CHAPTER I

INTRODUCTION

Lung cancer has been the most common cancer in the world. The 1990 global lung cancer incidence rate of 37.5 per 100,000 gave rise to 771.8 thousand cases. Between 1985 and 1990, lung cancer burden of the developing countries went up by 25% (from 261,000 to 327,100), while that of the developed countries went up by only 7% (Notani, 2001). By the year 2002, there were 1.35 million new lung cancer cases, representing 12.4% of all new cancers. It was also the most common cause of death from cancer, with 1.18 million deaths, or 17.6% of the world total. The age standardized incidence rate (ASR) and mortality rate in male and female are 35.5, 31.2, 12.1 and 10.3 respectively (Parkin, 2005).

In Thailand, cancer is defined as one of the major health problems and has been the most common cause of death since 1999 (Thailand., 2004). The incidence of lung cancer in Thailand is the second most common cancer in males after liver cancer and the fourth in females after cervix, breast and liver cancer. Interesting the incidence rate of lung cancer in Northern Thailand is higher in both sexs than other areas (Sriplung, 2005, Vatanasapt, 2002).

1.1. Risk factors

There are two major causes of lung cancer:

1.1.1. Environmental factors

The carcinogenic effects of a large number of environmental or industrial chemicals have first been described in humans. The environmental or life-

style factors are important contributors of lung cancer. For example, tobacco smoking, radon gas, fibers and dusts from industries such as asbestos and silica have been reported.

(1) Tobacco Smoke

Tobacco smoke contains irritants, oxidants, free radicals, and more than 50 carcinogenic agents. Polycyclic aromatic hydrocarbons (PAHs) and N-nitrosamines are the two major classes of tobacco-related inhaled carcinogens. The PAHs and N-nitrosamines exert their mutagenic/ carcinogenic action through the formation of DNA adducts (Dettterbeck, 2001). Worldwide geographic patterns of lung cancer incidence and mortality are very much influenced by tobacco smoking. For the year 2000, an estimated 85% of lung cancer in men and 47% of lung cancer in women is the consequence of tobacco smoking (Parkin, 2005).

(2) Radon

Radon is a colorless and odourless gas generated by the breakdown of radioactive radium, which in turn is the decay product of uranium, found in the earth's crust and also accumulates in indoor air. The radiation decay products ionize genetic material, causing mutations that sometimes turn cancerous (Dettterbeck, 2001). Radon exposure is the second major cause of lung cancer after smoking. It has been reported that the risk of developing lung cancer from residential radon exposure increases with radon concentration and exposure duration (Chen, 2005, Field, 2006) .

(3) Occupational exposure

Asbestos is a class of naturally occurring fibrous minerals that have been used extensively. More than 3000 products containing asbestos in some form have been identified, including fire-resistant cloth, cement wick, electrical compliance, water

pipes, floor and roofing tiles, and brake liner. Asbestos is a well recognized carcinogen that most likely acts as a tumor promoter and is the most frequent occupational cause of human lung cancers (Dettterbeck, 2001). Occupational exposure to asbestos review study also found the summary risk estimate of pleural mesothelioma for household exposure was 8.1 (95% confidence interval [CI]:5.3-12) and for neighborhood was 7.0 (95% CI: 4.7-11) (Bourdes, 2000).

Silica exposure occurs in mining and sandblasting, as well as in mansonry, concrete, gypsum, and pottery industries. Crystalline silica inhaled in the form of quartz of crystobalite from occupational sources was classified as carcinogenic to humans (Dettterbeck, 2001). A study of lung cancer in relation to silica exposure and silicosis in Tobi area, well known for traditional Japanese pottery, found that silica exposure significantly increased the lung cancer mortality in this area (Tsuda, 2002).

1.1.2. Genetic factors

Cancer is now regarded as an acquired genetic disease due to multiple alterations of genes that affect cell growth or differentiation. Genetic host factors can interact with environmental carcinogens, i.e., carcinogens in the diet, tobacco smoke and ambient air due to environmental or occupational sources, to place an individual at a greater or lesser risk of a particular cancer than another individual. Inherited susceptibilities to cancer result from variations in the genetic code that alter either protein expression, function, or localization (Harris, 2000, Lai, 1999, Shields, 2000).

The molecular epidemiology of lung cancer has received wide spread attention because the tobacco smoking etiology is well established, but only some smokers develop lung cancer whereas others do not. Benzo(a)pyrene (BaP) and most other

procarcinogenic PAHs presented in tobacco smoke require metabolic activation by phase I enzymes, especially CYP1A1, to become reactive (*e.g.*, BPDE) and subsequently to bind to DNA. Some of reactive intermediates are in turn detoxified by phase II enzymes, such as glutathione S-transferases (GSTs). The xenobiotic-metabolizing enzymes involved in such reactions exhibit polymorphic variation in humans. Such interindividual variation largely depends on genetic variation, diet, lifestyle, medication and other environmental factors acting as other modifiers.

At present, the majority of the large numbers of human genes involved in carcinogen metabolism are known to show genetic polymorphisms and these polymorphisms are assumed to contribute to individual susceptibility to develop lung cancer and other tobacco-related cancers. Adding to this concept, individual variation also exists in the capacity to repair DNA damage, and the genes involved in DNA repair show genetic polymorphisms (Husgafvel-Pursiainen, 2004). A highly simplified scheme of major gene products participating in the activation and inactivation of benzo(a)pyrene is shown in Figure. 1.

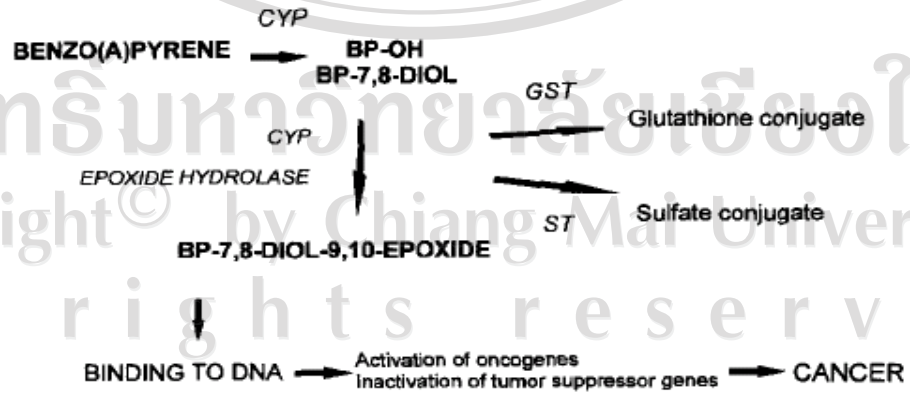

Figure 1.1 Activation and inactivation of BaP (From Raunio, 1995)

A genetic polymorphism is defined as a gene variant that is prevalent at least at a frequency of 1% in a given population. It appears that human can be differentially susceptible to carcinogenic insults depending on polymorphism in a variety of xenobiotic metabolizing enzyme. Polymorphism of genes involved in lung cancer risk are xenobiotic metabolism genes, which endcode enzyme cytochrome P450, glutathione-S-transferases, and gene which encode enzyme involved in phagocytosis (*e.g.* myeloperoxidase), DNA repair genes (*e.g.*, hOGG1), tumor suppressor genes (*e.g.*, p53) tumor development and invasion genes (*e.g.*, MMP-1)

(1) Cytochrome P450s (CYPs)

Cigarette smoke contains several thousand chemicals, of which about 50 compounds are known carcinogens, including PAHs, aromatic amines and N-nitrosamine compounds. Some of these compounds are reactive carcinogens, but most are procarcinogens, which need to be activated by phase I enzymes such as those encoded by the CYP supergene family, and converted into reactive carcinogens. All these reactive carcinogens can bind to DNA and form DNA adducts capable of inducing mutations and initiating carcinogenesis. CYPs are a multi-gene superfamily of mixed function monooxygenases. Based on sequence homology, the CYP superfamily is divided into 10 subfamilies, CYP1-CYP10 (Kiyohara, 2002). Two gene polymorphisms have been reported in CYP1A1 gene. One is in the 3' non-coding region, characterized by the T6235C transition (CYP1A1 MspI) and the other located in exon, which is caused by a point mutation from A to G (A4889G transition) resulting in a replacement of isoleucine by valine (Hayashi, 1991). The CYP1A1 polymorphism has been associated with alterations in regulation and transcript half-life, which result in elevated induction of the enzyme,

and thus, increased levels of activated intermediates (Cosma, 1993, Landi, 1994). The CYP1A1 homozygous genotype CYP1A1 Msp I or CYP1A1 Ile-Val polymorphism alone (Hayashi, 1991, Kawajiri, 1990) or combined with GSTM1 null has been associated with increased risk of lung cancer among Asian populations. (Hayashi, 1992, Kawajiri, 1990).

(2) Glutathione-S-transferases (GSTs)

Following phase I reaction, phase II enzymes such as GSTs are responsible for detoxification of activated forms PAH epoxides. GSTs are constitutively found in a wide variety of tissues, with different characteristic patterns of GST isozymes. GST genes form a superfamily of at least 13 genes consisting of five distinct families, named alpha (GSTA), sigma (GSTS), mu (GSTM), pi (GSTP) and theta (GSTT). Certain genes within the GSTM, GSTT and GSTP subfamilies (GSTM1, GSTT1 and GSTP1) are polymorphic in humans and the levels of individual enzymes expressed can be influenced by induction and by genetic polymorphism (Rebbeck, 1997). Carriers of homozygous deletions in the GSTM1 and GSTT1 genes have an absence of GSTM and GSTT enzyme activity, respectively. (Pemble, 1994, Sweeney, 2003). Most studies of GSTM1 and lung cancer have reported association between the null phenotype or null genotype and elevated risk (el-Zein, R, 1997, Ford, 2000, Persson, 1999).

(3) Myeloperoxidase gene

Myeloperoxidase (MPO) is a lysosomal phase I enzyme expressed predominantly in polymorphonuclear leukocytes, monocytes and macrophages. It converts a benzo(a)pyrene into highly reactive and carcinogenic metabolite benzo(a)pyrene 7,8-diol-9,10 epoxide (BPDE). It also activates tobacco smoke-

derived carcinogens such as PAHs, aromatic and heterocyclic amines and induces the formation of endogenous carcinogenic free radicals. The MPO gene is located on chromosome 17q23.1. A common G-463A transition in the gene promoter region resulting in its transcriptional activity decrease is known to reduce risk of lung cancer (Kiyohara, 2005).

(4) DNA repair genes

8-hydroxyguanine (8-OHG) is a major DNA lesion produced by oxygen a radical, one of the DNA adducts that can be induced by smoking. Primary defense mechanisms include antioxidants and enzymes such as glutathione peroxidase. The formation of 8-OHG in DNA causes G:C to T:A transversion, since 8-OHG pairs with adenine as well as cytosine. The human 8-oxoguanine DNA glycosylase 1 (hOGG1) gene encodes base excision repair proteins for 8-OHG in double-stranded DNA. The OGG1 protein possesses the ability to excise 8-OHG paired with cytosine (Boiteux, 2000, Kiyohara, 2002). Homozygous for the hOGG1-326Cys allele have a lower capacity to repair 8OHG than others (Lee, 2005). A number of studies have been shown that hOGG1 polymorphism play a role in susceptibility to lung cancer (Hung, 2005, Kohno, 2006, Le Marchand, 2002).

(5) p53 tumor suppressor gene

The p53 tumor suppressor protein is important in cell-cycle control, apoptosis, and DNA repair. This gene is a key and potent mediator of cellular response against genotoxic insults. A polymorphism at codon 72 of p53 gene resulting in a change of amino acid Arg to Pro has been reported. The two polymorphic variants differ in both their biochemical and biological properties. The Arg72 encoded p53 protein is more efficient in inducing apoptosis and suppressing transformation than Pro72 encoded

p53 protein (Thomas, 1999). A number of studies have demonstrated that Pro72 allele carriers have an increased risk of lung cancer (Fan, 2000, Jin, 1995, Wang, 1998, Wang, YC, 1999).

(6) Matrix Metalloproteinase-1 (MMP-1)

Matrix metalloproteinase is a family of metalloenzymes, the MMPs, is largely responsible for degradation of the extracellular matrix (ECM). MMP-1 is one of only a few MMPs capable of degrading interstitial collagen types I, II, and III, and MMP-1 is widely expressed at low levels in normal physiology. However, expression increases markedly in disease pathologies, and increased expression of MMP-1 has been associated with a poor prognosis in several cancers (Wyatt, 2002). An imbalance between MMPs and naturally occurring MMP inhibitors may cause excess extracellular matrix destruction, allowing cancer cells to invade surrounding tissues and metastasize, and permitting angiogenesis to occur (Egeblad, 2002). The human MMP-1 gene (MMP1) is located on chromosome 11q22-23. The insertion of extra guanine (G) at position -1607 creates a binding site (5'GGAT-3) for a group of transcription factors. The MMP1 promoter polymorphism is believed to be associated with enhanced gene transcription and, thus, increased enzyme activity (Rutter, 1998). The MMP-1 2G/2G genotype has been associated with significantly higher risk of lung cancer in one previous study, specifically for males, current smokers, and heavy smokers (Su, 2005, Zhu, 2001).

The role of genes and environment in the etiology of diseases has been at the center of an intent debate. In the case of lung cancer, several chemical exposures have been indicated as a cause of lung cancer beyond any reasonable doubt, while genetic predisposition seems to modulate their effects. As in the case of cigarette

smoking, smoking itself is a cause of lung cancer in that it increase the risk considerably, but it is neither a necessary nor a sufficient cause. Smoking may contribute to a causal complex together with other contributing factors. One of such factors is genetic predisposition, including the special type of metabolism. An epidemiologically relevant question is what proportion of cancers is attributable to genetic influence? Although it is well established the highly penetrant genes explain less than 5% of cancer, it is much less clear what proportion is attributable to low penetrant genes and their interaction with environmental exposure. It is believed that everyone may have a unique combination of polymorphic traits that modify genetic susceptibility and response to drugs, chemicals and carcinogens. Developments in molecular biology have led to growing interest in investigation of biological markers, which may increase predisposition to lung carcinogenesis. Therefore, the high-risk genotype of an individual could be determined easily. As there are a great numbers of carcinogen-activating and -detoxifying enzymes, the variation in their expression and the complexity of exposures to tobacco carcinogens, the existence of multiple alleles at loci of those enzymes may result in differential susceptibilities of individuals. Therefore, in this study the polymorphic of well-investigated genetic susceptibility genes, in relations to lung cancer development w examined. The results obtained from the investigation may enable us to understand the interplay of environmental and genetic polymorphisms at multiple loci and thus help identify individuals who are at increased risk for lung cancer.

II. LITERATURE REVIEWS

Lung cancer has been the most common cancer in the world since 1985 (Parkin, 1993) and by 2002, there were 1.35 million new cases, representing 12.4% of all new cancers.It was also the most common cause of death from cancer, with 1.18 million deaths, or 17.6% of the world total. Almost half (49.9%) of the cases occur in the developing countries of the world(Parkin, 2005).

Lung cancer initially occurs in normal bronchiolar epithelium, it arises as a result of the accumulation of multiple independent molecular lesion that affect critical genetic pathways, distinct from the random genetic damage that occurs in advanced neoplasm. Furthermore, it is believed that specific proto-oncogene and tumor-suppressor genes are the target of somatic mutations in lung cancer, resulting from the genotoxic effects of tobacco-smoke carcinogen such as BaP, nitrosamines, and irradiation such as radon gas (Carleton, 1995, Coleman, 2002).

2.1. Incidences of lung cancer in Thailand

In Thailand, the first population-based cancer registry was begun in Chiang Mai in 1986. At present, there are five registries actively working in different parts of the country. Provinces representative for the four regions of the country are Chiang Mai and Lampang in the North, Khon Kaen in the Northeast, Bangkok in the Central, and Songkhla in the South (Figure 1.2). The cancer registry shows that lung cancer is the most frequent malignancy for both sexes in northern region. The annual Age-Standardized incidence Rate (ASR) of lung cancer in males and females in Chiang Mai and Lampang, are 36.5, 53.5 or 25.1, 25.3 per 100,000 population respectively (Figure 1.3) (Sriplung, 2005). Tobacco smoking is believed to be the

main causative factor for the incidence of lung cancer in this area (Simarak, 1977), anothers are chronic benign respiratory diseases due to the fungus *Microsporum canis* (Nakachi, 1999, Suttajit, 1994) or exposure to indoor radon radiation (Wiwatanadate, 2001) or mutagenic environmental air (Vinitketkumnuen, 2002). However the fact that there are lower incidence rate of oral-pharyngeal cancer, which is also a cigarette-related cancer, than that in Songkhla (Sriplung, 2005) indicating a need for further studies.

Figure 1.2. Cancer Registries in Thailand (From Sriplung, 2005)

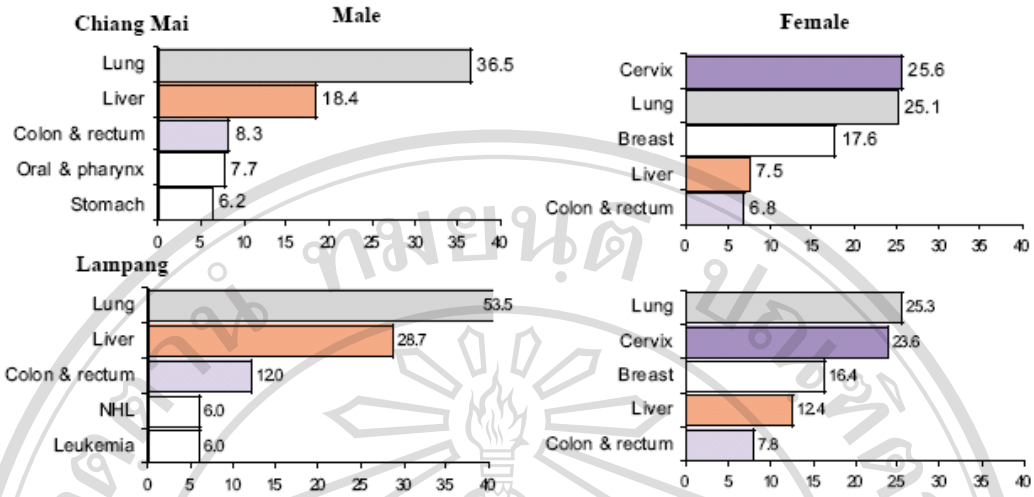

Figure 1.3. Leading Cancers among Northern Thai Population by Registry in Northern Thailand. (Age-standardized incidence rate per 100,000 population) (From Sriplung, 2005)

2.2. The relationship between polymorphisms of xenobiotic metabolizing enzymes and susceptibility to lung cancer

2.2.1. Cytochrome P450s (CYPs),

Cytochrome P450 is a group of monomeric heme containing enzymes; belong to the monooxygenase gene superfamily. Cytochrome P450 enzymes are grouped in families and subfamilies according to their amino acid sequence homology (Fig. 2). Families CYP1, CYP2, and CYP3 are the main layer in the oxidative metabolism of xenobiotics (Castell, 2005).

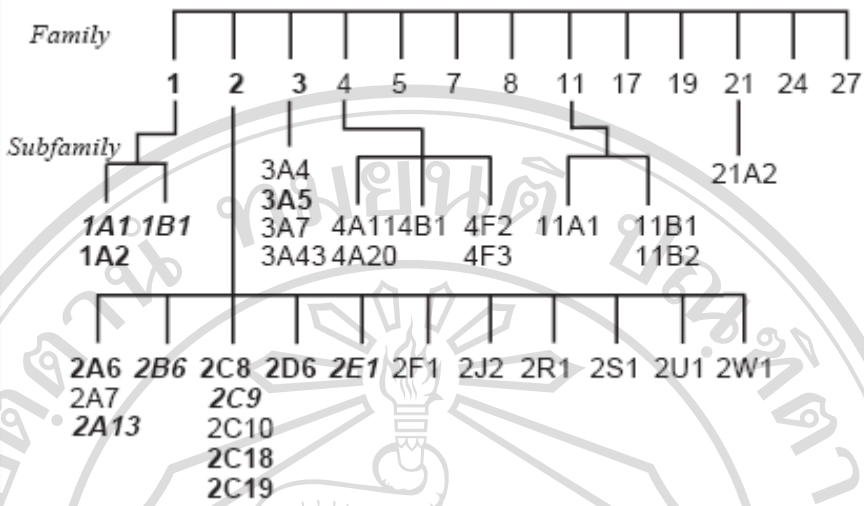

Figure 1.4. Human CYP involved in xenobiotic metabolism. Genes are classified into families and subfamilies according to the degree of nucleotide and amino acid sequence homology. Families 1, 2, and 3 are largely responsible of biotransformation of xenobiotics (bold). Italic typed indicates the predominant enzymes in human lung. (Figure from Castell, 2005)

(1) Cytochrome 1A1 [cytochrome P450, family 1, subfamily A, polypeptide1 (CYP1A1)]

CYP1A1 is located on chromosome 15q22-q24. CYP1A1 enzyme is involved in the activation of benzo(a)pyrene, other PAHs and aromatic amines, all of them are major classes of tobacco procarcinogens. The polymorphism in the 3' noncoding region (3'-UTR) of the CYP1A1 gene arises from a T→C transition, thus the allele bears a MspI cleavage site, in contrast with the wild-type variant. The human enzyme CYP1A1, which is well conserved is involved in the activation of major classes of tobacco procarcinogens, like PAHs and aromatic amines, and is present in many epithelial tissues. About 10% of the Caucasian population has a highly inducible form of the CYP1A1 enzyme (termed B[a]P-hydroxylase or previously arylhydrocarbon hydroxylase), which is associated with an increased risk for bronchial, laryngeal, and oral cavity tumors in smokers (Nebert, 1996). Beginning in

1973 studied on BaP hydroxylase inducibility and bronchogenic carcinoma (Kellermann, 1973), and the association of the genetic polymorphism of CYP1A1 and cancer started after cosegregation of the CYP1A1 high inducibility phenotype and polymorphism of the MspI restriction site (Petersen, 1991). The CYP1A1 Ile-Val mutation (at 462 exon 7) in the heme-binding region results in a 2-fold increase in microsomal enzyme activity and is in complete linkage disequilibrium in Caucasians with the CYP1A1 MspI mutation, which has also been associated experimentally with increased catalytic activity (Landi, 1994). There was a 3-fold elevation in CYP1A1 enzymatic activity in exon 7 variant genotypes. When Msp1 and exon 7 genotypes were combined, there was an increased CYP1A1 inducibility and enzymatic activity in subjects with the exon 7 polymorphism, and in subjects with both polymorphisms (Crofts, 1994). Although the Ile-Val mutation in the CYP1A1 allele did not increase activity *in vitro* (Zhang, 1996) it might be linked to other functional polymorphisms, for example in the regulatory region important for CYP1A1 inducibility. Smokers with the exon 7 Ile-Val mutation were found to have more PAH-DNA adducts in their WBCs than smokers without the variant (Mooney, 1997). The amount of these adducts is also elevated in cord blood and placenta of newborns with the CYP1A1-MspI polymorphism (Whyatt, 1998). In lung parenchymal tissue of smokers, the concentrations of BPDE and bulky (PAH)-DNA adducts were positively correlated with CYP1A1 enzyme activity (Alexandrov, 1992). Significantly ethnic differences in the frequency of homozygous variant alleles have been observed, CYP1A1 MspI alleles is found in 1.2% Caucasians, 14% in Asians, 5.9% in Africans and homozygous CYP1A1 Val alleles is found in 0.6% Caucasians, 4.9% in Asians. Both the MspI and Val alleles are rarer in Caucasian than in Asian populations (Garte,

2001). The relationship between CYP1A1 variants and lung cancer risk in various ethnic populations has been examined in several studies. Early Japanese studies pointed to an increased risk for lung cancer in association with both the CYP1A1 MspI variant (Kawajiri, 1990, Okada, 1994) and CYP1A1 Val alleles (Hayashi, 1991). The CYP1A1 genotype was particularly important at a low level of smoking and in the development of squamous cell carcinoma (Nakachi, 1991). However these findings were not confirmed in studies conducted in Norway (Tefre, 1991), the United States (Shields, 1993) and Sweden (Alexandrie, 1994).

(2) Cytochrome 2E1 (cytochrome P450, family 2, subfamily E, polypeptide 1)

The gene is located on 10q24.3-qter. The ethanol-inducible CYP2E1 metabolizes many known procarcinogens, including N'Nitrosanornicotine (NNN) 4-(methylnitrosamino)-1-(-3 pyridyl)-1 butanone (NNK) and other volatile nitrosamines found in tobacco smoke (Yamazaki, 1992). CYP2E1 is induced in mice exposed to cigarette smoke by inhalation (Villard, 1998). Two linked CYP2E1 polymorphisms in the 5'-flanking region have been detected with RsaI G-1259C and PstI C–1091T, which have been shown to affect CYP2E1 transcription level are located in the 5' flanking transcription region of this gene and appear to be in complete linkage disequilibrium with each other (c1, common allele; c2, rare allele) (Uematsu, 1991). Another DraI polymorphism is present in intron 6 of CYP2E1 gene, A T-7668A substitution in intron 6 of the CYP2E1 gene is revealed by a DraI RFLP (C, minor allele; D, common allele). Its regulation involves complex transcriptional and posttranscriptional mechanisms. The CYP2E1 genetic polymorphism varies significantly among different ethnic groups CYP2E1 RsaI/PstI 0.1% of Caucasians,

4.6% of Asians and homozygous CY2E1 DraI 0.8% of Caucasians, 9.4% of Asians(Garte, 2001). Although in Caucasians no relationship was found between *in vivo* activity of this enzyme and genotype, in Japanese the presence of the variant c2 alleles resulted in a significant reduction in the oral clearance of chlorzoxazone (clorzoxazone 6-hydroxylation is mediated by CYP2E1), after adjustment for age and sex. The mean activity in individuals with the c2/c2 genotype was significantly lower than that in individuals with either the homozygous wild-type or the heterozygote genotype (Marchand, 1999). The wild-type DraI genotype was associated with an increased risk for lung cancer studied in Japanese (Uematsu, 1991, Uematsu, 1992, Uematsu, 1994). Mexican-Americans, and mixed populations (Le Marchand, 1998, Wu, 1998). However conflicting results have been published concerning the RsaI/PstI mutation. The rare PstI/RsaI c2 allele has been associated with decreased risk for lung cancer (Le Marchand, 1998), and in one study the c2 allele frequency was significantly lower among cases than controls (Persson, 1993).

2.2.2. Glutathione-s-transferases (GSTs)

(1) Glutathione-S-transferase M1(GSTM1)

The GSTM subfamily is encoded by a 100-kb gene cluster at 1p13.3 arranged as 5'-GSTM4-GSTM2-GSTM1-GSTM5-GSTM3-3'(Fig.1.5) (Pearson, 1993)

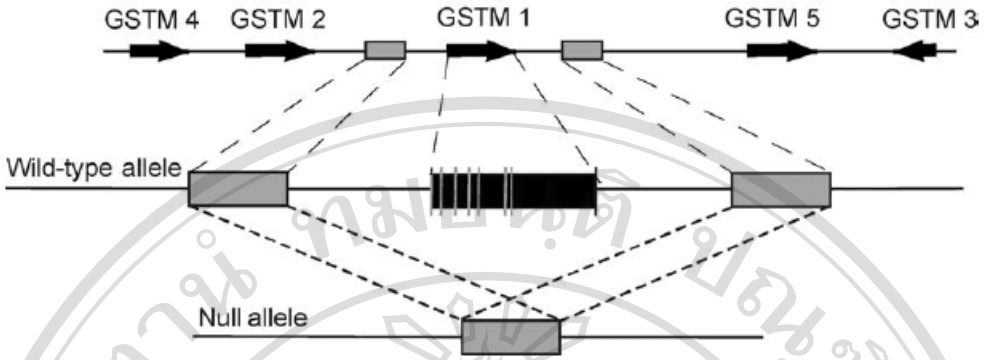

Figure 1.5 The GSTM1 gene is part of the Mu-class GST gene cluster at 1p13.3, which is arranged as 5'-GSTM4-GSTM2-GSTM1-GSTM5-GSTM3-3' (top of diagram). The GSTM1 gene (black box) consists of 8 exons, which range in size from 36 to 112 bp, while the introns vary from 87 to 2641 bp. The GSTM1 null allele arises by homologous recombination of the left and right 4.2-kb repeats, which results in a 16-kb deletion containing the entire GSTM1 gene [bottom of diagram (Figure from Parl, 2005]

GSTM1 is involved in degradation of active metabolites of PAHs (Ketterer, 1992). The GSTM1 locus is polymorphic owing to a gene deletion, which result in the virtual absent of enzyme activity in individuals with both alleles deletion (Seidegard, 1988). A meta-analysis of 30 studies involving over 10,000 individuals identified the GSTM1 null genotype in 53% of Caucasians (with a 42–60% range for individual studies). The frequency was similar in Asians, with a range of 42-54% but lower in African–Americans, 27% (16–36%) (Garte, 2001). Studies of the interactions between GSTs and xenobiotics also support the fact that exposures prior to cancer development (*e.g*, cigarette smoking or germ-line variants at other loci such as CYPIAI) are associated with cancer susceptibility in the presence of a GSTMI null genotype (Kihara, 1995, Nakachi, K. , 1993). The cause of K-ras gene mutation in smoker with lung adenocarcinoma may be in part an accumulation of BaP diole poxide which is not well detoxified in individual with GSTM1 null genotype

(Matsuzoe, 2001). Homozygous deletions of GST genes, alone or in combinations, have been associated with increased risk of lung cancer among Caucasians (Benhamou, 2002). GSTM1 null genotype is associated with increase in risk for lung cancer, [odds ratio (OR) = 1.27, 95% CI 0.91-1.77)] and increased two to six folds among heavy smokers (Nazar-Stewart, 2003).

(2) Glutathione-S-transferase T1 (GSTT1)

GSTT1 is a member of a superfamily of proteins that catalyze the conjugation of reduced glutathione to a variety of electrophillic and hydrophobic compounds. The GSTT subfamily consists of two genes, GSTT1 and GSTT2, which are located at 22q11.2 and separated by about 50 kb. The GSTT1 and GSTT2 genes have a similar structure, being composed of five exons with identical exon/intron boundaries (Figure 6) (Coggan, 1998, Landi, 2000, Whittington, 1999).

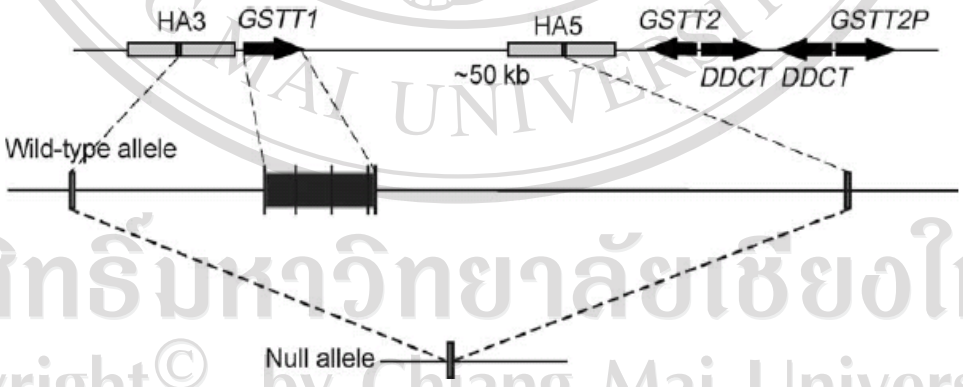

Figure 1.6 The GSTT1 gene is part of the Theta-class GST gene cluster at 22q11.2 (top of diagram). GSTT1 and GSTT2 are separated by approximately 50 kb. The GSTT1 gene (black box) consists of five exons, which range in size from 88 to 195 bp, while the introns vary from 205 to 2363 bp. The GSTT1 null allele arises by homologous recombination of the left and right 403-bp repeats, which results in a 54-kb deletion containing the entire GSTT1 gene (bottom of diagram) (Figure from Parl, 2005)

Both genes have five exons with identical intron/exon boundaries but share only 55% amino acid identity. The deletion of the GSTT1 gene does not include

GSTT2 (Coggan, 1998). Homozygous for a GSTT1 null allele are only about 20% of Caucasians, but more common in Asians, with frequencies ranging from 35 to 52% (Garte, 2001). Similar to GSTM1 null, the GSTT1 deletion is most likely caused by a homologous recombination event involving the left and right 403-bp repeats. The recombination results in a 54-kb deletion containing the entire GSTT1 gene . Smokers lacking a functional GSTT1 enzyme reportedly can not conjugate monohalomethanes found in tobacco smoke (Seidegard, 1988, Warwick, 1994) and may have greater susceptibility to chromosomal damage *via* sister chromatid exchanges (Kelsey, 1995, Schroder, 1995). Few studies of GSTT1 and lung cancer have been reported (el-Zein, R, 1997).

2.2.3. Myeloperoxidase (MPO)

Neutrophil recruitment into lung tissue occurs after exposure to variety of insults known to increase lung cancer risk, including tobacco smoke particles, infection, asbestos and ozone. Following immunological and/or chemical insults, neutrophils release MPO and undergo a respiratory burst, which is characterized by a massive increase in oxygen consumption and a consequent production of superoxide and other free radicals (Arnhold, 2004, Babior, 2000). MPO is present in the primary granules of neutrophils and catalyzes the production of the potent bacteriotoxic oxidizing agent hypochlorous acid (a one- and two-electron oxidant that can attack endogenous molecules including DNA) from hydroxyl radicals and chloride ions. The MPO gene is located on chromosome 17. G to A transition at position -463 in the promoter region of the MPO gene, which leads to the loss of a SP1 transcription binding site in an Alu hormone-responsive, has been shown to reduce MPO mRNA expression (Piedrafita, 1996). Myeloperoxidase (MPO) -463G reduces MPO activity

and DNA adduct levels in bronchoalveolar larvages of smokers. It is possible that possession of two copies of the A allele of the MPO gene reduces the risk of lung (Van Schooten, 2004).

The summary frequency of the A allele has been found to be 23.4% (95% CI =21.8 –25.0%) in Caucasians and 14.4% (95% CI = 11.3–17.6%) in Asians. The A allele was more frequently (1.6 times) observed in Caucasians than in Asians (Kiyohara, 2005). Possession of the A/A genotyp was associated with a decreased risk of lung cancer in Caucasians (OR=0.30, 95% CI=0.10-0.93) and 39% reduction (not statistically significant) in African-Americans compared with those with two G alleles (London, 1997). Populations with Caucasians, Japanese or Hawaiian ethnicity reported an overall 50% reduction in risk (95% CI=0.2-1.3) for those with the A/A genotype compared with those with the G/G genotype (Le Marchand, 2000). In a case-control study in the Berlin, possession of one or two A alleles was suggested as being a protective factor for cancer of the lung (OR=0.58, 95%CI=0.38-0.88) and larynx (OR=0.63, 95% CI=0.43-0.92) but not for cancer of the pharynx (OR=0.82, 95% CI=0.57-1.17) (Cascorbi, 2000). A allele carriers with heavy smoking and a history of asbestos exposure were lower risk for lung cancer (OR = 1.18; 95% CI 0.58-2.38) than that G/G genotype with the same exposure profile (OR = 2.19; 95% CI 1.16-4.11) (Schabath, 2002). In a case-control study nested within a Finnish clinical no evidence of an overall association between lung cancer risk and MPO genotype was observed (Misra, 2001). No evidence for reduce risk of genetic polymorphism in MPO on lung adenocarcinoma was found in Chinese population (Lu, 2002).

2.3. The relationship between polymorphisms of DNA repairing enzyme [(human Oxoguanine Glycosylase-1(hOGG-1)] and susceptibility to lung cancer

The OGG1 gene is located on chromosome 3p26.2. It consists of seven exons and six introns and encodes a 345 amino acid. This gene encodes a protein with DNA glycosylase and AP lyase activities that remove 8-OHG, an oxidatively damaged promutagenic base, from double-stranded DNA (Fig.7) (Boiteux, 2000, Shinmura, 2001)

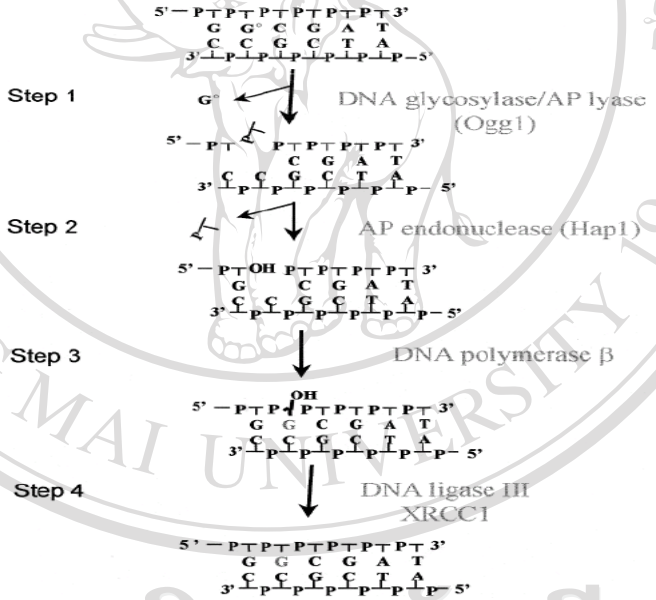

Figure 1.7. Diagram showing the role of hOGG1 in base-excision repairing system to remove 8-OHG in mammalian cells. (Figure from Boiteux S, 2000)

hOGG1 polymorphism at codon 326 associated with the amino acid change from Serine to Cysteine (Ser326Cys) (Kohno, 1998). The lower ability of hOGG1-326Cys protein than hOGG1-326Ser protein to prevent mutagenesis by 8-OHG has been shown by both *in vitro* (Dherin, 1999, Kohno, 1998) (Dherin C, 1999, Kohno,

1998) and *in vivo* (Yamane, 2004). It was also demonstrated that homozygotes for the hOGG1- 326Cys allele have a lower capacity to repair 8-OHG than others (Lee, 2005). Moreover other study found that hOGG activity was lower in peripheral blood mononuclear cells from lung cancer patients than control subjects and the estimate relative risk of lung cancer for smoker with low hOGG activity was higher for smoker with lower activity(Paz-Elizur, 2003). Frequencies of the hOGG1-326Cys allele are different among populations, and the frequency is higher in Asians (>40%) than in Caucasians (<20%). Several case-control studies have been undertaken on several populations, including Caucasians and Asians, and the hOGG1-326Cys allele encoding the less active OGG1 protein was reported to be associated with the risk of various cancers such as esophagus, prostate, orolaryng and nasopharyng cancers (Cho, 2003, Weiss, 2005). Case-control studies have also been undertaken on the association of this SNP with lung cancer risk. A recent meta-analysis showed that the overall OR of the homozygotes for the hOGG1- 326Cys allele against those for the OGG1-326Ser allele was 1.24 (95% CI = 1.01–1.53) suggesting that the SNP plays a role in susceptibility to overall lung cancer (Hung, 2005).

2.4 The relationship between polymorphisms of p53 tumor suppressor gene and susceptibility to lung cancer

The p53 gene is located on chromosome 17p13 and encodes a 53 kDa protein which plays a critical role in cell growth control. A few polymorphic sites have been identified in this gene, including the well-known Arg72Pro polymorphism at codon 72 in exon 4 [variant alleles encode either arginine (Arg-CGC) or proline (Pro-CCC)]. Two polymorphic variants differ in biochemical and biological properties in vitro.

The Arg72 encoded p53 protein is more efficient in inducing apoptosis and suppressing transformation than Pro72 encoded p53 protein (Thomas, 1999). Ethnic differences in the frequency of Arg72 and Pro72 alleles warrant analysis of the potential clinical relevance of this polymorphism in various populations. Several studies have demonstrated that Pro72 allele carriers have an increased risk of lung cancer (Fan, 2000, Jin, 1995) and other studies poorer prognosis in lung cancer (Tagawa, 1998, Wang, YC, 1999). However, the Pro/Pro genotype was not associated with elevated risk in older patients, nor with heavier smokers in African-Americans population (Jin, 1995). p53 genotyping studies in non-small cell lung cancer (NSCLC) have been performed in Asia (Murata, 1996, Tagawa, 1998, Wang, 1998) Sweden (Beckman, 1994, Birgander, 1995) and Chilean (Irarrazabal, 2003). The relation between Arg72Pro polymorphism and the risk of lung cancer has been previously investigated in the Chilean study (Irarrazabal, 2003). They found relation between the presence of the Pro allele and lung cancer risk in male smokers OR= 2.47 (95% CI: 1.34-4.54) for one single nucleotide polymorphic allele (Pro) and OR=3.88 (95% CI: 1.16-13.39) for Pro/Pro genotype. In another studied the p53 Arg72Pro polymorphism was associated with an increased risk of non-small cell lung cancer (NSCLC) (Szymanowska, 2006). The proportion of Pro homo/heterozygotes versus Arg homozygotes was significantly higher in NSCLC patients (54%) than in controls (46%, p = 0.034), OR was 1.28 (95% CI: 0.91-1.80). Moreover somatic p53 mutations in tumor cells were more frequently in Pro carriers (31%) than that in Arg homozygotes (20%, p = 0.06)

2.5. The relationship between polymorphisms of Matrix Metalloproteinase-1 (MMP-1) and susceptibility to lung cancer

One of the proteins that play an essential role in the dynamics of maintaining the cellular microenvironment is the MMPs family. Among the MMPs, MMP-1 is the most highly expressed interstitial collagenase degrading fibrillar collagens, the most abundant protein in the human body. Overexpression of MMP-1 has been demonstrated in tumor tissues and has been suggested to be associated with tumor invasion and metastasis (Chambers, 1997, Hewitt, 1991, Murray, 1996). Moreover, overexpression of MMP-1 has been found to be associated with an overall poor prognosis in colorectal and esophageal cancers (Murray, 1998, Murray, 1996). The expression of MMP-1 is partly regulated by the upstream promoter sequences of this gene (Buttice, 1996). The 2G allele of the MMP-1 -1607 1G/2G polymorphism located in a core recognition sequence of the transcription factor-binding site. Promoters containing the 2G allele display significantly higher transcriptional activity than 1G promoters (Rutter, 1998). Promoters containing the 2G allele was presented at a higher frequency in ovarian cancer patients than in healthy controls (Kanamori, 1999). This MMP-1 polymorphism may provide a mechanism for more aggressive matrix degradation, thereby facilitating tumor development. The association between the MMP1 2G/2G genotype and increased susceptibility to lung cancer has been reported in Caucasians (OR= 1.76; 95% CI= 1.29 –2.39), specifically for males, current smokers, (OR= 3 .16; 95% CI= 1.87–5.35) and heavy smokers, (OR= 2.55; 95% CI=1.61– 4.03) (Zhu, 2001). However, in others Caucasians studied there were

no associations between the MMP-1 genotypes and risk of lung cancer, with OR=1.15, CI= 0.94-1.40 but the 2G allele of the MMP-1 polymorphism is associated with higher risk of lung cancer in never-smokers and in males (Su, 2005). The study of association between polymorphisms of MMP-1 promoter and non-small cell lung carcinoma in North China showed that the variant alleles in cancer patients was not significantly different from the control group ($p>0.05$) (Fang, 2005).

2.6 Combination of susceptibility genotypes and the risk of lung cancer

The combined CYP1A1 variants (either MspI or Ilu-Val) and GSTM1 null genotype have been associated with a significantly increased risk for lung cancer CYP1A1 (Ilu462Val) +GSTM1 null OR=5.83 (95% CI=2.28–13.3) (Hayashi, 1992) GSTM1 null + CYP1A1 (MspI) OR=16 (95% CI=3.76–68.02) (Nakachi, K., 1993) GSTM1 null +CYP1A1 (Ilu462Val) OR=41 (95% CI=8.68–193.61) (Nakachi, K., 1993). For CYP1A1 (Ilu462Val) +GSTM1 null genotype studied in 1995 found OR=21.9, 95% CI=4.68–112.7) (Kihara, 1995). Single genotype variants were not significantly associated with lung cancer risk. However, inheritance of the combined GSTM1 and GSTT1 null genotypes showed a significant increase in lung cancer risk. OR= 2.32 (95% CI,1.01-6.04) (Cajas-Salazar, 2003). In Asian population, the combination of GSTT1 null and CYP1A1MspI genotypes present an increased risk of Sqamous Cell Carcinoma (SqCC) when compared to the risk associated with combination of respective wild-type counterparts OR=3.41 (95% CI=1.77-6.0) (Liang, 2004, Sobti RC, 2004). The risks of lung cancer associated with the GSTT1 null, GSTM1 null and CYP1A1 variant alleles were always smaller than the risk associated with the combination thereof. The CY2E1 PstI c2/c2 genotype were

combined with GSTM1 null was increase lung cancer risk OR= 2.88 (95%CI, 0.25-325) (El-Zein, RA, 1997). Few studies are only available indicating that the MPO polymorphism interacts with polymorphisms of other genes in affecting the lung cancer susceptibility. GSTT1 and GSTM1 null genotypes, when occurring in combination, are believed to increase the lung cancer risk almost 2.5-times (Cajas-Salazar, 2003). However, the MPO G/A genotype interfere with the GSTT1 and GSTM1 genotypes resulting in drastic lung cancer risk decrease(OR = 0.17) (Cajas-Salazar, 2003). Synergistic effect in increasing lung cancer risk was also reported for a combination of MPO G/G and CYP1A1 Ile462Val genotypes, resulting in almost 3-times higher risk of NSCLC compared to combination of the respective wild-type(Larsen, 2006).

2.7 Ethnic-dependent distribution of polymorphic genes

The frequencies of polymorphic genes in control populations have been reported to be different in various ethnic groups. In addition, intra-ethnic differences have been established (Garte, 2001). Significantly Ethnic differences in the frequency of homozygous variant alleles have been observed, CYP1A1 MspI alleles 1.2% of Caucasians, 14% 0f Asians, 5.9% Africans and homozygous CYP1A1 Val alleles 0.6% of Caucasians, 4.9% 0f Asians (Garte, 2001). Both the MspI and Val alleles are rarer in Caucasian than in Asian populations. The CYP2E1 RsaI/PstI 0.1% of Caucasians, 4.6% 0f Asians and homozygous CY2E1 DraI 0.8% of Caucasians, 9.4% 0f Asians (Garte, 2001). On the other hand, GSTM1 and GSTT1 deletion frequencies range from 42% to 60% and 13% to 26%, respectively, in Caucasians. GSTM1 null genotype was similar in Asians, 42-54% but lower in African–Americans, 27% (16–36%) (Garte, 2001). The summary frequency of MPO variants, A allele has been

found to be 23.4% (95% CI =21.8 –25.0%) in Caucasians and 14.4% (95% CI = 11.3–17.6%) in Asians, while the frequencies of the hOGG1-326Cys allele are higher in Asians (>40%) than in Caucasians (<20%) (Sugimura, 1999). A polymorphism at codon 72 of the p53 gene results in an arginine to proline substitution is common in African Americans and Caucasians. The Arginine allele is more frequent in the north of Europe and in the US, and the prevalence of the Proline allele increases near the Equator (Weston, 1997).S

Therefore, in this study the polymorphic frequency of the well-investigated lung cancer susceptibility genes were examined in Northern Thai population. Since there is a great numbers of genes reported to influence lung cancer risk, selection was made only for polymorphisms with a previous strong support, i.e. confirm by different groups of researchers. Ten polymorphisms were chosen including CYP1A1(MspI), CYP1A1(Ilu462Val), CYP2E1(PstI), CYP2E1(DraI), MPO(AciI) GSTM1, GSTT1, hOGG1(Ser326Cys), p53(Arg72Pro) and MMP-1(AluI). The results obtained from the investigation may enable us to understand the interplay of environmental and genetic polymorphisms at multiple loci and thus help identify individuals who are at increased risk for lung cancer.

III. OBJECTIVES

1. To investigate the polymorphic frequency of CYP1A1(MspI), CYP1A1(Ilu462Val), CYP2E1(PstI), CYP2E1(DraI), MPO(AciI) GSTM1, GSTT1, hOGG1(Ser326Cys), p53(Arg72Pro) and MMP-1(AluI) in northern Thai population.
2. To determine the association between genetic polymorphisms and the risk of lung cancer development in relation to smoking status and gender.