, total cases)	Asbestos Exposure (%, total cases)	Other Occupational exposure[a] (%, total cases)
Men (N=118)	63 (±13)	8 (7.92, 101)	14 (13.2, 106)	54 (50.9, 106)
Women (N=59)	66 (±12)	11 (22.0, 50)	0	8 (15.6, 52)
All (N=177)	64 (±13)	20 (13.2, 151)	14 (8.86, 158)	62 (39.2, 158)

[a] Includes silicon, petroleum, radon, nickel, chromium and others

Table 32. Packyear distribution of smoking history by gender ratio of lung cancer cases (N=177)

	0	1-20	21-40	40-60	≥ 61	Total[a]
Number of cases	19	18	45	30	39	151
Percent of total cases	12.6	11.9	29.8	19.9	25.8	100
Gender ratio (M/F)	0.7	1.6	3.1	2.0	2.6	2.0

[a] Missing information on 26 cases

Table 33. Subgroups of lung cancer cases by family history, age at diagnosis, smoking status, and histological cell type (N=177)

Group[a]	N	Mean age in years (±SD)
All cases	177	64 (±13)
1. ≥ one 1st-degree relative(s) with lung cancer	56	64 (±11)
2. ≥ two 1st-degree relative(s) with other cancer	40	68 (±10)
3. ≤ 50 years at diagnosis	30	43 (± 6)
4. lifetime nonsmokers[b]	16	66 (±18)
5. Uncommon tumor type[c]	14	62 (±18)
6. Sporadic	46	68 (± 8)
[a] 23 Cases fall in more than one group		
Group 1	15	
Group 2	12	
Group 3	12	
Group 4	8	
Group 5	1	

[b] < 100 cigarettes during their lifetime

[c] carcinoid tumor

Table 34. Gender, smoking status, and packyear distributions in each lung cancer subgroup

Group	Gender		Smoking status, N (%)		Mean Age (±SD)	Mean PY (±SD)
	Males N (%)	Females N (%)	Ever-smokers N (%)	Non-smokers N (%)		
1						
(N=56)	33 (58.9)	23 (41.1)	53 (94.6)	3 (5.4)	64 (±11)	50 (±30.6)
P-Value[a]	0.11				0.02	0.16
2						
(N=40)	28 (70.0)	12 (30.0)	36 (90.0)	4 (10.0)	68 (±10)	49 (±25.7)
P-Value[a]	0.69				0.69	0.13
3						
(N=30)	23 (76.7)	7 (23.3)	26 (86.7)	4 (13.3)	43 (± 6)	35 (±40.2)
P-Value[a]	0.79				0.00	0.02
4						
(N=16)	7 (43.8)	9 (56.2)	0	16 (100)	66 (±18)	0
P-Value[a]	0.03				0.62	
5						
(N=14)	9 (64.3)	5 (35.7)	9 (69.2)	4 (30.8)	62 (±18)	24 (±17.7)
P-Value[a]	0.48				0.18	0.00
6						
(N=46)	34 (73.9)	12 (26.1)	45 (100)	0	68 (± 8)	60 (±33.2)

[a] P-value were obtained by using χ^2-test, comparing values in each subgroup to those of sporadic group

Table 35. The distribution differences in gender, smoking status, age and packyear for each subgroup and combined group compared to sporadic cases with no history of occupational exposures

Group	Gender		Smoking status		Mean Age (±SD)	Mean PY (±SD)
	Males N (%)	Females N (%)	Ever-smokers N (%)	Non-smokers N (%)		
1						
(N=56)	33 (58.9)	23 (41.1)	53 (94.6)	3 (5.4)	64 (±11)	50 (±30.6)
P-Value[a]	0.66				0.002	0.29
2						
(N=40)	28 (70.0)	12 (30.0)	36 (90.0)	4 (10.0)	68 (±10)	49 (±25.7)
P-Value[a]	0.57				0.26	0.24
3						
(N=30)	23 (76.7)	7 (23.3)	26 (86.7)	4 (13.3)	43 (± 6)	35 (±40.2)
P-Value[a]	0.26				0.00	0.04
4						
(N=16)	7 (43.8)	9 (56.2)	0	16 (100)	66 (±18)	0
P-Value[a]	0.19				0.41	
5						
(N=14)	9 (64.3)	5 (35.7)	9 (69.2)	4 (30.8)	62 (±18)	24 (±17.7)
P-Value[a]	0.97				0.11	0.00
1-5[b]						
(N=45)	38 (84.4)	7 (15.6)	41 (91.1)	4 (8.9)	70 (±8)	42 (±24)
P-Value[a]	0.03				0.73	0.04
1-5[c]						
(N=59)	35 (59.3)	24 (40.7)	59 (100)	0	66 (±8)	44 (±30)
P-Value[a]	0.68				0.04	0.07
6[d]						
(N=33)	21 (63.6)	12 (36.4)	33 (100)	0	70 (± 7)	59 (±35.6)

[a] P-value were calculated by using χ^2-test, comparing number of cases of each subgroup to the sporadic group

[b] All cases in group 1 to 5 who were exposed to occupational exposures were combined

[c] All cases in group 1 to 5 who were not exposed to occupational exposures were combined

[d] All sporadic who were not exposed to occupational exposures

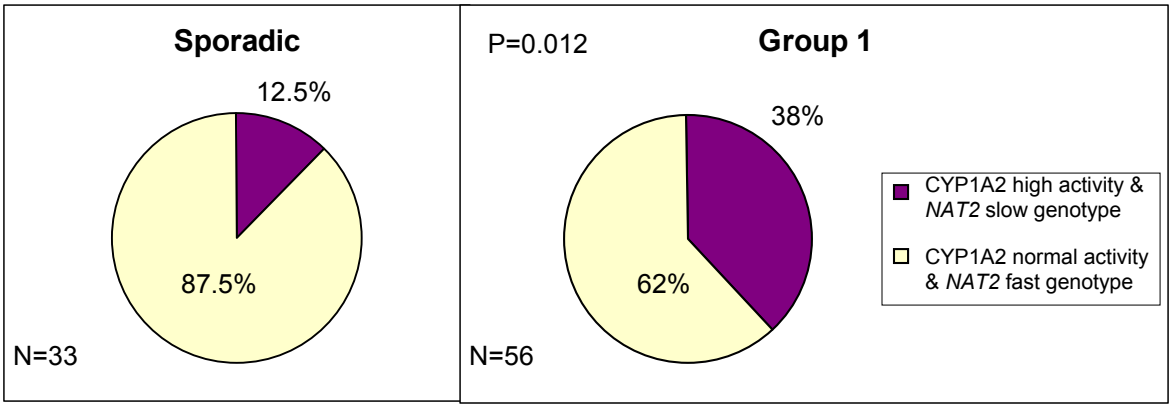


Figure 12. Percentage of combined CYP1A2 activity and *NAT2* genotype in lung cancer cases with at least one first-degree relative(s) with lung cancer compared to the sporadic group

Table 36. Gene-gene interaction for a combination of CYP1A2 and *NAT2* gene polymorphisms in relation to history of lung cancer in first-degree relative(s) of lung cancer probands

Combined genotypes		Number of individuals (%)		P-Value[c]
CYP1A2 and *NAT2*		CYP1A2 high activity & NAT2 slow acetylator	CYP1A2 normal activity & NAT2 fast acetylator	
Group 1	(N=56)	19[b] (38.0)	31 (62.0)	0.012
Sporadic[a]	(N=33)	4 (12.5)	28 (87.5)	

[a] Only sporadic who were not exposed to occupational exposures

[b] All 19 subjects were ever-smokers and 5 and 1 subjects from 19 were exposed to other occupational and asbestos exposures

[c] P-value were obtained by using χ^2-test

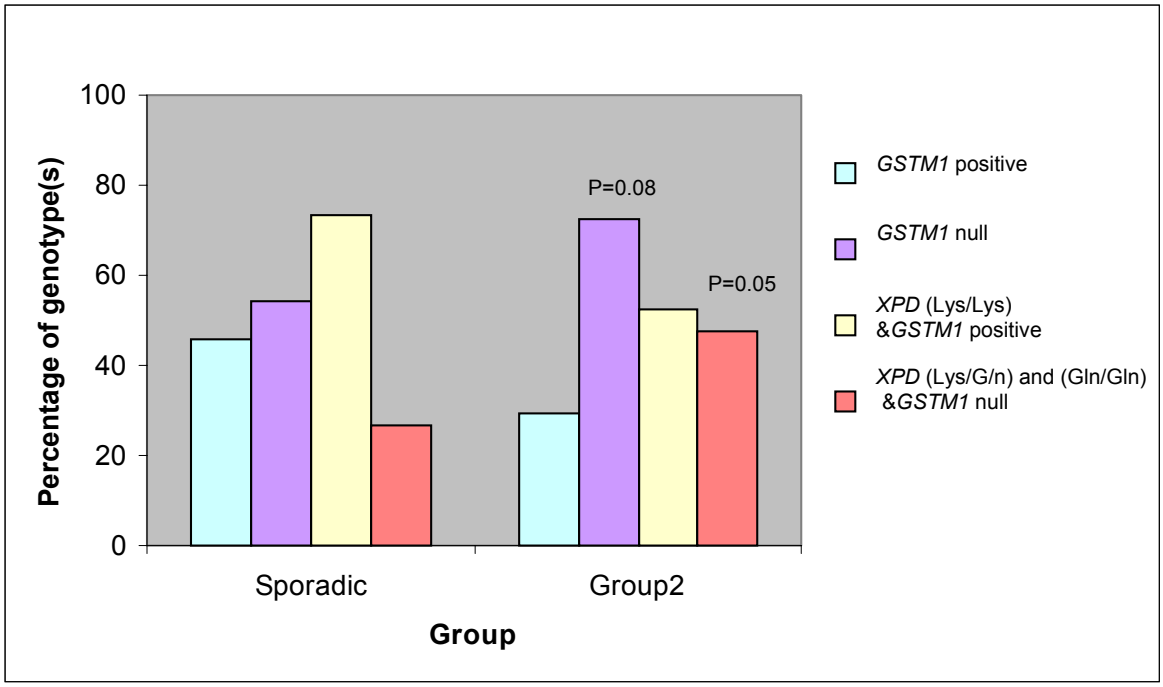


Figure 13. Interaction between *GSTM1*, *XPD* exon 23 genotypes and history of other cancer(s) in first-degree relative(s) of lung cancer probands

Table 37. Association of *GSTM1* null genotype and in combination of *XPD* exon 23 genotype on lung cancer development in relation to history of other cancer(s) in first-degree relatives of lung cancer probands

		Number of individuals (%)		P-value
GSTM1 genotype		*GSTM1* null	*GSTM1* positive	
Group 2	(N=40)	29[c] (72.5)	11 (29.5)	0.04[g], 0.08[h]
Sporadic[a]	(N=33)	16 (48.5)	17 (51.5)	
Sporadic[b]	(N=46)	25[d] (54.3)	21 (45.7)	
XPD exon 23 and *GSTM1*		*XPD* (Lys/Gln and Gln/Gln) & *GSTM1* null	*XPD* (Lys/Lys) & *GSTM1* positive	
Group 2	(N=40)	18[e] (47.4)	20 (52.6)	0.05[h]
Sporadic[b]	(N=46)	12[f] (26.7)	33 (73.3)	

[a] Only sporadic who were not exposed to occupational exposures
[b] All sporadics were included
[c] One and nine cases were exposed to asbestos and other occupational exposures were combined
[d] Three and six cases who were exposed to asbestos and other occupational exposures
[e] Four were lifetime nonsmokers and 2 from 4 were exposed to other occupational exposures
[f] Two and two were exposed to other occupational and asbestos exposures
[g] P-value were obtained by using χ^2-test, comparing number of individuals in each subgroup to those of sporadic[a]
[h] P-value were obtained by using χ^2-test, comparing number of individuals in each subgroup to those of sporadic[b]

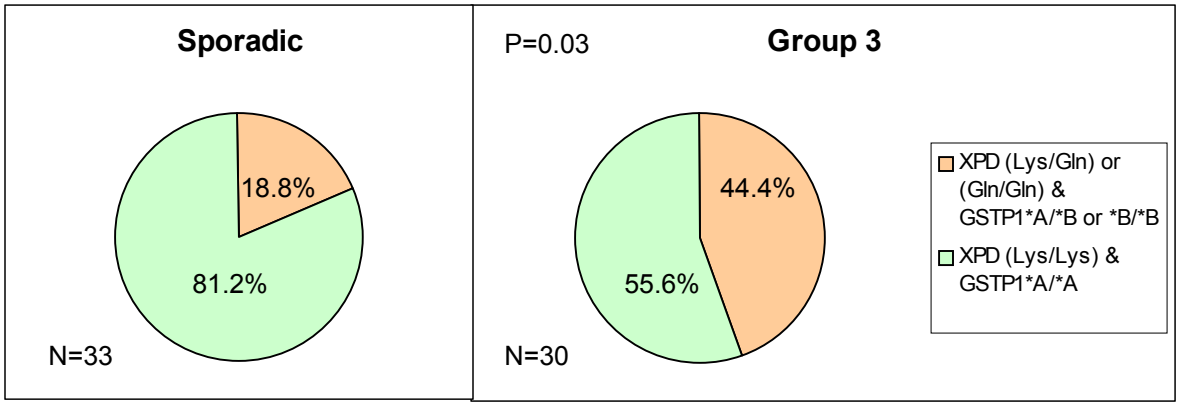


Figure 14. Interaction of *XPD* exon 23 and *GSTP1* genotypes in lung cancer cases with age at diagnosis less than or at 50 years old

Table 38. Association of combined *XPD* exon 23 and *GSTP1* genotypes and early age onset lung cancer cases

Combined genotypes		Number of individuals (%)		P-Value[c]
XPD exon 23 and *GSTP1*		*XPD* (Lys/Gln and Gln/Gln) & *GSTP1*A/*B* and **B/*B*	*XPD* (Lys/Lys) & *GSTP1*A/*A*	
Group 3	(N=30)	12[b] (44.4)	15 (55.6)	0.033
Sporadic[a]	(N=33)	6 (18.8)	26 (81.2)	

[a] Only sporadic who were not exposed to occupational exposures

[b] Six ever-smokers and 1 lifetime nonsmoker were exposed to other occupational exposures and 1 ever smoker was exposed to asbestos

[c] P-value were obtained by using χ^2-test

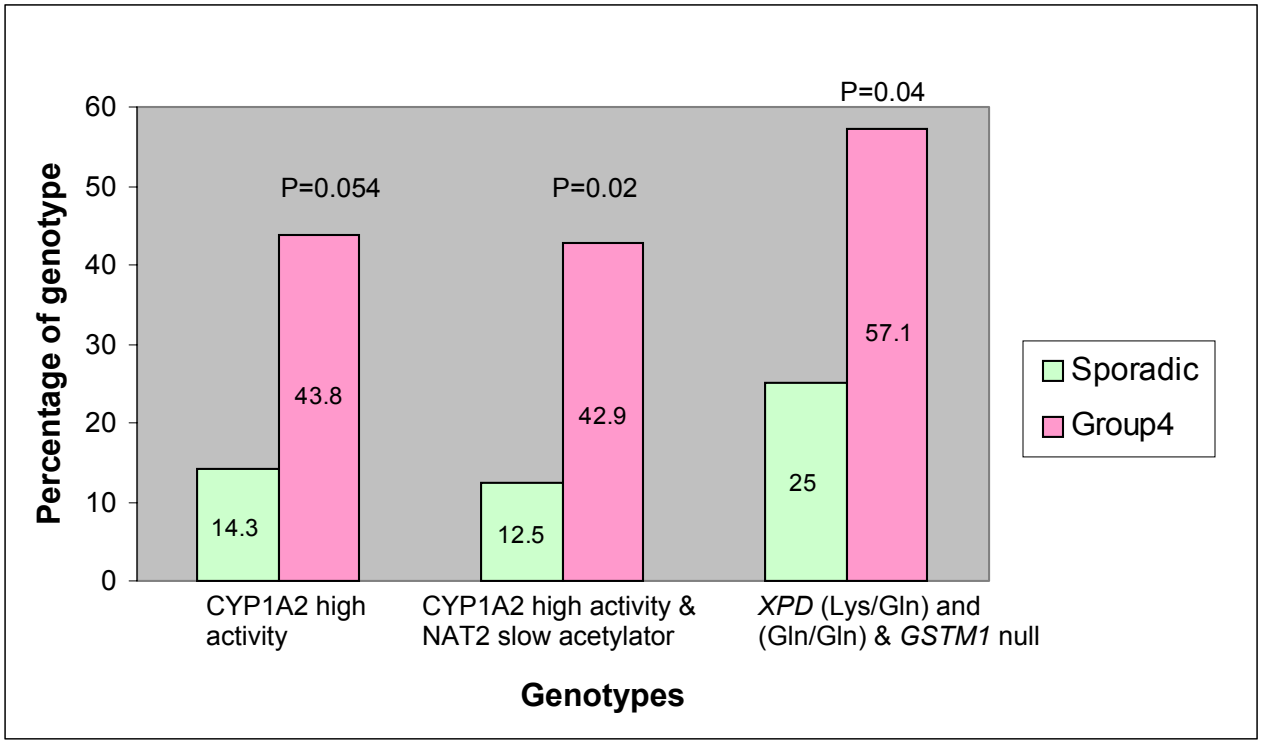


Figure 15. Single and multi gene loci polymorphisms and lung cancer risk among lifetime nonsmokers

Table 39. Interactions of polymorphisms at *GSTM1*, *NAT2* and *XPD* exon 23 genes among lifetime non-smoker lung cancer cases

Combined genotypes		Number of individuals (%)		P-Value[e]
CYP1A2 genotype		CYP1A2 high activity	CYP1A2 normal activity	
Group 4	(N=16)	12[b] (85.7)	2 (14.3)	0.054
Sporadic[a]	(N=33)	14 (43.8)	18 (56.2)	
CYP1A2 and *NAT2*		CYP1A2 high activity & slow acetylator	CYP1A2 high activity & fast acetylator	
Group4	(N=16)	6[c] (42.9)	8 (57.1)	0.022
Sporadic[a]	(N=33)	4 (12.5)	28 (87.5)	
XPD exon 23 and *GSTM1*		*XPD* (Lys/Gln and Gln/Gln) & *GSTM1* null	*XPD* (Lys/Lys) & *GSTM1* positive	
Group 4	(N=16)	8[d] (57.1)	6 (42.9)	0.035
Sporadic[a]	(N=33)	8 (25.0)	24 (75.0)	

[a] Only sporadic who were not exposed to occupational exposures
[b] Three cases were exposed to other occupational exposures
[c] Two were exposed to other occupational exposures
[d] Three were exposed to other occupational exposures
[e] P-value were obtained by using χ^2-test

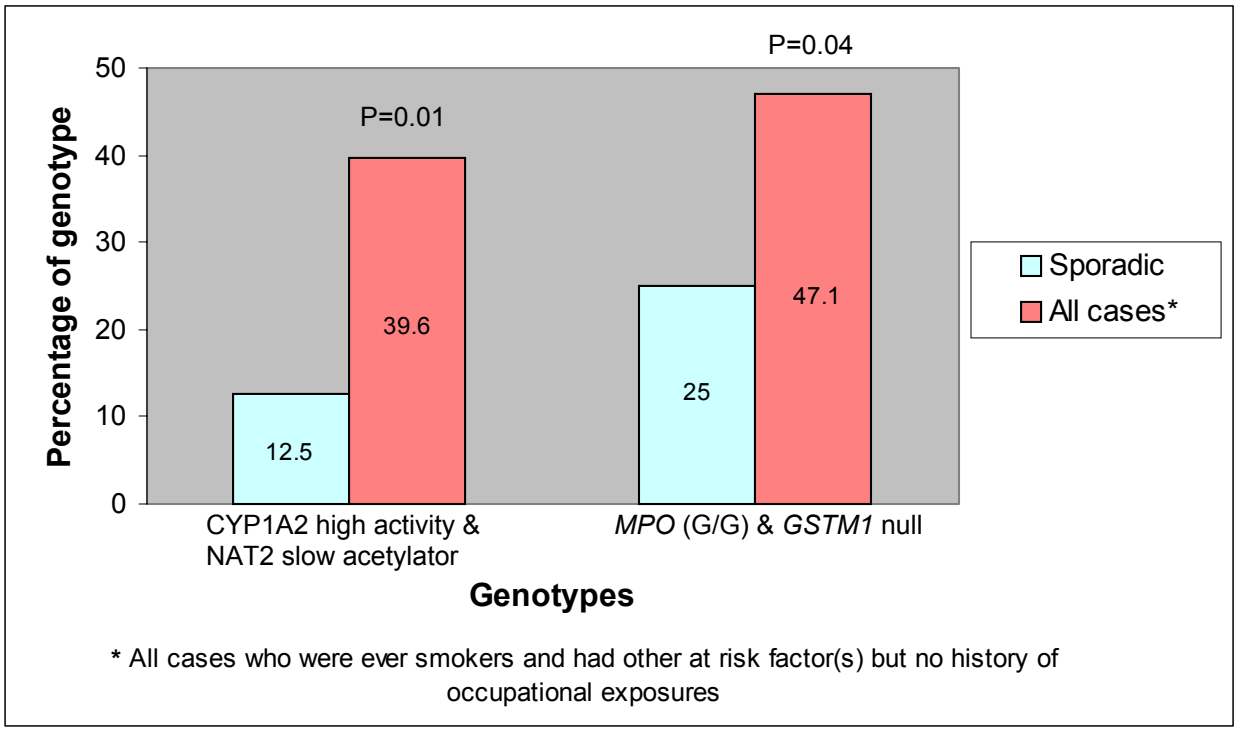


Figure 16. Interactions between multi at risk genotypes in relation to family history of cancer(s), and early age onset

Table 40. Associations of combined at risk genotypes and lung cancer risk among ever smoke lung cancer cases with family history of lung and other cancer(s), and early age onset

Combined genotypes	Number of individuals (%)		P-Value[c]
CYP1A2 and *NAT2*	CYP1A2 high activity & NAT2 slow acetylator	CYP1A2 high activity & NAT2 fast acetylator	
Group 1-5[a] (N=59)	21 (39.6)	32 (60.4)	0.008
Sporadic[b] (N=33)	4 (12.5)	28 (87.5)	
MPO and *GSTM1*	*MPO* (G/G) & *GSTM1* null	*MPO* (G/A and A/A) & *GSTM1* positive	
Group 1-5[a] (N=59)	24 (47.1)	27 (52.9)	0.044
Sporadic[b] (N=33)	8 (25.0)	24 (75.0)	

[a] All cases in group 1 to 5 who were not exposed to occupational exposures were combined
[b] Only sporadics who were ever smokers but had no history of occupational exposures
[c] P-value were obtained by using χ^2-test

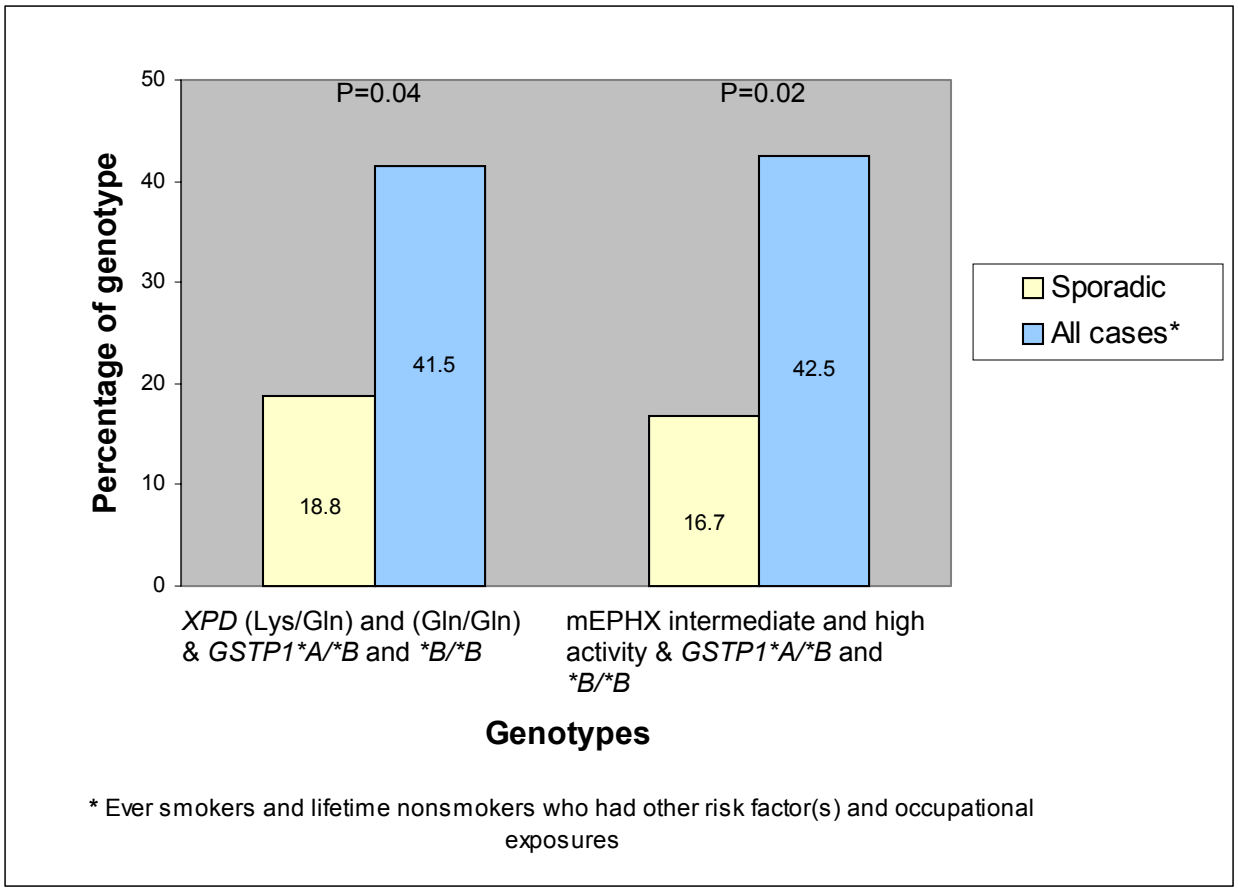


Figure 17. Associations of gene-gene interactions and lung cancer risk in relation to history of smoking, occupational exposures, family history of cancer(s) and early age onset

Table 41. Contribution of combined at risk genotypes in lung cancer development among lung cancer cases with history of smoking, occupational exposures, family history of lung and other cancer(s), and early age onset

Combined genotypes	Number of individuals (%)		P-Value[c]
XPD exon 23 and *GSTP1*	*XPD* (Lys/Gln and Gln/Gln) & *GSTP1*A/*B* and **B/*B*	*XPD* (Lys/Lys) & *GSTP1*A/*A*	
Group 1-5[a] (N=45)	17 (41.5)	24 (59.5)	0.038
Sporadic[b] (N=33)	6 (18.8)	26 (81.2)	
mEPHX and *GSTP1*	mEPHX high and intermediate activity & *GSTP1*A/*B* and **B/*B*	mEPHX low and very low activity & *GSTP1*A/*A*	
Group 1-5[a] (N=45)	17 (42.5)	23 (57.5)	0.021
Sporadic[b] (N=33)	5 (16.7)	25 (83.3)	

[a] There were eleven ever-smokers who were exposed to asbestos, thirty ever smokers and four lifetime nonsmokers who were exposed to others occupational exposures

[b] Only sporadic who were not exposed to occupational exposures

[c] P-value were obtained by using χ^2-test

4.3. mRNA expression

4.3.1. Lung sample characteristics

A total of twenty pairs of lung tumors and histological normal adjacent tissues were screened for CYP mRNA expression levels. These samples were obtained from three lifetime nonsmokers, three current smokers and fourteen former smokers. There were six women and fourteen men, with mean age ± SD of 71 ± 8 (range, 61-91 years) and mean packyear ± SD of 59 ± 31 (range, 14-130 packyear). Three patients were exposed to asbestos, and three patients had other occupational exposures. For histological cell type, eleven tumors were adenocarcinoma, and nine were squamous cell carcinoma. Each lung tumor sample was histologically examined by a pathologist who confirmed that it contained more than 80 percent tumor cells.

4.3.2. CYP mRNA expression in lung tumors and non-tumor tissues determined by RT-PCR FAF-ELOSA

The CYP1A1, CYP1B1, and CYP2E1 mRNA expression levels were quantitated by using the RT-PCR FAF-ELOSA method in lung tumors and corresponding nontumoral (histological normal adjacent) tissue samples. The β-actin gene was used as an endogenous control. The levels of CYP1A1, CYP1B1, and CYP2E1 mRNA expression in normal liver tissues sample were also quantitated as a reference of cytochrome P450 expression in humans. All three CYP mRNAs were detected in both lung tumors and non-tumors. The mean levels of CYPs relative to β-actin mRNA expression and their standard deviations are provided in Table 42. β-Actin mRNA was not detected in a few lung non-tumor and tumor tissues which resulted in missing data in the estimation of CYP mRNA quantitation. Moreover, there was not enough total RNA for some

samples to be used for determining all three CYP mRNA expression endpoints. Student's t-test was used to compare the mean level of CYP mRNA expression between lung tumor and histological normal adjacent tissues. The mean level of CYP1A1 mRNA quantity was similar between lung non-tumor and tumor tissues. Higher levels of CYP1B1 and CYP2E1 mRNA were observed in lung tumors compared to the mean level in non-tumors, although this was not statistically significant. However, the difference in mean level of CYP2E1 mRNA expression between tumor and non-tumor tissues was significant at P=0.04 after only matched tissue samples were included in comparison analysis (Table 43). The CYP2E1 mRNA was expressed in lung tumors more than in their corresponding non-tumors. Moreover, there was a trend of higher expression of CYP1B1 mRNA in lung tumors compared to the matched non-tumor tissue samples. The pattern of CYP1A1, CYP1B1, and CYP2E1 mRNA levels that expressed in two different tissue samples (liver and lung tissues) and in both lung tumors and non-tumors was showed in Fig. 18 and 19.

Table 42. CYP mRNA expression in lung histological normal adjacent and tumor tissues

CYP/β-actin (amol/amol)	Lung tissues	
	Non-tumor	Tumor
CYP1A1	0.1 ± 0.05[a]	0.18 ± 0.37[b,c]
CYP1B1	17.45 ± 11.68[d]	20.77 ± 20.24[e,f]
CYP2E1	1.98 ± 2.19[g]	3.40 ± 5.18[h,I]

[a] Mean ± SD of 11 samples
[b] Mean ± SD of 14 samples
[c] P = 0.44 compared to non-tumor tissues
[d] Mean ± SD of 9 samples
[e] Mean ± SD of 12 samples
[f] P = 0.47 compared to non-tumor tissues
[g] Mean ± SD of 16 samples including zeros where applicable
[h] Mean ± SD of 17 samples
[i] P = 0.16 compared to non-tumor tissues

Table 43. CYP mRNA expression in matched lung histological normal adjacent and tumor tissues

CYP/β-actin (amol/amol)	Lung tissues	
	Non-tumor	Tumor
CYP1A1	0.09 ± 0.05[a]	0.09 ± 0.14[b,c]
CYP1B1	15.08 ± 12.32[d]	29.99 ± 22.69[e,f]
CYP2E1	**2.03 ± 2.36[g]**	**5.13 ± 5.46[h,i]**

[a,b] Mean ± SD of 8 samples
[c] P = 0.89 compared to paired-non-tumor tissues
[d,e] Mean ± SD of 6 samples
[f] P = 0.27 compared to paired-non-tumor tissues
[g,h] Mean ± SD of 13 samples including zeros where applicable
[i] P = 0.039 compared to paired-non-tumor tissues

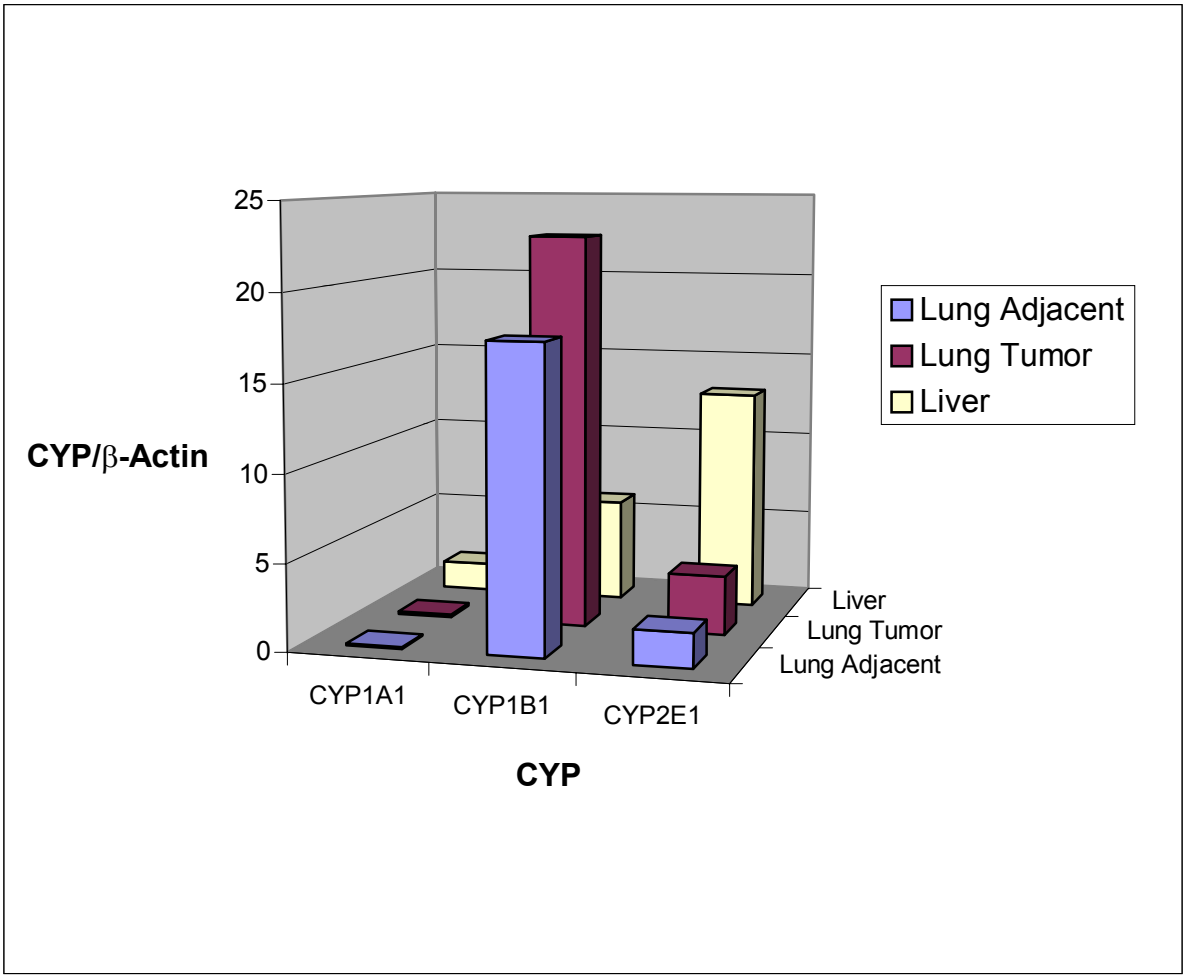
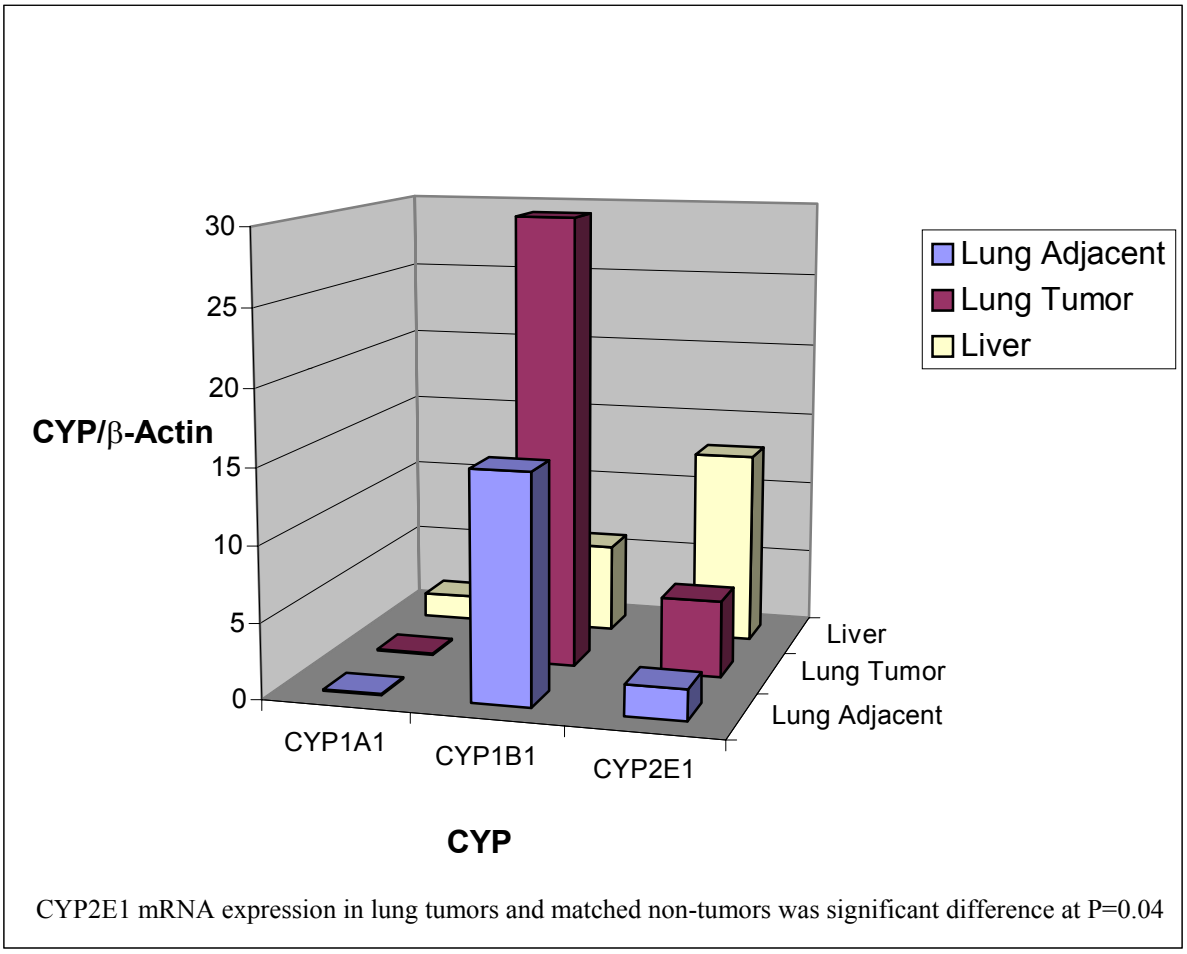


Figure 18. CYP mRNA expression in unpaired lung tumor and non-tumor tissue samples, and normal liver sample



CYP2E1 mRNA expression in lung tumors and matched non-tumors was significant difference at P=0.04

Figure 19. CYP mRNA expression in paired lung tumor and non-tumor tissue samples, normal liver sample

4.3.3. Comparison of the relative sensitivity of the RT-PCR FAF-ELOSA with the sensitivity of the TaqMan® assay for CYP1A1, CYP1B1, and CYP2E1 mRNA expression in normal liver and lung tissues

TaqMan assays for the relative quantitation of CYP1A1, CYP1B1, CYP2E1 and endogenous control β-GUS mRNA in human tissues were developed. The comparative C_T method is used to achieve the same result for relative quantitation as the standard curve that was used in the FAF-ELOSA analysis. A validation experiment was performed to assure that efficiencies of target and reference are approximately equal before using the $\Delta\Delta C_T$ method for quantitation. Then the relative expression of CYP1A1, CYP1B1, and CYP2E1 to endogenous control β-actin mRNA in the same liver tissue as determined by FAF-ELOSA were quantitated by TaqMan® assays (Fig. 20). The starting amount of total RNA used in FAF-ELOSA and TaqMan® assays were 1000 ng and 105 ng, respectively. The expression pattern of mRNA levels of CYP1A1, CYP1B1, and CYP2E1 relative to β-actin determined by FAF-ELOSA (Fig. 21) was similar to that of TaqMan.

After the percentage of each CYP mRNA expression to total CYP that quantitated by both FAF-ELOSA and TaqMan® assays were calculated and plotted in the same graph (Fig. 22), the levels of CYP1B1 mRNA expression was equal. However, the relative quantity of CYP1A1 mRNA expression detected by TaqMan® assay was slightly lower than that of FAF-ELOSA and vice versa for the CYP2E1. The differences in those levels of CYP1A1, and CYP2E1 mRNA expression determined by FAF-ELOSA and TaqMan® can be explained by the difference in the percentage values of the coefficient of variance (%CV) from both methods. The percentage of coefficient of variance for measuring CYP1A1 mRNA expression among three replications at each concentration by TaqMan® was between 0.19 and 1.43 which is smaller than the value

obtained by FAF-ELOSA (9.49-50) as shown in Fig. 23 and 24. In comparison with FAF-ELOSA for mRNA quantitation, the TaqMan® assay was associated with not only significantly lower %CV, but also could be performed using substantially lower starting amount of total RNA. The TaqMan® assay is a very sensitive method that could detect the CYP mRNA expression with very low starting amount of total RNA. Figure 25 shows the levels of CYP mRNA expression in normal lung tissue with only 1 ng starting amount of lung cDNA. In contrast, the FAF-ELOSA required 1000 nanogram of starting amount of RNA.

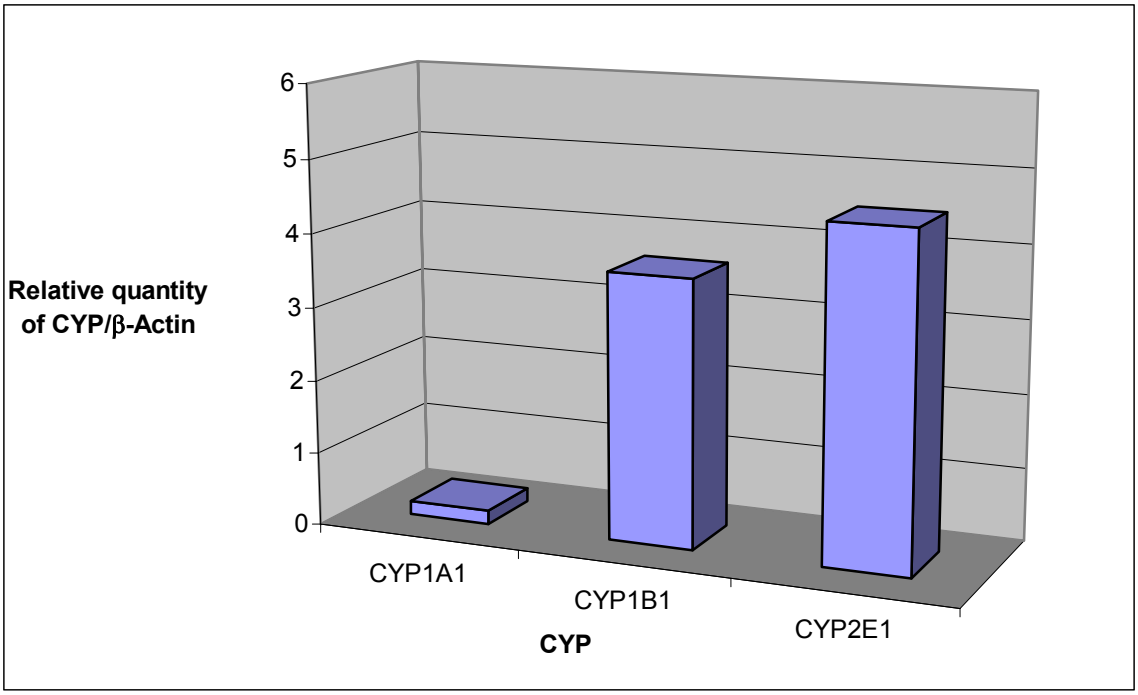


Figure 20. The relative quantity of CYP1A1, CYP1B1, and CYP2E1 to β-actin mRNA determined by the TaqMan® assay

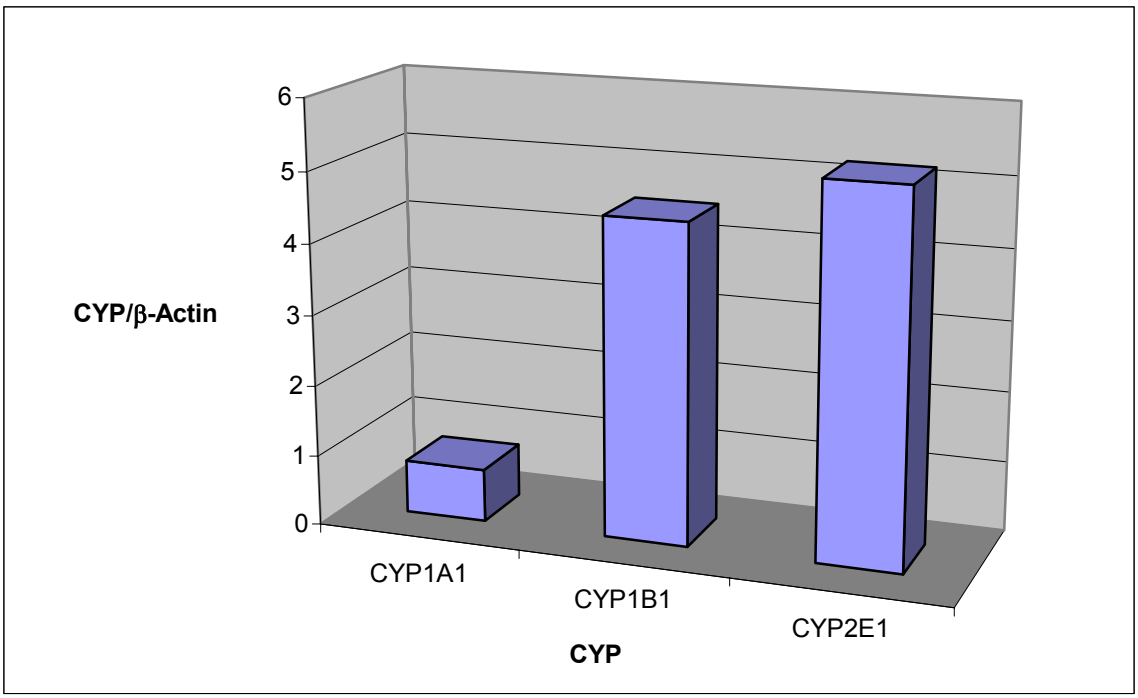


Figure 21. The relative quantity of CYP1A1, CYP1B1, and CYP2E1 to β-actin mRNA determined by FAF-ELOSA

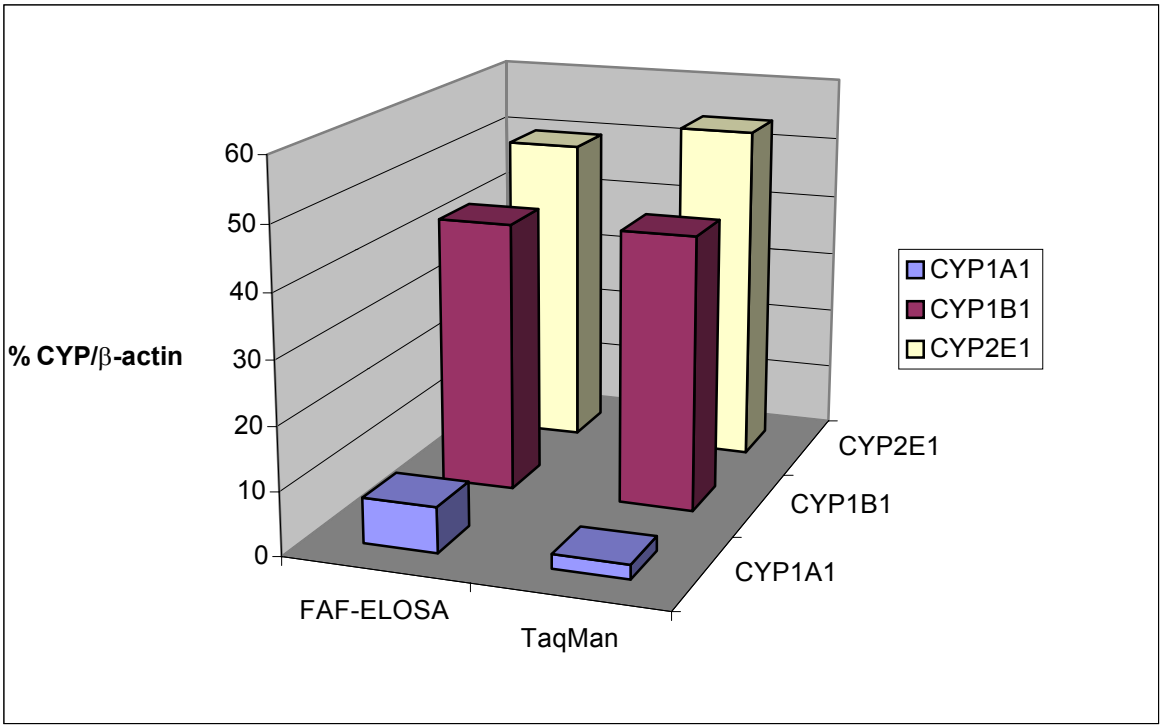


Figure 22. The pattern of CYP mRNA expression in liver tissue quantitated by both FAF-ELOSA and TaqMan® assays

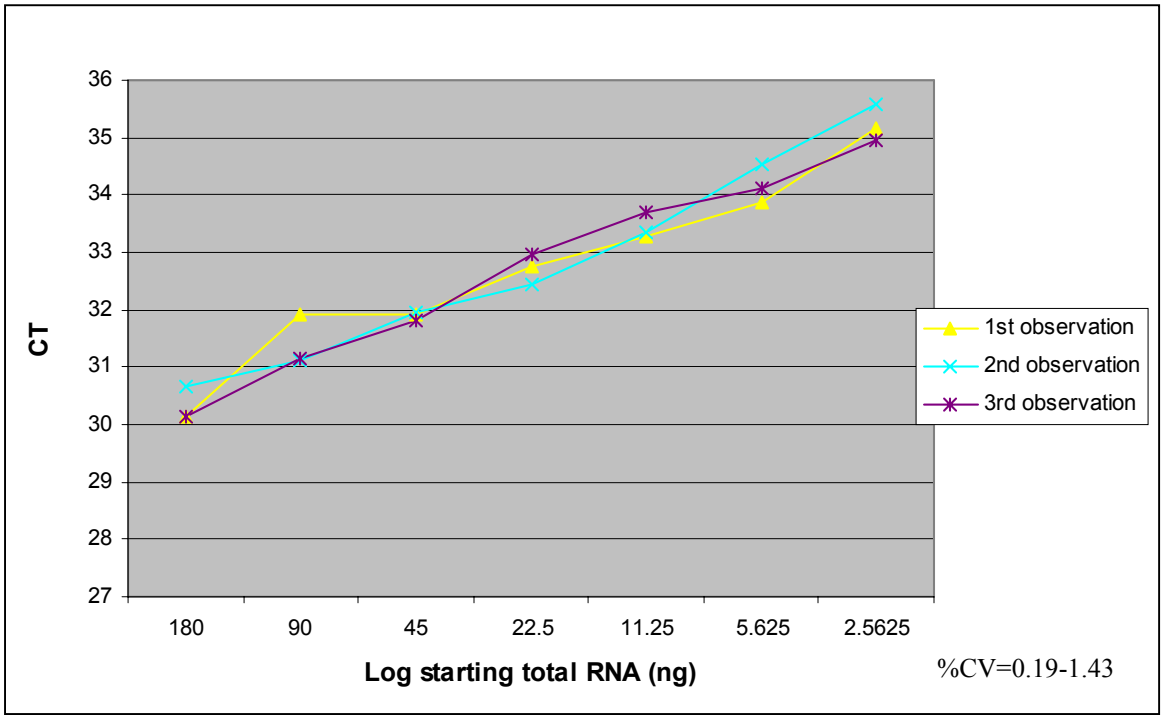


Figure 23. Plot of CT numbers from each observation and log starting total RNA (ng) of TaqMan® CYP1A1 PCR

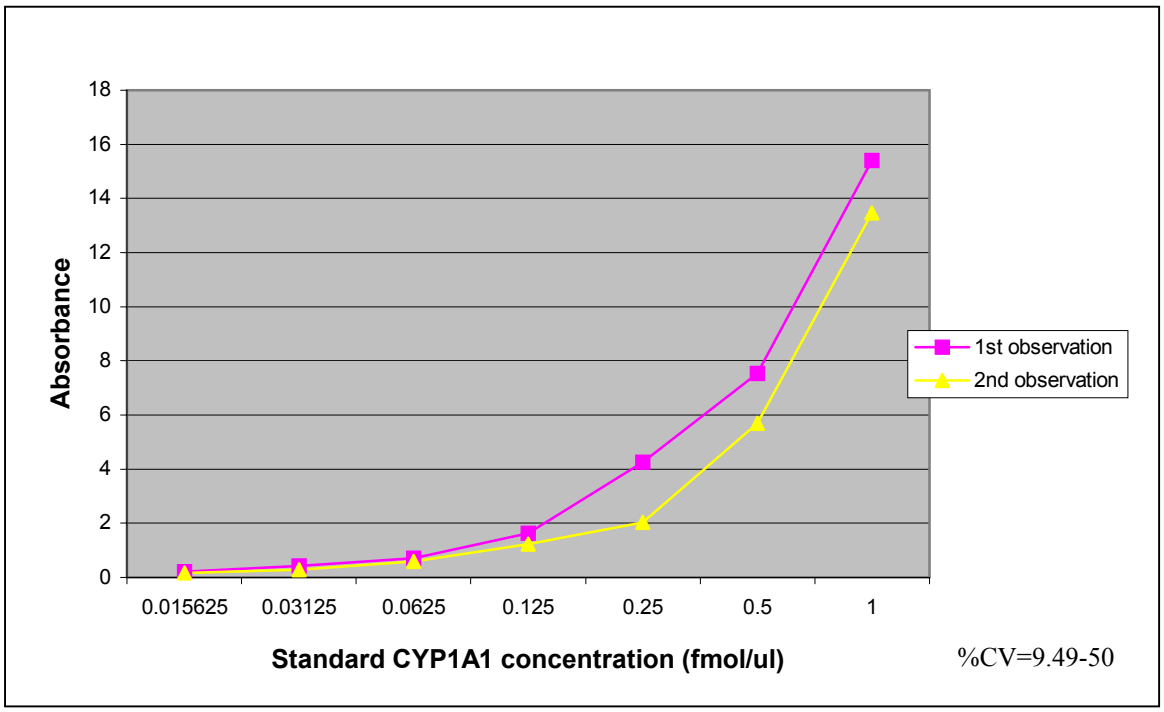


Figure 24. Plot of absorbance and CYP1A1 standard concentration for each observation in FAF-ELOSA assay

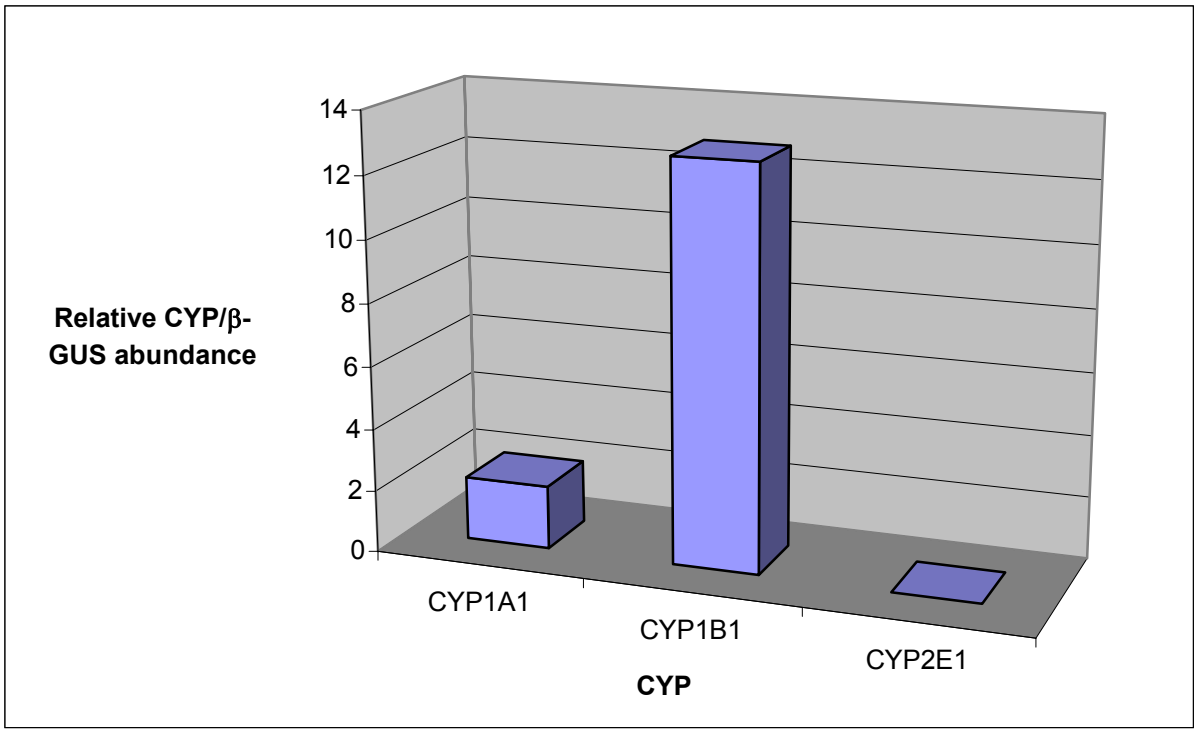


Figure 25. CYP1A1, CYP1B1, and CYP2E1 mRNA expression in normal lung tissue quantitated by TaqMan® assay with 1 nanogram of starting total RNA

DISCUSSION

Lung cancer has become a major public heath problem and a leading cause of cancer deaths for both men and women in the US. Even though tobacco and cigarette smoking are well recognized as a major cause of lung cancer, there is evidence that passive exposure, other occupational exposures such as asbestos, radon, arsenic, chromium, nickel, dietary carcinogens, and host factors also contribute to the risk of lung cancer. These host factors and other environmental exposures influence individual susceptibility to cigarette smoking on lung cancer risk. A large number of molecular epidemiological studies have focused on the study of host susceptibility variability such as differences in metabolism of cigarette smoking components, and DNA repair and risk of lung cancer. Many previous studies have evaluated the effects of a single gene polymorphism in metabolic enzymes or DNA repair genes on lung cancer risk. There are also several reports about the relationship of combined at risk genotypes for phase I and phase II metabolic enzyme genes and lung cancer susceptibility.

This research is the first molecular epidemiological study that investigated the association of lung cancer risk with a panel of genetic polymorphisms in metabolic enzymes that are involved in tobacco carcinogen metabolism and also of a DNA repair gene involved in the nucleotide excision pathway that repairs the bulky DNA adducts caused by polycyclic aromatic hydrocarbons in tobacco smoke. This is a major strength of this study that allowed analyses on gene-gene or gene-gene-gene interactions of several biologically plausible combinations among or between metabolic enzymes and DNA repair genes on lung cancer risk. Furthermore, a combination of other risk factors for lung cancer such as family history, early age onset, and

occupational exposures and genetic polymorphisms in metabolic enzymes and DNA repair genes was examined for possible of gene-gene and gene-environment interactions contributing to increased risk of lung cancer.

A population based case control study and case only analysis was conducted to determine whether any genetic variants in drug metabolizing enzymes and DNA repair gene represent lung cancer susceptibility biomarkers. There were some limitations in the case control analysis in this present study. Cases and controls were recruited from different areas, which might introduce a confounding effect due to differences in environmental exposures into the analyses of relative risk determination. Furthermore, the mean age of controls at the entry of the study was lower than those of cases, although this confounding effect was adjusted in all of the logistic regression models. Given the significant difficulty and the costs in selecting appropriate control subjects in the case control study, a recently proposed study design, a case only analysis, has been introduced as an approach to study the role of genetic factors in disease without controls (286). It has been used increasingly as an efficient and valid methodology for studying gene-environment interactions (287). Moreover, the number of subjects required in a case only study is substantially smaller than the number that would be required for a case control study for the same power to measure of gene-environment interaction (288). A case only analysis was used in this study not only to reduce the needed power to detect the effects of genetic and environmental exposures on lung cancer outcome but also to evaluate the contribution of genetic polymorphisms in metabolic enzymes and DNA repair genes in lung cancer risk among high risk lung cancer cases such as cases who had family history of lung and other cancers, and early age onset compared to cases who had none of those risk factors.

5.1. Case control study

In the case control study, associations between single gene loci polymorphisms in *CYP1A2, CYP1B1*, *mEPHX, GSTM3, GSTP1*, and *XPD* genes and lung cancer risk were observed. Moreover, significant contributions of gene-gene and gene-environment interactions of several at risk genotypes combinations in increasing risk of lung cancer were found.

5.1.1. Single at risk genotypes and lung cancer risk

5.1.1.1. *CYP1A2* variant alleles

The *CYP1A2*1C* and *CYP1A2*1F* variant allele frequencies were screened as a surrogate measure of CYP1A2 enzyme activity. Predicted levels of CYP1A2 activity was classified based on in vitro differences in enzyme activity, with approximate 23% reduction for the *CYP1A2*1C/*1C* genotype and 36% increases for the *CYP1A2*1F*1F* genotype (148-150). An association between risk of lung cancer and predicted high CYP1A2 activity was found. The risk increased approximately two-fold in ever smokers with predicted high CYP1A2 activity after adjustment for age, race, packyear, and occupational exposure. This finding was in agreement with the previous report of higher inducibility of CYP1A2 activity by tobacco smoke (150). Moreover, a significant increased risk of lung cancer was found in ever smokers with high CYP1A2 activity and *NAT2* slow genotype (age, race, and packyear adjusted OR=5.01; 95%CI, 1.25-20.02). CYP1A2 is involved in the metabolic activation of several carcinogens found in tobacco smoke and in high-temperature cooked meats such as aromatic and heterocyclic amines, and nitroaromatic compounds. Therefore, individuals who regularly smoke and consume high-temperature cooked meat in combination of having 'at risk' genotypes may be at greater risk of lung cancer. It is possible to hypothesize that among smokers who had higher inducible *CYP1A2*

genotype with or without the *NAT2* slow genotype, might have higher level of reactive metabolites of tobacco carcinogenic compounds formed in liver and transported to lung tissues and resulting in DNA adduct and carcinogenic process in lung epithelium cells.

CYP1A2 is also involved in the metabolism of endogenous substrates including estrogens, which may be associated with increasing risk of lung adenocarcinoma among non-smoking women. An association between high CYP1A2 activity and increasing risk of lung adenocarcinoma among nonsmoker females was not investigated here, given the low number nonsmokers in this study population. However, there was a trend of increasing risk of lung cancer among females and lung adenocarcinoma cases who were ever smokers and had high CYP1A2 activity (P=0.12 and 0.09, respectively). These observations could explain the higher risk among smoking females when compared to smoking males observed in previous studies. The contribution of high CYP1A2 activity in increased risk of lung cancer among smokers observed in this study suggests the importance of polymorphisms in the *CYP1A2* gene on lung cancer development.

5.1.1.2. *CYP1B1*3* variant allele

CYP1B1 is expressed predominantly in extrahepatic tissues such as lungs and mammary glands and has been known to be an important enzyme involved in the metabolism of polycyclic aromatic hydrocarbon, arylamines, and endogenous steroid hormones. It has been suggested that polymorphisms in the coding region of the *CYP1B1* gene may alter the catalytic activity of CYP1B1 and may determine interindividual differences in susceptibility toward lung cancer. The Leu432Val polymorphism in exon 3 of *CYP1B1* gene screened in the present study is located in the catalytically important heme-binding domain of the CYP1B1 enzyme. The role of *CYP1B1*3* variant allele as a lung cancer susceptibility biomarker was evaluated.

There were no statistically significant differences in the genotype frequencies for either cases or controls in the total population or in specific subgroups. There was an overall relationship between *CYP1B1* (**1/*3* and **3/*3*) genotypes with an interaction term between *CYP1B1* (**1/*3* and **3/*3*) genotypes and transformed packyear (packyear+packyear square) and lung cancer risk. The lung cancer risk was increased 2.6-fold among individuals who were ever smokers and had the 'at risk' *CYP1B1* genotype after adjusting for age, race, packyear, and occupational exposures. This significant association between *CYP1B1* (**1/*3* and **3/*3*) genotypes including a gene-environment interaction term and lung cancer risk has not been reported previously. An earlier study reported no relationship between *CYP1B1*3* genotype and lung cancer (99). The difference in results may in part be attributed to a gene-environment interaction (genotype-smoking interaction) that was found to be a significant interaction term adjusted for in the present logistic regression models.

The statistically significant interaction between *CYP1B1*3* genotype and transformed packyear was also observed among females and in adenocarcinoma subgroup (P = 0.02). However, the risk of lung cancer was only borderline significantly associated with the *CYP1B1*3* genotype in both subgroups (age, race, smoking status, packyear and a gene-packyear interaction term adjusted OR=3.69; 95%CI, 0.94-14.54, and adjusted OR=2.34; 95%CI, 0.99-5.53 in female and adenocarcinoma subgroups, respectively). The significant risk of lung cancer associated with the *CYP1B1*3* genotype which we observed for cases who were ever smokers and had history of occupational exposures might be explained by the gene-environment interaction and/or the altered function of CYP1B1 by genetic polymorphism. CYP1B1 is known to be induced by PAHs found in tobacco and occupational or environmental pollutants (257,263,289). Willey *et al.* (257) observed the higher levels of CYP1B1 expression in human

lung of smokers compared to those of nonsmokers. Although the effect of the *CYP1B1*3* genotype on catalytic activity of CYP1B1 towards PAHs, and aryl amines has not been fully evaluated, it is possible that the Leu432Val substitution may affect the CYP1B1 protein structure resulting in changes in its function. Therefore, mutation and DNA damage that already occurred in lung tissue of cases from exposures to asbestos, chromium, silicon, radon together with that caused by higher level of CYP1B1 induced by PAHs found in cigarette smoke and/or altered catalytic activity of CYP1B1 by Leu432Val substitution towards cigarette carcinogens may accelerate the pathway of lung cancer development.

Another interesting observation in this present study was the increased lung cancer risk by the *CYP1B1*3* genotype found among females and in the adenocarcinoma subgroup. Although the associations were not statistically significant, they were consistent with several recent reports of higher incidence of lung adenocarcinoma in females and the involvement of female sex hormones in higher risk among smoking females compared to smoking males. CYP1B1 has been reported to catalyze 17β-estradiol 4-hydroxylation to give 4-hydroxyestradiol, a pathway which has been suggested to be involved in the development of breast cancer in human (97,290,291). Shimada *et al.* (160) reported the relationship between the *CPY1B1*3* variant allele and higher activity in catalyzing activation of 17β-estradiol 4-hydroxylation by CYP1B1. Therefore, individuals who carry the genetic polymorphism in exon 3 of the *CYP1B1* gene may have higher activity of CYP1B1 involved in estrogen metabolism resulting in high levels of potentially carcinogenic 4-hydroxyestradiol and ultimate cancer development. Therefore, a significant increase in CYP1B1 protein levels by cigarette smoke induction in female smokers who also had *CYP1B1*3* genotype may result in higher levels of DNA damage from cigarette carcinogen reactive metabolites and 4-hydroxyestradiol contributing to lung carcinogenesis. This could in

part explain the differences in risk of lung cancer between female and male smokers. Furthermore, it implicates the potentially important role of CYP1B1 in lung carcinogenesis and of the *CYP1B1*3* genotype as a lung cancer susceptibility biomarker.

5.1.1.3. *mEPHX* variant alleles

Two allelic variants of *mEPHX* that are located within the coding region of the *mEPHX* gene at exons 3 and 4 and are involved with decreasing and increasing mEPHX activity by 40% and 25%, respectively, were genotyped to predict hydrolase enzyme activities (high, intermediate, low, and very low). Interindividual variation in mEPHX activity is likely to contribute to lung cancer susceptibility because of its dual role in the detoxification and activation of procarcinogens, depending on the substrate. For example, it is considered an important detoxification enzyme in catalyzing the hydrolysis of highly reactive epoxide intermediates to less reactive and more water soluble trans-dihydrodiol derivatives. An example of the activation role of mEPHX is the generation of trans-dihydrodiols from polycyclic aromatic hydrocarbons, which are then substrates for further metabolism to more reactive carcinogenic polycyclic hydrocarbon diol epoxides such as BPDE.

Previous studies exploring the relationship between *mEPHX* genotypes and lung cancer risk have yielded inconsistent results. Among five Caucasian studies, there were two studies that found no significant relationship between *mEPHX* genotype and lung cancer risk (229,292). Two other studies found the high mEPHX enzymatic activity had an elevated lung cancer risk (227,293). Zhou *et al.* (225) reported the effects of very low activity of *mEPHX* genotype on increasing risk of lung cancer among nonsmokers and light smokers but on decreased lung cancer risk among heavy smokers.

In this study, the relationship between predicted mEPHX high and intermediate activity and increasing risk of lung cancer was observed only among females. The lung cancer risk increased approximately 2.7 fold after adjusting for age, race, smoking status, and packyear. These female ever smokers were moderate smokers with mean packyear of 48.25±33.23 (range, 15-160). There were two individuals who had packyears of 100 and 160. This finding is in the same direction as that of Benhamou *et al.* (227) and Zhao *et al.* (293), observing on increased risk of lung cancer for high mEPHX enzyme activity among smokers. The mEPHX enzyme is highly expressed in several tissues including lung and can be induced by cigarette smoking. Therefore, there is the potential for increases in the activation of polycyclic aromatic hydrocarbon compounds found in cigarette smoke by mEPHX enzyme among smokers who also have high activity of mEPHX enzyme, thereby resulting in higher concentrations of polycyclic aromatic hydrocarbon reactive metabolites and DNA adducts in the lung.

An association between high activity of mEPHX enzyme and lung cancer risk was not found in all cases, and other subgroups. In contrast, there was a trend towards a protective effect of high mEPHX enzyme activity on lung cancer risk in the male subgroup. The packyear distribution among males was not remarkably different from that of females although males were more likely to be moderate to heavy smokers. Differences in associations between genders can result from other environmental and occupational exposures, for example a higher ratio of occupational exposures history among males, additional allelic variants that modulate overall metabolic enzyme activity, and endogenous hormones. The observations in this study suggest the possibility of both variants in *mEPHX* gene as a lung cancer susceptibility biomarker.

5.1.1.4. *Glutathione S-transferase* (*GST*) variant alleles

Glutathione S-transferases are involved in the detoxification of several reactive tobacco related metabolites and carcinogens. Polymorphisms in the GST Mu, Theta, and Pi class genes, which result in variation in glutathione S-transferase enzyme activity may influence the susceptibility to lung cancer. The effect of interindividual differences in GST enzyme activity related to the *GSTM1* null genotype in modulating the lung cancer susceptibility has been extensively studied. To a lesser degree, studies have also focused on the relationship between *GSTT1* null, *GSTM3*A/*A*, *GSTP1*B*, and *GSTP1*C* and lung cancer risk. Again however, the results were inconsistent. Neither *GSTM1*, nor *GSTT1* null genotypes contributes significantly to the risk of lung cancer in the present study. The prevalence of *GSTM1* null genotype was similar among cases (57.6%) and controls (58%). The *GSTT1* null genotype was underrepresented in cases (22%) compared to controls (34%). These could explain the lack of an association between *GSTM1* and/or *GSTT1* null genotypes and increased risk of lung cancer that observed in this study as well as in other several studies (184,185,241,294).

The genetic polymorphisms of *GSTM3*, and *GSTP1* and lung cancer susceptibility have been recently been studied since their proteins are abundant in human lung tissues. The *GSTM3*B* variant allele has been postulated to result in increase GSTM3 transcription. It has been reported that this allele is linked with increased GSTM3 expression and associated with decreased risk for lung adenocarcinoma (240). There are a few other studies evaluating the role of this polymorphism and lung cancer risk but no significant association was found (184,185). This finding either is or is not a consequence of the similar distribution of *GSTM3* genotype between the lung cancer cases and the controls. The prevalence of the *GSTM3*B*B* genotype observed in our controls (0.04) in the present study was similar to that observed in previous studies among Caucasians (181,185,284). Conversely, there was a significant higher frequency of the

*GSTM3*A/*A* genotype among cases compared to controls in this study population. The *GSTM3*A/*A* genotype have shown increased overall lung cancer risk (age, race, smoking status, packyear, and occupational exposures adjusted OR=1.84; 95%CI, 1.03-3.31). Moreover, the risk of lung cancer was increased 3-fold among females. These results suggest the potential of *GSTM3*A/*A* genotype as a modifier of lung cancer susceptibility.

Polymorphisms in exon 5 and 6 of *GSTP1* result in two genetic variants, *GSTP1*B* and *GSTP1*C*. The variant allele *GSTP1*B*, has been shown to be associated with reduced GST activity in lung tissue (113). Ryberg *et al.* (190) found a significant increasing risk of lung cancer among male lung cancer patients who had *GSTP1*A/*B* and *GSTP1*B/*B* genotype. The correlation between *GSTP1*B* genotype and lung cancer risk was not observed for other Caucasian populations among males or both genders (184,185,191,294). A significant increased risk of lung cancer among male ever smokers who had *GSTP1*A/*B* and *GSTP1*B/*B* genotype (age, race, and smoking status adjusted OR=2.09; 95%CI, 1.03-4.25) observed in the present study was similar to a 1.7-fold increase in the risk of lung cancer that reported by Ryberg *et al.* (190). This finding indicates *GSTP1*B* genotype may also contribute to lung cancer susceptibility.

GSTM1, GSTM3, GSTP1, and GSTT1 are all known to catalyze detoxification of active intermediates of PAHs, with considerable overlap in substrate specificity for GSTM1 and GSTP1. Tissue specificity of expression of the GST enzymes is observed. GSTM1 and GSTT1 are mainly expressed in liver, with very low levels expressed in the lung. In contrast, GSTM3 and GSTP1 are highly expressed in lung. The relative level of combined GSTs expressed in a target tissue may influence detoxification capability. Lung tissues are exposed to carcinogens from both inhalation and bloodstream distribution. Lack of GSTM1 and GSTT1 activities may

therefore increase reactive metabolite presence in the lung. Variability in the expression of GSTM3 and GSTP1 due to genetic polymorphisms in the lung may also influence the detoxification of carcinogens. This data taken together suggests that both *GSTM3* and *GSTP1* genotypes may be important in lung cancer etiology.

5.1.1.5. *XPD* exon 23 variant allele

The *XPD* gene encodes for XPD proteins involved in the nucleotide excision repair (NER) pathway that removes DNA bulky adducts induced by exposure to tobacco related carcinogenic compounds. Several genetic polymorphisms in the *XPD* gene have been identified. There are four amino acid substitution variants that exist in coding regions of *XPD* gene. The effects of these polymorphisms on DNA repair capacity have not been fully explored. However, there is evolutionary conservation at the sites of these amino acid substitutions suggesting a functional relevance. Furthermore, NER consists of several proteins that work together as a large protein complex in a DNA repair process. It is likely that the amino acids changes could affect the XPD protein structure, and therefore bindings with other proteins in the NER pathway, and alter DNA repair efficiency. Consequently, it has been hypothesized that genetic polymorphisms in *XPD* gene may influence risk of lung cancer. Two polymorphic variants of *XPD* that have been studied as possible modulators of cancer risk are the polymorphisms at Asp312Asn in exon 10, and Lys751Gln in exon 23 because of their high allele frequencies in the population (0.42, and 0.30, respectively). In this study, we only screened the *XPD* exon 23 variant allele.

The relationship between the *XPD* exon 23 polymorphism and DNA repair capacity has been studied but there are inconsistent results. Lunn *et al.* (233) first reported the *XPD*-751Lys common allele may alter the XPD protein product resulting in suboptimal repair of X-ray-induced DNA damage in 31 Caucasian women. Similarly, an association between *XPD*-

751Lys/Lys genotype and increased risk of basal cell carcinoma was observed by Dybdahl *et al.* (295). In contrast, the *XPD*-751Gln variant allele was found to be associated with less optimal DNA repair capacity and increased prostate cancer risk in study 66 cases and 54 controls by Hu *et al.* (296). Sturgis *et al.* (297) also reported the effect of the variant allele of *XPD* exon 23 on increased risk of squamous cell carcinoma of head and neck. Similarly, Spitz *et al.* (298) reported a correlation of the *XPD*-751Gln variant genotype to reduced DNA repair capacity in a study on DNA repair phenotype and *XPD* genotypes among 360 healthy controls and 341 lung cancer cases. The observed lower DNA repair capacity associated with the *XPD*-751Gln variant allele suggests a functional significance of this polymorphism.

To date, no overall relationship between the *XPD* (Lys751Gln) polymorphism and lung cancer risk has been reported in several studies (234,298-300). The present study is the first study found an overall significant relationship between *XPD*-751Gln variant genotype and increasing risk of lung cancer. After adjusting for age, race, smoking status, and/or packyear, and/or occupational exposures, the *XPD* (Lys751Gln) polymorphism contributed a 2-3-fold increased risk of lung cancer in all cases and in specific subgroups of gender, histological subtypes, and no history of occupational exposures. No significant associations between the *XPD*-751Gln variant genotype and lung cancer risk has been reported in previous studies (234,298-300) (Table 44). However, in the study of Davis-Beabes *et al.* (299), the percentages of *XPD* Gln/Gln homozygous variant allele were higher in cases compared to that of controls, but the difference did not reach statistical significance. A slight increase in the percentage of the homozygous variant allele among cases was also observed in two other studies by Spitz *et al.* (298) and Zhou *et al.* (300). It is important to note that the *XPD*-751Gln variant allele frequency of 0.29 among controls observed in this study was identical to that of 0.29 first reported by Shen

et al. (121), 0.26 by Lunn *et al.* (233), and 0.33 by Spitz *et al.* (298). Therefore, our significant findings of an elevated risk in lung cancer cases with the *XPD*-751Gln variant allele in this study were not due to unusual allelic frequencies.

Table 44. Genotype frequencies and lung cancer risk in relation to the *XPD* exon 23 polymorphism observed in different studies

Genotype	Cases N (%)	Controls N (%)	Adjusted OR[b] (95%CI)	Reference
XPD exon 23				
Lys/Lys	67 (37.6)	197 (43.5)	1.00	(299)
Lys/Gln	77 (43.3)	198 (43.7)	0.97 (0.62-1.52)	
Gln/Gln	34 (19.1)	58 (12.8)	1.34 (0.74-2.42)	
Lys/Gln+Gln/Gln	111 (62.4)	256 (56.5)	1.06 (0.70-1.61)	
Lys/Lys	141 (41.3)	159 (44.2)	1.00	(298)
Lys/Gln	153 (44.9)	162 (45.0)	1.07 (0.78-1.47)	
Gln/Gln	47 (13.8)	39 (10.8)	1.36 (0.84-2.20)	
Lys/Lys	428 (39.2)	499 (40.2)	1.00	(300)
Lys/Gln	498 (45.6)	575 (46.4)	0.98 (0.8-1.2)	
Gln/Gln	166 (15.2)	166 (13.4)	1.06 (0.8-1.4)	
Lys/Lys	77 (40.3)	97 (54.8)	1.00	Present study
Lys/Gln	87 (45.6)	59 (33.3)[a]	2.11[c] (1.21-3.67)	
Gln/Gln	27 (14.1)[d]	21 (11.9)	1.64[c] (0.74-3.64)	
Lys/Gln+Gln/Gln	114 (59.7)	80 (45.2)	2.56[e] (1.45-4.51)	

[a] Significant difference from genotype frequency of cases at P = 0.006 obtained from two-sided χ^2 test
[b] Adjusted OR for age, sex, and smoking status
[c] Adjusted OR for age, race, and smoking status
[d] Not significant higher than that of controls at P = 0.14 obtained from two-sided χ^2 test
[e] Adjusted OR for age, race, smoking status (packyear), and occupational exposures

5.1.2. Gene-gene and gene-environment interactions and lung cancer risk

It is clear that multiple genes are involved in cigarette smoke carcinogen metabolism. An imbalance of enzyme activities resulting from genetic polymorphisms in the genes involved in activation, detoxification, and repair may influence risk of lung cancer. Therefore, it is plausible

that interactions among or between metabolic and repair genes would be expected to affect the risk of lung cancer. Although a number of studies have evaluated gene-gene interactions for metabolic genes, this is the first study to report a significant interaction between the DNA repair enzyme, *XPD* (Lys/Gln and Gln/Gln), and 'at risk' genotypes for activation or detoxification of cigarette carcinogen metabolism. Moreover, the significant associations between combined predicted high and intermediate mEPHX activity and other 'at risk' genotypes in lung cancer risk that occurred only among female lung cancer patients were first observed in this study. The joint effects of a combination between two different 'at risk' genotypes and/or cigarette smoke exposure on increased risk of lung cancer observed in this present study included additive, synergistic, and multiplicative effects. An additive effect is described by the combined effect of two different at risk genotypes that is equal to the sum of their individual effects. If the effect of two combined genotypes on an increased risk of lung cancer appears to be greater than the effect genotype individually, or the sum of individual effects, this is called a synergistic effect. A multiplicative effect is defined by the joint effect of two main factors, which is the product of their effects.

The *XPD* (Lys/Gln and Gln/Gln) genotype, possibly reflecting reduced DNA repair capacity interacted with the *CYP1B1* (**1/*3* and **3/*3*) genotype, which may influence the catalytic activity of CYP1B1 enzyme toward PAHs and aryl amines and has been shown to be involved with higher rate of 17β-estradiol 4-hydroxylation, was found to be associated with 2-3-fold increased risk of lung cancer in all cases, adenocarcinomoa and squamous cell carcinoma subgroups. The interaction between these two 'at risk' genotypes contributed to an imbalance of activation and repair pathways of carcinogenic compound metabolism. The increase in risk of lung cancer conferred by the interaction effect of the *XPD* (Lys/Gln and Gln/Gln) and *CYP1B1*

(**1/*3* and **3/*3*) genotypes compared to the risk determined by each individual gene was stronger among females (OR=5.6) than in all cases, adenocarcinoma, and squamous cell carcinoma subgroups.

The joint effect between the *XPD* and *CYP1B1* 'at risk' genotypes among females appears to be more than an additive effect, i.e., a synergistic effect. In contrast, additive effects of these two polymorphic genes were indicated in all cases, and other subgroups. This could be a result of an important role of CYP1B1 in estrogen metabolism that may enhance the carcinogenic process caused by cigarette smoking. Estrogen metabolism can influence the carcinogenic process by the formation of catechol estrogens, the 2-OH and 4-OH derivatives, mainly by CYP1A1 and CYP1B1 respectively (301). Generally these two catechol estrogens are mainly inactivated by *O*-methylation and other possible conjugations such as glucuronidation and sulfation. If these conjugations are incomplete, the 2-OH and 4-OH catechol estrogens are oxidized to semiquinones and quinones (302). The quinones are reactive eletrophilic metabolites, which are capable of forming DNA adducts. Further DNA damage results from quinone-semiquinone redox cycling generated by enzymatic reduction of catechol estrogen quinones to semiquinones and subsequent auto-oxidation back to quinones (301). In addition, estrogen 4-hydroxylation that primarily catalyzed by CYP1B1 had received particular attention because of the fact that the 2-OH and 4-OH catechol estrogens differ in carcinogenicity (301). The 4-OH catechol estrogen significantly increased 8-hydroxydeoxyguanosine levels and induced DNA single strand break whereas 2-OH catechol estrogen had negligible effects (301). Furthermore, comparison of the corresponding catechol estrogen quinones showed that the quinones that derived from 4-OH derivatives produced two to three orders of magnitude higher levels of depurinating adducts than the those quinones derived from 2-OH derivatives (302). Given the carcinogenic potential from

estrogen metabolism and the findings that support the causative role of 4-OH catechol estrogens in carcinogenesis, it implicates the synergistic action of estrogen-mediated carcinogenesis on lung cancer risk related to cigarette smoke exposure and the joint effect of the *XPD* and *CYP1B1* 'at risk' genotypes among females.

Similarly, a significant association for increased risk of lung cancer was the combination of the *XPD* (Lys/Gln and Gln/Gln) genotype and predicted high and intermediate mEPHX activity observed in all cases, non occupational exposures, female and adenocarcinoma subgroups. The odds ratio of 5.81 obtained among females with these two combined genotypes was nearly twice greater than the value of 2.72 for the *XPD* (Lys/Gln and Gln/Gln) genotype alone and of 2.41 with predicted high and intermediate mEPHX activity. This result indicates more than an additive effect but less than a multiplicative effect. In contrast, the 2-2.4-fold increased in risk with combined genotypes in all cases, adenocarcinoma and non-occupational exposures subgroups probably resulted mainly from the effect of the *XPD* at risk genotype as the odds ratio determined by combined genotypes were similar to those predicted by the *XPD* (Lys/Gln and Gln/Gln) genotype alone. Moreover, there was no overall relationship between 'at risk' genotypes of *mEPHX* and lung cancer risk in all cases, adenocarcinoma, and non- occupational exposures subgroups. Interestingly, the relative risk was increased to six-fold among lung adenocarcinoma cases who had combined genotypes of the *XPD* (Lys/Gln and Gln/Gln) and predicted intermediate mEPHX activity. The interaction effect between these two 'at risk' genotypes was more than a multiplicative effect. Several magnitudes increased in risk of lung cancer among these lung cancer cases who were heavy ever smokers could be a result of an imbalance of activation and repair pathways of cigarette smoke carcinogenic compound

metabolism, genetic effect, which further interacted with environmental factor, cigarette smoke exposure.

Further, a significant increase in lung cancer risk determined by the interaction between the *XPD* (Lys/Gln and Gln/Gln) and the *MPO* (G/G) genotypes was observed among females, squamous cell carcinoma, and non occupational exposures subgroups. Again, this is likely a result of an increased activation of cigarette carcinogen metabolism and decreased DNA repair capacity. However, the odds ratio values determined from the gene-gene interaction were not different from those values predicted by the *XPD* (Lys/Gln and Gln/Gln) genotype alone. Therefore, it is likely that genetic polymorphism in the *XPD* exon 23 gene shows a stronger effect on modulating lung cancer risk than the effect of polymorphism in *MPO* gene when these two polymorphisms were combined.

The effects of gene-gene interaction between genes, which encode the detoxification and DNA repair enzymes on increased lung cancer risk were first observed in this present study. The *GSTP1*B* variant allele has been shown to be associated with reduced enzyme activity and decreased detoxifying capacity against the reactive intermediates occurred during cigarette carcinogen metabolism. Moreover, suboptimal DNA repair capacity that related to the genetic polymorphism in the *XPD* exon 23 further accelerates the carcinogenic process initiated by the reactive intermediates. The additive effect of combined polymorphic genes in the *XPD* and *GSTP1* on 2.8-fold increased in lung cancer risk was found among males. The increase in risk modulated by the effect of interaction between these two at risk genotypes in all cases who had or had no history of occupational exposures appears to be mainly effect of the *XPD* (Lys/Gln and Gln/Gln) genotype alone. The higher proportion of heavy smokers among males than females in this study population together with the distribution of females in the groups of all cases with or

without history of occupational exposures may explain why the combined two polymorphic variants of the *XPD* and *GSTP1* showed significant additive effect on lung cancer risk only among male population.

Another combination of GST and DNA repair enzymes that has been hypothesized to influence the host susceptibility to lung cancer and first examined in this study was the interaction between the *XPD* (Lys/Gln and Gln/Gln) and *GSTM3*A/*A* genotypes. GSTM3 is highly expressed in human lung tissues. The polymorphism in the *GSTM3* gene called the *GSTM3*B* variant allele has been shown to increase GSTM3 transcription and result in higher level of GST enzyme. Therefore, individuals who carry a wild type *GSTM3*A/*A* allele, that is not inducible, may have lower level of GST enzyme in their lung tissues compared to individuals who carry a *GSTM3* polymorphic variant allele. In this study population, there was an overrepresentation of *GSTM3*A/*A* at risk allele among cases. Also, the *XPD* Gln751 allele was over represented among cases. Together, the combined effect of these two at risk genotypes on increased risk of lung cancer were significantly pronounced in all cases and each specific subgroup.

The higher degree of risk conferred by the combination of these two 'at risk' genotypes compared to the risk estimated by individual gene separately was indicated in all cases and among females. This result was in accordance with the hypothesis that the combination of the putative 'at risk' genotypes could result in higher risk. This is also supported by certain biological findings. The interaction effect observed in adenocarcinoma and non occupational exposures subgroup is likely due to the main effect of *XPD* 'at risk' genotype. In contrast, the *GSTM3*A/*A* genotype demonstrated significant modifying the risk of lung cancer contributed by the *XPD* (Lys/Gln and Gln/Gln) genotype in squamous cell carcinoma and male subgroup

after combined these two at risk genotypes together. This finding indicates the combination of different genetic polymorphisms may have additive effects on lung cancer risk.

A significant association was observed for concurrent decreases in the level of GST enzymes as a result of polymorphism in the *GSTP1* gene and the *GSTM3*A/*A* allele and susceptibility to squamous cell carcinoma. Smokers having both deficient *GSTM3* and *GSTP1* genotypes had a significant increased risk 2.75-fold to lung squamous cell carcinoma. This significant increase in risk of lung cancer linked to the joint effect of *GSTP1*B* and *GSTM3*A/*A* genotypes gives some support to the hypothesis of combination of at risk *GST* genotypes appear to be important determinants of lung cancer rather than considered separately. Moreover, this finding may be in accordance with other studies that indicating the role of GSTs in proximal or bronchial parts of the lung where squamous cell carcinomas are usually located (190,240). The bronchial lung is exposed to cigarette carcinogens directly by inhaled cigarette smoke and also to reactive metabolites from liver metabolism by systemic circulation. The GSTM3 and GSTP1 are highly expressed in human lung tissues. Therefore, deficiencies in lung GSTs associated with *GSTP1*B* and *GSTM3*A/*A* genotypes would affect the host susceptibility to lung cancer especially for squamous cell carcinoma histological subtype.

To date, the associations between the interactions of *mEPHX* and *CYP1B1*, *MPO*, and *GST* polymorphisms and lung cancer risk have not been explored. This study first found the contribution of combined these 'at risk' genotypes to lung cancer risk among females. Two significant interactions among enzymes involved in the activation pathway of cigarette carcinogens metabolism that affected the risk of lung cancer only among females that observed in this study were the combinations of predicted high and intermediate mEPHX activity with *CYP1B1* (**1/*3* and **3/*3*) or *MPO* (G/G) genotypes.

Both CYP1B1 and mEPHX catalyze an activation step in the metabolism of PAH compounds found in cigarette smoke. Increasing activities of both enzymes caused by genetic polymorphisms could potentially affect a balance of the carcinogen metabolism pathways. A nine-fold increased risk of lung cancer after adjusting for age, race, smoking status, and packyear among females who carried *CYP1B1*3* genotype together with having high and intermediate mEPHX activity suggests more than a multiplicative effect for joint effect of these two polymorphic genes as lower degree of risk were observed when studied separately. Gene-gene, gene-environment (*CYP1B1*3* and transformed packyear), and gene-gene-environment (*mEPHX* and *CYP1B1*3* and transformed packyear) interactions that appeared to be significant first and second order of interaction terms adjusted in logistic regression model show this highest risk of lung cancer among females with these combinations.

CYP1B1 has been reported as a very active enzyme in catalyzing the activation of PAHs such as B[*a*]P, and B[*a*]P-7,8-dihydrodiol (160,303). For B[*a*]P metabolism, the presence of mEPHX shifts formation away from phenols to diol metabolites (303). The 7,8-dihydrodiol metabolite is further metabolized by the P450 enzymes of the CYP1 family including CYP1B1 to the mutagenic (±)-BPDE species which is extremely reactive and can bind to macromolecules (303). Both CYP1B1 (257) and mEPHX (257,304) are expressed in human lung tissues. Therefore, the joint effect of smoking induced CYP1B1 expression, altered function of CYP1B1 by polymorphism at position 432, and higher mEPHX activity can result in higher levels of DNA adducts, DNA damage and lung cancer development. Moreover, this carcinogenesis process can be enhanced by the effect of the *CYP1B1*3* genotype on increasing catalytic activity towards 17β-estradiol 4-hydroxylation as described previously.

An increasing in the rate of the mutagenic B[*a*]P diol-epoxide formation can also be influenced by the combined effects of higher mEPHX and MPO activities in lung tissues. MPO is expressed in neutrophils and metabolically activates a wide range of tobacco smoke carcinogens and environmental pollutants to DNA-damaging metabolites including PAHs, aromatic amines, and heterocyclic amines (305). Neutrophil recruitment into lung tissue occurs after exposure to a variety insults such as tobacco smoke particulates, asbestos, and infection. A single base substitution (G to A) within the Alu element 463 base pair preceding the *MPO* gene confers less transcriptional activity compared to the allele that contains a G residue (119). Possession of the *MPO* (G/G) genotype have been shown to be associated with an increased the risk of lung cancer (231,306).

In this study, lung cancer risk associated with predicted high and intermediate mEPHX activity observed among females was enhanced by the effect of the *MPO* (G/G) genotype. An increasing in risk from 2.8-fold to be 5.5-fold after adjusting age, race, smoking status, and packyear suggests the joint effect of these two at risk genotypes in a multiplicative way. Exposures to tobacco smoke particulates and/or other chemicals leads to the recruitment of neutrophils into the lung and the releasing of MPO enzyme from these recruited neutrophils. Compared to female cases who had predicted low or very low activity of mEPHX together with the *MPO* (G/A or A/A) genotypes which related to low level expression of MPO, female cases who had more activity of mEPHX enzyme and normal amount of MPO in lung tissues possibly contributed to greater rate of B[*a*]P-dihydrodiol expoxide formation which lead to higher level of DNA adducts, and DNA damage, mutation, and lung cancer development. In addition to an important role of MPO in enhancing the risk of lung cancer associated with high mEPHX activity through activation of procarcinogens, production of free radicals by MPO under

respiratory burst might possibly be another explanation. The consistency of agreement between previous reports and this study suggests the contribution of MPO in lung cancer etiology.

Among females, the present study first found the significant modulation of lung cancer risk associated the combined effect of high and intermediate mEPHX activity and the *GSTM3*A/*A* genotype. When studied separately, each of these 'at risk' genotypes associates with an increased risk of lung cancer. An additive effect of having higher capacity of mEPHX enzyme to catalyze B[*a*]P-7,8-epoxide to the dihydrodiol and a lower level of GSTM3 enzyme to detoxify both B[*a*]P-7,8-epoxide and B[*a*]P-7,8-diol-9,10-epoxide potentially lead to an accumulation of reactive metabolites in lung tissues and further react with DNA in target cells which may cause lung cancer development. According to this, individuals who have high and intermediate mEPHX activity and lower level of GSTM3 are likely to be more susceptible to lung cancer induced by cigarette smoke carcinogens compared to individuals who have deficient mEPHX activity and the protective *GSTM3* genotype.

A significant interaction between combined three 'at risk' genotypes such as *mEPHX*, *GSTM1*, and *GSTM3* influenced the risk of lung cancer was also found among this female subgroup. This is an interesting finding about the effect of a gene-gene-gene interaction on the risk of lung cancer. The polymorphic in the *mEPHX* gene resulting in high and intermediate mEPHX activity interacted with the *GSTM1* null and *GSTM3*A/*A* genotypes that causes low level of GST enzymes could easily shift a balance of activation and detoxification pathway of cigarette smoke carcinogens and therefore place these females at higher risk of lung cancer.

Similarly, another significant modulation of lung cancer risk by a combination of three at risk genotypes observed in the present study was a 2-fold increased in relative risk for individuals who had a combination of the *XPD* exon 23, *GSTM1*, and *GSTM3* at risk genotypes after

adjusting for age, race, smoking status, packyear, and history of occupational exposures. The effect of the combined these three polymorphic genes on increased risk of lung cancer could probably be a result of reduced capacity in detoxification and DNA repair that possibly leaded to higher level of reactive metabolites, DNA adducts, DNA damage and cancer development.

Regarding the lower degree of relative risk for the two combinations of three 'at risk' genotypes (*mEPHX*, *GSTM1*and *GSTM3*; OR=3.04 and *XPD*, *GSTM1* and *GSTM3*; OR=2.02) compared to those odds ratio obtained from two genes combination (*mEPHX* and *GSTM3*; OR=3.74 and *XPD* and *GSTM3*; OR=3.18), an explaination might be the lack of relationship between *GSTM1* null genotype alone and lung cancer risk in this study population. However, as described earlier, GSTM1 is mainly expressed in liver and is involved in the detoxification of PAH compounds metabolism. Deficiencies in liver GSTM1 enzyme possibly affect the level of PAHs reactive metabolites that recirculate to lung and lung epithelial cells. Therefore, the combined effect of *GSTM1* null along with *GSTM3*A/*A* genotypes has been hypothesized to associate with the level of GSTs indicated the capacity of detoxifying lung carcinogens. The results agreed with the hypotheses that individuals who had lower level of GST enzymes from combined *GSTM1* null and *GSTM3*A/*A* genotypes together with higher activation or lower DNA repair capacity are at higher risk of lung cancer compared to those who had reference genotypes.

5.2. Case only analysis

The effects of the single and combined polymorphic genes on higher risk of lung cancer observed in our case control study were further evaluated in a case only analysis, for example, among cases who had family history of cancer(s), and early age onset. The analyses were performed in all cases and in each particular subgroup based on type of risk factors. The results

of the case only analysis were in agreement with the findings from case control study. Furthermore, it also indicated additive effects of other risk factor(s) such as family history of cancer(s) and occupational exposures to the effects of lung cancer susceptibility genes such as metabolic and DNA repair genes on greater risk of lung cancer development.

The combined effect between genetic polymorphisms in metabolic genes, family history of lung cancer and/or occupational exposures on increasing risk of lung cancer was found in the first subgroup, for example, the group of lung cancer cases who had at least one first-degree relative(s) with lung cancer. The mean age at diagnosis among cases in this group was significantly lower than those of sporadic. The distribution of packyear in both groups was similar. Cases who had a family history of lung cancer were all ever smokers and thirty-two percent were exposed to occupational exposures. Similarly, all sporadic cases were ever smokers but did not have history of family with lung cancer and occupational exposures. There was a significant overrepresentation in the number of cases who had predicted high CYP1A2 activity and slow *NAT2* genotype in this group compared to the sporadic cases. The age adjusted OR of 5.3 (95%CI, 1.18-23.28) was found among these at risk lung cases who were ever smokers, had predicted high CYP1A2 activity and slow *NAT2* genotype together with family history of lung cancer and/or occupational exposures. This result was in agreement with the finding of overall 5-fold increased risk of lung cancer among cases who had two combined at risk genotypes in a case control study. As discussed previously, both CYP1A2 and NAT2 enzymes are involved in the metabolism of carcinogens found in cigarette smoke, combustion products, high temperature cooked meats, and other occupational exposures. Having genetic polymorphisms in both genes could shift the balance between activation and detoxification to higher level of DNA adducts, ultimately resulting in lung cancer development.

High-risk ever smokers from lung cancer families showed higher percent of individuals (38%) who carried these at risk genotypes compared to 12.5% of the sporadic cases. This finding supported the hypothesis that there is a genetic predisposition to lung cancer. An estimated 90% of lung cancer is attributable to cigarette smoking however there were approximately only 10% of smokers develops lung cancer (307). A genetic predisposition to lung cancer may contribute to interindividual susceptibility to cigarette carcinogen metabolism and familial aggregation. The result obtained in this analysis indicated that smokers with metabolic enzyme polymorphisms in combination with a family history of lung cancer and/or environmental risk factor (occupational exposures) are likely to be at greater risk of lung cancer compared to smokers who had no family history of lung cancer and occupational exposures. According to the similar results from both the case control and case only analyses, combination of predicted high CYP1A2 activity and *NAT2* slow genotype contributed to host susceptibility of lung cancer.

Among at risk lung cancer cases in group 2 or cases who had at least two first-degree relatives with other cancer type(s), borderline associations between single gene loci polymorphism at the *GSTM1* gene or in combination with *XPD* (Lys/Gln and Gln/Gln) genotype and increased risk of lung cancer were observed. Twenty-nine of forty cancer family cases had *GSTM1* null genotype compared to twenty-five of forty-six sporadic cases. The crude odds ratio was 2.22 (95%CI, 0.90-5.47). Similarly, there was higher proportion of family cancer cases who carried both the *GSTM1* and *XPD* exon 23 at risk genotypes with the relative risk of 2.5 (95%CI, 0.99-6.20) compared to the sporadic cases. The mean age, and packyear, and the distribution of number of individuals who had history of occupational exposures among high-risk cases and sporadic cases who had at risk genotypes were similar. There were four family cancer cases who were lifetime nonsmokers, two of which had occupational exposures. Therefore, these results

suggest that metabolic enzyme and DNA repair gene polymorphisms in combination with family history of other cancer(s) is likely to contribute to the greater risk of lung cancer development.

Lung cancer cases who had age at diagnosis less than or at 50 years old were classified into the third at risk lung cancer case group. There was a significantly higher number of cases who had the combined *XPD* exon 23 and *GSTP1* 'at risk' genotypes in the early age onset subgroup compared to the sporadic cases. There were two lifetime nonsmokers among these cases, and one had occupational exposure. In contrast, all sporadics who had *XPD* (Lys/Gln and Gln/Gln) and *GSTP1* (**A/*B* and **B/*B*) genotypes were ever smokers and did not have occupational exposures. Also the sporadic cases had higher packyear histories compared to the early age onset lung cancer cases. However, fifty percent of the ever smokers among the early age onset lung cancer cases had history of occupational exposures. According to this observation, carrying metabolic enzyme and DNA repair at risk genotypes together with history of smoking, and/or occupational exposures might result in individuals developing lung cancer at early ages.

It is interesting that there were twelve of fourteen lifetime nonsmoker lung cancer cases who had the higher inducible *CYP1A2* allele (*CYP1A2*1F/*1F*) and wild type genotype for the *CYP1A2*1C* allele that resulted predicted high CYP1A2 activity compared to fourteen from thirty two of the sporadic cases, although this was not quite statistically significant (P=0.054). However, a significant association between two combined at risk genotypes for *CYP1A2* and *NAT2* gene and increased risk of lung cancer was observed among lifetime nonsmoker cases (OR=5.3; 95%CI, 1.52-18.59). Moreover, there was a significant overrepresentation of cases who carried both *GSTM1* null and *XPD* (Lys/Gln and Gln/Gln) genotypes among lifetime nonsmokers group compared to the sporadic cases. The distribution of age and gender of nonsmoker lung cancer cases was similar to those of sporadic cases. There were three cases who

had history of occupational exposures and two of them had family history of other cancer(s). There was one lifetime nonsmoker who had at least one first-degree relative(s) with lung cancer.

Among the twelve nonsmoker cases who had high CYP1A2 activity, there were two cases who had occupational exposures and family history of other cancer(s). There was also one case who had family history of lung cancer and another case who was only exposed to occupational exposures. Furthermore, six and eight of these twelve nonsmokers were slow acetylators and had combined *GSTM1* and *XPD* exon 23 at risk genotypes, respectively. Taken together, it is possible to suggest that passive smoking and/or other environmental such as combustion and charbroiled foods and/or occupational exposures in combination with having higher enzyme activity in activation, lower capacity in detoxifying carcinogenic reactive metabolites and repair DNA damage put these lifetime nonsmokers at risk of lung cancer development.

Since twenty-three lung cases fell in more than one subgroup, all cases were combined in order to further evaluate the influence of genetic polymorphisms in metabolic and/or DNA repair genes on greater risk of lung cancer among the entire group of higher-risk cases. Furthermore, 'at risk' lung cancer cases were categorized into two all case groups based on history of occupational exposures for examining an additive effect of occupational exposures together with other risk factors for lung cancer development.

Among high risk cases who did not have history of occupational exposures, there were two combinations of combined 'at risk' genotypes that were shown to be significantly associated with a 5- or 3-fold increased risk of lung cancer compared to the sporadic cases who also did not exposed to occupational exposures. These were the combined predicted high CYP1A2 activity, and *NAT2* slow genotypes and the *MPO* (G/G), and *GSTM1* null genotypes (P=0.01 and 0.03, respectively). The mean age and packyear among the 'at risk' lung cancer cases who had these at

risk genotype combinations and sporadic cases were similar. Interestingly, there were nine of fifty-nine high-risk lung cancer cases who had the combination of all four genotypes; high CYP1A2 activity, *NAT2* slow, *MPO* (G/G), and *GSTM1* null genotypes while no sporadic cases had all four. These findings are in agreement with results from the case control study supporting a role of genetic polymorphisms in the genes that encode for enzymes involved in cigarette carcinogen metabolism as lung cancer susceptibility biomarkers. The data also suggests an effect of metabolic gene polymorphisms in a combination other risk factors such as family history of lung and other cancers and early age onset on higher risk of lung cancer development.

The effect of occupational exposures in combination with genetic polymorphisms in metabolic and DNA repair genes and family history of lung and other cancers on early age onset of lung cancer was observed among all high-risk cases who ever had a history of occupational exposures. There were significant overrepresentations of high-risk lung cancer cases with combined *GSTP1* (**A/*B* and **B/*B*) genotype combined with *XPD* (Lys/Gln and Gln/Gln) genotype or predicted mEPHX high and intermediate activity compared to sporadic cases who did not have both family history of lung or other cancers and occupational exposures. Not only did the combination of carrying metabolic and DNA repair 'at risk' genotypes with family history of lung or other cancers, and occupational exposures influence risk of lung cancer by 3 to 4 times, but it also affected the age at onset compared to sporadic cases. The mean age among cases who had at risk genotypes at *GSTP1* and *XPD* exon 23 or *mEPHX* genes together with ever exposed to occupational exposures and had family history of cancers (59±16 and 55±14, respectively) appeared to be significantly lower than sporadic cases who carried the same combination of at risk genotypes but did not have occupational exposures and family history of cancers (72±11 and 70±5; P=0.04 and 0.002, respectively).

Moreover, there were twelve 'at risk' lung cancer cases compared to only three sporadic cases who had the combination of predicted high and intermediate mEPHX activity, *GSTP1* (**A/*B* and **B/*B*), and *XPD* (Lys/Gln and Gln/Gln) genotypes with a statistical significant difference at P=0.04. Similarly, taking together the risk factors of having family history of lung and other cancer(s), occupational exposures, and at risk genotypes at *mEPHX*, *GSTP1*, and *XPD* exon 23 genes was shown to be associated with early age at onset among cases with these risk factors compared to the sporadic cases (mean age 52±13 and 73±2, respectively). Lung cancer is clearly a multistep carcinogenesis process that requires several genetic and/or epigenetic events in order to drive the carcinogenic pathway to show a malignant phenotype. The results observed were in agreement with this concept. For example, individuals with a family history of cancer at risk lung cancer cases might have some genes already mutated in the germline that provides hereditary predisposition to lung carcinogenesis. Cellular transformation can be further caused by environmental factors such as exposed to asbestos, chromium, silica and other occupational exposures, cigarette smoke, combustion, and high temperature cooked meat. Then, polymorphisms in metabolic activation and detoxification and DNA repair enzymes play important roles in the acquisition of susceptibility to those environmental factors and could possibly enhance the entire lung carcinogenic process. Therefore, the joint effect between genetic host factors (interindividual variations in carcinogen metabolism and DNA repair and familial aggregation) and environment factors (environmental and occupational exposures) potentially influence individual to develop lung cancer at early age.

The results obtained from both case control and case only analyses were in agreement and indicated an association between metabolic enzyme, DNA repair enzyme genetic polymorphisms and lung cancer risk, particularly in a combination with other risk factors such as cigarette

smoking, occupational exposures, an early age onset, and family history of lung and other cancers.

5.3. mRNA expression

In order to answer the key question about tissue specific chemical carcinogenesis, it is important to determine the capacity of the actual target tissue to convert procarcinogens to reactive metabolites. In the present study, the determination of CYP mRNA expression levels in lung tissue was performed to evaluate the potential for local bioactivation and as smoking exposure biomarker. The levels of CYP1A1, CYP1B1, and CYP2E1 mRNA expression were measured in twenty pairs of lung tumors and histologically normal adjacent lung tissues. The results on CYP expression in both lung tumors and histological normal adjacent lung tissues obtained in this study were in agreement with those presented in earlier studies (255-257,263,265,308,309,309).

In this study, CYP1A1 was expressed at low levels in both lung tumors and histologically normal tissue. The levels of expression among tumors were not statistically significantly different from their corresponding adjacent histological non-tumor samples for both former smokers and nonsmokers. Furthermore, the levels of CYP1A1 mRNA that was expressed in lung tumor and histologically normal adjacent lung tissues of former smokers were similar to those of nonsmokers. This finding is probably due to the cessation of cigarette smoking more than 3 months prior to the beginning of the study. As reported by Petruzzelli *et al.* (255) that the effect of cigarette smoking on inducible CYP1A1 expression in lung tissues could last up to 60 days after cessation of smoking. Also, McLemore *et al.* (256) reported that the decrease in CYP1A1 mRNA levels was noted as early as 2 weeks after cessation of cigarette smoking and completely disappeared in normal lung tissue within 6 weeks after smoking cessation. There were three pairs

of lung tumors and histological normal adjacent lung tissues that obtained from current smokers in this study. However, there were insufficient total RNA amounts to measure the level of CYP1A1 expression in lung tissue samples from one of these current smokers. β-Actin mRNA was not detectable in the remaining two samples.

There was interindividual variation in the amount of constitutive CYP1A1 mRNA in histological normal adjacent lung tissue. The range of CYP1A1 expression in eleven histological normal adjacent lung tissues was 0.04- to 0.21 amol/amol β-actin. The levels of CYP1A1 mRNA in histological normal adjacent lung tissues of four former smokers and two nonsmokers were between 0.04 and 0.07 amol/amol β-actin, which in accordance with the results of previous studies. The levels of expression in histological normal adjacent lung tissues of two current smokers were higher than those of former smokers and nonsmokers (0.12 and 0.14 amol CYP1A1/amol β-actin). However, one nonsmoker had showed a level of CYP1A1 mRNA in histological normal adjacent lung tissue as those of current smokers (0.13 amol/amol β-actin). This may either be due to misreporting of smoking history and/or greater inducibility of constitutive levels due to carrying the variant *CYP1A1* genotype as well as possible exposures to petroleum and other occupational carcinogenic PAH compounds. This could also explain the detected levels of 0.15 and 0.21 amol CYP1A1/amol β-actin in histological normal adjacent lung tissues of two former smokers who were exposed to petroleum and occupational pollutants and also had the inducible *CYP1A1* genotype.

In comparing the levels of CYP1A1 mRNA expression between eight pairs of lung tumors and their corresponding histological normal adjacent tissues obtained from five former smokers and three nonsmokers, the mean level of CYP1A1 mRNA detected among tumors was not different from the mean value of histological normal adjacent lung tissues as shown in Table 42.

However, CYP1A1 expression in lung tumors is most likely to be lower than that observed in the corresponding histologically normal adjacent lung tissues (observed in six from eight pairs of lung samples). Furthermore, after the two pairs of lung samples that showed higher CYP1A1 mRNA in tumors compared to their histological normal adjacent tissues were excluded, the mean level of CYP1A1 mRNA was 2.3-fold lower (P=0.02) in tumors (0.032 ± 0.01) compared to the matched histological normal adjacent lung tissues (0.075 ± 0.03). This is in accordance with previous studies that showed 3-fold lower of CYP1A1 expression in tumors compared to the corresponding normal tissues (308) and a decreased level of AHH in non small cell lung cancers versus normal lung tissue (310). This decrease in CYP1A1 mRNA expression in lung tumors may be related to altered CYP1A1 gene regulatory pathways. Higher levels of CYP1A1 expression that was observed in lung tumors compared to those of matched histological normal adjacent tissues in both a former smoker with a history of petroleum and asbestos exposures and another not exposed to other environmental inducers might also be related to altered CYP1A1 gene expression such as activation the CYP1A1 gene in tumor cells. Further studies are required to examine an expression of CYP1A1 in lung tumor cells because the CYP1A1 regulation may be altered in tumor tissue and may be important with respect to lung carcinogenesis.

Both histological normal adjacent lung tissues and lung tumors from former smokers and lifetime nonsmokers in the absence of cigarette smoking and other environmental pollutants displaying constitutive expression of the CYP1A1 mRNA indicates the expression of CYP1A1 in lung tissue and the possible existence of local bioactivation of lung carcinogens by CYP1A1 in specific target lung tissue, although the presence of mRNA can not be regarded as a direct indication of the existence of corresponding CYP1A1 proteins. However, detectable mRNA can potentially be translated to functional protein in the specific target tissue. Higher levels of

CYP1A1 mRNA expression detected in histologically normal adjacent lung tissue from current smokers and former smokers and lifetime nonsmokers who had exposed to other environmental inducers suggests an important role of CYP1A1 mRNA as a cigarette exposure biomarker. No correlation between the expression of CYP1A1 mRNA in histological normal adjacent lung and lung tumor tissues is evident in the present study. This lack of correlation should be investigated for changes occurring in tumor cells that may lead to altered expression of CYP1A1 in lung tumors further in future studies. Also, the role of CYP1A1 mRNA expression as a cigarette smoking exposure biomarker should to be further evaluated in larger numbers of lung tumors and normal lung from current smokers.

The present study also demonstrated the expression of CYP1B1 in both lung tumors and histological non-tumors at the mRNA level. There was considerable interindividual variation in expression of CYP1B1 mRNA regardless of tissue types (tumors or non-tumors), smoking status, and subjects' lung function. For example, the lowest mRNA level of CYP1B1 was detected in histologically normal adjacent lung tissue obtained from a current smoker. A weak association between the levels of detected CYP1B1 mRNA in histologically normal adjacent lung tissues and airway obstructive diagnosis from the tested lung function results was observed. CYP1B1 mRNA was also detected in the histologically normal adjacent lung tissues obtained from lifetime nonsmokers and former smokers who either had or did not have history of exposures to other environmental and occupational pollutants and/or carcinogenic compounds. These findings were in accordance with the previous reports for CYP1B1 gene expression in bronchial epithelial cells of nonsmokers (252), current smokers (257) and in lung tumors and non-tumors of current and former smokers (263). Regarding the observation of CYP1B1 expression in lung tissues of two lifetime nonsmokers who had reported no history of other environmental and occupational

exposures, this could due to passive smoking exposure and inhalation of non-cigarette procarcinogens such as engine exhaust. Also this might possibly be very low constitutive expression of the *CYP1B1* gene in lung tissue and was in agreement with the previous report that CYP1B1 is expressed constitutively in normal adult lung tissue (156,252).

The detectable level of CYP1B1 mRNA in histological normal adjacent lung tissues of former smokers might also be explained by the exposures to other environmental inducers and tobacco smoke and/or the mechanisms of CYP1B1 expression induced by cigarette smoke. There are important questions that remain to be answered and cannot be addressed in this study such as how soon the effect of smoking induction starts, how long does it last, is there possible cumulative effect of smoking on CYP1B1 expression and interindividual differences in the level of induction. Human *CYP1B1* gene expression induced by polycyclic aromatic hydrocarbons has been well documented and involves the Ah receptor for increasing transcriptional of the *CYP1B1* gene (262,311). However, non-Ah receptor mediated pathways of transcriptional regulation and post-transcriptional mechanisms may also contribute to regulating expression of *CYP1B1* gene. Since human CYP1B1 mRNA contains multiple polyadenylation sites, it has been suggested that there is cell-type-specific alternative processing of CYP1B1 mRNA, which may regulate the amount and/or the ability of the final transcript to be translated (311,312). Furthermore, the stability of CYP1B1 mRNA that can be influenced by any other alternative processing may be responsible for the regulation of the *CYP1B1* gene expression (312).

Conversely, the presence of CYP1B1 mRNA in histological normal lung tissue of former smokers may represent constitutive expression of CYP1B1 gene in lung tissue. The great CYP1B1 interindividual variability observed in this study may not be based on genetic polymorphism in the *CYP1B1* gene as there was not an association between the *CYP1B1*3*

variant allele and the constitutive amount of CYP1B1 mRNA detected in histological normal adjacent lung tissues. In the present study, the plasma nicotine and/or cotinine levels were not measured in any of blood samples from patients whose lung samples collected in order to correlate the self-report smoking status and cigarette smoke exposure. However, Spivack *et al*. (263) reported the expression of CYP1B1 in lung tumors and non-tumors of former smokers with no detectable nicotine and cotinine present in their plasma samples. Furthermore, none of these subjects had reported exposures to other inhaled toxicants and medications. Taking the results from the earlier study and this study, it appears that CYP1B1 is expressed in human lung tissue at the mRNA level with a wide interindividual variation in expression. This also indicates a potential role of CYP1B1 mRNA expression as a susceptibility biomarker of lung cancer for preferential bioactivation of procarcinogens in lung tissue.

As shown in Table 42, the mean level of CYP1B1 mRNA measured in histological normal adjacent lung tissue was not statistically significantly different from the observed mean value of unmatched lung tumors. However, there was a trend of higher expression in level of CYP1B1 mRNA in tumors compared to their matched histological normal adjacent (measured in six pairs of lung tissue samples) (Table 43). There were two of six pairs lung tissue samples showed lower level of CYP1B1 mRNA detected in tumors compared to their corresponding histological normal adjacent tissues. After the values obtained from these two pairs of lung samples were excluded, an increasing in the mean level of CYP1B1 mRNA expression in lung tumors compared to that of their matched histological normal adjacent tissues did reach statistical significance (P=0.01). The mechanism of the discrepancy between the observed levels of CYP1B1 mRNA expression among lung tumors possibly resulted from the differences in constitutive levels of CYP1B1 mRNA, regulatory control of CYP1B1 expression by AhR and non-AhR-dependent factors, and

cell type specific regulating expression of CYP1B1 as described previously. However, the finding of a tendency towards overexpression of CYP1B1 mRNA expression in lung tumors could potentially be important in determining utility of CYP1B1 mRNA as a smoking and/or environmental exposure biomarker. Moreover, it is likely that there is a mechanistic link between cigarette smoke and/or environmental procarcinogens exposures and lung carcinogenesis but it should be evaluated in future studies.

In the present study, CYP2E1 appeared to be expressed in both lung tumors and histologically normal adjacent lung tissue with several orders of differences in detected mRNA levels across individuals. The range of the CYP2E1 mRNA levels detected in histological normal adjacent lung tissues was between 0 to 8 amol CYP2E1/amol β-actin (among former smokers who had reported none of environmental and occupational exposures) and was between 0.22 to 4.5 amol CYP2E1/amol β-actin (among current smokers, nonsmokers and former smokers who had exposed to petroleum and other environmental and/or occupational exposures). Higher expression of CYP2E1 in histological normal adjacent lung tissues of current smokers (2-3 amol CYP2E1/amol β-actin) compared to nonsmokers (0.8 amol CYP2E1/amol β-actin) with no history of environmental and/or occupational CYP2E1 inducers was observed which indicated the induction of *CYP2E1* gene expression by cigarette smoke. The detection of CYP2E1 mRNA in histological normal adjacent lung tissues of former smokers and nonsmokers who were exposed to petroleum and/or other environmental compounds might be a consequence of the CYP2E1 induction effect of polycyclic aromatic hydrocarbons and benzene. The mechanism of CYP2E1 induction by cigarette smoke has been thought to involve protein and/or mRNA stabilization. Therefore, it is likely that the CYP2E1 mRNA observed in histological normal adjacent lung tissues of current smokers and former smokers and nonsmokers who were exposed

to cigarette smoke or petroleum and other compounds in this study may be due to mRNA stabilization. Further studies are required to examine the mechanistic effect of CYP2E1 induction by cigarette smoke and other agents.

Similar to the observed results of CYP1A1 and CYP1B1 mRNA expression, the finding of a wide variation of CYP2E1 mRNA expression across individuals in this study was not likely to be influenced by differences in the ability for mRNA production in their lung tissues since the diagnosed lung function as mild or borderline airway obstruction or normal airflow was weakly associated with the levels of mRNA expressed in histological normal adjacent lung tissues. However, the limitation of these lung tissue samples clinical and demographic information on other contributing factors to variability in CYP2E1 mRNA expression such as diet and alcohol intake might be an explanation. Nevertheless, these results suggest the presence of interindividual susceptibility to local bioactivation of lung procarcinogens as indicated by the variation of the constitutive amount of CYP2E1 expressed in lung tissues. Furthermore, it indicates the potential of CYP2E1 mRNA as cigarette smoke and/or other CYP2E1 inducers exposure biomarker as the observation of CYP2E1 expression induced by cigarette smoking and/or petroleum and other environmental and occupational exposures.

The mean level of CYP2E1 mRNA measured in sixteen histological normal adjacent lung tissues was not statistically different from the mean mRNA level detected in seventeen unpaired lung tumors, which is likely to be an effect of an interindividual variation in the CYP2E1 mRNA levels. However, the significant elevation of CYP2E1 mRNA in lung tumors was shown after comparing the expression of CYP2E1 among thirteen matched tumors and non-tumors (Table 43). An overexpression of CYP2E1 mRNA in tumors suggests a mechanistic link between environmental procarcinogens and/or cigarette smoke exposure and lung carcinogenesis. Also,

an alteration of CYP2E1 gene expression is likely to occur in the neoplastic cells during lung cancer development. The function of CYP2E1 in activation of procarcinogens found in cigarette smoke and environmental pollutants, the presence of constitutive CYP2E1 in lung tissues, and inducibility of CYP2E1 by cigarette smoke, benzene, and other compounds indicates the potential for a significant contribution of CYP2E1 for local bioactivation of lung procarcinogens in the target lung tissue and in lung carcinogenesis.

According to these results for CYP mRNA quantitation, CYP1A1, CYP1B1, and CYP2E1 are expressed in both lung tumors and non-tumors. Compared to *CYP1A1* gene expression, CYP1B1 appears to show several orders of magnitude greater levels of expressed mRNA consistent with a previous report (263). The significant difference in the detected mRNA expression levels between CYP1A1 and CYP1B1 is not likely to be an effect of analytical variability as cDNA from a normal liver sample was used as a control across all plates over time for normalization of any plate-to-plate variation. The level of CYP mRNA expression in relation to an endogenous gene (β-actin) was used in order to minimize differences in the initial input amount of the total RNA. Also, the effect of probe variability was also normalized for each CYP mRNA determination. The higher levels of CYP1B1 mRNA compared to those of CYP1A1 detected in histological normal lung tissues from current smokers, former smokers and nonsmokers could be explained by the differences in the degree and/or mechanisms and/or long lasting effects of both CYPs mRNA induced by cigarette smoke and/or other environmental and occupational compounds and/or by the differences in the constitutive amount that expressed in lung tissue.

CYP1A1 and CYP1B1 are expressed predominantly in extrahepatic tissues including lung and they participate principally in the activation of carcinogenic polycyclic aromatic

hydrocarbons. However, it has been shown that CYP1B1 is able to catalyze the activation of both polycyclic aromatic hydrocarbons and aryl amines. Also, CYP1B1 has shown the higher catalytic activities compared to CYP1A1 toward several PAHs including (±)B[*a*]P-diol, and dibenzo[*a,l*]pyrene-11,12-diol and several aryl amines including 2-aminoanthracene, and 6-aminochrysene (156). Moreover, It has been demonstrated that CYP1B1 is a catalytically efficient 17β-estradiol 4-hydroxylase, and a slight 2-hydroxylase (290). According to the capacity of CYP1B1 in activation of a wide range of carcinogenic chemicals and endogenous estrogen hormone and the finding of highly expressed of CYP1B1 in lung tissue in this study, it suggests an important of another CYP1 family member, CYP1B1, for local bioactivation of procarcinogens, interindividual susceptibility to lung cancer and lung cancer risk especially among females in addition to the well established role of CYP1A1 in polycyclic aromatic hydrocarbons metabolism and as cigarette exposure biomarker.

The present study has demonstrated the existence of the constitutive CYP2E1 mRNA in lung tissues among current, former smokers and nonsmokers, which was in accordance with the previous report (309). Also, Willey *et al.* (252) and Crawford *et al.* (313) reported that CYP2E1 was expressed in bronchial epithelial cells of nonsmokers. Additionally, *CYP2E1* gene expression induced by cigarette smoke and/or benzene and other environment inducers was indicated in this study. These findings support the role of CYP2E1 mRNA as a cigarette smoke and other environmental inducers exposure biomarker and as an interindividual lung cancer susceptibility biomarker for local procarcinogen bioactivation.

CYP2E1 activates low molecular weight carcinogens found in cigarette smoke and environmental pollutants such as vinyl choride, benzene, and some nitrosamines including NNK and NNAL, all of which have been postulated to be responsible for lung adenocarcinoma.

Although there were no observable differences in the levels of CYP2E1 mRNA in both histological normal adjacent lung tissues and lung tumors between lung adenocarcinoma and squamous cell carcinoma analyzed lung tissue samples in this study, the presence of constitutive levels of CYP2E1 mRNA indicated pulmonary NNK and other cigarette and/or environmental carcinogen biotransformation, which is believed to play a role in lung carcinogenesis. Interindividual variation in constitutive levels CYP2E1 mRNA, CYP2E1 mRNA induction by exposed to other CYP2E1 inducers, mechanisms of regulatory system for CYP2E1 expression, an involvement of procarcinogens activation by CYP2E1 and other cytochrome P450s in liver, which bronchial epithelial cells can be exposed to these blood-borne carcinogenic reactive metabolites, the expression of other activation enzymes such as CYP2A6 and CYP2B6 that also involved in catalyzing the nitrosamines compounds in the lung, and cytochrome P450 independent pathways for NNK metabolism in human lung may contribute to the negative association of CYP2E1 mRNA levels in lung tissues and their corresponding histological adenocarcinoma cell type.

In order to be able to quantify the relative low constitutive amount of mRNA and screen the gene expression of several metabolic enzymes in large numbers human lung samples, a reliable method that requires a small amount of biological sample that is sensitive, reproducible, fast and precise is necessary. The limitation of the mRNA results obtained from this study was primarily due to the limited yields of total RNA in some lung tissue samples as well as the relative high initial amount of cDNA needed for measuring mRNA expression by the RT-PCR FAF-ELOSA technique. The real-time quantitative RT-PCR technique, TaqMan assay, offers higher sensitivity, rapid, accuracy, and high-through capability over other conventional RNA quantification methods. Therefore, the TaqMan based assays for quantification of CYP1A1,

CYP1B1, CYP2E1, and endogenous β-GUS mRNA expression were developed in the present study. A comparison of relative sensitivity of the RT-PCR FAF-ELOSA and the TaqMan assays for measuring CYP1A1, CYP1B1, and CYP2E1 mRNA in the same normal liver sample was performed. Since there was limited amount of total RNA extracted from 20 pairs of lung tumors and histological normal adjacent tissues, the results obtained only the comparison of the mRNA levels in liver sample detected by two methods.

Using the relative levels of CYP mRNA expression to endogenous β-actin, estimated levels of CYP1A1, CYP1B1, and CYP2E1 mRNA/β-actin of liver cDNA sample detected by both TaqMan assay and RT-PCR FAF-ELOSA were similar as shown in Fig. 17 and 18. A similar pattern of these three CYP expression levels in the same liver sample between the two techniques was demonstrated (Fig. 22) although there was a slight difference in the measured levels of CYP1A1 and CYP2E1 between two methods. As the liver cDNA were analyzed in triplicate for each concentration by TaqMan® assay, the results showed that the C_T values were reproducible with intra-assay coefficients of variations of 0.19-1.43% (Fig. 23). In contrast, with a duplication of each concentration analyzed by FAF-ELOSA, the results showed intra-assay coefficients of variations of 9.49-50% (Fig. 24). This finding indicate higher reproducibility of the TaqMan® assay compared to FAF-ELOSA and could explain the differences in the levels of CYP1A1 and CYP2E1 mRNA in the same liver sample detected by both methods.

Another important aspect of quantitative mRNA measurement is the need for a reliable RT-PCR method to show a high sensitivity. For the TaqMan® assay, the starting amount of total RNA analyzed was 105 ng (an average of three concentrations of 180, 90 and 45 ng) compared to 1000 ng for FAF-ELOSA, constituting a 10-fold higher sensitivity for TaqMan® assay than the FAF-ELOSA. A 20-fold higher sensitivity in the quantification of CYP1A1, CYP1B1, and

CYP2E1 mRNA at the lowest starting amount of template 45 ng of liver cDNA with the TaqMan® assay compared to 1000 ng used in the FAF-ELOSA method was also shown. Furthermore, the TaqMan® assay demonstrated high sensitivity in the quantification of CYP1A1, CYP1B1, and CYP2E1 relative to endogenous β-GUS mRNA in normal human lung tissue compared to reference normal liver tissue with an early detectable C_T of 17 only with a 1 ng starting amount of lung and liver cDNA. In comparison with FAF-ELOSA method, TaqMan® assay therefore has been shown to achieve its potential increase in sensitivity, reproducibility, rapid and reliability in quantification of gene expression.

CONCLUSIONS

This study has evaluated the contribution of several genetic polymorphisms in metabolic and DNA repair genes that are involved in cigarette smoke carcinogen metabolism and in the nucleotide excision repair pathway in lung cancer risk The panel of genes screened include *CYP1A1, CYP1A2*, *CYP1B1*, *CYP2D6*, *CYP2E1*, *mEPHX*, *MPO*, *GSTM1*, *GSTM3*, *GSTT1*, *GSTP1*, *NAT2** and *XPD*. Significant increases in risk of lung cancer were observed for subjects with multiple genetic polymorphisms. The additive or multiplicative effect of the combinations of different genetic polymorphisms on increased risk of lung cancer emphasized the importance of screening a panel of metabolic and DNA repair enzyme polymorphisms for determining risk of lung cancer. The central hypothesis in this study to characterize the interactions between genetic and environmental factors in lung cancer risk was confirmed in both case/control analysis and case only study. The genetic risk factors identified included a single 'at risk' genotype or the combinations of two or three 'at risk' genotypes, family history of lung cancer and other cancer(s), and early age onset. The significant environmental risk factors consisted of history of cigarette smoking and occupational exposures such as asbestos, petroleum, nickel, chromium, silicon, radon, and others. The gene-environment interactions not only resulted in an increasing risk of lung cancer but also an earlier age of onset of lung cancer development.

Local bioactivation of cigarette smoke and other environmental carcinogens has been shown to be another important determinant of interindividual susceptibility to lung cancer by the study of CYP1A1, CYP1B1, and CYP2E1 gene expression in target lung tissues. High but variable constitutive levels of CYP1B1 mRNA expression in lung tissues were observed in the present

study. This data, taken together with information presented in the literature that CYP1B1 plays a significant role in catalyzing the activation of various carcinogens found in cigarette smoke and the environment, suggests an important role of CYP1B1 in lung carcinogenesis. The role of CYP2E1 mRNA expression as lung cancer susceptibility biomarker and cigarette and/or other environmental carcinogens exposure biomarker has been supported by the observations of interindividual variation in constitutive amounts of CYP2E1 mRNA and also the detectable levels of CYP2E1 mRNA induced by cigarette smoke and/or other environmental carcinogens that were observed in histological normal adjacent lung tissues. The significant increase in the levels of CYP2E1 mRNA expression in lung tumors compared to their corresponding histological normal adjacent tissues among current and former smokers and nonsmokers who were exposed to petroleum and/or other environmental exposures further suggest a mechanistic link between environmental carcinogens exposures and lung cancer development. These results provide evidence for the potential of local procarcinogen bioactivation in lung tissues by CYP1B1 and CYP2E1. Future studies are required to explore the mechanisms of cigarette smoke and/or other environmental carcinogens induce CYP1B1, and CYP2E1 mRNA expression.

Based on overlapping substrate specificity of CYP enzymes and an involvement of several activation and detoxification enzymes in the metabolism of procarcinogens found in cigarette smoke and environment together with the presence of local bioactivation of lung carcinogens in lung tissues suggested by the findings in the present study, the importance of determining endogenous factors for host susceptibility to lung cancer is indicated. There are other metabolic enzymes such as CYP2A6, CYP2B6, mEPHX, GSTs, and NADPH oxidoreductase that have been reported to express in human lung and metabolize major carcinogenic compounds in cigarette smoke and environmental pollutants (252,313). Measurement of the expression of those

genes in target lung tissues is therefore likely to also be an important determinant of metabolic capabilities and consequences host factors in susceptibility of lung cancer.

Therefore, individual susceptibility to lung cancer determined by endogenous host factors such as genetic polymorphisms in metabolic and DNA repair genes, family history of lung and other cancers, early age onset, and interindividual differences in capacity of local procarcinogens bioactivation could interact with each other or interact with individual exposures or other exogenous factors such as cigarette smoke, environmental carcinogens and occupational exposures in modifying lung cancer risk. Taken together, the results in the present molecular epidemiological study indicated the contribution of both gene-gene and gene-environment interactions in lung carcinogenesis. Furthermore, it supports the usefulness of the multi-risk factor paradigm, gene-gene and gene-environment interactions, for improved human cancer risk assessment. Moreover, the gene-gene and gene-environment interactions characterized in this study could be very useful in public health intervention strategies for lung cancer such as the identification of individuals who would most benefit from modification of hazardous lifestyles, rational therapeutic strategies and the reduction of involuntary exposure to carcinogens by regulation. Not only has this study explored several lung cancer susceptibility biomarkers that can be used to identify individual lung cancer risk, but it also demonstrates the significant power of current molecular genetic technologies, which can be used to facilitate massive population screening in a rapid, efficient, cost effective fashion to enhance human cancer risk assessment.

Although the relative levels of encoded proteins and their functional significance cannot be predicted by the determination of the polymorphic genes and the abundance of mRNA expression as well as lung cancer risk is also influenced by other contributors such as other metabolic and DNA repair polymorphic variant alleles, mutation in oncogenes and tumor

suppressor genes, and genomic instability, it is likely that the findings in this study have provided the evidence to suggest the important role of interindividual variation in host susceptibilities to lung cancer and how they modify the effects of cigarette smoke and/or other environmental carcinogens and/or occupational exposures. Further functional assays for these polymorphic encoded proteins and the mechanisms related to lung cancer risk are required to be elucidated in order to accomplish the ultimate goal in cancer research for lung cancer prevention.

BIBLIOGRAPHY

1. Shields PG, Harris CC. Genetic predisposition to lung cancer. In: Roth JA, Cox JD, Hong WK, eds. *Lung Cancer*. Boston: Blackwell Scientific Publications, 1990: 3-13.

2. Pershagen G. Passive smoking and lung cancer. In: Samet JM, ed. *Epidemiology of lung cancer*. New York: Marcel Dekker, Inc., 1994: 109-130.

3. Hughes JM, Weill H. Asbestos and man-made fibers. In: Samet JM, ed. *Epidemiology of lung cancer*. New York: Marcel Dekker, Inc., 1994: 185-205.

4. Leonard A. Recent advances in arsenic mutation and carcinogenesis. Current Topics in Environmental Toxicology & Chemistry 8: 443-451, 1985.

5. Barrett JC, Lamb PW, Wang TC, Lee TC. Mechanisms of arsenic induced cell transformation. Biological Trace Element Research 21: 421-429, 1989.

6. Harley NH, Harley JH. Potential lung cancer risk from indoor radon exposure. Ca A Cancer Journal for Clinicians 40: 265-275, 1990.

7. Goldsmith DF. Silica exposure and pulmonary cancer. In: Samet JM, ed. *Epidemiology of lung cancer*. New York: Marcel Dekker, Inc., 1994: 245-298.

8. Coultas DB. Other occupational carcinogens. In: Samet JM, ed. *Epidemiology of lung cancer*. New York: Marcel Dekker, Inc., 1994: 299-333.

9. Jain M, Barch JD, Howe GR, Risch HA, Miller AB. Dietary factors and risk of lung cancer: results from a case-control study, Toronto, 1981-1985. International Journal of Cancer 45: 287-293, 1990.

10. Byers TE, Graham S, Haughey BP, Marshall JR, Swanson MK. Diet and lung cancer risk: findings from the western New York diet study. American Journal of Epidemiology 125: 351-363, 1987.

11. Tokuhata GK, Lilienfeld AM. Familial aggregation of lung cancer in humans. Journal of the National Cancer Institute 30: 289-312, 1963.

12. Tokuhata GK, Lilienfeld AM. Familial aggregation of lung cancer among hospital patients. Public Health Reports 78: 277-283, 1963.

13. Lynch HT, Fain PR, Albano WA, Ruma T, Black L, Lynch J, Shonka M. Genetic/epidemiological findings in a study of smoking-associated tumors. Cancer Genetics & Cytogenetics 6: 163-169, 1982.

14. Lynch HT, Kimberling WJ, Markvicka SE, Biscone KA, Lynch JF, Whorton E, Mailliard J. Genetics and smoking-associated cancers. A study of 485 families. Cancer 57: 1640-1646, 1986.

15. Ooi WL, Elston RC, Chen VW, Bailey-Wilson JE, Rothschild H. Increased familial risk for lung cancer. Journal of the National Cancer Institute 76: 217-222, 1986.

16. Wynder EL, Graham EA. Tobacco smoking as a possible etiologic factor in bronchiogenic carcinoma: a study of six hundred and eighty-four proved cases. Journal of the American Medical Association 143: 329-346, 1950.

17. Levin ML, Goldstein H, Gerhardt PR. Cancer and tobacco smoking: a preliminary report. Journal of the American Medical Association 143: 336-338, 1950.

18. Doll R, Hill AB. A study of aetiology of carcinoma of the lung. British Medical Journal 2: 1271-1286, 1952.

19. US Department of Health and Human Services. Smoking and health. 1103. 1964. Washington, DC, US Government Printing Office. Report of the Advisory Committee to the Surgeon General.

20. US Department of Health and Human Services. The health consequences of smoking : cancer and chronic lung disease in the workplace. DHHS publication no. (PHS) 85-50207. 1985. Washington, DC, US Government Printing Office. A report of the Surgeon General.

21. Saracci R. The interactions of tobacco smoking and other agents in cancer etiology. Epidemiologic Reviews 9: 175-193, 1987.

22. Doll R, Peto R. *Effects on health of exposure to asbestos*. London: H.M.Stationary Office, 1985.

23. National Research Council CotBEoIR. Health risks of radon and other internally deposited alpha-emitters. 1988. Washington, DC, National Academy Press. National Academy of Sciences Report: BEIR IV.

24. Fraumeni JF, Blot WJ. Lung and pleura. In: Schottenfeld D, Fraumeni JR, eds. *Cancer Epidemiology and Prevention*. Philadelphia: W.B.Saunders, 1982: 564-582.

25. Tomatis L, Aitio A, Wilbourn J, Shuker L. Human carcinogens so far identified. Japanese Journal of Cancer Research 80: 795-807, 1989.

26. Archer VE, Wagoner JK, Lundin FE. Uranium mining and cigarette smoking effects on man. Journal of Occupational Medicine 15: 204-211, 1973.

27. Hammond EC, Selikoff IJ, Seidman H. Asbestos exposure, cigarette smoking and death rates. Annuals of the New York Acadamy of Sciences 330: 473-490, 1979.

28. Tomatis L. Outdoor air pollution and lung cancer. Annuals of Oncology 2: 265-267, 1991.

29. Vena JE. Air pollution as a risk factor for lung cancer. American Journal of Epidemiology 116: 42-56, 1982.

30. Speizer FE. Assessment of the epidemiologic data relating lung cancer to air pollution. Environmental Health Perspectives 47: 33-42, 1983.

31. Samet JM, Humble CG, Skipper BE, Pathak DR. History of residence and lung cancer risk in New Mexico. American Journal of Epidemiology 125: 800-811, 1987.

32. Svensson C, Pershagen G, Klominek J. Lung cancer in women and type of dwelling in relation to radon exposure. Cancer Research 49: 1861-1865, 1989.

33. Frank AL. The epidemiology and etiology of lung-cancer. Clinics in Chest Medicine 3: 219-228, 1982.

34. Enterline PE, Marsh GM, Esmen NA. Respiratory-disease among workers exposed to man-made mineral fibers. American Review of Respiratory Disease 128: 1-7, 1983.

35. Ayesh R, Idle JR, Ritchie JC, Crothers MS, Hetzel MR. Metabolic oxidation phenotypes as markers of lung cancer susceptibility. Nature 312: 169-170, 1984.

36. Samet JM, Humble CG, Pathak DR. Personal and family history of respiratory disease and lung cancer risk. American Review of Respiratory Disease 134: 466-470, 1986.

37. Skillrud DM, Offord KP, Miller RD. Higher risk of lung cancer in chronic obstructive pulmonary disease. A prospective, matched, controlled-study. Annals of Internal Medicine 105: 503-507, 1986.

38. Nelson N, Levine RJ, Albert RE, Blair AE, Griesemer RA, Landrigan PJ, Stayner LT, Swenberg JA. Contribution of formaldehyde to respiratory cancer. Environmental Health Perspectives 70: 23-35, 1986.

39. Buffler PA, Cooper SP, Stinnett S, Contant C, Shirts S, Hardy RJ, Agu V, Gehan B, Burau K. Air pollution and lung cancer mortality in Harris County, Texas, 1979-1981. American Journal of Epidemiology 128: 683-699, 1988.

40. Kabat GC. Recent developments in the epidemiology of lung-cancer. Seminars in Surgical Oncology 9: 73-79, 1993.

41. Samet JM. The epidemiology of lung cancer. Chest 103: S20-S29, 1993.

42. Miller AB. Trends in cancer mortality and epidemiology. Cancer 51: 2413-2418, 1983.

43. Anonymous. Lung cancer among women. Journal of the American Medical Association 252: 2806, 1984.

44. Devesa SS, Diamond EL. Socioeconomic and racial differences in lung cancer incidence. American Journal of Epidemiology 118: 818-831, 1983.

45. Mason TJ. The descriptive epidemiology of lung cancer. In: Samet JM, ed. *Epidemiology of lung cancer*. New York: Marcel Dekker, Inc., 1994: 51-69.

46. Burns DM. Tobacco smoking. In: Samet JM, ed. *Epidemiology of lung cancer*. New York: Marcel Dekker, Inc., 1994: 15-49.

47. Brunnemann KD, Hoffman D. The pH of tobacco smoke. Food Cosmetics Toxicology 12: 115-124, 1974.

48. Gori GB, Benowitz NL, Lynch CJ. Mouth versus deep airways absorption of nicotine in cigarette smokers. Pharmacology Biochemistry & Behavior 25: 1181-1184, 1986.

49. Wu-Williams AH, Samet JM. Lung cancer and cigarette smoking. In: Samet JM, ed. *Epidemiology of lung cancer*. New York: Marcel Dekker, Inc., 1994: 71-108.

50. Bross ID, Gibson R. Risks of lung cancer in smokers who switch to filter cigarettes. American Journal of Public Health 58: 1396-1403, 1968.

51. Wynder EL, Mabuchi K, Beattie EJ. The epidemiology of lung cancer. Journal of the American Medical Association 213: 2221-2228, 1970.

52. Hammond EC, Garfinkel L, Seidman H, Lew EA. Tar and nicotine content of cigarette smoke in relation to death rates. Environmental Research 12: 263-274, 1976.

53. Rimington J. The effects of filters on the incidence of lung cancer in cigarette smokers. Environmental Research 24: 162-166, 1981.

54. Wynder EL, Kabat GC. Effect of low-yield cigarette smoking on lung cancer risk. Cancer 62: 1223-1230, 1988.

55. Schoenberg JB, Wilcox HB, Mason TJ, Bill J, Stemhagen A. Variation in smoking-related lung cancer risk among New Jersey women. American Journal of Epidemiology 130: 688-695, 1989.

56. Lubin JH, Blot WJ. Assessment of lung cancer risk factors by histologic category. Journal of the National Cancer Institute 73: 383-389, 1984.

57. Cummings KM, Giovino G, Mendicino G. Cigarette advertising and black-white differences in brand preference. Public Health Reports 102: 698-701, 1987.

58. Lubin JH, Blot WJ, Berrino F, Flamant R, Gillis CR, Kunze M, Schmahl D, Visco G. Modifying risk of developing lung cancer by changing habits of cigarette smoking. British Medical Journal 288: 1953-1956, 1984.

59. Becklake MR. Asbestos-related diseases of lung and pleura: current clinical issues. American Review of Respiratory Disease 126: 187-194, 1982.

60. Lakowicz JR, Bevan DR, Riemer SC. Transport of a carcinogen, benzo[*a*]pyrene, from particulates to lipid bilayers: a model for the fate of particle-adsorbed polynuclear aromatic hydrocarbons which are retained in the lungs. Biochimica et Biophysica Acta. 629: 243-258, 1980.

61. Menard H, Noel L, Khorami J, Jouve JL, Dunnigan J. The adsorption of polyaromatic hydrocarbons on natural and chemically modified asbestos fibers. Environmental Research 40: 84-91, 1986.

62. Kelsey KT, Christiani DC, Little JB. Enhancement of benzo[*a*]pyrene-induced sister chromatid exchanges in lymphocytes from cigarette smokers occupationally exposed to asbestos. Journal of the National Cancer Institute 77: 321-327, 1986.

63. Hesterberg TW, Barrett JC. Induction by asbestos fibers of anaphase abnormalities: mechanism for aneuploidy induction and possibly carcinogenesis. Carcinogenesis 6: 473-475, 1985.

64. Oshimura M, Barrett JC. Chemically induced aneuploidy in mammalian cells: mechanisms and biological significance in cancer. Environmental Mutagenesis 8: 129-159, 1986.

65. Okui T, Fujiwara Y. Inhibition of human excision DNA repair by inorganic arsenic and the co-mutagenic effect in V79 Chinese hamster cells. Mutation Research 172: 69-76, 1986.

66. Saracci R, Boffetta P. Interactions of tobacco smoking and other causes of lung cancer. In: Samet JM, ed. *Epidemiology of lung cancer*. New York: Marcel Dekker, Inc., 1994: 465-493.

67. Darby SC, Samet JM. Radon. In: Samet JM, ed. *Epidemiology of lung cancer*. New York: Marcel Dekker, Inc., 1994: 219-243.

68. Speizer FE, Samet JM. Air pollution and lung cancer. In: Samet JM, ed. *Epidemiology of lung cancer*. New York: Marcel Dekker, Inc., 1994: 131-150.

69. Gao YT, Blot WJ, Zheng W, Ershow AG, Hsu CW, Levin LI, Zhang R, Fraumeni JF. Lung cancer among Chinese women. International Journal of Cancer 40: 604-609, 1987.

70. Cagle PT, Mody DR, Schwartz MR. Estrogen and progesterone receptors in bronchogenic carcinoma. Cancer Research 50: 6632-6635, 1990.

71. Annegers JF, Malkasian GD. Patterns of other neoplasia in patients with endometrial carcinoma. Cancer 48: 856-859, 1981.

72. Adami HO, Persson I, Hoover R, Schairer C, Bergkvist L. Risk of cancer in women receiving hormone replacement therapy. International Journal of Cancer 44: 833-839, 1989.

73. Sporn MB, Roberts AB. Role of retinoids in differentiation and carcinogenesis. Journal of the National Cancer Institute 73: 1381-1386, 1984.

74. Le Marchand L, Yoshizawa CN, Kolonel LN, Hankin JH, Goodman M.T. Vegetable consumption and lung cancer risk: A population-based case-control study in Hawaii. Journal of the National Cancer Institute 81: 1158-1164, 1989.

75. Tockman MS. Lung cancer: chemoprevention and intermediate effect markers. International Agency for Research on Cancer Scientific Publications (Lyon) 154: 257-270, 2001.

76. Goodman GE. Prevention of lung cancer. Critical Review of Oncology & Hematology 33: 187-197, 2000.

77. Kroes R, Beems M, Bosland GS, Bunnik SJ, Sinkeldam EJ. Nutritional factors in lung, colon, and prostrate carcinogenesis in animal models. Federation Proceedings 45: 136-141, 1986.

78. Birt DF, Pour PM. Effects of the interaction of dietary fat and protein on N-nitrosobis (2-oxopropyl) amine-induced carcinogenesis and spontaneous lesion in Syrian golden hamsters. Journal of the National Cancer Institute 74: 1121-1127, 1985.

79. Hinds MW, Kolonel LN, Lee J, Hankin JH. Dietary cholesterol and lung cancer risk among men in Hawaii. American Journal of Clinical Nutrition 37: 192-193, 1983.

80. Shaw GL, Falk RT, Dickle LW, Mason TJ, Buffler PA. Lung cancer risk associated with cancer in relatives. Journal of Clinical Epidemiology 44: 429-437, 1991.

81. Mc Duffie HH. Clustering of cancer in families of patients with primary lung cancer. Journal of Clinical Epidemiology 44: 69-76, 1991.

82. Osann KE. Lung cancer in women: the importance of smoking, family history of cancer, and medical history of respiratory disease. Cancer Research 51: 4893-4897, 1991.

83. Sellers TA, Bailey-Wilson JE, Elston RC, Wilson AF, Elston GZ, Ooi WL, Rothschild H. Evidence for Mendelian inheritance in the pathogenesis of lung cancer. Journal of the National Cancer Institute 82: 1272-1279, 1990.

84. Schwartz AG, Yang P, Swanson GM. Familial risk of lung cancer among nonsmokers and their relatives. American Journal of Epidemiology 144: 554-562, 1996.

85. Yang P, Schwartz AG, McAllister AE, Swanson GM, Aston CE. Lung cancer risk in families of nonsmoking probands: heterogeneity by age at diagnosis. Genetic Epidemiology 17: 253-273, 1999.

86. Paine AJ. Heterogeneity of cytochrome P450 and its toxicological significance. Human & Experimental Toxicology 14: 1-7, 1995.

87. Johansson I, Yue QY, Dahl ML, Heim M, Sawe J, Bertilsson L, Meyer UA, Sjoqvist F, Sundberg MI. Genetic analysis of the interethnic differences between Chinese and Caucasians in the polymorphic metabolism of debrisoquine and codeine. Pharmacogenetics 40: 553-556, 1991.

88. Sundberg MI, Johansson I, Persson I, Yue QY, Dahl ML, Bertilsson L, Sjoqvist F. Genetic polymorphism of cytochromes P450: interethnic differences and relationship to incidence of lung cancer. Pharmacogenetics 2: 264-271, 1992.

89. Taioli E, Crofts F, Trachman J, Bayo S, Toniolo P, Garte SJ. Racial differences in CYP1A1 genotypes and function. Toxicology Letters 77: 357-362, 1995.

90. Marez D, Legrand M, Sabbagh N, Guidice JM, Spire C, Lafitte JJ, Meyer UA, Broly R. Polymorphism of the cytochrome P450 CYP2D6 gene in a European population: characterization of 48 mutations and 53 alleles, their frequencies and evolution. Pharmacogenetics 7: 193-202, 1997.

91. Inoue K, Asao T, Shimada T. Ethnic-related differences in the frequency distribution of genetic polymorphisms in the *CYP1A1* and *CYP1B1* genes in Japanese and Caucasian populations. Xenobiotica 30: 285-295, 2000.

92. Kawajiri K, Nakachi K, Imai K, Yoshii A, Shinoda N, Watanabe J. Identification of genetically high risk individuals to lung cancer by DNA polymorphism of the cytochrome P450 *1A1* gene. Federation of European Biochemical Societies Letters 263: 131-133, 1990.

93. Hayashi S, Watanabe J, Nakachi K, Kawajiri K. Genetic linkage of lung cancer associated *MspI* polymorphism with amino acid replacement in the heme-binding region of the human cytochrome P450 *CYP1A1* gene. Journal of Biochemistry 110: 407-411, 1991.

94. Crofts F, Cosma GN, Taioli E, Currie DC, Toniolo P, Garte SJ. A Novel CYP1A1 gene polymorphism in African-Americans. Carcinogenesis 14: 1729-1731, 1993.

95. Nakajima M, Yokoi T, Mizutani M, Shin S, Kadlubar FF, Kamataki T. Phenotyping of CYP1A2 in Japanese population by analysis of caffeine urinary metabolites: absence of mutation prescribing the phenotype in the CYP1A2 gene. Cancer Epidemiology, Biomarkers & Prevention 3: 413-421, 1994.

96. Yokoi T, Sawada M, Kamataki T. Polymorphic drug metabolism: studies with recombinant Chinese hamster cells and analyses in human populations. Pharmacogenetics 5: S65-S69, 1995.

97. Bailey LR, Roodi N, Dupont WD, Parl FF. Association of cytochrome P450 *1B1* (*CYP1B1*) polymorphism with steroid receptor status in breast cancer. Cancer Research 58: 5038-5041, 1998.

98. Stoilov I, Akarsu AN, Alozie I, Child A, Barsoum-Homsy M, Turacli ME, Or M, Lewis RA, Ozdemir N, Brice G, Aktan SG, Chevrette L, Coca-Prados M, Sarfarazi M. Sequence analysis and homology modeling suggest that primary congenital glaucoma on 2p21 results from mutations disrupting either the hinge region or the conserved core structures of cytochrome P4501B1. American Journal of Human Genetics 62: 573-584, 1998.

99. Watanabe J, Shimada T, Gillam EMJ, Ikuta T, Suemasu K, Higashi Y, Gotoh O, Kawajiri K. Association of *CYP1B1* genetic polymorphism with incidence to breast and lung cancer. Pharmacogenetics 10: 25-33, 2000.

100. Gough AC, Miles JS, Spurr NK, Moss JE, Gaedigk A, Eichelbaum M, Wolf CR. Identification of the primary gene defect at the cytochrome P450 CYP2D locus. Nature 347: 773-776, 1990.

101. Kagimoto M, Heim M, Kagimoto K, Zeugin T, Meyer UA. Multiple mutations of the human cytochrome P450IID6 gene (*CYP2D6*) in poor metabolisers of debrisoquine. Journal of Biological Chemistry 265: 17209-17214, 1990.

102. Daly AK, Armstrong M, Monkman SC, Idle ME, Idle JR. The genetic and metabolic criteria for the assignment of debrisoquine hydroxylation (cytochrome P450IID6) phenotypes. Pharmacogenetics 1: 33-41, 1991.

103. Hayashi S, Watanabe J, Kawajiri K. Genetic polymorphisms in the 5′-flanking region change transcriptional regulation of the human cytochrome P450IIEI gene. Journal of Biochemistry 110: 559-565, 1991.

104. Uematsu F, Kikuchi H, Motomiya M, Abe T, Sagami I, Ohmachi T, Wakui A, Kanamuru R, Watanabe M. Association between restriction fragment length polymorphism of the human cytochrome CYP450IIEI gene and susceptibility to lung cancer. Japanese Journal of Cancer Research 82: 254-256, 1991.

105. Strange RC, Jones PW, Fryer AA. Glutathione S-transferase: genetics and role in toxicology. Toxicology Letters 112-113: 357-363, 2000.

106. Board PG. Biochemical genetics of glutathione-*S*-transferase in man. American Journal of Human Genetics 33: 36-43, 1981.

107. Seidegard J, Vorachek WR, Pero RW, Pearson WR. Hereditary differences in the expression of the human glutathione transferase active on trans-stilbene oxide are due to

a gene deletion. Proceedings of the National Academy of Sciences of the United States of America 85: 7293-7297, 1988.

108. Pemble S, Schroeder KR, Spencer SR, Meyer DJ, Hallier E, Bolt HM, Ketterer B, Taylor JB. Human glutathione-S-transferase-theta (GSTT1) - cDNA cloning and the characterization of a genetic-polymorphism. Biochemical Journal 300: 271-276, 1994.

109. Wiencke JK, Pemble S, Ketterer B, Kelsey KT. Gene deletion of glutathione transferase theta 1: correlation with induced genetic damage and potential role in endogenous mutagenesis. Cancer Epidemiology, Biomarkers & Prevention 4: 253-260, 1995.

110. Board PG, Weber GC, Coggan M. Isolation of a cDNA clone and localization of the human glutathione-S-transferase 3 genes to chromosome bands 11q13 and 12q13-14. Annuals of Human Genetics 53: 205-213, 1989.

111. Zimniak P, Nandur B, Pikula S, Bandorowics-Pikula J, Singhal S, Srivastava S, Awashthi S, Awasthi Y. Naturally occurring human glutathione *S*-transferase GSTP1-1 isoforms with isoleucine and valine in position 105 differ in enzymatic properties. European Journal of Biochemistry 224: 893-899, 1994.

112. Nelson HH, Wiencke JK, Christiani DC, Cheng TJ, Zuo ZF, Schwartz BS, Lee BK, Spitz MR, Wang M, Xu XP, Kelsey KT. Ethnic-differences in the prevalence of the homozygous deleted genotype of glutathione-S-transferase-theta. Carcinogenesis 16: 1243-1245, 1995.

113. Watson MA, Stewart RK, Smith GBJ, Massey TE, Bell DA. Human glutathione *S*-transferase P1 polymorphisms: relationship to lung tissue enzyme activity and population frequency distribution. Carcinogenesis 19: 275-280, 1998.

114. Evans D. *N*-acetyltransferase. In: Kalow W, ed. *Pharmacogenetics of Drug Metabolism*. New York: Pergamon, 1992: 95-178.

115. Vatsis KP, Weber WW. Structural heterogeneity of Caucasian N-acetyltransferase at the NAT1 gene locus. Archives of Biochemistry and Biophysics 301: 71-76, 1993.

116. Hirvonen A. Polymorphic NATs and cancer predisposition. In: Vineis P, Malats N, Lang M, d'Errico A, Caporaso N, Cuzick J, Boffetta P, eds. *Metabolic polymorphisms and susceptibility to cancer*. Lyon: Intenational Agency for Research on Cancer, 1999: 251-270.

117. Klebanoff SJ. Antimicrobial mechanisms in neutrophilic polymorphonuclear leukocytes. Seminars in Hematology 12: 117-142, 1975.

118. Hunninghake GW, Crystal RG. Cigarette-smoking and lung destruction-accumulation of neutrophils in the lungs of cigarette smokers. American Review of Respiratory Disease 128: 833-838, 1983.

119. Piedrafita FJ, Molander RB, Vansant G, Orlova EA, Pfahl M, Reynolds WF. An Alu element in the myeloperoxidase promoter contains a composite SP1-thyroid hormone-retinoic acid response element. Journal of Biological Chemistry 271: 14412-14420, 1996.

120. Berwick M, Vineis P. Markers of DNA repair and susceptibility to cancer in humans: an epidemiological review. Journal of the National Cancer Institute 92: 874-897, 2000.

121. Shen MR, Jones IM, Mohrenweiser H. Nonconservative amino acid substitution variants exist at polymorphic frequency in DNA repair genes in healthy humans. Cancer Research 58: 604-608, 1998.

122. Fan F, Liu C, Tavare S, Arnheim N. Polymorphisms in the human DNA repair gene XPF. Mutation Research Genome 406: 115-120, 1999.

123. Khan SG, Metter EJ, Tarone RE, Bohr VA, Grossman L, Hedayati M, Bale SJ, Emmert S, Kraemer KH. A new xeroderma pigmentosum group C poly(AT) insertion/deletion polymorphism. Carcinogenesis 21: 1821-1825, 2000.

124. Garte SJ. The role of ethnicity in cancer susceptibility gene polymorphisms: the example of *CYP1A1* . Carcinogenesis 19: 1329-1332, 1998.

125. Jaiswal AK, Gonzalez FL, Nebert DW. Human dioxin-inducible cytochrome P1-450: complementary DNA and amino acid sequence. Science 228: 80-83, 1985.

126. Jaiswal AK, Gonzalez FL, Nebert DW. Human P1-450 gene sequence and correlation of mRNA with genetic differences in benzo[*a*]pyrene metabolism. Nucleic Acids Research 13: 4503-4520, 1985.

127. Kawajiri K, Watanabe J, Gotoh O, Tagashira Y, Sogawa K, Fujii-Kuriyama Y. Structure and drug inducibility of the human cytochrome P450-c gene. European Journal of Biochemistry 159: 219-225, 1986.

128. Smart J, Daly AK. Variation in induced CYP1A1 levels: relationship to CYP1A1, Ah receptor and GSTM1 polymorphisms. Pharmacogenetics 10: 11-24, 2000.

129. Spurr NK, Gough AC, Stevenson K, Wolf CR. Msp-1 polymorphism detected with a cDNA probe for the P-450 I family on chromosome 15. Nucleic Acids Research 15: 5901, 1987.

130. Zhang ZY, Fasco MJ, Huang L, Guengerich FP, Kaminsky LS. Characterization of purified human recombinant cytochrome P4501A1-Ile462 and -Val462: assessment of a role for the rare allele in carcinogenesis. Cancer Research 56: 3926-3933, 1996.

131. Persson I, Johansson I, Ingelman-Sundberg M. In vitro kinetics of two human CYP1A1 variant enzymes suggested to be associated with interindividual differences in cancer susceptibility. Biochemical Biophysics Research Communications 231: 227-230, 1997.

132. Cascorbi I, Brockmoller J, Roots I. A C4887A polymorphism in exon 7 of human *CYP1A1*: population frequency, mutation linkages, and impact on lung cancer susceptibility. Cancer Research 56: 4965-4969, 1996.

133. Chevalier D, Allorge D, Lo-Guidice JM, Cauffiez C, Lhermitte M, Lafitte JJ, Broly F. Detection of known and two novel (M331I and R464S) missense mutations in the human CYP1A1 gene in a French Caucasian population. Human Mutation 17: 355, 2001.

134. Petersen DD, MeKinney CE, Ikeya K, Smith HH, Bale AE, McBride OW, Nebert DW. Human *CYP1A1* gene: cosegregation of the enzyme inducibility phenotype and an RFLP. American Journal of Human Genetics 48: 720-725, 1991.

135. Kiyohara C, Hirohata T, Inutsuka S. The relationship between aryl hydrocarbon hydroxylase and polymorphisms in the *CYP1A1* gene. Japanese Journal of Cancer Research 87: 18-24, 1996.

136. Tefre T, Ryberg D, Haugen A, Nebert DW, Skaug V, Brogger A, Borresen AL. Human *CYP1A1* (cytochrome P450) gene: lack of association between the *Msp* I restriction fragment length polymorphism and incidence of lung cancer in a Norwegian population. Pharmacogenetics 1: 20-25, 1991.

137. Hirvonen A, Husgafvel-Pursiainen K, Karjalainen A, Anttila S, Vainio H. Point-mutational *Msp* I and *Ile-Val* polymorphisms closely linked in the *CYP1A1* gene: lack of association with susceptibility to lung cancer in a Finnish study population. Cancer Epidemiology, Biomarkers & Prevention 1: 485-489, 1992.

138. Kawajiri K, Nakachi K, Imai K, Watanabe J, Hayashi S. The *CYP1A1* gene and cancer susceptibility. Critical Review of Oncology & Hematology 14: 77-87, 1993.

139. Taioli E, Crofts F, Trachman J, Demopoulos R, Toniolo P, Garte SJ. A specific African-American CYP1A1 polymorphism is associated with adenocarcinoma of the lung. Cancer Research 55: 472-473, 1995.

140. London SJ, Daly AK, Fairbrother KS, Holmes C, Carpenter CL, Navidi WC, Idle JR. Lung cancer risk in African-American in relation to a race-specific CYP1A1 polymorphism. Carcinogenesis 18: 1203-1214, 1995.

141. Xu X, Kelsey KT, Wiencke JK, Wain JC, Christiani DC. Cytochrome P450 CYP1A1 MspI polymorphism and lung cancer susceptibility. Cancer Epidemiology, Biomarkers & Prevention 5: 687-692, 1996.

142. Hamada GS, Sugimura H, Suzuki I, Nagura K, Kiyokawa E, Iwase T, Tanaka M, Takahashi T, Watanabe S, Kino I. The heme-binding region polymorphism of cytochrome P4501A1 (Cyp1A1), rather than the RsaI polymorphism of IIE1 (CypIIE1), is associated with lung cancer in Rio de Janeiro. Cancer Epidemiology, Biomarkers & Prevention 4: 63-67, 1995.

143. Sugimura H, Wakai K, Genka K, Nagura K, Igarashi H, Nagayama K, Ohkawa A, Baba S, Morris BJ, Tsugane S, Ohno Y, Gao CM, Li ZY, Takezaki T, Tajima K, Iwamasa T. Association of Ile462Val (Exon7) polymorphism of cytochrome P4501A1 with lung cancer in the Asian population: further evidence from a case-control study in Okinawa. Cancer Epidemiology, Biomarkers & Prevention 7: 413-417, 1998.

144. Ishibe N, Wiencke JK, Zuo ZF, McMillan A, Spitz MR, Kelsey KT. Susceptibility to lung cancer in light smokers, associated with *CYP1A1* polymorphisms in Mexican- and African-Americans. Cancer Epidemiology, Biomarkers & Prevention 6: 1075-1080, 1997.

145. Ikeya K, Jaiswal AK, Owens RA, Jones JE, Nebert DW, Kimura S. Human CYP1A2: sequence, gene structure, comparison with the mouse and rat orthologous gene, and differences in liver 1A2 mRNA expression. Molecular Endocrinology 3: 1399-1408, 1989.

146. Quattrochi LC, Tukey RH. The human cytochrome Cyp1A2 gene contains regulatory elements responsive to 3-methylcholanthrene. Molecular Pharmacology 36: 66-71, 1989.

147. Welfare MR, Aitkin M, Bassendine MF, Daly AK. Detailed modelling of caffeine metabolism and examination of the CYP1A2 gene: lack of a polymorphism in CYP1A2 in Caucasians. Pharmacogenetics 9: 367-375, 1999.

148. Nakajima M, Yokoi T, Mizutani M, Kinoshita M, Funayama M, Kamataki T. Genetic polymorphism in the 5'-flanking region of human CYP1A2 gene: effect on the CYP1A2 inducibility in humans. Journal of Biochemistry 125: 803-808, 1999.

149. Chida M, Yokoi T, Fukui T, Kinoshita M, Yokota J, Kamataki T. Detection of three genetic polymorphisms in the 5'-flanking region and intron 1 of human CYP1A2 in the Japanese population. Japanese Journal of Cancer Research 90: 899-902, 1999.

150. Sachse C, Brockmoller J, Bauer S, Roots I. Functional significance of a C->A polymorphism in intron I of the cytochrome P450 CYP1A2 gene tested with caffeine. British Journal of Clinical Pharmacology 47: 445-449, 1999.

151. Chevalier D, Cauffiez C, Allorge D, Lo-Guidice JM, Lhermitte M, Lafitte JJ, Broly F. Five novel natural allelic variants-951A>C, 1042G>A (D348N), 1156A>T (I368F), 1217G>A (C406Y) and 1291C>T (C431Y)-of the human CYP1A2 gene in a French Caucasian population. Human Mutation 17: 355-356, 2001.

152. Huang JD, Guo WC, Lai MD, Guo YL, Lambert GH. Detection of a novel cytochrome P-450 1A2 polymorphism (F21L) in Chinese. Drug Metabolism and Disposition 27: 98-101, 1999.

153. Ohgaki H, Kusama K, Matsukura N, Morino K, Hasekawa H, Sato S, Takayama S, Sugimura T. Carcinogenicity in mice of a mutagenic compound, 2-amino-3-methylimidaxo[4,5-*f*]quinoline, from broiled sardine, cooked beef and beef extract. Carcinogenesis 5: 921-924, 1984.

154. Ohgaki H, Hasekawa H, Suenaga M, Sato S, Takayama S, Sugimura H. Carcinogenicity in mice of a mutagenic compound, 2-amino-3,8-dimethylimidazo[4,5-*f*]quinoxaline (MeIQx) from cooked foods. Carcinogenesis 8: 665-668, 1987.

155. Wilson VL, Weston A, Manchester DK, Glenwood ET, Roberts DW, Kadlubar FF, Wild CP, Montesano R, Willey JC, Mann DL, Harris CC. Alkyl and aryl carcinogen adducts detected in human peripheral lung. Carcinogenesis 10: 2149-2153, 1989.

156. Shimada T, Hayes CL, Yamazaki H, Amin S, Hecht SS, Guengerich FP, Sutter TR. Activation of chemically diverse procarcinogens by human cytochrome P-450 1B1. Cancer Research 56: 2979-2984, 1996.

157. Tang YM, Wo YY, Stewart J, Hawkins AL, Griffin CA, Sutter TR, Greenlee WF. Isolation and characterization of the human cytochrome P450 CYP1B1 gene. Journal of Biological Chemistry 271: 28324-28330, 1996.

158. McLellan RA, Oscarson M, Hidestrand M, Leidvik B, Jonsson E, Otter C, Ingelman-Sundberg M. Characterization and functional analysis of two common human cytochrome P450 1B1 variants. Archives of Biochemistry and Biophysics 378: 175-181, 2000.

159. Stoilov I, Akarsu AN, Sarfarazi M. Identification of three different truncating mutations in cytochrome P4501B1 (CYP1B1) as the principal cause of primary congenital glaucoma (Buphthalmos) in families linked to the GLC3A locus on chromosome 2p21. Human Molecular Genetics 6: 641-647, 1997.

160. Shimada T, Watanabe J, Kawajiri K, Sutter TR, Guengerich FP, Gillam EMJ, Inoue K. Catalytic properties of polymorphic human cytochrome P4501B1 variants. Carcinogenesis 20: 1607-1613, 1999.

161. Legrand-Andreoletti M, Stucker I, Marez D, Galais P, Cosme J, Sabbagh N, Spire C, Cenee S, Lafitte JJ, Beaune P, Broly F. Cytochrome P450 *CYP2D6* gene polymorphism and lung cancer susceptibility in Caucasians. Pharmacogenetics 8: 7-14, 1998.

162. Caporaso NE, Tucker MA, Hoover RN, Hayes RB, Pickle LW, Issa HJ, Muschik GM, Green-Gallo L, Buivys D, Aisner S, Resau J, Trump BF, Tollerud D, Weston A, Harris A, Harris CC. Lung cancer and the debrisoquine metabolic phenotype. Journal of the National Cancer Institute 82: 1264-1271, 1990.

163. Hirvonen A, Husgafvel-Pursiainen K, Anttila S, Karjalainen A, Pelkonen O, Vainio H. PCR-based *CYP2D6* genotyping for Finnish lung cancer patients. Pharmacogenetics 3: 19-27, 1993.

164. Wolf CR, Smith CAD, Gough AC, Moss JE, Vallis KA, Howard G, Carey FJ, Mills K, Mcnee W, Carmichael J, Spurr NK. Relationship between the debrisoquine hydroxylase polymorphism and cancer susceptibility. Carcinogenesis 13: 1035-1038, 1992.

165. Wolf CR, Smith G. Cytochrome P450 CYP2D6. In: Vineis P, Malats N, Lang M, d'Errico A, Caporaso N, Cuzick J, Boffetta P, eds. *Metabolic polymorphisms and susceptibility to cancer*. Lyon: Intenational Agency for Research on Cancer, 1999: 209-229.

166. London SJ, Daly AK, Leathart JBS, Navidi WC, Carpenter CL, Idle JR. Genetic polymorphism of *CYP2D6* and lung cancer risk in African-Americans and Caucasians in Los Angeles County. Carcinogenesis 18: 1203-1214, 1997.

167. Umeno M, McBride OW, Yang CS, Gelboin HV, Gonzalez FJ. Human ethanol-inducible P450IIE1: complete gene sequence, promoter characterization, chromosome mapping, and cDNA-directed expression. Biochemistry 27: 9006-9013, 1988.

168. McBride OW, Umeno M, Gelboin HV, Gonzalez FJ. A Taq I polymorphism in the human P450IIE1 gene on chromosome 10 (CYP2E). Nucleic Acids Research 15: 10071, 1987.

169. Brockmoller J, Cascorbi I, Kerb B, Roots I. Combined analysis of inherited polymorphisms in arylamine *N*-acetyltransferase 2, glutathione *S*-transferase M1 and T1, microsomal epoxide hydrolase, and cytochrome P450 enzymes as modulators of bladder cancer risk. Cancer Research 56: 3915-3925, 1996.

170. Hu Y, Hakkola J, Oscarson M, Ingelman-Sundberg M. Structural and functional characterization of the 5'-flanking region of the rat and human cytochrome P450 2E1 genes: identification of a polymorphic repeat in the human gene. Biochemical Biophysics Research Communications 263: 286-293, 1999.

171. McCarver DG, Byun R, Hines RN, Hichme M, Wegenek W. A genetic polymorphism in the regulatory sequences of human CYP2E1: association with increased chlozoxazone hydroxylation in the presence of obesity and ethanol intake. Toxicology & Applied Pharmacology 152: 276-281, 1998.

172. Hu Y, Oscarson M, Johansson I, Yue QY, Dahl ML, Tabone M, Arinco S, Albano E, Ingelman-Sundberg M. Genetic polymorphism of human CYP2E1: characterization of two variant alleles. Molecular Pharmacology 51: 370-376, 1997.

173. Fairbrother KS, Grove J, de Waziers I, Steimel DT, Day CP, Crespi CL, Daly AK. Detection and characterization of novel polymorphisms in the CYP2E1 gene. Pharmacogenetics 8: 543-552, 1998.

174. Watanabe J, Hayashi S, Nakachi K, Imai K, Suda Y, Sekine T, Kawajiri K. PstI and RsaI RFLPs in compltete linkage disequilibrium at the CYP2E gene. Nucleic Acids Research 18: 7194, 1990.

175. Persson I, Johansson I, Berglingo H, Dahl ML, Seidegard J, Rylander R, Rannung A, Hogberg J, Sundberg MI. Genetic polymorphism of cytochrome P450 *2E1* in a Swedish population. Relationship to incidence of lung cancer. Federation of European Biochemical Societies Letters 319: 207-211, 1993.

176. Kato S, Shields PG, Caporaso NE, Sugimura H, Trivers GE, Tucker MA, Trump BF, Weston A, Harris CC. Analysis of cytochrome P450 2EI genetic polymorphisms in relation to human lung cancer. Cancer Epidemiology, Biomarkers & Prevention 3: 515-518, 1994.

177. Kato S, Shields PG, Caporaso NE, Hoover RN, Trump BF, Sugimura H, Weston A, Harris CC. Cytochrome P450IIEI genetic polymorphisms, racial variation, and lung cancer risk. Cancer Research 52: 6712-6715, 1992.

178. Uematsu F, Ikawa S, Kikuchi H, Sagami I, Kanamuru R, Abe T, Satoh K, Motomiya M, Watanabe M. Restriction fragment length polymorphism of the human CYP2E1 (cytochrome P450IIE1) gene and susceptibility to lung cancer: possible relevance to low smoking exposure. Pharmacogenetics 4: 58-63, 1994.

179. d'Errico A, Malats N, Vineis P, Boffetta P. Review of studies of selected metabolic polymorphisms and cancer. In: Vineis P, Malats N, Lang M, d'Errico A, Caporaso N, Cuzick J, Boffetta P, eds. *Metabolic polymorphisms and susceptibility to cancer*. Lyon: Intenational Agency for Research on Cancer, 1999: 323-393.

180. Strange RC, Fryer AA. The glutathione S-transferases: influence of polymorphism on cancer susceptibility. In: Vineis P, Malats N, Lang M, d'Errico A, Caporaso N, Cuzick J, Boffetta P, eds. *Metabolic polymorphisms and susceptibility to cancer*. Lyon: Intenational Agency for Research on Cancer, 1999: 231-249.

181. Inskip A, Elexperu-Camiruaga J, Buxton N, Dias PS, Macintosh J, Campbell D, Jones PW, Yengi L, Talbot JA, Strange RC, Fryer AA. Identification of polymorphism at the glutathione S-transferase, GSTM3 locus: evidence for linkage with *GSTM1*A*. Biochemical Journal 312: 713-716, 1995.

182. Benhamou S, Lee W, Alexandrie AK, Boffetta P, Brockmoller J, Clapper ML, Deakin M, Dolzan V, Bouchardy C, Gaspari L, Kremers P, Marchand LL, London SJ, Nazar-Stewart V, Persson I, Romkes M, Ryberg D, Saarikoski ST, Stucker I, To-Figueras J, Taioli E, Brennan P. Meta-analysis of effects of glutathione *S*-transferase M1 polymorphism and smoking on lung cancer risk. 2001.

183. Anttila S, Hirvonen A, Vainio H, Husgafvel-Pursiainen K, Hayes JD, Ketterer B. Immunohistochemical localization of glutathione *S*-transferases in human lung. Cancer Research 53: 5643-5648, 1993.

184. Saarikoski ST, Voho A, Reinikanen M, Anttila S, Karjalainen A, Malaveille C, Vainio H, Husgafvel-Pursiainen K, Hirvonen A. Combined effect of polymorphic *GST* genes on individual susceptibility to lung cancer. International Journal of Cancer 77: 516-521, 1998.

185. Jourenkova-Mironova N, Wikman H, Bouchardy C, Voho A, Dayer P, Benhamou S, Hirvonen A. Role of glutathione *S*-transferase *GSTM1*, *GSTM3*, *GSTP1* and *GSTT1* genotypes in modulating susceptibility to smoking-related lung cancer. Pharmacogenetics 8: 502, 1998.

186. Deakin M, Elder J, Hendrickse C, Peckham D, Baldwin D, Pantin C, Wild N, Leopard P, Bell DA, Jones P, Duncan H, Brannigan K, Alldersea J, Fryer AA, Strange RC. Glutathione S-transferase GSTT1 genotypes and susceptibility to cancer: studies of interactions with GSTM1 in lung, oral, gastric and colorectal cancers. Carcinogenesis 17: 881-884, 1996.

187. Kelsey KT, Spitz MR, Zuo ZF, Wiencke JK. Polymorphisms in the glutathione S-transferase class *mu* and *theta* genes interact and increases susceptibility to lung cancer in minority populations (Texas, United States). Cancer Causes and Control 8: 554-559, 1997.

188. Ali-Osman F, Akande O, Antoun G, Mao J, Buolamwin J. Molecular cloning, characterization and expression in *Escherichia coli* of full-length cDNAs of three human glutathione S-transferase Pi gene variants. Journal of Biological Chemistry 15: 10004-10012, 1997.

189. Hu X, O'Donnel R, Srivastava SK, Xia H, Zimniak P, Nanduri B, Bleicher RJ, Awasthi S, Awasthi YC, Ji X, Singh SV. Active site architecture of polymorphic forms of human glutathione *S*-transferase P1-1 accounts for their enantioselectivity and disparate activity in the glutathione conjugation of 7,8-dihydroxy-9,10-oxy-7,8,9,10-tetrahydrobenzo[*a*]pyrene. Biochemical Biophysics Research Communications 235: 424-428, 1997.

190. Ryberg D, Skaug V, Hewer A, Phillips DH, Harries LW, Wolf CR, Ogreid D, Ulvik A, Vu P, Haugen A. Genotypes of glutathione transferase M1 and P1 and their significance for lung DNA adduct levels and cancer risk. Carcinogenesis 18: 1285-1289, 1997.

191. Harris MJ, Coggan M, Langton L, Wilson SR, Board PG. Polymorphism of the Pi class glutathione S-transferase in normal populations and cancer patients. Pharmacogenetics 8: 27-31, 1998.

192. Hein DW, Doll MA, Fretland AJ, Leff MA, Webb SJ, Xiao GH, Devanaboyina U-S, Nangju NA, Feng Y. Molecular genetics and epidemiology of the *NAT1* and *NAT2* acetylation polymorphisms. Cancer Epidemiology, Biomarkers & Prevention 9: 29-42, 2000.

193. Wikman H, Thiel S, Jager B, Schmezer P, Spiegelhalder B, Edler L, Dienemann H, Kayser K, Schulz V, Drings P, Bartsch H, Risch A. Relevance of *N*-acetyltransferase 1 and 2 (*NAT1, NAT2*) genetic polymorphisms in non-small cell lung cancer susceptibility. Pharmacogenetics 11: 157-168, 2001.

194. Deguchi T. Sequences and expression of alleles of polymorphic arylamine *N*-acetyltransferase of human liver. Journal of Biological Chemistry 267: 18140-18147, 1992.

195. Vatsis KP, Martell KJ, Weber WW. Diverse point mutations in the human gene for polymorphic N- acetyltransferase. Proceedings of the National Academy of Sciences of the United States of America 88: 6333-6337, 1991.

196. Lin JJ, Han C-Y, Link BK, Hardy S. Slow acetylator mutations in the human polymorphic *N*-acetyltransferase gene in 786 Asians, Blacks, Hispanics, and Whites: application to metabolic epidemiology. American Journal of Human Genetics 52: 827-834, 1993.

197. Martinez C, Agundez JAG, Olivera M, Martin R, Ladero JM, Benitez J. Lung cancer and mutations at the polymorphic NAT2 gene locus. Pharmacogenetics 5: 207-214, 1995.

198. Woolhouse NM, Qureshi MM, Bayoumi RA. A new mutation C759T in the polymorphic *N*-acetyltransferase (NAT2) gene. Pharmacogenetics 7: 83-84, 1997.

199. Agundez JAG, Olivera M, Ladero JM, Lescure-Rodriguez A, Ledesma MC, Diaz-Rubio M, Meyer UA, Benitez J. Increased risk for hepatocellular carcinoma in NAT2-slow acetylators and CYP2D6-rapid metabolizers. Pharmacogenetics 6: 501-512, 1996.

200. Leff MA, Fretland AJ, Doll MA, Hein DW. Novel human *N*-acetyltransferase 2 alleles that differ in mechanism for slow acetylator phenotype. Journal of Biological Chemistry 274: 34519-34522, 1999.

201. Ferguson RJ, Doll MA, Rustan TD, Gray K, Hein DW. Cloning, expression, and functional characterization of two mutant (NAT2 191 and NAT2 341/803) and wild-type human polymorphic *N*-acetyltransferase (NAT2) alleles. Drug Metabolism and Disposition 22: 371-376, 1994.

202. Bell DA, Taylor JA, Butler MA, Stephens EA, Wiest J, Brubaker LH, Kadlubar FF, Lucier GW. Genotype-phenotype discordance for human arylamine N- acetyltransferase (NAT2) reveals a new slow-acetylator allele common in African-Americans. Carcinogenesis 14: 1689-1692, 1993.

203. Agundez JAG, Olivera M, Martinez C, Ladero JM, Benitez J. Identification and prevalence study of 17 allelic variants of the human NAT2 gene in a white population. Pharmacogenetics 6: 423-428, 1996.

204. Lin HJ, Han C-Y, Lin BK, Hardy S. Ethnic distribution of slow acetylator mutations in the polymorphic *N*-acetyltransferase (NAT2) gene. Pharmacogenetics 4: 125-134, 1994.

205. Shishikura K, Hohjoh H, Tokunaga K. Novel allele containing a 190C>T nonsynonymous substitution in the N-acetyltransferase (NAT2) gene. Human Mutation 15: 581, 2000.

206. Blum M, Grant DM, McBride W, Heim M, Meyer UA. Human arylamine *N*-acetyltransferase genes: isolation, chromosomal localization, and functional expression. DNA Cell Biology 9: 193-203, 1990.

207. Ohsako S, Deguchi T. Cloning and expression of cDNAs for polymorphic and monomorphic arylamine *N*-acetyltransferases from human liver. Journal of Biological Chemistry 265: 4630-4634, 1990.

208. Doll MA, Jiang W, Deitz AC, Rustan TD, Hein DW, Hein DW. Identification of a novel allele at the human NAT1 acetyltransferase locus. Biochemical Biophysics Research Communications 233: 584-591, 1997.

209. Hughes NC, Janezic SA, Mcqueen KL, Jewett MAS, Castranio T, Bell DA, Grant DM. Identification and characterization of variant alleles of human acetyltransferase NAT1 with defective function using *p*-aminosalicylate as an in-vivo and in-vitro probe. Pharmacogenetics 8: 55-66, 1998.

210. Hubbard A, Moyes C, Wyllie AH, Smith CAD, Harrison DJ. *N*-Acetyltransferase 1: two polymorphisms in coding sequence identified in cororectal cancer patients. British Journal of Cancer 77: 913-916, 1998.

211. de Leon JH, Vatsis KP, Weber WW. Catalytic activities of human *N*-acetyltransferase (NAT1) variants expressed in COS-1 cells. Faseb Journal 10: A456, 1996.

212. Butcher NJ, IIett KF, Minchin RF. Functional polymorphism of the human arylamine *N*-acetyltransferase type 1 gene caused by C190T and G560A mutations. Pharmacogenetics 8: 67-72, 1998.

213. Lin HJ, Probst-Hensch NM, Hughes NC, Sakamoto GT, Louie AD, Kau IH, Lin BK, Lee DB, Lin J, Frankl HD, Lee ER, Hardy S, Grant DM, Haile RW. Variants of *N*-acetyltransferase NAT1 and a case-control study of colorectal adenomas. Pharmacogenetics 8: 269-281, 1998.

214. Deitz AC, Doll MA, Fretland AJ, Hein DW. *Homo sapiens N*-acetyltransferase NAT1 (allele NAT1*18A) gene, complete cds. Genbank AF032677. 1997.

215. Deitz AC, Doll MA, Fretland AJ, Hein DW. *Homo sapiens N*-acetyltransferase NAT1 (allele NAT1*18B) gene, complete cds. Genbank AF032678. 1997.

216. Deitz AC, Fretland AJ, Leff MA, Hein DW. *Homo sapiens N*-acetyltransferase-1 (NAT1) gene, NAT1*26A allele, complete cds. Genbank AF071552. 1998.

217. Deitz AC, Fretland AJ, Leff MA, Doll MA, Hein DW. *Homo sapiens N*-acetyltransferase-1 (NAT1) gene, NAT1*26B allele, complete cds. Genbank AF067408. 1998.

218. Smelt VA, Upton A, Adjaye J, Payton MA, Johnson N, Mardon HJ, Sim E. Expression of arylamine N-acetyltransferase in preterm placentas and in human preimplantation embryos. 2000.

219. Lo-Guidice JM, Marez D, Barat F, Spire C, Chevalier D, Broly F. Human *N*-acetyltransferase 1 (NAT1) gene, NAT1*28 allele. Genbank AF082904. 1999.

220. Lo-Guidice JM, Marez D, Barat F, Spire C, Chevalier D, Broly F. Human *N*-acetyltransferase 1 (NAT1) gene, NAT1*29 allele. Genbank AF082903. 1999.

221. Cascorbi I, Brockmoller J, Mrozikiewicz PM, Bauer S, Loddenkemper R, Roots I. Homozygous rapid arylamine *N*-acetylatransferase (NAT2) genotype as a susceptibility factor for lung cancer. Cancer Research 56: 3961-3966, 1996.

222. Burgess EJ, Trafford JA. Acetylator phenotype in patients with lung carcinoma: a negative report. European Journal of Respiratory Disease 67: 17-19, 1985.

223. Nyberg F, Hou S-M, Hemminki K, Lambert B, Pershagen G. Glutathione *S*-transferase M1 and *N*-acetyltransferase 2 genetic polymorphisms and exposure to tobacco smoke in nonsmoking and smoking lung cancer patients and population controls. Cancer Epidemiology, Biomarkers & Prevention 7: 875-883, 1998.

224. Bouchardy C, Mitrunen K, Wikman H, Husgafvel-Pursiainen K, Dayer P, Benhamou S, Hirvonen A. *N*-acetyltransferase NAT1 and NAT2 genotypes and lung cancer risk. Pharmacogenetics 8: 291-298, 1998.

225. Zhou W, Thurston SW, Liu G, Xu LL, Miller DP, Wain JC, Lynch TJ, Su L, Christiani DC. The interaction between microsomal epoxide hydrolase polymorphisms and cumulative cigarette smoking in different histological subtypes of lung cancer. Cancer Epidemiology, Biomarkers & Prevention 10: 461-466, 2001.

226. Hassett C, Aicher L, Sidhu JS, Omiecinski CJ. Human microsomal epoxide hydrolase: genetic polymorphism and functional expression *in vitro* of amino acid variants. Human Molecular Genetics 3: 421-428, 1994.

227. Benhamou S, Reinikanen M, Bouchardy C, Dayer P, Hirvonen A. Association between lung cancer and microsomal epoxide hydrolase genotypes. Cancer Research 58: 5291-5293, 1998.

228. Persson I, Johansson I, Lou Y-C, Yue Q-Y, Duan L-S, Bertilsson L, Ingelman-Sundberg M. Genetic polymorpism of xenobiotic metabolizing enzymes among Chinese lung cancer patients. International Journal of Cancer 81: 329, 1999.

229. Smith CAD, Harrison DJ. Association between polymorphism in gene for microsomal epoxide hydrolase and susceptibility to emphysema. The Lancet 350: 630-633, 1997.

230. Mallet WG, Mosebrook DR, Trush MA. Activation of (+/-)-trans-7,8-dihydroxy-7,8-dihydrobenzo[*a*]pyrene to diolepoxides by human polymorphonuclear leukocytes or myeloperoxidase. Carcinogenesis 12: 521-524, 1991.

231. London SJ, Lehman TA, Taylor JA. Myeloperoxidase genetic polymorphism and lung cancer risk. Cancer Research 57: 5001-5003, 1997.

232. Schabath MB, Spitz MR, Zhang X, Delclos GL, Wu X. Genetic variants of myeloperoxidase and lung cancer risk. Carcinogenesis 21: 1163-1166, 2000.

233. Lunn RM, Helzlsouer KJ, Parshad R, Umbach DM, Harris EL, Sanford KK, Bell DA. XPD polymorphisms: effects on DNA repair proficiency. Carcinogenesis 21: 551-555, 2000.

234. Butkiewicz D, Rusin M, Enewold L, Shields PG, Chorazy M, Harris CC. Genetic polymorphisms in DNA repair genes and risk of lung cancer. Carcinogenesis 22: 593-597, 2001.

235. Hoffmann D, Hoffmann I. The changing cigarette, 1950-1995. Journal of Toxicology & Environmental Health 50: 307-364, 1997.

236. Hayashi S, Watanabe J, Kawajiri K. High susceptibility to lung cancer analyzed by combined genotypes of CYP450 *1A1* and glutathione-*S*-transferase μ-genes. Japanese Journal of Cancer Research 83: 866-870, 1992.

237. Nakachi K, Imai K, Hayashi S, Kawajiri K. Polymorphisms of the CYP1A1 and glutathione S-transferase genes associated with susceptibility to lung cancer in relation to cigarette dose in a Japanese population. Cancer Research 53: 2994-2999, 1993.

238. Kihara M, Kihara M, Noda K. Risk of smoking for squamous and small call carcinomas of the lung modulated by combinations of CYP1A1 and GSTM1 gene polymorphisms in a Japanese population. Carcinogenesis 16: 2331-2336, 1995.

239. Alexandrie AK, Ingelman-Sundberg M, Seidegard J, Tornling G, Rannung A. Genetic susceptibility to lung cancer with special emphasis on *CYP1A1* and *GSTM1*: a study on host factors in relation to age at onset, gender and histological cancer types. Carcinogenesis 15: 1785-1790, 1994.

240. Anttila S, Luostarinen L, Hirvonen A, Elovaara E, Karjalainen A, Nurminen T, Hayes JD, Vainio H, Ketterer B. Pulmonary expression of glutathione *S*-transferase M3 in lung cancer patients: association with *GSTM1* polymorphism, smoking, and asbestos exposure. Cancer Research 55: 3305-3309, 1995.

241. Jourenkova N, Reinikanen M, Bouchardy C, Pursiainen KH, Dayer P, Benhamou S, Hirvonen A. Effect of glutathione-*S*-transferase *GSTM1* and *GSTT1* genotypes on lung cancer risk in smokers. Pharmacogenetics 7: 515-518, 1997.

242. Kihara M, Kihara M, Kazumasa N. Lung cancer risk of the *GSTM1* null genotype is enhanced in the presence of the *GSTP1* mutated genotype in male Japanese smokers. Cancer Letters 137: 53-60, 1999.

243. Nakachi K, Imai K, Hayashi S, Watanabe J, Kawajiri K. Genetic susceptibility to squamous cell carcinoma of the lung in relation to cigarette smoking dose. Cancer Research 51: 5177-5180, 1991.

244. Nakachi K, Hayashi S, Kawajiri K, Imai K. Association of cigarette smoking and CYP1A1 polymorphism with adenocarcinoma of the lung by grades of differentiation. Carcinogenesis 16: 2209-2213, 1995.

245. Kato S, Bowman ED, Harrington AM, Blomeke B, Shield PG. Human lung carcinogen-DNA adduct levels mediated by genetic polymorphisms in vivo. Journal of the National Cancer Institute 87: 902-907, 1995.

246. London SJ, Daly AK, Cooper J, Navidi WC, Carpenter CL, Idle JR. Polymorphism of glutathione *S*-transferase M1 and lung cancer risk among African-Americans and Caucasians in Los Angeles County, California. Journal of the National Cancer Institute 87: 1246-1252, 1995.

247. Lan Q, He X, Costa DJ, Tian L, Rothman N, Hu G, Mumford JL. Indoor coal combustion emissions, GSTM1 and GSTT1 genotypes, and lung cancer risk: a case-control study in Xuan Wei, China. Cancer Epidemiology, Biomarkers & Prevention 9: 605-608, 2000.

248. Godschalk RWL, Dallinga JW, Wikman H, Risch A, Kleinjans JCS, Bartsch H, Schooten F-JV. Modulation of DNA and protein adducts in smokers by genetic polymorphisms in *GSTM1*, *GSTT1*, *NAT1* and *NAT2*. Pharmacogenetics 11: 389-398, 2001.

249. Rojas M, Cascorbi I, Alexandrov K, Kriek E, Auburtin G, Mayer L, Kopp-Schneider A, Roots I, Bartsch H. Modulation of benzo[*a*]pyrene diolepoxide-DNA adduct levels in human white blood cells by *CYP1A1*, *GSTM1* and *GSTT1* polymorphism. Carcinogenesis 21: 35-41, 2000.

250. Guengerich FP, Shimada T. Oxidation of toxic and carcinogenic chemicals by human cytochrome P-450 enzymes. Chemical Reserach in Toxicology 4: 391-407, 1991.

251. Shimada T, Yun CH, Yamazaki H, Gautier JC, Beaune PH, Guengerich FP. Characterization of human lung microsomal cytochrome P-450 1A1 and its role in the oxidation of chemical carcinogens. Molecular Pharmacology 41: 856-864, 1992.

252. Willey JC, Coy E, Brolly C, Utell MJ, Frampton MW, Hammersley J, Thilly WG, Olson D, Cairns K. Xenobiotic metabolism enzyme gene expression in human bronchial epithelial and alveolar macrophage cells. American Journal of Respiratory Cell and Molecular Biology 14: 262-271, 1996.

253. Raucy JL, Ingelman-Sundberg M, Carpenter S, Rannung A, Rane A, Franklin M, Romkes M. Drug metabolizing enzymes in lymphocytes. Journal of Biochemistry & Molecular Toxicology 13: 223-226, 1999.

254. Sutter TR, Tang YM, Hayes CL, Wo YY, Jabs EW, Li X, Yin H, Cody CW, Greenlee WF. Complete cDNA sequence of a human dioxin-inducible mRNA identifies a new gene subfamily of cytochrome P450 that maps to chromosome 2. Journal of Biological Chemistry 269: 13092-13099, 1994.

255. Petruzzelli S, Camus AM, Carrozzi L, Ghelarducci L, Rindi M, Menconi G, Angeletti CA, Ahotupa M, Hietanen E, Aitio A, Saracci R, Bartsch H, Giuntini C. Long-lasting effects of tobacco smoking on pulmonary drug-metabolizing enzymes: a case-control study on lung cancer patients. Cancer Research 48: 4695-4700, 1988.

256. McLemore TL, Adelberg S, Liu MC, McMahon NA, Yu SJ, Hubbard WC, Czerwinski M, Wood TG, Storeng R, Lubet RA, Eggleston JC, Boyd MR, Hines RN. CYP1A1 gene in patients with lung cancer: evidence for cigarette smoke-induced gene expression in normal lung tissue and for altered gene regulation in primary pulmonary carcinomas. Journal of the National Cancer Institute 82: 1333-1339, 1990.

257. Willey JC, Coy EL, Frampton MW, Torres A, Apostolakos MJ, Hoehn G, Schuermann WH, Thilly WG, Olson DE, Hammersley JR, Crespi CL, Utell MJ. Quantitative RT-PCR measurement of cytochrome P450 1A1, 1B1, and 2B7, microsomal epoxide hydrolase, and NADPH oxidoreductase expression in lung cells of smokers and nonsmokers. American Journal of Respiratory Cell and Molecular Biology 17: 114-124, 1997.

258. Rojas M, Camus AM, Alexandrov K, Husgafvel-Pursiainen K, Anttila S, Vainio H. Stereoselective metabolism of (-)-benzo[*a*]pyrene-7,8-diol by human lung microsomes and peripheral blood lymphocytes: effect of smoking. Carcinogenesis 13: 933, 1992.

259. Alexandrov K, Rojas M, Geneste O, Castegnaro M, Camus AM, Petruzzelli S. An improved fluorometric assay for dosimetry of benzo[*a*]pyrene diol-epoxide-DNA adducts in smokers' lung: comparisons with total bulky adducts and aryl hydrocarbon hydroxylase activity. Cancer Research 52: 6253, 1992.

260. Bauer E, Guo Z, Ueng YF, Bell LC, Zeldin D, Guengerich FP. Oxidation of benzo[*a*]pyrene by recombinant human cytochrome P450 enzymes. Chemical Reserach in Toxicology 8: 136-142, 1995.

261. Shou M, Korzekwa KR, Crespi CL, Gonzalez FJ, Gelboin HV. The role of 12 cDNA-expressed human, rodent, and rabbit cytochromes P450 in the metabolism of benzo[*a*]pyrene and benzo[*a*]pyrene *trans*-7,8-dihydrodiol. Molecular Carcinogenesis 10: 159-168, 1994.

262. Dogra S, Whitelaw ML, May BK. Transcriptional activation of cytochrome P450 genes by different classes of chemical inducers. Clinical and Experimental Pharmacology and Physiology 25: 1-9, 1998.

263. Spivack SD, Hurteau GJ, Reilly AA, Aldous KM, Ding X, Kaminsky LS. CYP1B1 expression in human lung. Drug Metabolism and Disposition 29: 916-922, 2001.

264. Murray GI, Taylor MC, McFayden MCE, McKay JA, Greenlee WF, Burke MD, Melvin WT. Tumor-specific expression of cytochrome CYP1B1. Cancer Research 57: 3026-3031, 1997.

265. Raunio H, Hakkola J, Hukkanen J, Lassila A, Paivarinta K, Pelkonen O, Anttila S, Piipari R, Boobis A, Edwards RJ. Expression of xenobiotic-metabolizing CYPs in human pulmonary tissue. Experimental in Toxicological Pathology 51: 412-417, 1999.

266. Villard PH, Seree E, Lacarelle B, Therenefenoglio MC, Barra Y, Attolini L, Bruguerole B, Durand A, Catalini J. Effect of cigarette smoke on hepatic and pulmonary

cytochromes P450 in mouse: evidence for CYP2E1 induction in lung. Biochemical Biophysics Research Communications 202: 1731-1737, 1994.

267. Foiles PG, Akerkar SA, Carmella SG, Kagan M, Stoner GD, Resau JH, Hecht SS. Mass spectrometric analysis of tobacco-specific nitrosamine-DNA adducts in smokers and nonsmokers. Chemical Reserach in Toxicology 4: 364-368, 1991.

268. Mustonen R, Schoket B, Hemminki K. Smoking-related DNA adducts: ^{32}P-postlabeling analysis of 7-methylguanine in human bronchial and lymphocyte DNA. Carcinogenesis 14: 151-154, 1993.

269. Petruzzelli S, Tavanti LM, Celi A, Giuntini C. Detection of N^7-methyldeoxyguanosine adducts in human pulmonary alveolar cells. American Journal of Respiratory Cell and Molecular Biology 15: 216-223, 1996.

270. Landgraf A, Reckmann B, Pingoud A. Direct analysis of polymerase chain reaction products using enzyme-linked immunosorbent assay techniques. Analytical Biochemistry 198: 86-91, 1991.

271. Carcillo JA, Parise RA, Romkes-Sparks M. Comparison of the enzyme-linked oligonucleotide sorbent assay to the ^{32}P-labeled PCR/Southern blotting technique in quantitative analysis of human and rat RNA. PCR Methods and Applications 3: 292-297, 1994.

272. Bieche I, Olivi M, Champeme M-H, Vidaud D, Lidereau R, Vidaud M. Novel approach to quantitative polymerase chain reaction using real-time detection: application to the detection of gene amplification in breast cancer. International Journal of Cancer 78: 661-666, 1998.

273. Lie YS, Petropoulos CJ. Advances in quantitative PCR technology: 5' nuclease assays. Current Opinion in Biotechnology 9: 43-48, 1998.

274. Wang T, Brown MJ. mRNA quantitation by real time TaqMan polymerase chain reaction: validation and comparison with RNase protection. Analytical Biochemistry 269: 198-201, 1999.

275. Lee LG, Connell CR, Bloch W. Allelic discrimination by nick-translation PCR with fluorogenic probes. Nucleic Acids Research 21: 3761-3766, 1993.

276. Orlando C, Pinzani P, Pazzagli M. Developments in quantitative PCR. Cinical Chemistry and Laboratory in Medicine 36: 255-269, 1998.

277. Lyamichev V, Brow MAD, Dahlberg JE. Structure-specific endonucleolytic cleavage of nucleic acids by eubacterial DNA polymerases. Science 260: 778-783, 1993.

278. Gibson UEM, Heid CA, Williams PM. A novel method for real time quantitative RT-PCR. Genome Research 6: 995-1001, 1996.

279. Yajima T, Yagihashi A, Kameshima H, Kobayashi D, Furuya D, Hirata K, Watanabe N. Quantitative reverse transcription-PCR assay of the RNA component of human telomerase using the TaqMan fluorogenic detection system. Clinical Chemistry 44: 2441-2445, 1998.

280. Goldsworthy SM, Stockton PS, Trempus CS, Foley JF, Maronpot RR. Effects of fixation on RNA extraction and amplification from laser capture microdissected tissue. Molecular Carcinogenesis 25: 86-91, 1999.

281. Godfrey TE, Kim SH, Chavira M, Ruff DW, Warren RS, Gray JW, Jensen RH. Quantitative mRNA expression analysis from formalin-fixed, paraffin-embedded tissues using 5' nuclease quantitative reverse transcription-polymerase chain reaction. Journal of Molecular Diagnostics 2: 84-91, 2000.

282. Landi MT, Bertazzi PA, Shield PG, Clark G, Lucier GW, Garte SJ, Cosma G, Caporaso NE. Association between CYP1A1 genotype, mRNA expression and enzymatic activity in humans. Pharmacogenetics 4: 242-246, 1994.

283. Bell DA, Thompson CL, Taylor JA, Miller CR, Perera F, Hsieh LL, Lucier GW. Genetic monitoring of human polymorphic cancer susceptibility genes by polymerase chain reaction: application to glutathione transferase mu. Environmental Health Perspectives 98: 113-117, 1992.

284. Park JY, Schantz SP, Stern JC, Kaur T, Lazarus P. Association between glutathione *S*-transferase pi genetic polymorphisms and oral cancer risk. Pharmacogenetics 9: 497-504, 1999.

285. Lancaster JM, Brownlee HA, Bell DA, Futreal PA, Marks JR, Berchuck A, Wiseman RW, Taylor JA. Microsomal epoxide hydrolase polymorphism as a risk factor for ovarian cancer. Molecular Carcinogenesis 17: 160-162, 1996.

286. Begg CB, Zhang ZF. Statistical analysis of molecular epidemiology studies employing case-series. Cancer Epidemiology, Biomarkers & Prevention 3: 173-175, 1994.

287. Khoury MJ, Flanders WD. Nontraditional epidemiologic approaches in the analysis of gene-environment interaction: case-control studies with no controls! American Journal of Epidemiology 144: 207-213, 1996.

288. Yang Q, Khoury MJ, Flanders WD. Sample size requirements in case-only designs to detect gene-environment interaction. American Journal of Epidemiology 146: 713-720, 1997.

289. Eltom SE, Zhang L, Jefcoate CR. Regulation of cytochrome P450 1B1 in mouse hepa-1 variant cell lines: a possible role for aryl hydrocarbon receptor nuclear translocator (ARNT) as a suppressor of *CYP1B1* gene expression. Molecular Pharmacology 55: 594-604, 1999.

290. Hayes CL, Spink DC, Spink BC, Cao JO, Walker NJ, Sutter TR. 17ß-Estradiol hydroxylation catalyzed by human cytochrome P450 1B1. Proceedings of the National Academy of Sciences of the United States of America 93: 9776-9781, 1996.

291. Liehr JG, Ricci MJ. 4-Hydroxylation of oestrogens as marker of human mamary tumors. Proceedings of the National Academy of Sciences of the United States of America 93: 3294-3296, 1996.

292. London SJ, Smart J, Daly AK. Lung cancer risk in relation to genetic polymorphisms of microsomal epoxide hydrolase among African-Americans and Caucasians in Los Angeles County. Lung Cancer 28: 147-155, 2000.

293. Zhao H, Spitz MR, Gwyn KM, Wu X. Microsomal epoxide hydrolase polymorphisms and lung cancer risk in non-Hispanic whites. Molecular Carcinogenesis 33: 99-104, 2002.

294. Risch A, Wikman H, Thiel S, Schmezer P, Edler L, Drings P, Dienemann H, Kayser K, Schulz V, Spiegelhalder B, Bartsch H. Glutathione S-transferase M1, M3, T1, and P1 polymorphisms and susceptibility to non-small-cell lung cancer subtypes and hamartomas. Pharmacogenetics 11: 757-764, 2002.

295. Dybdahl M, Vogel U, Frentz G, Wallin H, Nexo BA. Polymorphisms in the DNA repair gene XPD: correlations with risk and age at onset of basal cell carcinoma. Cancer Epidemiology, Biomarkers & Prevention 8: 77-81, 1999.

296. Hu JJ, Hall C, Torti FM, McCullough DL, Lohman K, Grossman L. Deficient DNA repair and human prostate cancer risk. Proceedings of the American Association in Cancer Research 41: 360, 2000.

297. Sturgis EM, Zheng R, Li L, Castillo EJ, Eicher SA, Chen M, Strom SS, Spitz MR, Wei Q. XPD/ERCC2 polymorphisms and risk of head and neck cancer: a case-control analysis. Carcinogenesis 21: 2219-2223, 2000.

298. Spitz MR, Wu X, Wang Y, Wang L-E, Shete S, Amos CI, Guo Z, Lei L, Mohrenweiser H, Wei Q. Modulation of nucleotide excision repair capacity by XPD polymorphisms in lung cancer patients. Cancer Research 61: 1354-1357, 2001.

299. David-Beabes GL, Lunn RM, London SJ. No association between *XPD* (Lys751Gln) polymorphism or the *XRCC3* (Thr241Met) polymorphism and lung cancer risk. Cancer Epidemiology, Biomarkers & Prevention 10: 911-912, 2001.

300. Zhou W, Liu G, Miller DP, Thurston SW, Xu LL, Wain JC, Lynch CJ, Su L, Christiani DC. Gene-environment interaction for the *ERCC2* polymorphisms and cumulative cigarette smoking exposure in lung cancer. Cancer Research 62: 1377-1381, 2002.

301. Hanna IH, Dawling S, Roodi N, Guengerich FP, Parl FF. Cytochrome P450 *1B1* (*CYP1B1*) pharmacogenetics: association of polymorphisms with functional differences in estrogen hydroxylation activity. Cancer Research 60: 3440-3444, 2000.

302. Cavalieri EL, Stack DE, Devanesan PD, Todorovic R, Dwivedy I, Higginbotham S, Johansson SL, Patil KD, Gross ML, Gooden JK, Ramanathan R, Cerny RL, Rogan GE. Molecular origin of cancer: catechol estrogen-3,4-quinones as endogenous tumor initiators. Proceedings of the National Academy of Sciences of the United States of America 94: 10937-10942, 1997.

303. Kim JH, Stansbury KH, Walker NJ, Trush MA, Strickland PT, Sutter TR. Metabolism of benzo[*a*]pyrene and benzo[*a*]pyrene-7,8-diol by human cytochrome P450 1B1. Carcinogenesis 19: 1847-1853, 1998.

304. Nakajima T, Elovaara E, Anttila S, Hirvonen A, Camus AM, Hayes JD, Ketterer B, Vainio H. Expression and polymorphism of glutathione *S*-transferase in human lungs: risk factors in smoking-related lung cancer. Carcinogenesis 16: 707-711, 1995.

305. Williams JA. Single nucleotide polymorphisms, metabolic activation and environmental carcinogenesis: why molecular epidemiologists should think about enzyme expression. Carcinogenesis 22: 209-214, 2001.

306. Cascorbi I, Henning S, Brockmoller J, Gephart J, Meisel C, Muller JM, Loddenkemper R, Roots I. Substantially reduced risk of cancer of the aerodigestive tract in subjects with variant -463A of the myeloperoxidase gene. Cancer Research 60: 644-649, 2000.

307. Shields PG. Molecular epidemiology of lung cancer. Annuals of Oncology 10: S7-S11, 1999.

308. Toussaint C, Albin N, Massaad L, Grunenwald D, Parise O, Morizet J, Gouyette A, Chabot GG. Main drug- and carcinogen-metabolizing enzyme systems in human non-small cell lung cancer and peritumoral tissues. Cancer Research 53: 4608-4612, 1993.

309. Hukkanen J, Hakkola J, Anttila S, Piipari R, Karjalainen A, Pelkonen O, Raunio H. Detection of mRNA encoding xenobiotic-metabolizing cytochrome P450s in human bronchoalveolar macrophages and peripheral blood lymphocytes. Molecular Carcinogenesis 20: 224-230, 1997.

310. Sabadie N, Richter-Reichelm HB, Saracci R, Mohr U, Bartsch H. Interindividual differences in oxidative benzo[*a*]pyrene metabolism by normal and tumorous surgical lung specimens from 105 lung cancer patients. International Journal of Cancer 27: 417-425, 1981.

311. Murray GI, Melvin WT, Greenlee WF, Burke MD. Regulation, function, and tissue-specific expression of cytochrome P450 CYP1B1. Annuals Review of Pharmacological and Toxicology 41: 297-316, 2001.

312. Shehin SE, Stephenson RO, Greenlee WF. Transcriptional regulation of the human CYP1B1 gene: evidence of involvement of an aryl hydrocarbon receptor response element in constitutive expression. Journal of Biological Chemistry 275: 6770-6776, 2000.

313. Crawford EL, Weaver DA, DeMuth JP, Jackson CM, Khuder SA, Frampton MW, Utell MJ, Thilly WG, Willey JC. Measurement of cytochrome P450 2A6 and 2E1 gene expression in primary human bronchial epithelial cells. Carcinogenesis 19: 1867-1871, 1998.