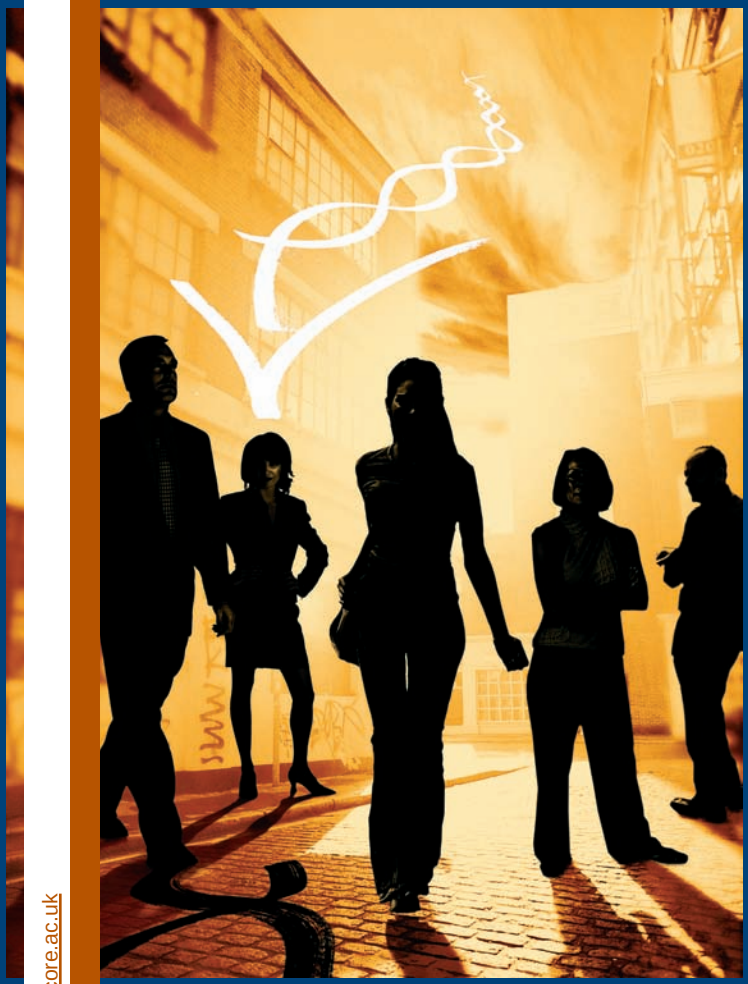
nu Voho

enetic Variation Effect on the Risk of Cancers f Lung and Oropharynx



Finnish Institute of Occupational Health

People and Work Research Reports 71

People and Work

Address Finnish Institute of Occupational Health
Topeliuksenkatu 41 a A
FIN-00250 Helsinki
Tel. +358-9-47 471
Fax +358-9-477 5071
www.ttl.fi

Cover design	Mari Pakarinen
Cover picture	Jaakko Haaparanta
ISBN in print	951-802-637-8
ISBN PDF	951-802-636-6
ISSN	1237-6183
Press	Tampereen yliopistopaino, 2005

Finnish Institute of Occupational Health
University of Helsinki

Anu Voho

Genetic Variation - Effect on the Risk of Cancers of Lung and Oropharynx

ACADEMIC DISSERTATION

To be presented, with the permission of Faculty of Biosciences, University of Helsinki, for public discussion in the Auditorium 2041 of Biocenter 2, Viikinkaari 5, Helsinki, on November 25th, 2005, at 12 o´clock.

Genetic Variation – Effect on the Risk of Cancers of Lung and Oropharynx

Anu Voho

People and Work
Research Reports 71

Department of Biological and Environmental Sciences
Faculty of Biosciences
University of Helsinki, Finland

Department of Industrial Hygiene and Toxicology
Finnish Institute of Occupational Health
Helsinki, Finland

Helsinki 2005

TABLE OF CONTENTS

ABSTRACT

Humans are known to differ from each other in their responses to environmental exposures. During recent years these differences have been attributed to interindividual variations in the genes encoding for enzymes involved in the metabolism of exogenous agents. The individual variations in metabolic capacity have been suggested to be an important modifier of individual susceptibility to environmentally indused diseases such as cancer.

Tobacco smoking is known to cause many types of malignancies, of which the highest relative risks are seen for cancers of lung and the upper aerodigestive tract. Most of the carcinogenic chemicals in tobacco smoke need to be metabolized before they can bind to cellular macromolecules and exert their harmful effects. Cytochrome P450 (CYP) enzymes oxidise these chemicals to highly reactive intermediates that can be further metabolised by glutathione-*S*-transferases (GSTs) and microsomal epoxide hydrolases (EPHX1). The activity of these enzymes modulates the amount of chemical binding to DNA, and polymorphisms in their genes have been associated with a modified risk of tobacco-related cancers in smokers.

In this study the interplay between tobacco smoke and polymorphism in *GSTM1*, *GSTM3*, *GSTP1*, *GSTT1* and *EPHX1* detoxification enzymes in relation to lung and oropharyngeal cancer risk was studied. The case-control study populations consisted of a Finnish lung cancer study and a French oral and pharyngeal cancer study. In addition, in a novel study desing, a large control population distributed in five different age strata was used to study the differences in genotype frequencies across age groups.

The *GST* genotypes were not associated with increased lung cancer risk when studied alone. However, *GSTM1* genotype frequencies were different across different age-groups and *GSTP1* genotypes were affected by smoking status. *GSTP1* and *GSTT1* genotypes were associated with marginally increased oral and pharyngeal cancer risk.

Increased lung cancer risk was seen for the carriers of *EPHX1* wildtype diplotypes. The effect was stronger in combination with *GSTM1* null genotype and showed a response to cumulative tobacco consumption. The highest risks were seen in heavy smokers. In the cancer-free control group, an overpresentation of protective *GST* genotypes was seen in higher age-groups in individuals having high cumulative smoking dose compared to lighter smokers.

The results indicate that studied gene polymorphisms enzymes do modulate the cancer susceptibility of smokers. However, future studies are needed before any firm conclusions can be drawn. The future seems very interesting but challenging; we are just starting to understand the complex interplay between environmental and genetic factors in the individual susceptibility to environmentally induced diseases.

LIST OF ORIGINAL PUBLICATIONS

This dissertation is based on the following original publications, referred to in the text by the Roman numerals I–V as indicated below.

I Saarikoski, S.T., **Voho, A.,** Reinikainen, M., Anttila, S., Karjalainen, A., Malaveille, C., Vainio, H., Husgafel-Pursiainen, K., Hirvonen, A. Combined effect of polymorphic *GST* genes on individual susceptibility to lung cancer. *Int. J. Cancer,* 77, 516–521 (1998).

II Jourenkova-Mironova, N., **Voho, A.,** Bouchardy, C., Wikman, H., Dayer, P., Benhamou, S., Hirvonen, A. Glutathione *S*-transferase *GSTM1*, *GSTM3*, *GSTP1* and *GSTT1* genotypes and the risk of smoking-related oral and pharyngeal cancers. *Int. J. Cancer,* 81, 44–48 (1999).

III **Voho, A.,** Impivaara, O., Järvisalo, J., Metsola, K., Vainio, H., Hirvonen, A. Distribution of Glutathione *S*-transferase *M1*, *P1* and *T1* genotypes in different age-groups of Finns without diagnosed cancer. *Cancer Prev and Det,* in press.

IV **Voho, A.,** Metsola, K., Anttila, S., Impivaara, O., Järvisalo, J., Vainio, H., Husgafvel-Pursiainen, K., Hirvonen, A. *EPHX1* gene polymorphisms and individual susceptibility to lung cancer. *Cancer Letters,* in press.

V **Voho, A.,** Saarikoski, S.T., Metsola, K., Impivaara, O., Anttila, S., Järvisalo, J., Vainio, H., Husgafvel-Pursiainen, K., Hirvonen, A. Combined effect of *EPHX1* and *GST* genes on individual susceptibility to lung cancer. *Lung Cancer,* Manuscript.

ABBREVIATIONS

AC	adenocarcinoma
ARE	antioxidant response element
B[a]P	benzo[a]pyrene
BPDE	7,8-diol-9,10-epoxide
BRI	biologically reactive intermediate
CI	confidence interval
CDNB	1-chloro-2,4-dinitrobenzene
CNP	copy number polymorphism
cSNP	coding single nucleotide polymorphism
CYP	cytochrome P450
DCNB	1,2-dichloro-4-nitrobenzene
DMBA	7,12-dimethylbenz[a]anthracene
EPHX	epoxide hydrolase
GSH	glutathione
GST	glutathione *S*-transferase
HNC	head and neck cancer
MAPK	mitogen-activated protein kinase
NAT	*N*-acetyltransferase
NNK	4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone
NNN	*N′*-nitrosonornicotine
OR	odds ratio
PAH	polyaromatic hydrocarbon
RFLP	restriction fragment length polymorphism
ROS	reactive oxygen species
rSNP	regulatory single nucleotide polymorphism
SCC	squamous cell cancer
SCE	sister chromosome exchange
SCLC	small cell lung cancer
SD	standard deviation
SNP	single nucleotide polymorphism
TSNA	tobacco specific nitrosamine
TSO	stilbene oxide
UAT	upper aerodigestive tract
UDP	uridine-5′-diphosphate
XME	xenobiotic metabolizing enzyme
XRE	xenobiotic response element

INTRODUCTION

We are now living in the post-genomic era, which offers huge possibilities for disease prevention, identification and treatment. Though the first phase of the human genome project is completed, further projects are on-going including the validation and clarification of genetic variation in the human genome.

There is variation in every 300 bp in almost all of our genes. Most human sequence variation is attributable to single nucleotide polymorphisms (SNPs), with the rest being attributable to insertions or deletions of one or more bases, repeat length polymorphisms and rearrangements (http://www.ncbi.nlm.gov/SNP). Inherited differences in DNA sequence contribute to phenotypic variation, influencing an individual's risk of disease and response to environmental factors.

In most cases, the variations in our genome do not by themselves evoke the disease. Instead, the interplay between the genes and the environment leads to variation in an individual disease susceptibility. Chemical or radiation exposure, hormones and inflammation can lead to cancer development. For instance, although about 90% of lung cancers are attributable to smoking, only 10–15% of smokers will develop lung cancer. This implies that the exposure alone is not adequate to lead to cancer formation.

In this study we have assessed the effect of genetic variation on the risk of tobacco related cancers in lung and oropharynx. Tobacco smoke is a complex mixture containing some 4000 constituents, including about 70 carcinogens. These carcinogens have been shown to play an important role in tobacco-associated cancers. The majority of the carcinogenic chemicals in tobacco smoke need to be metabolically activated before they are able to bind to cellular macromolecules. Xenobiotic metabolism is the main metabolic pathway both in activating and inactivating these chemicals.

A wide interindividual variation has been observed in the effectiveness of metabolism, in many cases this being attributable to inherited genetic variation. It has been hypothesized that it may cause differences in the individual risk of cancer. The activity of xenobiotic metabolizing enzymes (XME) may modulate the amount of chemical binding to DNA and to the other cellular macromolecules leading to disruptions in DNA, signaling networks, and the control of cell cycle. These events may gradually lead to cancer development.

One of the goals of genetic research is to identify those genes that contribute to disease. Genome-based knowledge may be translated into health benefits, e.g., by applying the results to the prevention, diagnosis and treatment of disease [1]. However, much still needs to be done before this goal can be reached. To date, the results on the disease susceptibility cannot be directly linked to disease risk of individuals. The complexity and multifactorial biological pathways need to be resolved before recommendations can be issued. There are also many ethical questions which have to be considered.

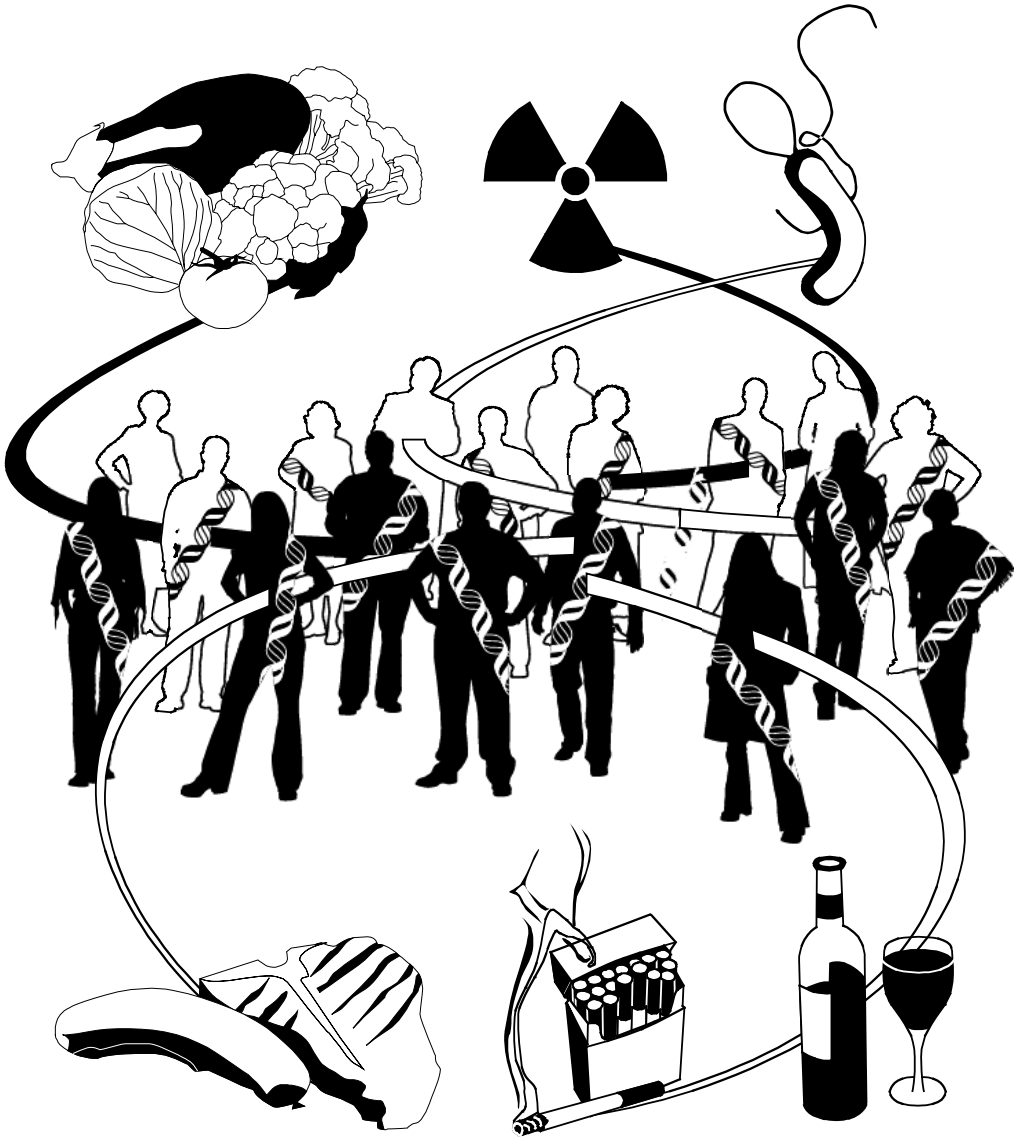


Figure 1. Human cancer risk is mediated by different factors. The effects of alcohol, food, radiation, smoking and infections on individual cancer proneness may be mediated by individual genetic factors.

REVIEW OF THE LITERATURE

Tobacco-induced cancers

Tobacco smoking is an enormous environmental exposure. It has been estimated that there are 1200 million smokers in the world [2]. Globally it is estimated that 47% of men and 12% of women smoke. In the developed countries, the corresponding figures are 42% for men and 24% for women. In Finland, the respective numbers are 27% for men and 20% for women and in France 39% for men and 27% for women [3].

In the developed countries, more than one-third of all deaths of middle-year men (35–69 years) are attributable to smoking. For women in the same age range, the respective percentage increased from a mere 2% in 1955 to 13% in 1995, and continues to increase rapidly [2]. In Finland, about 20% of all male deaths in middle age are caused by tobacco, whereas about 5% of female deaths are attributable to tobacco [2].

For tobacco users, cancer is one of the main causes of death. Tobacco-related cancer constitutes 16% of the total annual incidence of cancer cases – and 30% of cancer-related deaths – in the developed countries. In populations where cigarette smoking has been common for several decades, 90% of lung cancers and 15–20% of other cancers are attributable to tobacco [4–6]. In addition to lung cancer, cancers of the upper aerodigestive tract (UAT), lower urinary tract and pancreas are causally related to tobacco smoking [7]. Also the risk of cancers of the kidney, stomach, liver, colon, oesophagus, bone marrow and cervix are increased in smokers [3]. The relative risk of a smoker to develop cancer when compared to never smokers depends on the cancer site (Table 1) [8].

Table 1. Relative risk of different cancer sites in smokers versus non-smoker. Modified from [8].

Cancer site	Relative risk (men)	Relative risk (women)
Lung	22.4	11.9
Cancers of the bladder and other urinary organs	2.9	2.6
Pancreatic cancer	2.1	2.3
UAT cancers	24.5	5.6
Oesophageal cancers	7.6	10.3
Kidney cancer	3	1.4

In addition to its carcinogenic effects, tobacco smoke has a number of other pathogenic properties. Cigarette smoking is a major independent risk factor for coronary heart disease and is an important risk factor for other cardiovascular diseases. It is also the most important cause of cough, sputum production, chronic bronchitis, chronic obstructive pulmonary disease and asthma. Smoking increases the risk of peptic ulcer and has detrimental effects on the gastric mucosa. It is a risk for Crohn´s disease and type 2 diabetes mellitus [3].

Lung cancer

Lung cancer is the most common cancer accounting for 12.3% of all cancers and 17.8% of all cancer deaths [8]. It is estimated that there are about 1.2 million new cases and over one million deaths annually [3, 9]. Lung cancer is almost invariably associated with poor prognosis. Tobacco smoking has been proven to be the most important etiological factor of lung cancer for both sexes. The causal relationship between tobacco and lung cancer was established already in the 1950s [3, 10]. There is a strong male predominance

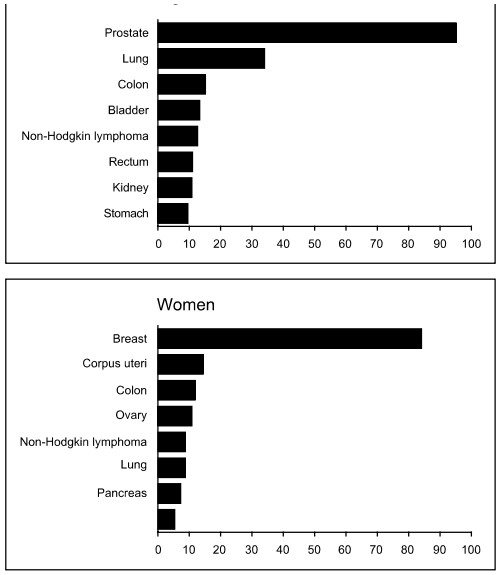


Figure 2.
Cancer incidence in men and women in Finland in 2003. Age-standardized rate. (ASR) Finland in (www.cancerregisty.fi).

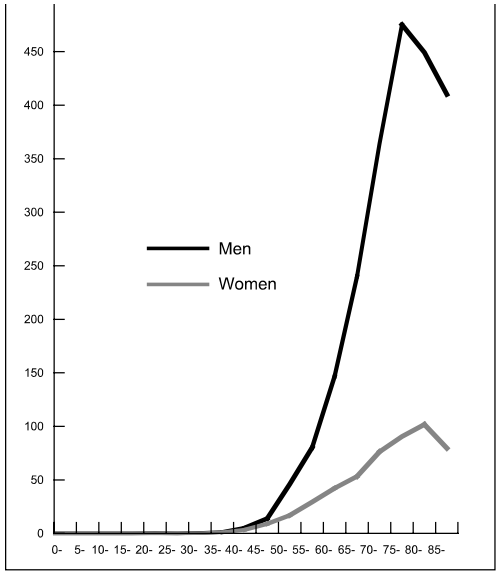


Figure 3.
Age-specific incidence rate of lung cancer per 100 000 among men and women in 2003 (www.cancerregistry.fi).

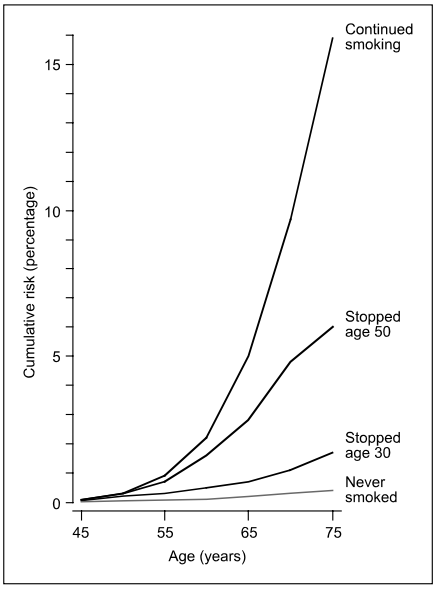


Figure 4. Cumulative risk (%) of death of lung cancer in men. Stopping smoking reduces the risk of death of lung cancer [3].

in lung cancer, due to the different smoking habits of men and women. Lung cancer is the second most common cancer among men and the sixth most common for women in Finland (Figure 2). The highest incidence of lung cancer occurs at around 70 years of age and incidence declines after 80 years (Figure 3) [11].

The lifetime probability of a cigarette smoker to develope lung cancer has increased over time. According to the latest calculations, the lifetime probability is 24.1% and 11.0% for men and women smokers, respectively [3]. The most important parameter related to how smoking affect the lung cancer risk is the duration of regular smoking. The risk also increases with the number of cigarettes smoked per day and the depth of inhalation. Smoking cessation reduces the risk of lung cancer within 1–4 years of quitting smoking and the magnitude of the reduction in relative risk increases with increasing time since cessation (Figure 4) [3, 12].

Occupational exposure to compounds like asbestos, arsenic, silica and synthetic mineral fibers has been shown to affect the risk of lung cancer. Other risk factors for lung cancer are radon and ionizing radiation [13]. Evidence from epidemiological studies indicate synergistic effects between asbestos and smoking and radon and smoking in causing lung cancer [3, 14, 15]. Environmental pollution and certain previous respiratory diseases also seem to raise the risk of developing lung cancer [3].

There appears to be no gender differences in the susceptibility to lung cancer; when the smoking habits among men and women are equivalent, their lung cancer rates are also the same [16]. Some ethnic differences between lung cancer susceptibility have been noted. However, a recent study reported similar risks for blacks and whites when they had similar smoking habits [17]. However, there is some evidence that nicotine metabolism may contribute to differences in the intensity of smoking, this being different between Caucasians and Asians. Thus Asians smoke fewer cigarettes per day and thus have a lower probability of suffering lung cancer [18].

Genetic susceptibility may also alter the risk of lung cancer. There is evidence about genetic differences that affect the activity of carcinogen metabolism and tendency to become dependent on nicotine [19]. There is also debate about the role of family history on the risk of lung cancer; first-degree blood relatives of people under 59 years of age with a history any type of malignant disease have been shown to have increased lung cancer risk [20, 21].

Lung cancers can be classified into two major entities based on their clinicopathological characteristics. Small cell lung cancer (SCLC) accounts for 15–25%, and non-small cell lung cancer (NSCLC) for 75–85% of lung cancers. The latter consists of three major histologic types of lung cancer, *i.e.*, squamous cell carcinoma (SCC), adenocarcinoma (AC) and large-cell carcinoma [9]. There is an association and exposure-response relationship between tobacco smoke and all histological types of lung cancer. However, notable shifts over time have been observed in the incidence rates of the various histological types of lung cancer. In the initial decades of the smoking related epidemic of lung cancer, SCC was the most common type of lung cancer among smokers. However, subsequently the incidence rate of AC has increased steadily both in the general population and among smokers [22–24] The reasons for this are unclear; there are no known risk factors other than smoking for AC of the lung that might explain the increase in incidence. One possible contributory factor may be related to advances in methods to detect tumours in the distal airways [3]. Another explanation could be that changes in the formulation of cigarettes have led to a shift in histological type og lung cancers [25].

Oral and pharyngeal cancers

The upper aerodigestive tract (UAT) cancers consist of oral and pharyngeal cancers together with larynx and oesophagus cancers. The incidence of oral and pharyngeal cancers varies widely throughout the world, France being one of the high risk areas [11].

Risk factors for head and neck cancers (HNC) include tobacco smoking and use of any form of smokeless tobacco, heavy alcohol drinking, chewing of betel nut, and human papillomavirus infection. Epidemiological studies also suggest a role for diet in this context [7, 26]. In Western countries, the main risk factors are tobacco and alcohol consumption, which have been shown to account for 75–90% of cases. The risk increases rapidly with

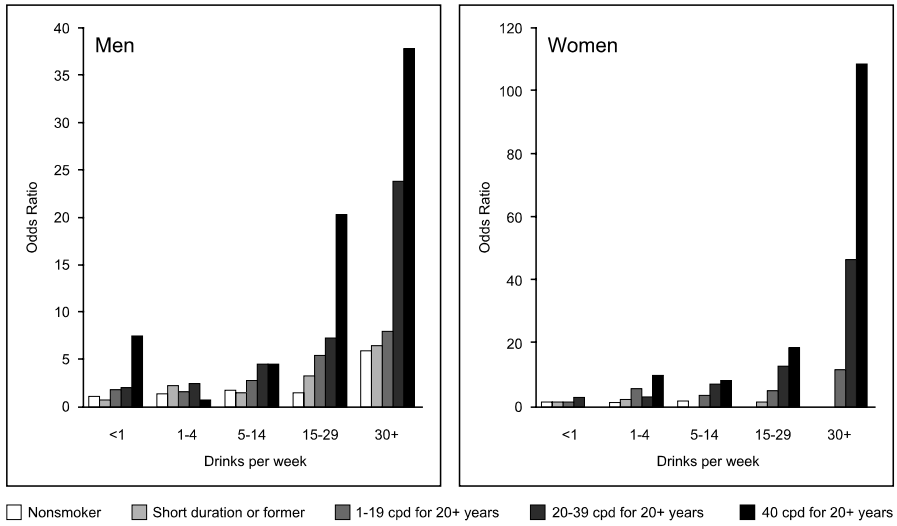


Figure 5. Synergism in the use of alcohol and smoking in the risk of oropharyngeal cancer incidence. Effect of alcohol consumption and cigarette smoking compared to those not drinking or smoking. Modified from Thun et al. [32].

increasing tobacco and alcohol consumption, with evidence of a synergistic effect of these two exposures (Figure 5) [3, 26–29]. However, the mechanism of the interaction is unclear. An increase in risk has been observed with both duration and quantity of tobacco consumption, and cessation of smoking has been shown to lower the risk [3, 28–31]. Tobacco and alcohol cannot explain the large differences between different Western countries and furthermore only a fraction of exposed individuals develop neoplastic lesions. Other risk factors may interact as well as differences in genetic susceptibility [30].

Components of tobacco smoke

Mainstream tobacco smoke is an aerosol containing some 4000 constituents including about 70 carcinogens. The chemicals in the mainstream smoke are distributed between the particulate and vapour phase. The particulate matter (tar) of mainstream smoke contains some 3500 chemicals of which nicotine is the most abundant [3]. The major classes of carcinogens in tobacco smoke are polyaromatic hydrocarbons (PAHs), tobacco-specific nitrosamines (TSNAs), and aromatic amines. Chemicals such as benzene and heavy metals,

Table 2. Carcinogens and tobacco-induced cancers. Modified from Hecht, 2003 [33].

Cancer type	Likely carcinogen involvement found in tobacco smoke
Lung	PAH, NNK, 1,3-butadiene, isoprene, ethylene oxide, ethyl carbamate, aldehydes, benzene, metals
Oral cavity	PAH, NNK, NNN
Laryngeal	PAH
Nasal	Nitrosamines, aldehydes
Bladder	Aromatic amines
Oesophagus	Nitrosamines
Liver	Nitrosamines, furan
Cervical	PAH, NNK

independently established as being carcinogenic to humans, are also present in tobacco smoke. Strong carcinogens such as PAH, nitrosamines and aromatic amines, occur in smaller amounts in tobacco smoke (1–200 ng per cigarette) than weak carcinogens such as acetaldehyde (nearly 1 mg per cigarette) [33]. The total amount of carcinogens in cigarette smoke ranges from one to three mg per cigarette (similar to the amount of nicotine, which is 0.5–1.5 mg per cigarette) [33]. The country of origin and type of the product play major roles in determining the chemical composition of tobacco [3].

Cancer causation by tobacco smoke is not attributable to any one chemical component, but an overall effect of the complex mixture of chemicals in smoke [8]. However, different tissue specificities of some chemicals are suggested (Table 2) [33].

Polyaromatic hydrocarbons

Tobacco smoke contains numerous PAHs, of which at least ten species are carcinogenic [3, 34]. In addition to being present in tobacco smoke, PAHs are found in broiled foods and polluted environments. PAHs are diverse group of strong, locally acting carcinogens [33]. PAHs are known to induce tumours of the UAT and lung when administered by inhalation [35]. Several studies have quantified PAHs in lung tissue and the levels have been reported to be higher in smokers than in non-smokers [36, 37]. PAH-DNA adducts have been detected in human lung samples, and the fact that mutations in the tumor suppressor gene (*TP53*) can lead to lung and larynx tumours is support for a role for presence of PAH

in the development of these cancers [33, 38]. Benzo[a]pyrene (B[a]P) is the best known member of the PAHs. It is slowly absorbed by the tracheal epithelium and extensively metabolized there, which partially explains its local carcinogenic effect on lung [39].

One of the metabolic pathways of B[a]P is shown in figure 7. It is ultimately converted to a 7,8-diol-9,10-epoxide (BPDE). This diol-epoxide is highly carcinogenic and can react with DNA to form an adducts. Also other PAHs are activated through a bay-region diol-epoxide mechanism, representing the main mechanism of carcinogenesis of PAHs [40]. Dibenz[*a,h*]anthracene and 5-methylchrysene have been shown to induce lung tumours in mouse [33, 41]. However, the contribution to tobacco components is not so straightforward since smoke is a complex mixture of carcinogens, co-carcinogens and other factors. For example, low molecular weight PAHs have been shown to inhibit the tumorigenicity of higher molecular weight PAHs in the mouse skin-painting model, presumably occupying the liver enzymes required for the metabolic activation of the more tumorigenic higher molecular weight PAHs [34].

The levels of PAHs in cigarette smoke have decreased in parellel to the increase in concentrations of nitroamines. This is becasue nitrate concentrations in tobacco increased over the period from 1959 to 1997 due to the use of tobacco blends that contain higher levels of air-cured tobacco, the use of reconstituted tobacco and other factors [33]. The changing histology of lung cancer, from SCC to AC may be linked to the reduction in PAH and increased concentrations of 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK); B[a]P is thought to induce SCC and NNK primarily AC of the lung [33].

N-Nitrosamines

N-Nitroamines are a large group of carcinogens that induce tumours in a wide variety of animal species and tissues [33]. They are present in small amounts in foods and can be formed endogenously. Tobacco products are, however, the most widespread and largest source of exposure to these carcinogens [3]. Tobacco smoke contains volatile N-nitrosamines such as *N*-nitrosodimethylamine and *N*-nitrosopyrrolidine as well as TSNAs such as *N´*-nitrosonornicotine (NNN) and NNK. NNK and NNN are chemically related to nicotine and other tobacco alkaloids and are therefore found only in tobacco products or related materials. They can thus be used as biomarkers of tobacco exposure [33].

NNN, NNK, and the derivative of NNK called 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanol (NNAL), are strong carcinogens in commonly used rodents models (i.e. rat, mouse and hamster). NNN causes tumours of the oesophagus and nasal cavity in rats, whereas the principal target of NNK in rodents is the lung. NNK is the only tobacco smoke carcinogen that induce lung tumours systemically in all three commonly used rodent models [3, 42]. DNA adducts derived from NNK and NNN are present at higher levels in lung tissue from lung cancer patients than in lung specimens from controls [43–45]. NNK as well as

B[a]P are considered to be major lung carcinogens [46]. On the basis of animal studies, PAH, NNK and NNN are the likely causes of oral cancer in smokers [33].

Other compounds

Aromatic amines were first identified as carcinogens in exposed workers in the dye industry. They are also abundant in broiled foods. Tobacco smoke contains four carcinogenic aromatic amines [3]. Of these, 4-aminobiphenyl and 2-naphthylamine are well established human bladder carcinogens. Aromatic amines are moderately carcinogenic in various tissues, including breast and colon. In laboratory animals, they are known to cause tumours at a wide variety of sites [3, 47, 48].

Tobacco smoke contains six species of carcinogenic heterocyclic hydrocarbons, eight *N*-heterocyclic amines, and aldehydes like formaldehyde and acetaldehyde which induce respiratory tract tumours in rodents when administered by inhalation. These are weaker carcinogens than PAHs and TSNAs, but they are present in thousands of times greater abundance in tobacco smoke [3]. Tobacco smoke contains phenolic compound, like catechol. There are many volatile hydrocarbons, e.g butadiene and benzene, present in considerable quantities in tobacco smoke [3]. Butadiene is a multi-organ carcinogen, which is particularly potent in mice, whereas benzene is known to cause leukemia in humans [3, 47]. Tobacco smoke contains ten carcinogenic miscellaneous organic compounds. Processed tobaccos contain numerous metals and more than ten metalloids such as nickel, chromium and cadmium that are human carcinogens [3, 47].

Cigarette smoke can also cause oxidative damage, probably because it contains free radicals such as nitric oxide (NO) and mixtures of hydroquinones, semiquinones and quinones [3, 46, 49–51]. NO is a highly active molecule that can be oxidized or complexed with other biomolecules depending on the microenvironment. Cigarette smokers may be exposed to an increased burden of free radicals also due the elevated levels of pulmonary macrophages and neutrophils; these are cells that can generate reactive oxygen species (ROS) which are frequently found in the lungs of cigarette smokers [52].

Tobacco smoke contains tumour promoters (phenolics), co-carcinogens (catechol and related compounds), toxic agents (acrolein and other aldehydes) and free radical species (nitric oxide and others). Most of the compounds listed above are thought to exert their carcinogenic effects through classical genotoxic mechanisms, e.g., the formation and persistence of DNA adducts with consequent miscoding. Non-genotoxic (epigenetic) mechanisms such as cytotoxicity, changes in gene expression via hypermethylation and genomic instability are other mechanisms of carcinogenesis that could operate after exposure to the compounds present in tobacco smoke [3]. The many carcinogens in tobacco smoke can exert effects on both early and late stages of cancer development [8].

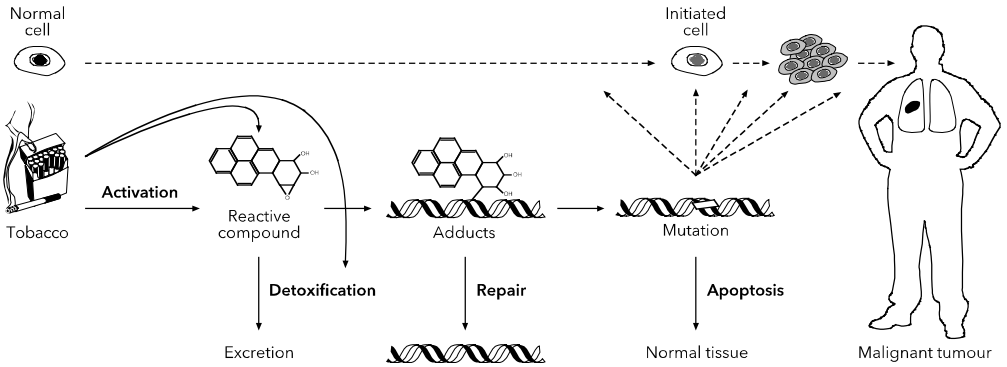


Figure 6. Simplified role of tobacco smoke in lung carcinogenesis.

Biological responses to tobacco smoke

The relationship between environmental exposure and carcinogenesis involves at least three systems; the metabolic pathways, the pathways that control the recognition and repair of induced DNA and protein damage and the control of cell cycle checkpoints. Tobacco smoke is a complex mixture of compounds and the pathways affected by tobacco smoke are numerous and interconnected [53]. Smoking can act as a cancer initiator and promoter and long-term chronic exposure to tobacco smoke is believed to be necessary for carcinogenesis (Figure 6).

Biotransformation

Metabolism of most environmental carcinogens involves multiple pathways [54]. Metabolic activation and detoxification play crucial roles in the metabolism of tobacco carcinogens. Xenobiotic metabolism occurs primarily in the liver, but the prerequisite enzymes are present in lung and other "target" organs. The initial step of metabolic activation of tobacco carcinogens is usually carried out by the cytochrome P450 (CYPs) family of enzymes which oxygenate the carcinogenic substrate [55]. This reaction may lead to the formation of biologically reactive intermediates (BRIs) that can react with DNA or other macromolecules to form covalent binding products known as adducts. BRIs formed by activating enzymes can undergo further transformations catalyzed by inactivating enzymes like glutathione-*S*-transferases (GSTs), uridine-5′-diphosphate (UDP)-glucuranosyltransferases, epoxide hydrolases (EPHX), N-acetyltransferases (NATs) and sulfotransferases which normally

detoxify BRIs to more water-soluble and readily excretable forms [56]. However, the reactions may also act in the other direction to yield BRIs or active radicals [54].

The concentrations of BRIs that can attack cellular macromolecules to initiate mutagenic or carcinogenic events will be determined by the final balance between the production of BRI by activation (phase I) pathways and their elimination by inactivation (phase II) pathways. The ultimate carcinogenic BRIs usually have a very short half-life. Furthermore, the tissue specificity of a carcinogen is thought to be determined by production of BRIs within the target tissue [54]. The balance between metabolic activation and detoxification varies between individuals. Interindividual differences in the activities of XMEs are therefore likely to affect an individual´s cancer risk [57, 58].

An example of carcinogen metabolism is the conversion of B[a]P to the highly carcinogenic BPDE (Fig 7). Also TSNAs, aromatic amines, and many other tobacco components are directed to biotransformation [54]. Numerous CYP enzymes participate in NNK activation and glucuronidation is thought to be the most important detoxification pathway in humans [42].

Tobacco smoke induces several XMEs in human tissues. For instance, expression of several CYPs is increased with active cigarette smoking (Figure 7) [59]. The induction of expression of the different classes of enzymes may depend on the duration of exposure [60–64].

Chemical carcinogenesis

DNA adducts

The induction of DNA damage, frequently due to the formation of chemically stable adducts, is an early and essential step in the sequence of events by which genotoxic carcinogens initiate the carcinogenic process. Higher prevalences or elevated concentrations of carcinogen-DNA adducts have been found in many human tissues as well as carcinogen-protein adducts in the blood of smokers [59, 70–76]. The DNA adducts represent an integrated marker of exposure to carcinogenic compounds. The amount of DNA adducts may reflect the ability of the individual to activate/detoxificate carcinogens, and to repair DNA damage.

Smokers exhibit higher levels of PAH- and aromatic-DNA adducts in blood and lung tissue than non-smokers [70, 77–81]. Highly carcinogenic BPDE form adducts with exocyclic N^2 of guanine (dG-N^2-BPDE). NNK, on the other hand, forms methyl and pyridyloxobutyl DNA adducts. The levels of DNA adducts, and specifically BPDE-adducts, are higher in oral tissue and buccal mucosa of smokers compared to non-smokers [82–85]. Also, PAH, 4-aminobiphenyl, malondialdehyde, 1,N2-propanodeoxyguanosine and 8-hydroxydeoxyguanosine adducts are detected in oral cells [26]. Higher levels of adducts formed by PAHs, TSNAs, aromatic amines, ethylene, and benzene have been

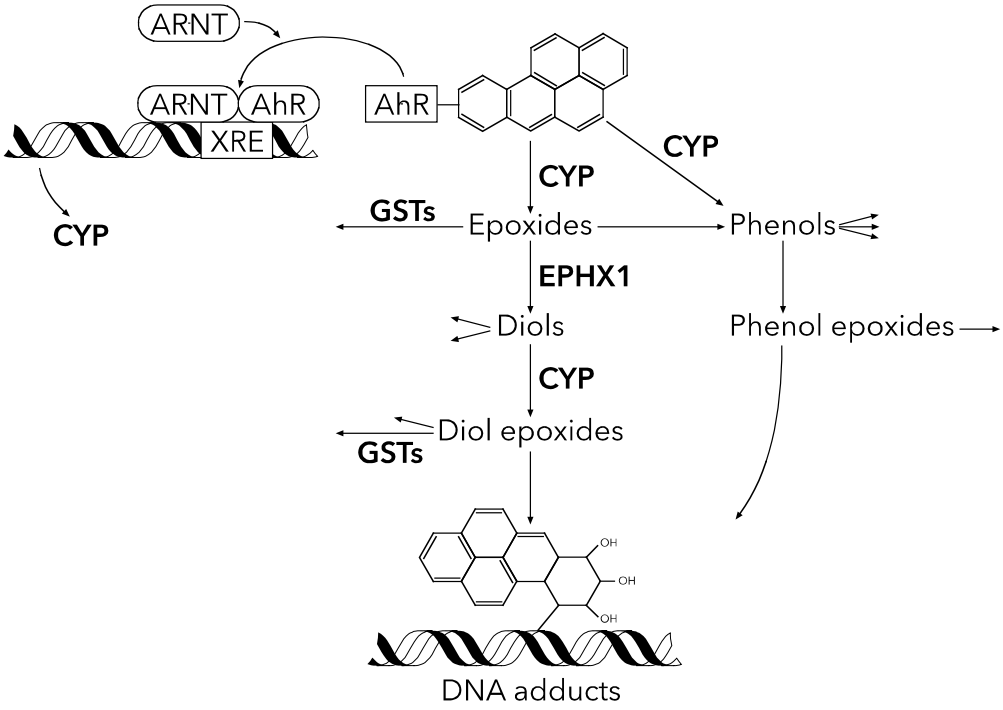


Figure 7. Outline of some cellular responses to B[a]P. B[a]P mediates XME induction through an AhR-dependent pathway. After diffusing into the cell, B[a]P binds with AhR and is translocated into the nuclei, where B[a]P-AhR heterodimers form complexes with Ah receptor nuclear translocator (Arnt) proteins [65, 66]. The complex then transactivates the CYP genes via an interaction with xenobiotic response element (XRE) in the promoter region of the genes. B[a]P is oxidized by CYPs [67, 68] and epoxide hydrolases (EPHX1s) to highly reactive diol epoxides, such as BPDE, as well as phenol epoxides which can form DNA adducts. GSTs participate in a detoxification pathway making the compounds more water-soluble and readily excretable. Modified from Hecht, 2002 and Luch, 2005 [25, 69].

found in the haemoglobin of smokers compered to non-smokers [3]. Also acetaldehyde, the primary metabolite of ethanol, yields a number of unstable and stable adducts with proteins and DNA [86, 87]. Higher levels have been observed in heavy alcohol user HNC patients compared to controls [30]. The amount of adducts has been negatively associated with the consumption of fruit and vegetables [88, 89].

Oxidative damage

Reactive oxygen and nitrogen species formed in cigarette smoke and the cigarette tar promoted neutrophil induction are the primary causes of the DNA strand breaks. ROS can induce DNA single-strand breaks (SSB) and formation of 8-OHdG adducts in cultured human lung cells and in rodents [90–96]. The mean level of 8-OHdG in lung tissues was significantly higher in smokers than in non-smokers [97]. Higher levels of oxidized proteins have also been found in smokers than in non-smokers [98]. Reactive nitrogen species have been shown to be involved in the formation of DNA lesions [99]. However, the role of oxidative damage in development of specific-tobacco induced cancer remains unclear.

DNA repair

Cells have DNA repair systems that can remove adducts and restore DNA to its normal structure. At least four pathways of DNA repair operate on specific types of damaged DNA [100]. For example BPDE-induced DNA damage is effectively removed by the nucleotide excision repair (NER) pathway. Alcohol derived acetaldehyde adducts, on the other hand, are removed through base-excision repair (BER). However, the DNA repair systems are not completely error-free, some adducts escape repair and persist in DNA [101, 102]. A preferential accumulation of the mutation on the non-transcribed strand of DNA is the consequence of the repair of adducts on the transcribed strand and not in nontranscribed strand by transcription-coupled repair systems (TCR) [103]. This phenomenon results in the accumulation of certain types of mutation on the non-transcribed strand, such as G-T transversions in lung cancer [104]. DNA repair enzyme levels can be also affected by recent exposure [105], for example some tobacco smoke constituents may inhibite repair enzymes [3]. Smokers have been shown to have lower levels of folic acid than nonsmokers which may impact on the repair capacity [106]. Interindividual differences in the capacity for DNA repair have also been observed [100]. These persistent adducts can cause miscoding and lead to mutations which can activate oncogenes such as *KRAS* or inactivate tumor-suppressor genes such as *TP53* [80, 107–109].

Genetic changes

Many genetic abnormalities occur during the process of lung cancer induction, including loss of heterozygosity, microsatellite alterations, mutations in *RAS* oncogenes, *MYC* amplification, *BCL-2* expression, mutations in *TP53*, *RB*, *CDKN2* and *FHIT* tumor suppressor genes, expression of telomerase activity etc. [3]. In head and neck cancers, mutations in *RB* and *CDKN2* genes frequently occur as does *cyclin D1* amplification [30].

Mutations in the smoking-associated tumours have been identified both in oncogenes and in tumor suppressor genes. The gene most frequently found to be mutated in smok-

ing-associated lung tumours is *TP53* [104, 110]. The *TP53* mutation seems to occur at an early stage in many types of cancer that are caused directly by exogenous carcinogens [3, 111]. The frequency of *TP53* mutations bears a direct correlation with the number of cigarettes smoked. Mutations in *KRAS* and *TP53* result directly from the reaction of these genes with metabolically activated carcinogens. For instance, the B[a]P metabolite BPDE has been shown to inflict mutations in the *TP53* gene [38, 109, 112]. There is also evidence that endogenous oxyradicals and nitrogenoxyradicals have an impact on the *TP53* mutation spectra [111].

Thirty percent of *TP53* mutations in lung tumours of smokers are GC→TA transversions. This is the primary class of base substitution induced by PAHs that form bulky DNA adducts [113–115]. There is a precise correlation between the mutational hot spots and sites of DNA adducts remaining after the cells have been exposed to BPDE and other diol epoxides, and have subsequently undergone a period of DNA repair [109, 116]. These mutations are targeted at methylated CpG sites [117]. Most of the mutations are found in the nontranscribed DNA strand resulting from preferential binding and slow repair of the BPDE adducts formed on the nontranscribed strand [118].

Mutations at the *KRAS* gene occur in about 30% of the lung adenocarcinomas of smokers and are primarily GC→TA transversions [118]. In HNC patients, an intermediate *TP53* mutation profile between tobacco-related and non-related cancers occurs. The degree of similarity with tobacco-related cancers increases from the oral cavity to the pharynx and larynx [30].

Deletions and other genome rearrangements are associated with carcinogenesis and inheritable diseases. The fractional allelic loss or gain occurred at a much higher frequency in lung tumours of smokers (48%) than in those of nonsmokers (11%). This indicates that different genetic alterations play a role in the development of smoking associated lung cancer compared to those implicated in the development of lung cancer in non-smokers [3, 119]. Also microsatellite instability in colon tumours [120] and chromosome 9 alterations in bladder tumours [121] have been associated with cigarette smoking.

The frequency of sister chromatid exchanges (SCEs) in peripheral lymphocytes is higher in smokers than in nonsmokers [3]. It is suggested that PAHs and the neutral fractions and weakly acidic, semivolatile components of cigarette-smoke condensate were the most potent inducers of SCE [122, 123]. The acid fraction included phenolic compounds such as catechol and hydroquinone, compounds known to generate free radicals. PAHs alone were not adequate for the mutagenic activity of the condensate [124, 125]. Exposure of rodents to cigarette smoke has been consistently shown to produce micronuclei in bone marrow, blood and in pulmonary alveolar macrophages [126–128].

Epigenetic changes

Epignetic refers to a heritable change in gene expression without any alterations in the primary DNA sequence of the gene. For example, methylation usually silences gene

expression. Normally, about 70% of all CpG dinucleotides in the mammalian genome are methylated. Hypomethylation across the genome or hypermethylation in the CpG islands are a hallmark of most cancers. The major epigenetic aberration occurring in cancer development is the silencing of tumor suppressor genes through the hypermethylation of the CpG islands in the promoter regions [129, 130].

Knudson`s two hit hypothesis proposes that two successive genetic hits are needed to transform a normal cell into a tumour cell [131], that is, two inactivating hits are required to cause loss of function of tumor suppressor genes. The functional significance of tumor suppressor gene hypermethylation has been highlighted in several studies showing that tumor suppressor gene hypermethylation can act as the second inactivating hit of a tumor suppressor gene following the first-hit gene mutation [132]. However, epigenetic changes can occur in the absence of a genetic lesion and also biallelically if the coding sequences are unaltered. Genes affected by epigenetic silencing in cancer include genes related to cell-cycle control, DNA repair, apoptosis, and metastatic potential. The patterns of CpG methylation are non-random and show distinct tumour-type specificities, suggesting that each tumour may display a distinguishable methylation subtype in the genome [130].

B[a]P has been shown to interact with SAM-dependent methyltransferases like glycine *N*-methyltransferase (GNMT). GNMT enzyme activity was reduced by nearly 50% in the presence of B[a]P. B[a]P may thus affect DNA methylation via interactions with DNA methyltransferases and in this way contribute to a carcinogenic pathway [133].

Cancer initiation

Mutations that activate oncogenes or inactivate tumor suppressor genes accumulate during the development of cancers of the oropharynx [30] and lung [25]. These changes lead to self-sufficiency in growth signals, insensitivity to anti-growth signals, evasion of apoptosis, tissue invasion and metastasis, sustained angiogenesis and limitless replicative potential [3, 134]. The disruption can occur through genetic or epigenetic changes. Smoking is associated with some of the genetic and epigenetic changes affecting these major pathways. Thus the effects of tobacco smoke as a carcinogen can be viewed as the result of both genetic and epigenetic changes. Epigenetic changes can trigger a complex suppressive cellular stress response and genetic changes may endow some cells with a capacity to escape normal immunosuppression [3].

Genetic variation

Even though the DNA sequence of any two people is 99.9 percent identical, a tremendous amount of variability exists in the human genome. Inherited differences in the sequence of DNA among individuals are called genetic variation. Genetic variation accounts for

differences between people, such as eye color and blood group. It also plays a role in determining how an individual responds to the environment, and in that way the individuals risk of developing a disease [135].

Single DNA base differences that occur at a frequency of >1% in the population are called single nucleotide polymorphisms (SNPs). SNPs account for 90–95% of all variant sites in human genome [135]. SNPs can be intronic, exonic, or can be located upstream or downstream from the gene, *i.e.*, in the 5′- flanking region and 3′-untranslated region, respectively. Other types of genetic variation include deletions, insertions and repeat variants. Recently variation in the gene copy number was also found to be widespread in human genome. This phenomenon is called copy number polymorphism (CNP) [136]. The variation in DNA can alter the function of the protein, or expression, or have a gene-dosage effect thus producing changes in phenotype.

It is estimated that there are approximately 11 million SNPs in the human genome with a minor allele frequency of 1% and approximately 5 million SNPs at the 10% minor allele frequency [135]. One common SNP ($q \geq 0.10$) occurs every 500–600 bases, and one polymorphic SNP ($q \geq 0.01$) occurs, on average, every 150–250 bases [135]. Many of these SNPs remain to be discovered.

Many public data bases gather data about genetic variation and many programs are working to confirm, identify, characterize and quantify the genetic variation [137, 138]. The dbSNP database currently includes 10 million human SNPs which have been mined from different libraries and individual DNA samples (http://www.ncbi.nlm.nih.gov/projects/SNP/).

Genetic susceptibility to cancer

Given that tobacco-related cancers develop in only a small fraction of the individuals exposed to tobacco smoke, it is tempting to hypothesize that genetic factors may render some smokers more susceptible to cancer. Genetic factors may modify the individual susceptibility in the risk of tobacco-related cancers via effects on smoking behaviour, carcinogen metabolism and detoxification, DNA repair, cell cycle control and other cellular responses (Figure 8) [139]. The strong linkage of alcohol in the development of HNCs mean that it is advisable to search for inherited changes related to the propensity to drink as well as to alcohol biotransformation in any examination of the role of genetic variation in the risk of these cancers [26].

In general, cancer susceptibility genes can be divided to high-penetrance genes, in which specific genetic changes are associated with high disease risks, and to low-penetrance genes in which the risk attributable to the genetic change is low. Although the individual risk associated with the at-risk genotypes of low-penetrance genes is much lower than that of the high-penetrance genes, the prevalence in the population of low-penetrance alleles

is much higher the and consequently also the proportion of cancer cases attributable to low-penetrance alleles is anticipated to be much larger than that of high-penetrance alleles [140].

The existence of functional genetic polymorphisms of XMEs that have a significant impact on the expression and catalytic activity of the enzymes has been documented. It has also been demonstrated that either inhibition of XME-catalyzed carcinogen activation or induction of XME-catalyzed carcinogen inactivation may result in a reduction of tumour occurrence. To date, the bulk of data exists points to the existence of many XME genes as being potential modifiers of lung and oropharyngeal cancers. These XMEs include CYP1, CYP2 and CYP3 genefamilies, members of the GST genefamily, NATs and alcohol metabolizing enzymes like alcohol and aldehyde dehydrogenases [30, 141, 142]. The genes and their products have been shown to act as effect-modifiers with respect to dietary elements, tobacco use, occupation, other exposures and other genes [143].

Coding region SNPs (cSNPs) and and regulatory region SNPs (rSNPs) are considered to be the most important SNPs in disease susceptibility. cSNPs are located in the coding exon regions of the gene and they may change the structure of the protein e.g an enzyme, to increase/decrease the activity or change the interaction with other proteins. There are an estimated 50,000 non-synonymous cSNPs and 20,000 non-conservative cSNPs genome-wide that are potentially functional variants [135]. rSNPs affect gene expression throught various signaling pathways and molecules. Regulation regions of a gene, and thus rSNPs, can be located up to 10kb upstream of the gene, in intron regions and within or near the 3′ ends of genes [144]. An unknown number of SNPs in non-coding rSNP may be functional variants [135].

Glutathione-S-transferases

GSTs, which were described already in the early 1970s [145], are crucial enzymes in the protection of nucleophilic centres in DNA and protein from modification by electrophiles [145, 146]. GSTs catalyze the conjugation of reduced glutathione (GSH) with electophilic groups of a wide variety of compounds, neutralizing their electrophilic sites, and rendering the products more water-soluble [147]. GST conjugates are transported to the kidneys, the glutamate and glycine residues are removed and the amino group of cysteine is acetylated before the resulting mercapturic acid is excreted in urine [54]. In some cases, however, glutathione conjugation can lead to the formation of conjugates that may be more toxic or reactive than the parent compounds [148–151].

GSTs are a superfamily of ubiquitous, multifunctional enzymes currently consisting of close to 20 human cytosolic forms, and five membrane bound forms. Based on sequence homology and immunological crossreactivity, human cytosolic GSTs have been grouped into eight families, designated as Alpha, Kappa, Mu, Omega, Pi, Sigma, Theta, and Zeta

[152, 153]. It is possible that GSTs have arisen from a single common ancestor and their substrate specificities and diversities have been reshaped by gene duplication, gene recombination and accumulation of mutations. These proteins have been found in virtually all living species examined, including plants, animals and bacteria [154]. GST composition varies among tissues and the particular combination of GSTs expressed in a tissue may influence its detoxifying capability [155].

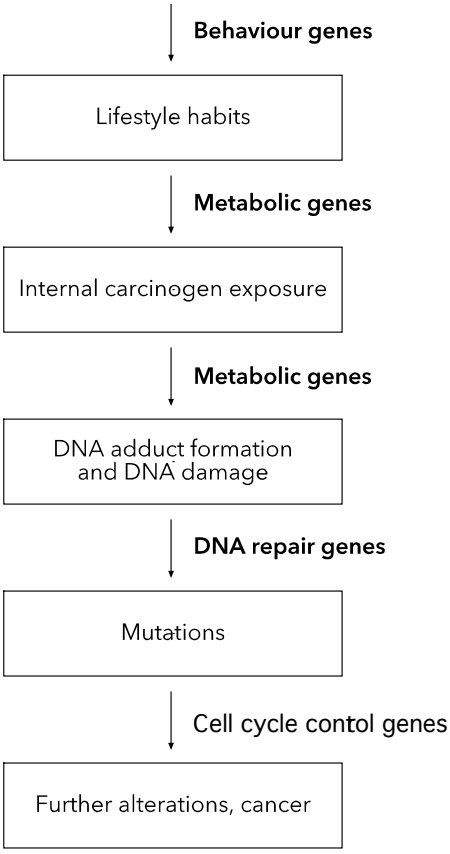


Figure 8. Genetic factors may modify the individual susceptibility to the risk of tobacco-related cancers via various pathways. Differences in smoking behaviour may be mediated by genes participating in nicotine metabolism. Metabolic capacity may modulate via the differences in metabolic genes which may specify the amount of internal dose and thus the DNA adduct formation and the amount of DNA damage. DNA repair gene polymorphim may define the amount of mutations and differences in cell cycle control genes may define further alterations. Modified from Brennan and Boffetta [30].

Soluble GSTs exist as dimeric proteins, with subunit molecular weights of approximately 25 kDa. Each subunit of a dimeric enzyme has an active site composed of two distinct functional regions: a hydrophilic G-site, which binds the physiological substrate glutathione, and an adjacent H-site which provides a hydrophobic environment for the binding of stucturally diverse electrophilic substrates [156].

Substrates

GSTs catalyze the general function: GSH + R-X → GSR + HX. The enzyme brings the substrate into close proximity with GSH by binding both GSH and the electrophilic substrate to the active site of the protein, and allowing the activation of the sulhydryl group on GSH, thereby facilitating the nucleophilic attack of GSH on the electrophilic substrate (R-X) [145, 155]. The electrophilic functional center of the substrates can be a carbon, nitrogen or sulfur. A large number of diverse chemicals serve as substrates for GSTs. Several reactive endogenous molecules derived from prostaglandin metabolism and lipid peroxidation serve as substrates for certain GSTs [154]. Chlorinated nitrobenzenes (1-chloro-2,4-dinitrobenzene (CDNB) and 1,2-dichloro-4-nitrobenzene (DCNB)) have been used as standard substrates to measure GST activity. However, theta class GSTs do not catalyze these reactions and the specific activities towards CDNB and DCNB varies between the different isoforms.

Many epoxide carcinogens are detoxified by GSTs. The carcinogenic epoxide of B[a]P (BPDE) is efficiently detoxified by GSTs, with especially high and selective activity exhibited by GSTP1 [157]. The *trans*-isomer of stilbene oxide (TSO) is exclusively conjugated by GSTM1. GSTT1 catalyzes the activation of small bifunctional electrophilic molecules such as dichloromethane, ethylene bromide, and butediene diepoxide [158, 159]. Certain quinone metabolites of endogenous molecules may also be substrates for GSTs [155]. GSTs are also involved in the metabolism of lung cancer chemopreventive agents such as isothiocyanates [160].

Expression

GSTs are expressed in rather large amounts, accounting for as much as 4% of total soluble protein in the liver. *GSTM1* is most strongly expressed in the liver and to a much lower extent in lung and colon. It is also expressed in lymphocytes, heart, liver, stomach, adrenal gland and kidney [161–164]. Also *GSTT1* is highly expressed in liver. It is also found in erythrocytes, lung, kidney, brain, skeletal muscles, heart, small intestine and speen. Unlike other GSTs, the activity of GSTT1 is undetectable in lymphocytes [165, 166]. GSTP1, the dominant GST present in the lung, is also widely expressed in other types of epithelial tissue and in lymphocytes [162]. *GSTM3* is expressed in testis, brain, lung and lymphocytes [162, 167, 168].

Induction of GSTs by xenobiotics is mediated via several different transcriptional mechanisms. Most of the functional studies have been conducted in animals. The rat *GSTA2* gene contains a glucocorticoid-response element (GRE), a xenobiotic response element (XRE), and an antioxidant-response element (ARE) [169]. The GRE mediates induction by glucocorticoids, the XRE by planar aromatic hydrocarbons, and the ARE by phenolic antioxidants [155].

In humans, GSTs and glutathione synthesis have been shown to be up-regulated by electrophiles through the Nfr2 transcription factor via the cis-acting ARE [170]. Compounds that up-regulate GSTs via this pathway have been shown to possess anticarcinogenic properties [171–173]. Since many naturally-occurring plant products (phytochemicals) are known to induce GSTs, the role of dietary induction of GSTs has been studied as a mechanistic explanation for the anticarcinogenic effects of fruits and vegetables [174]. Chemopreventive compounds have been shown at low concentrations to activate mitogen-activated protein kinase (MAPK) pathways leading to induction of phase II detoxifying enzymes and other cellular defense enzymes via the ARE. However, at higher concentrations, these agents activate the MAPK and caspase pathways leading to apoptosis [175].

The expression of *GSTP1* is significantly increased in many human tumors and human cell lines resistant to chemotherapeutic agents. It can be induced by oxidative stress, probably via the NFκB response element [154]. Both transcriptional; functional AP-1 and SP1 response elements have been identified in the 5′-regulatory region of human *GSTP1* gene [176] repressors [177] and post-transcriptional mechanisms are involved in the regulation of human GTSP1 protein levels [178]. Also the methylation status of a CpG island in the regulatory region of the gene may modulate the expression of the *GSTP1* [179]. The regulation of mu and theta class GSTs is poorly understood.

There is no clear evidence of GST induction by tobacco smoke [3, 180]. Regulation of *GST* expression varies between different tissues. However, due to the modulation of expression by diet and xenobiotics and due to the fact that some individuals lack the *GSTM1* and *GSTT1* gene it has been difficult to predict accurately the extent of expression of any single *GST* gene in a given tissue [155]. However, there is some information on the relative expression of different *GST* classes in various tissues. Alpha class *GSTs* are relatively highly expressed in liver, kidney, and testis; *GSTP1* is expressed in lung and brain; *GSTM1* is expressed at high levels in the liver; *GSTM2* is expressed at the highest level in brain; *GSTM3* is expressed in testis; *GSTT1* is expressed predominantly in liver and kidney and to a lesser extent in other organs [154, 168, 181].

GSTM1 and GSTM3 in cancer proneness

GSTs have been considered as being important in cancer development because of their critical role in providing protection of DNA against damage and adduct formation [182, 183]. The mu family of GSTs consists of five members called GSTM1–GSTM5. The GSTM

enzymes have multiple substrates, including epoxides of PAH such as BPDE, acrolein and other unsaturated carbonyls generated by lipid peroxides and oxidative damage to DNA as well as constituents of plant foods, *e.g.*, isothiocyanates [184, 185]. In addition to having different specific activities toward substrates and partially different tissue specific expression patterns, these enzymes also have overlapping substrate specificities [162, 186].

In humans, the Mu class genes are clustered together on chromosome region 1p13.3 and are arranged as 5' - GSTM4 - GSTM2 - GSTM1 - GSTM5-3′ (plus strand) and 3′-GSTM3 -5′ in the inverted direction (minus strand). The *GSTM1* gene consists of eight exons encoding a protein consisting of 217 amino acids [187–190]. The *GSTM3* gene consists of nine exons encoding a protein of 224 amino acids, sharing 72% amino acid sequence similarity with the GSTM1 [191].

Homozygous deletion of the *GSTM1* gene (null genotype) results in total lack of enzyme activity in about 50% of Caucasians [192]. The frequency of the null genotype is similar in Asians, but lower in Africans (about 30 %) [193]. The *GSTM1* gene deletion is most likely caused by a homologous recombination involving the two almost identical 4.2-kb regions flanked by the *GSTM1* gene [194]. The frequency of the positive allele is about 0.23 in Caucasians and 0.41 in African-Americans [195].

In addition to deletion polymorphism, a total of 50 SNPs spanning the *GSTM1* gene region are currently listed in NCBI dbSNP database (http://www.ncbi.nlm.gov/SNP). Two of these SNPs are non-synonymous changes. One of these changes is the $Lys^{173}Asn$ amino acid replacement (SNP; rs 1065411) resulting from a C to G base change in exon 7. The variant allele was previously desingnated as *GSTM1*B*. The amino acid change does not appear to affect the enzyme function [196]. One other SNP is the $Thr^{210}Ser$ change (SNP; rs 449856) resulting from an A to T nucleotide substitution in exon 8. This SNP is possibly a database error as the same change has been observed in the *GSTM2* gene [197]. A total of 14 SNPs are in the promoter region, 21 are located in the intron regions, four are synonymous, and the remaining nine are located in the 3′-untranslated region (http://www.ncbi.nlm.gov/SNP). A duplication of the *GSTM1* gene has also been observed in Saudi Arabians. These subjects exhibited a ultrarapid GSTM1 enzyme activity [198].

To date, 26 SNPs in *GSTM3* gene region have been listed in the dbSNP database (http://www.ncbi.nlm.gov/SNP). Two non-synonymous changes, $Asn^{128}Lys$ (SNP; rs1803687) and $Val^{224}Ile$ (SNP; rs7483), have been reported. The frequency of the former one is not known while the frequency of the minor allele (Ile^{224}) of the latter is about 0.3 in Caucasians. The $Val^{224}Ile$ change is located in the last exon of the gene. In experimental analysis, the Ile^{224} variant (*GSTM3*C*) tended to show increased specific activity and catalytic efficiency with CDNB [197]. Four of the SNPs are in the putative promoter region. One of these SNPs, A/C (SNP; rs1332018) is located 63 bp upstream of the translation start site and it has recently been shown to have a strong effect on gene expression [199]. Fourteen of the SNPs are located in the 3′-untranslated region, one of the SNPs is synonymous, and five are in the intron regions. One of these SNPs is the deletion of 3 bp (-/AGG) in intron 6 (SNP;

rs1799735), resulting in the formation of a recognition sequence for the YY1 transcription factor in the latter [200]. This polymorphism has been designated as *GSTM3*B* allele [200]. The minor allele frequency is about 0.15 in Caucasians [201–203].

Multiple lines of evidence from molecular epidemiological studies suggest that *GSTM1* is involved in cancer susceptibility [143, 204–206]. An association between *GSTM1* null genotype and lung cancer has been found in many populations [207–211], although some studies have also failed to find any significant association between the *GSTM1* null genotype and lung cancer risk [212]. The first meta-analysis (Table 3) comprising of 12 studies showed a moderate increase in the risk of lung cancer, with an odds ratio of 1.41 (95% CI, 1.23–1.61). The risk was somewhat lower when only the results from genotyping analysis were considered 1.34 (95% CI, 1.16–1.55) [206]. In a subsequent meta-analysis of 23 studies, the lung cancer risk associated with the *GSTM1* null genotype was 1.13 (95% CI, 1.04–1.25). The results were based on the genotype data from 3098 cases and 5580 controls [143]. The most recent meta-analysis of 43 studies including over 18000 individuals shows a similar effect, *i.e.*, an OR of 1.17 (95% CI, 1.07–1.27) [213]. However, results from pooled analysis of 3940 cases and 5515 controls did not support the role of *GSTM1* null in lung cancer proneness (OR 1.08; 95% CI, 0.98–1.18) [213]. Diet seems to modulate the results; a weaker association between years of smoking and lung cancer risk was observed in *GSTM1* null carriers who were subjeted to α-tocopherol supplementation than in those without supplementation [214].

In interpreting the GSTM1 genotype data, it has to be borne in mind that *GSTM1* is only weakly expressed in pulmonary tissue [146, 162, 215], the major mu class enzyme in the lung is GSTM3. Interestingly, the expression of GSTM3 has been shown to be greater in *GSTM1* positive than in *GSTM1* null smokers [216]. Therefore it has been hypothesized that the *GSTM1* gene has two effects on lung cancer, one direct through its expression in liver and its effects on the hepatic metabolism of tobacco smoke carcinogens and the other indirect, due to an effect of the *GSTM1* genotype on expression of pulmonary *GSTM3* [216]. However, a recent study does not support the theory that the GSTM1 polymorphism plays an important modifying role in *GSTM3* expression [199]. Although linkage dysequilibrium has been noted between the *GSTM1*A* and *GSTM3*B* alleles [200, 217], no association between *GSTM3*B* polymorphism and lung cancer has been observed [201, 217, 218].

The combination of the *CYP1A1* and *GSTM1* variant genotypes has been of great interest, since the combined effect of increased activation and decreased detoxification of PAH has been hypothesized to pose a particularly high lung cancer risk. An increased risk of lung cancer with the combination of variant *CYP1A1* and *GSTM1* genotypes has been observed in many populations; the strongest effects have been shown in Asians [63, 219–221]. A pooled analysis among Caucasian non-smokers, detected no risk for *GSTM1* alone, but this mutation was associated with an increased lung cancer risk when it occurred

Table 3. Meta and pooled analysis on lung cancer and *GSTM1* genotype.

First authors year	Case/ Control (n)	Type of study	Number of studies included	Study population	OR (95% CI) for *GSTM1*
McWilliams, 1995 [206]	1593/2135 phenotyping 304/312	Meta-analysis	12	Caucasians Japanese	All 1.41 (1.23-1.61) Caucasian 1.17 (0.98-1.40) Japanese 1.60 (1.25-2.13 Phenotyping All 1.80 (1.29-2.50) Genotyping All 1.34 (1.16-1.55)
Houlston, 1999 [143]	3593/6131 phenotyping 495/551	Meta-analysis	23	Caucasians Asian African-Americans Mexican-Americans	Phenotyping 2.12 (1.43-3.13) Genotyping 1.13 (1.04-1.25) Caucasian genotyping 1.08 (0.97-1.22) Asian genotyping 1.38 (1.12-1.69)
Benhamou, 2002 [213]	7463/10789	Meta-analysis	43	Caucasians Asians mixed others	All 1.17 (1.07-1.27) Caucasian 1.10 (1.01-1.19) Asian 1.33 (1.06-1.67)
	3940/5515	Pooled analysis	21	Caucasians Asian	1.08 (0.98-1.18)

in combination with *CYP1A1* Ile/Val polymorphism (OR 4.67; 95% CI, 2.00–10.09) [222].

A meta-analysis of 26 studies on HNC including oral, pharyngeal and laryngeal cancers found an increased risk for *GTSM1* null carriers (OR 1.23; 95% CI, 1.06–1.42) [223]. Pooled analysis comprising 2334 cases and 2766 controls pointed to a somewhat higher cancer proneness for *GSTM1* null carriers (OR 1.32; 95% CI, 1.07–1.62) [223]. Decreased [203, 224, 225] and increased [226] risks of UAT cancers associated with *GSTM3*B/B* genotype have been observed.

GSTP1 in cancer proneness

The GSTP1 enzyme has been of particular interest as a lung cancer susceptibility marker due to its high expression in the lung, and its role as one of the main detoxifiers of the highly carcinogenic compound, BPDE [157, 162, 227].

The GSTP class appears to contain only one or two distinct genes in most species. Thus it is less complicated than the mu class multi-gene family. In humans, *GSTP1* is mapped to chromosome 11q13. The translated protein of 209 amino acids is the product of a single gene. The gene is 2.8 kb long and contain seven exons [228].

To date, a total of 62 variants in the *GSTP1* gene region have been described (http://www.ncbi.nlm.gov/SNP). These variants include five non-synonymous SNPs, of which the $Ile^{105}Val$ (SNP; rs947894) resulting from an A to G base change in exon 5, is the most common. The frequency of the minor allele Val^{105} as (*GSTP1*2*) is about 0.20. The $Ile^{105}Val$ amino acid change affects substrate specificity to the point of distinguishing between planar and nonplanar substrates [228–232]. The effect of this polymorphism appears to depend on which substrate is present. A *GSTP1*2* homozygotes have been reported to be more susceptible to the effects of carcinogens that share a structural similarity to CDNB but less susceptible to the effects of PAH diol epoxides [232–234]. The residue at site 105 appears to define the geometry of the hydrophobic substrate-binding site such that enzyme activity toward small substrates will be greater with isoleucine at 105 where a larger substrates such as PAH have better access when valine is at position105 [229–231, 235].

The $Ala^{114}Val$ polymorphism (SNP; rs1799811) results from a C to T transition in exon 6. The frequency of minor allele Val^{114} (*GSTP1*3*) is about 0.10 in Caucasians. *GSTP1*2* and *GSTP1*3* alleles have been shown to be in linkage disequilibrium. $Ala^{114}Val$ polymorphism may modify the enzyme activity in the presence of Val^{105} allele [232]. However, no clear effects on enzyme activity have been described [233].

The rest of the non-synonymous SNPs in *GSTP1* either have not been validated or have a very low frequency. The gene also contains five synonymous SNPs. Twenty three of the SNPs are located in the 5′-untranslated region and five in the 3′-untranslated region. The rest of the SNPs are located in the intron regions.

The data on the association between the *GSTP1* polymorphism and lung cancer risk has been conflicting (Table 4). Most of the studies haave detected no statistically significant associations [201, 217, 218, 236–245]. However, a few studies have indicated that the *GSTP1*2* allele is a risk factor for lung cancer [210, 246, 247] and many studies have observed increased risk linked to *GSTP1*2* allele when it is present in combination with other *GST* gene polymorphisms [201, 210, 218, 238, 246].

The *GSTP1*2* allele has been associated with an increased risk of UAT cancers in some, but not all, studies. Meta-analysis of nine studies on HNC including oral, pharyngeal and laryngeal cancers did not show an increased risk for *GSTP1*2* allele-carriers, OR 1.10 (95% CI, 0.92–1.31) [223]. Pooled analysis comprising 2334 cases and 2766 controls showed similar results, OR 1.15 (0.86–1.53) [223]. However, in combination with *GSTM1* and *GSTT1* null genotypes, the OR was 2.06 (95% CI, 1.11–3.81) [223].

Table 4
Studies on the *GSTP1* genotypes and lung cancer risk.

First author, year	Case/ Control (n) ethnicity	Ever smokers Case/ Control (%)	Mean Age Case/ Control	Cancer subtypes (%) of all cancers	*2-allele freq	*GSTP1*1/*2* Case/Control *GSTP1*2/*2* Case/Control freq. (%)	OR (95% CI) for *GSTP1*	Genotype combinations and comments
Ryberg, 1997 [210]	138/297 Caucasian	100/65	62/54	SCC 49 AC 30	0.288	45.7/39.4 15.9/9.1	*GSTP1*2*-carrying 1.7 (1.1-2.6)	SCC: *GSTP1*2* -carrying 2.5 (1.4-4.4); *GSTM1* null and *GSTP1*2*-carrying 2.0 (1.3-3.1)
Jourenkova-Mironova, 1998 [218]	150/172 Caucasian hospital	100/100	58/55	SCC 65 SCLC 35	0.333	44.0/37.2 11.3/12.8	*GSTP1*2*-carrying 1.1 (0.7-1.8)	>35PY: *GSTM1* null and *GSTP1*2*-carrying 2.4 (1.1-5.1); *GSTM3 AA* and *GSTP1*2*-carrying 2.1 (1.0-4.2); *GSTM1* null and *GSTM3 AA* and *GSTP1*2*-carrying 2.7 (1.2-6.0)
Harris, 1998 [236]	184/199 Caucasian (Australia)	94/ND	66/39	SCC 44 AC 18	0.344	41.0/50.8 14.6/9.0	*GSTP1*2*-carrying 0.8 (0.6-1.3)	
To-Figueras, 1999 [237]	164/132 Caucasians	98/100	59/46	SCC 32 AC 26 SCLC 35	0.311	38/41 11/11	*GSTP1*2*-carrying 1.2 (0.7-2.1)	*GSTM1* null and *GSTP1*2*-carrying 1.4 (0.9-2.4)
Kihara, 1999 [238]	382/257 Japanese	94/72	62/56	SCC 33 AC 46 SCLC 21	0.158	22.5/25.3 4.7/3.1	*GSTP1*2*-carrying 0.9 (0.7-1.3)	*GSTM1* null and *GSTP1*2*-carrying 2.6 (1.3-5.3)
Kiyohara, 2000 [245]	86/88 Japanese	68.6/45.5	ND	SCC 28 AC 47	0.148	26.7/29.5 2.3/-	*GSTP1*2*-carrying 1.0 (0.4-5.5)	
Risch, 2001 [201]	389/353 Caucasian hospital	88.9/66	60.9/55.3	SCC 44 AC 39	0.313	44.9/42.8 9.8/9.9	*GSTP1*2*-carrying 1.1 (0.8-1.5)	*GSTT1* present and *GSTP1*2*-carrying 2.2 (1.1-4.5)
Stücker, 2002 [246]	251/264 Caucasian hospital	97/97	60/60	SCC 46 AC 24 SCLC 19	0.303	40.2/45.5 12.0/7.6	*GSTP1*2/*2* vs *GSTP1*1*-carrying 2.0 (1.0-4.1)	*GSTM1* null and *GSTP1*2/*2* 2.8 (0.9-8.5), SCLC: for *GSTP1*2/*2* vs *GSTP1*1*-carrying 3.6 (1.3-9.6); *GSTM1* null and *GSTP1*2/*2* 7.0 (1.6-30.2)

Table 4 (contd)

First author, year	Case/ Control (n) ethnicity	Ever smokers Case/ Control (%)	Mean Age Case/ Control	Cancer subtypes (%) of all cancers	*2-allele freq	*GSTP1*1/*2* Case/Control *GSTP1*2/*2* Case/Control freq. (%)	OR (95% CI) for *GSTP1*	Genotype combinations and comments
Lewis, 2002 [239]	94/165 Caucasian hospital	98/87	59.5/67.4	SCC 34 AC 12 SCLC 16	0.331	57.0/49.0 6.5/8.6	*GSTP1*2*-carrying 1.3 (0.8-2.2)	
Miller, 2002 [248]	767/927 Caucasian	96/64	67/61	SCC 24 AC 44	0.332	42/46 12/10		*GSTM1* null and *GSTP1*2*-carrying 1.3 (0.9-2.1)
Wang, Y 2003 [240]	362/419 Caucasian hospital	87.5/91.4	60.9/60.4	ND	0.335	49.2/46.1 9.7/10.5	*GSTP1*2*-carrying 1.1 (0.8-1.5)	Ala114Val: GSTP1*3-carrying 1.4 (1.0-1.9)
Wang, J 2003 [241]	112/119 Chinese	42.9/40.3	56.5/54.5	AC 100	0.151	39.3/28.6 0.9/0.8	*GSTP1*2*-carrying 1.6 (0.9-2.8)	
Lin, 2003 [242]	198/332 Taiwanese	57.6/47	64/58	SCC 42 AC 53	ND	*GSTP1*2*-carrying 37.4/31.9	*GSTP1*2*-carrying 1.4 (0.9-2.0)	SCC: *GSTP1*2*-carrying 1.6 (0.96-2.7)
Nazar-Stewart, 2003 [243]	274/501 >95% Caucasian	96/72.2	range 18-74	SCC 29 AC 35 SCLC 19	ND	*GSTP1*2*-carrying 59.1/58.1	*GSTP1*2*-carrying 1.4 (0.9-2.0)	
Reszka, 2003 [249]	138/165 Caucasian hospital	95.7/75.6	59.7/59.6	SCC 44 AC 9 SCLC 25	0.264	43.5/46.7 3.6/3.0	*GSTP1*2*-carrying 1.1 (0.6-1.8)	
Miller, 2003 [247]	1042/1161 not mentioned	94/64	66/60	SCC 22 AC 39	0.330	42.3/45.2 12.3/10.4	current smokers *GSTP1*2/*2* 1.95 (1.03-2.95)	non-smokers *GSTP1*2/*2* 0.86 (0.32-2.24) ex-smokers *GSTP1*2/*2* 1.19 (0.78-1.82) Risk associated with PYs
Schneider, 2004 [244]	446/622 Caucasian hospital	89.3/75.4	64.4/63.6	SCC 41 AC 25 SCLC 15		41.7/40.8 13.9/11.3	*GSTP1*2*-carrying 1.0 (0.8-1.4)	
Sorensen, 2004 [217]	256/269 Caucasian	97.3/96.6	range 50-65	SCC 22 AC 33 SCLC 20	0.335	40.9/47.8 13.0/9.0	*GSTP1*1/*2* 0.8 (0.6-1.1) *GSTP1*2/*2* 1.2 (0.8-2.2)	interaction with *GSTT1*

GSTT1 in cancer proneness

The human GST theta subfamily includes two members, GSTT1 and GSTT2. GSTT1 is a homodimeric enzyme consisting of protein of 239 amino acids encoded by a single gene [151, 250–252].

GSTT1 exhibits different catalytic activity compared with the other GSTs. The differences between GSTT1 and other GSTs derive from its early evolutionary divergence. *GSTT1* is suggested to be the ancient progenitor *GST* gene [159]. GSTT1 uses GSH at higher concentrations and the rate of enzymatic activity is many times higher than is the case with the other GSTs. It also has lower affinity for glutathione-conjugates and also releases the products easily. GSTT1 can dehalogenate and biotransform molecules to mutagenic electrophilic compounds. It shows also activity towards molecules with epoxide groups, and has peroxidase activity towards hydroperoxides.

Like *GSTM1*, *GSTT1* exhibits genetic polymorphism due to a null allele, resulting from gene deletion [151]. There are two 18 kb regions, HA3 and HA5, with >90% homology flanking the *GSTT1* gene. HA3 and HA5 contain two identical 403-bp repeats, which were identified as deletion/junction regions of the *GSTT1* null allele [253]. The *GSTT1* null allele has been thus most likely caused by a homologous recombination involving the left and right 403-bp repeats. About 20% of Caucasians are homozygous for a *GSTT1* null allele. From 49% to 65% of Asians but less than 10% of Mexican-Americans lack the gene [254, 255]. Determination of GSTT1 enzyme activity in erythrocytes showed a trimodal phenotypic distribution corresponding to the positive/positive, positive/null, null/null genotypes. The frequency of the null allele is about 0.43 [253].

In addition to deletion polymorphism, a total of 32 SNPs spanning the *GSTT1* gene region are currently listed in NCBI dbSNP database (http://www.ncbi.nlm.gov/SNP). Five of these are non-synonumous changes. These SNPs have not, however, been validated or are present at very low frequencies. Eight SNPs have been described in the promoter region, but none of these have been validated. 14 are located in the intron regions, two are synonymous and the reamining three are located in the 3′-untranslated region.

In addition to two functional genes a pseudogene has been identified in humans for *GSTT2*. The genes are found on chromosome 22q11.2 [159]. Both genes have five exons with identical intron/exon boundaries but share only 55% amino acid identity.

The *GSTT1* null genotype has been suspected to confer either decreased or increased risk of cancer in relation to the source of environmental exposure [159]. This is possible since GSTT1 is involved in the metabolism of monohalomethanes and ethylene oxide found in tobacco smoke. PAHs are likely to be minor substrates for GSTT1. No changes in the level of BPDE-adducts have been observed in *GSTT1* null individuals [256]. Furthermore, no support for any association between *GSTT1* genotype and lung cancer risk has been observed (Table 5) [201, 218, 239, 241, 243–246, 257–262]. A recent meta-analysis of 21 studies on HNC including oral, pharyngeal and laryngeal cancers reported

Table 5
Selected studies on *GSTT1* null genotype and lung cancer risk.

First author, year	Case/control ethnicity	Ever smokers Case/ Control (%)	GSTT1 null freq. (%) case/control	OR (95% CI) for GSTT1 null
Deakin, 1996 [265]	108/509 Caucasian	100/51	15.7/18.5	1.9 (1.3-2.8)
Kelsey, 1997 [259]	60/146 Mex-American 108/132 African-American	ND	17/12 25/22	1.5 (0.7-3.5) 1.2 (0.7-2.2)
El-Zein, 1997 [260]	54/50 ND	100/100	22/14	1.75 (0.56-5.52)
To-Figueras, 1997 [257]	160/192 and 120	98/ND and 100	24/19.2 and 22.5	1.4 (0.8-2.3) and 1.1 (0.6-2.1)
Jourenkova, 1997 and 1998 [218, 261]	150/172 Caucasian hospital	100/100	18.0/15.7	1.2 (0.6-2.3)
Malats, 2000 [258]	122/121 mixed	31.1/21.7	26/36	0.6 (0.3-1.2)
Kiyohara, 2000 [245]	86/88 Japanese	68.6/45.5	54.6/44.3	2.0 (0.8-5.1)
Risch, 2001 [201]	389/353 Caucasian hospital	88.9/66	12.8/18.8	0.70 (0.45-1.09)
Stücker, 2002 [246]	251/264 Caucasian hospital	97/97	19.4/15.1	0.80 (0.4-1.3)
Lewis, 2002 [239]	94/165 Caucasian hospital	98/87	21.8/19.6	1.15 (0.60-2.21)
Wang, J 2003 [241]	112/119 Chinese	42.9/40.3	47.3/45.4	1.08 (0.63-1.83)
Nazar-Stewart, 2003 [243]	274/501 >95% Caucasian	96/72.2	19.0/18.0	1.07 (0.73-1.65)
Schneider, 2004 [244]	446/622 Caucasian hospital	89.3/75.4	16.8/18.5	0.88 (0.59-1.32)
Alexandrie, 2004 [262]	524/530 Caucasian	94.1/56.2	13.0/14.0	0.85 (0.5-1.4)
Sorensen, 2004 [217]	256/269 Caucasian	97.3/96.6	14.6/6	2.40 (1.31-4.41)

a borderline increased risk in *GSTT1* null carriers, OR 1.17 (95% CI, 0.98–1.40) [223]. Pooled analysis comprising 1929 cases and 1830 controls pointed to a somewhat higher cancer proneness for *GSTT1* null carriers, OR 1.25 (95% CI, 1.00–1.57) [223].

One of the early studies found that with high isothiocyanate consumption levels, the *GSTT1* null genotype was protective against lung cancer [263]. The protective effect was increased in combination with *GSTM1* null genotype. Conversely, when the isothiocyanate consumption was low *GSTT1* null genotype appeared to pose an increased risk of lung cancer [263, 264], and the risk was higher in combination with *GSTM1* null genotype. This is an example of how dietary factors can modulate the role of genes in cancer proneness. However, did not allow make any statement on the role of *GSTT1* in cancer susceptibility.

Microsomal epoxide hydrolase

Epoxides are organic three-membered oxygen compounds that arise from oxidative metabolism of endogenous, as well as xenobiotic compounds via chemical and enzymatic oxidation processes. Epoxide hydrolases catalyze the hydration of epoxides to their corresponding dihydrodiol products. In general, this hydration leads to the formation of more stable and less reactive intermediates. In humans, sytosolic (EPHX2) and microsomal (EPHX1) epoxide hydrolases participate in the metabolism of xenobiotics. EPHX1 substrates are generally quite specific for this hydrolase, with little or no activity being exhibited towards the other epoxide hydrolases [266].

EPHX1 is a smooth endoplasmic reticulum enzyme. It is expressed relatively ubiquitously in most tissues and in many species [267, 268]. The location of this enzyme on hepatic ER membrane is known to be similar to the P450 enzymes; the N-terminal region is anchored to the membrane and almost the entire protein molecule is located in the cytosol [269–271]. It is probable that CYP and EPHX1 enzymes cooperate via protein-protein interactions, meaning that a metabolite produced by CYP can be directly transferred to the other enzymes participating in the subsequent metabolism [272, 273].

Substrates

EPHX1 plays an important role both in the activation and detoxification of exogenous chemicals. It catalyzes the hydrolysis of reactive epoxides generated by CYP enzymes to *trans*-dihydrodiols which are more water-soluble dihydrodiol derivatives [266, 272]. In certain instances, the initial *trans*-dihydrodiol metabolites are further activated by CYP to form highly electrophilic and reactive dihydrodiol-epoxides that form covalent adducts with DNA [274]. Biotransformation of B[a]P is an example of that kind of reaction.

B[a]P is first oxidized by CYP enzymes and then further transformed by EPHX1 and CYP enzymes to BPDE, which is the ultimate carcinogen of B[a]P. 7,12-dimethylbenz [a]anthracene (DMBA) is a prototype for the PAH class of carcinogens. Similarly to BaP, DMBA is activated in conjunction with CYPs to the ultimate carcinogenic metabolite [275]. In experimental analysis, EPHX1 null mice were found to be highly resistant to DMBA-induced carcinogenesis, which indicates that EPHX1 may be a critical enzyme in the pathway leading to carcinogenic activation of PAHs [276].

Expression

The *EPHX1* gene is expressed in many human tissues. However, its activity levels are much lower in lymphocytes than in liver or lung [277]. In rabbits, a heterogenous expression pattern of EPHX1 among various cell types of lung have been observed [278]. Alternative promoters probably define the basis of tissue-specific expression for EPHX1 [279].

The *EPHX1* core promoter region has several putative transcription factor binding sites [280]. *EPHX1* steady-state expression levels are likely to reflect a variation in the interindividual response capacity [281]. The *EPHX1* constitutive expression is possibly regulated by CCAAT/enhancer binding protein alpha (C/BPalpha) interacting with DNA bound nuclear factor Y (NF-Y). The *EPHX1* core promoter region also contains a putative binding site for Nrf2 on the ARE that may be involved in the inducible expression of *EPHX1* by xenobiotics, possibly with a similar model of induction as takes place with the CYP enzymes [282, 283]. However, the mechanism of *EPHX1* expression is not completely defined. Prototypic chemical agents did not markedly perturb levels of mRNA for these enzymes.

EPHX1 in cancer proneness

EPHX1 activity has been shown to vary between individuals. Because of its participation in B[a]P activation and its strong expression levels in bronchial epithelial cells [277], the high activity EPHX1 variants of the *EPHX1* gene are thought to pose an increased lung cancer risk. However, the dual role of the EPHX1 enzyme in activation and detoxification makes its role in disease susceptibility more complex and the effect may depend on which of the many environmental agents the individual is exposed.

The *EPHX1* gene is located in chromosome 1q42.1 [284]. The gene contains nine exons, eight of which are coding [285]. The translated protein of 455 amino acids is the product of a single gene [286], although alternatively spliced non-coding regions of exon 1 have been reported [287].

A total of 111 SNPs spanning the *EPHX1* gene region are currently listed in the NCBI dbSNP database (http://www.ncbi.nlm.gov/SNP). Eight of these are non-synonumous changes, while most of the SNPs are located in the intron regions. The most common of

the validated non-synonymous changes is the T to C substitution located in exon 3, and resulting in a $Tyr^{113}His$ amino acid change (SNP; rs1051740). The frequency of the His^{113} minor allele is about 0.3 in Caucasians [288–290]. The second of the validated common SNPs is the $His^{139}Arg$ amino acid change in exon 4 (SNP; rs2234922) resulting in an A to G substitution. The frequency of the Arg^{139} minor allele is 0.2 in Caucasians [192]. The rest of the non-synonymous SNPs either have not been validated or have a very low frequency.

Of the synonymous changes, $Lys^{119}Lys$ (SNP; rs2292566), resulting from a G to A substitution, has a minor allele frequency of 0.12 in Caucasians (A allele) [291]. This SNP has not any functional consequences, but the close proximity to $Tyr^{113}His$ polymorphism has been shown to cause erroneous genotyping results for the $Tyr^{113}His$ locus with conventional PCR-RFLP -methods. A linkage between the codon 119 A allele and the Tyr^{113} allele has been observed [291, 292].

Functional consequences of the $Tyr^{113}His$ and $His^{139}Arg$ polymorphism have been widely studied in experimental settings [293–296]. In the first evaluation, the activity of four EPHX1 constructs was measured in microsomal S9 fraction [293]; the activity was greatest in the Tyr^{113}/Arg^{139} construct. The higher activity levels were based on the higher amount of enzyme. However, the RNA expression levels were almost identical in all of the constructs. This indicated that the stability of the enzyme variants were different and that the Tyr^{113}/Arg^{139} protein exhibits the highest enzyme activity. However, the most recent study measuring spesifically the EPHX1 enzyme activity, found the highest activity for the Tyr^{113}/His^{139} protein [295]. RNA and protein levels were not measured. Activity assays in the liver microsomal samples revealed no differencies in the activity of the different alleles [294, 295]. Predicted activity categories based on the results of the first study have been used in the evaluation of the role of EPHX1 activity on the disease proneness [293, 297]. The EPHX1 activity was measured in lymphocytes and the activity was highest in the putatively high activity group [296].

The association between lung cancer and *EPHX1* polymorphism has been studied in several studies (Table 6). Most of the studies that evaluated the $Tyr^{113}His$ polymorphism with the appropriate genotyping methods have found a decreased risk for His^{113} allele [288–290]. In most of the previous studies, however, the $Tyr^{113}His$ polymorphism has been analyzed with a method which is disturbed by the synonymous $Lys^{119}Lys$ polymorphism and thus they may have overestimated the numbers of homozygotes for His^{113} [292, 297–304]. This artefact may well have affected the outcome of a meta-analysis of seven published studies comprising of 2078 cases and 3081 controls that did not indicate an association for $Tyr^{113}His$ polymorphism; the results were based mostly on incorrect genotyping analyses [305]. A pooled analysis, however, found a decreased risk for homozygous His^{113} genotype (OR 0.70; 95%CI 0.51–0.96) [305].

Results from studies on the $His^{139}Arg$ polymorphism have been inconsistent (Table 6). Only three studies found a significant association between $His^{139}Arg$ polymorphism and lung cancer risk. Two of these studies indicated the Arg^{139} allele was a risk factor [303, 304]

Table 6
Studies on the *EPHX1* genotypes and lung cancer risk.

First author, year	Case/ Control (n) ethnicity	Ever smokers Case/ Control (%)	Mean Age Case/ Control	His^{113} allele freq Control	$Tyr^{113}His^{113}$ and $His^{113}His^{113}$ Case/Control freq. (%)	OR (95% CI) for $Tyr^{113}His$	Arg^{139} allele freq	$His^{139}Arg^{139}$ and $Arg^{139}Arg^{139}$ Case/Control freq. (%)	OR (95% CI) for $His^{139}Arg$	OR (95% CI) for EPHX1 activity
Persson, 1999 [310]	76/122 Chinese	26/ND	ND/ND	0.42	45/48 27/18	Tyr^{113} /His^{113} 1.1 (0.6-2.1) His^{113} /His^{113} 1.8 (0.8-3.9)	0.10	27/15 0/2	Arg^{139}allele-carrying 1.8 (0.9-3.6)	ND
London, 2000 [288]	182/458 Caucasian	95/66	64/63	0.28	45/40.2 8.2/8.1	Tyr^{113} /His^{113} 1.4 (0.9-2.1) His^{113} /His^{113} 1.0 (0.5-2.1)	0.19	27.5/29.7 3.8/4.4	His^{139}/Arg^{139} 0.9 (0.6-1.4) Arg^{139}/Arg^{139} 0.6 (0.2-1.8)	for high vs low 0.6 (0.4-1.2)
	155/242 Af-Am			0.21	31/31.8 0.6/5	Tyr^{113} /His^{113} 0.9 (0.6-1.4) His^{113} /His^{113} 0.1 (0.01-0.6)	0.29	45/43 10/8	His^{139}/Arg^{139} 1.1 (0.7-1.7) Arg^{139}/Arg^{139} 1.1 (0.5-2.5)	for very slow vs all others 0.1 (0.01-0.63)
To-Figueras, 2001[289]	176/187 Caucasian	98/100	60/50	0.31	40.0/45.5 4.6/8.0	Tyr^{113} /His^{113} 0.8 (0.5-1.3) His^{113} /His^{113} 0.4 (0.3-0.7)	0.18	30.3/28.9 1.7/3.7	His^{139}/Arg^{139} 1.1 (0.7-1.9) Arg^{139}/Arg^{139} 0.6 (0.3-0.9)	for high vs low 1.5 (0.9-2.4)
Gsur, 2003 [290]	277/496 Caucasian	89/59	64/63	0.32	41.2/43.9 5.8/10.9	Tyr^{113} /His^{113} 0.8 (0.5-1.1) His^{113} /His^{113} 0.4 (0.2-0.8)	0.18	28.9/29.0 4/3.2	His^{139}/Arg^{139} 1.0 (0.7-1.5) Arg^{139}/Arg^{139} 1.8 (0.8-4.4)	ND
Smith, 1997 [298]	50/203 Caucasian	100/ND	range 32-82/18-65	0.31	40/49 10/6	His^{113} /His^{113} 1.6 (0.6-4.8) His^{113}allele-carrying 1.0 (0.6-1.6)	0.15	32/26 2/2	Arg^{139}/Arg^{139} 1.4 (0.1-13.3) Arg^{139}allele-carrying 1.3 (0.7-2.3)	for very slow vs all others 1.9 (0.6-5.9)
Benhamou, 1998 [297]	150/172 Caucasian	100/100	58.4/55	0.40	30.7/44.8 14.7/18	Tyr^{113} /His^{113} 0.5 (0.3-0.8) His^{113} /His^{113} 0.5 (0.3-1.0)	0.15	35.3/28.5 2.0/1.2	Arg^{139}allele-carrying 1.2 (0.7-2.0)	for high vs low 2.7 (1.3-5.3)

Table 6 (contd)

EPHX1	Case/ Control (n) ethnicity	Ever smokers Case/ Control (%)	Mean Age Case/ Control	His^{113} allele freq Control	$Tyr^{113}His^{113}$ and $His^{113}His^{113}$ Case/Control freq. (%)	OR (95% CI) for $Tyr^{113}His$	Arg^{139}allele freq	$His^{139}Arg^{139}$ and $Arg^{139}Arg^{139}$ Case/Control freq. (%)	OR (95% CI) for $His^{139}Arg$	OR (95% CI) for EPHX1 activity
Lin, 2000 [311]										
Yoshikawa, 2000 [292]	71/107 Japanese	ND	ND	0.43	49/48 17/20	Tyr^{113} /His^{113} 1.0 (0.5-2.0) His^{113} /His^{113} 0.8 (0.4-2.0)	0.17	34/26 3/4	His^{139}/Arg^{139} 1.4 (0.7-2.8) Arg^{139}/Arg^{139} 0.7 (0.1-4.2)	ND
Yin, 2001 [299]	84/84 Chinese	53.6/ 53.6	60.3/ 60.9	0.44	64/55 18/17	Tyr^{113} /His^{113} 1.9 (0.9-4.0) His^{113} /His^{113} 1.7 (0.7-4.5)	0.10	21/17 1/1	His^{139}/Arg^{139} 1.4 (0.6-3.0) Arg^{139}/Arg^{139} 1.1 (0.1-17.3)	ND
Wu, 2001 [300]	60/76 Mex-Am	90/ 64.5	64.7/ 62.3	0.34	51.0/45.3 9.8/10.9	His^{113}allele-carrying 1.5 (0.6-3.7)	0.06	25.0/9.6 5.4/1.4	Arg^{139}allele-carrying 3.6 (1.3-10.4)	ND
	78/72 Af-Am	96.1/ 65.3	59.4/ 59.1	0.23	33.9/32.3 4.6/6.4	His^{113}allele-carrying 2.0 (0.8-5.2)	0.27	50.7/50.7 8.0/1.4	Arg^{139}allele-carrying 0.9 (0.4-2.2)	ND
Zhou, 2001 and 2002 [301, 302]	974/1142 >95% Caucasian	94/ 65	65.1/ 58.7	0.34	34.1/31.1 18.2/18	Tyr^{113} /His^{113} 1.2 (0.96-1.4) His^{113} /His^{113} 1.1 (0.9-1.4)	0.18	30.8/28.6 3.2/3.8	His^{139}/Arg^{139} 1.1 (0.9-1.3) Arg^{139}/Arg^{139} 0.9 (0.5-1.4)	for very slow vs all others 1.00 (0.7-1.3)
Zhao, 2002 [303]	181/163 Caucasian	92/ 90	63.3/ 61.	0.32	35/3x 12.3/14.4	His^{113}allele-carrying 0.9 (0.6-1.4)	0.18	37.3/29.3 6.0/3.8	Arg^{139}allele-carrying 1.6 (1.0-2.5)	for high vs low 1.7 (0.9-3.3)
Cajas-Salazar, 2003 [304]	110/119 Caucasian	100/ 100	59.9/ 57.3	0.26	34/44 5/4	Tyr^{113} /His^{113} 0.7 (0.4-1.2) His^{113} /His^{113} 1.1 (0.3-4.6)	0.18	37/32 5/2	His^{139}/Arg^{139} 1.3 (0.7-2.4) Arg^{139}/Arg^{139} 6.3 (1.0-38.3)	for high vs low 2.5 (1.1-5.7)

but one found it to be a protecting factor for lung cancer [289]. The results from meta- and pooled analysis did not indicate any association for His^{139}Arg polymorphism [305].

High predicted EPHX1 activity has been associated with increased lung cancer risk (Table 6) in some [288, 297, 304], but not in all studies [288, 289, 298, 301–303]. In a meta-analysis, the predicted EPHX1 activity levels was not associated with an increased risk of lung cancer [305]. Similar results were observed in a pooled analysis [305].

An interaction between the *EPHX1* polymorphism and cumulative smoking exposure has been observed; the higher activity genotypes, compared with the very low activity genotype, were protective against lung cancer in nonsmokers but represented a significant risk factor in heavy smokers [301]. An additive effect of *NAT2* and *EPHX1* genotypes in modifying lung cancer risk has also been observed [302], as well as a joint effect betwen the *GSTP1* and *EPHX1* genotypes; subjects who are homozygous for both *GSTP1*1* and *EPHX1* Tyr^{113} alleles have been shown to exhibit an increased risk of lung cancer (OR 2.34; 95%CI 1.21–4.52) [289].

An increase in the risk for upper aerodigestive tract cancers was observed for subjects with the Tyr^{113} variant in two studies [306, 307], while no significant association was found in another two studies [308, 309]. No association was found for the His^{139}Arg polymorphism [306–309].

The predicted high activity genotypes were significantly increased in oropharyngeal and laryngeal cancers and a positive interaction was found between the EPHX1 activity and *GSTM3* genotype in susceptibility to laryngeal cancer [307]. And increased risk in heavy smokers with predicted high activity phenotype was observed in one study [306]. But no association was found in two other studies [308, 309].

Validity of association studies

The basic approach used in an association study is straightforward; to test the involvement of a SNP in a specific condition, allele frequencies are compared in affected and un-affected individuals. Individuals to be associated with the genetic component should be matched for their exposure to environmental factors to enable proper assessment of the risk. However, the reproducibility of the association studies has been low. There is still much work to be done in optimizing epidemiological, statistical and laboratory approaches to achieve more credible outcomes [312].

Meta-analyses have many potential biases, including inconsistent outcomes from the individual studies [143, 206, 213, 313]. The potential sources of bias include; variation in the distribution of histological types of lung cancer among the studies, differences in the racial distribution between case and control groups among the studies and differences in the determination of genotypes (or phenotypes) among studies. Errors in genotype determination can lead to bias in the estimation of genotype effects and gene-environment

interactions [314]. Other potential sources of bias include variation in gender distribution among studies, variation in age or no data on the age provided in the studies as well as variation in exposure between studies.

In many cases, lifetime exposure to tobacco exposure is imprecisely quantified. Although the duration and intensity of smoking have been shown to influence risk, most molecular epidemiologic studies still classify tobacco exposure crudely as ´ever´ versus ´never´ or according to pack-years (PYs) smoked and e.g. not considering the important effects of cessation on the risk of tobacco-related cancers [32].

Variation in the source of (cases and) controls may affect the results. Also a publication bias is likely to occur. More valid and precise conclusions regarding a particular exposure-disease relation are expected in pooled analysis because of their use of common definitions, coding, cutpoints for variables and adjustment for the same confounders [213]. However, the factors mentioned above also influence pooled analysis.

Studies conducted thus far have been quite small. The need for larger studies is recognized, and larger individual studies have been initiated and consortiums of several institutions have been created to conduct collaborative studies. However, the undertaking use of these studies on a large scale is very expensive since the amount of SNPs in the genome is massive. Future association studies will involve genotyping of many markers and evaluation of many individuals. Large studies might, however, include population admixtures, which can lead to false positive results [315].

Quite often inadequate emphasis on *a priori* hypothesis is observed. In some early molecular epidemiology studies, there was a lack of careful consideration linking the candidates XME with the existence of specific environmental carcinogens [54]. Future studies should therefore focus on a pathway-based approach rather than focusing on a few genotypes; carcinogenesis is a multifactorial and multistage process and it is unlikely that a single gene have a major effect on phenotype [19]. The polygenic model for cancer susceptibility indicates that most of the genetic susceptibility to common cancers results from the combined effects of many genetic variants, each of which has a modest effect individually.

Future trends

At present it is known that the genome contains a huge amount of genetic variation. This gives us new opportunities but also challenges. The major challenges for the next few years are the evaluation of the functionality of coding variants and the role of the variation in the regulatory regions. The selection of the XMEs and the genetic polymorphisms that are most relevant for cancer molecular epidemiological studies also remain a major challenge [19].

Recently computational [316–318] and experimental approaches have been developed to select the target SNPs that most likely affect phenotypic differences [319–321]. The

advantages of combined genotyping and phenotyping assays should not be forgetten; new phenotyping assays have appeared [19], and molecular phenotyping may be very important for studying the expression of the human genome sequence [135].

One novel approach is to focus on haplotypes [135]. Haplotype is defined as a group of associated SNPs that do not segregate randomly. The number of haplotypes for a given gene is strongly correlated with the number of SNPs [322]. Even the haplotype blocks may contain a large number of SNPs and a few SNPs may be sufficient to uniquely identify the haplotypes in a block.

The specific SNPs that identify the haplotypes are called tag SNPs [323]. The haplotype map should be valuable in reducing the number of SNPs required to examine the entire genome for association with a phenotype from the 10 million SNPs that are known to exist to roughly 500,000 tag SNPs. This will make genome scan approaches to finding regions with genes that affect diseases much more efficient and comprehensive, since effort will not be wasted typing more SNPs than necessary and all regions of the genome can be included [323]. However, this can be make the biology more difficult to analyse since all the SNPs in the associated blocks need to be analyzed [135]. Haplotypes have inherent advantages when the contributing SNPs are not directly observed and when there is no additive phenotypic effect of consecutive SNPs. However, it is not known if the haplotypes offer a convenient way to define variation. Currently we do not know how large are the differencies in the haplotype structures between populations.

Functionality and validation of polymorphism reported in the databases must be evaluated before conducting future studies. It is clear that many of the SNPs reported in the SNP databases do not represent real polymorphism. A recent study on the SNPs in *GSTM* class genes failed to confirm many of the polymorphisms reported in the databases [197]. It is also unknown how many of the SNPs remain to be discovered and how much information is lost using the known SNPs [324].

Current approaches are based on the ´common disease: common variant´ hypothesis, in which common polymorphism or haplotypes are tested. It may, however, be possible that much of the cancer risk is due to rarer alleles. Virtually all susceptibility alleles identified so far have frequencies of less than 1%. These include both high-penetrance and low penetrance variants [324].

Ethical aspects

The completion of the Human Genome Project has drasmatically expanded the knowledge of our genome [325]. A great number of gene tests can be developed based information in the public domain. However, a discussion of the ethics in both scientific and public forums should be encouraged. As the human molecular epidemiology studies include genetic research among healthy people, a discussion of the ethical implications is necessary.

The declaration of Helsinki (www.wma.net) and other guidelines and recommendations, like UNESCOs `Universal Declaration on Human Genome and Human Rights´ (www.unesco.org), are guidelines for genetic research in humans. Informed consent, a individual written specific consent for a particular study, is obligatory in these studies. In Finland, the legislation also demands provision of informed consent (Medical Research Act No. 488/1999, www.finlex.fi). The consent form should include information of the background of the study, the purpose of the study, methods and procedures, potential general benefits as well as benefits for the person in question. It provides the individual with an unequivocal autonomy in his/her decision whether or not to participate in the study [326].

On the other hand, determination of the putative individual susceptibility could protect people from potential health risks and help them to seek early diagnosis, treatment and preventive therapy. Defining genetic susceptibility to environmental cancer may aid in the decisions about the level of acceptable daily intake -value (ADI) for a carcinogen and to protect also the most sensitive individuals in a population [326].

Predictive testing has different implications depending on the age of individual and the potential to be able to treat the disease [326]. As the intervention strategies may be limited, the psychological and social consequences of genetic-testing for later-onset diseases need to be discussed. Genetic discrimination, like the denial of health or life insurance or employment, based on individuals´s genetically-determined risk of developing serious diseases is a major threat [327–329]. Family members carrying a predisposing allele may also suffer from increased anxiety and depression from being made aware about this inherited trait [8]. Other limitations in the analysis of susceptibility are the poor predictive value of most tests. Currently, most methods are also still in the developmental phase and cannot be used in screening for individual susceptibility [326].

Presently, the genetic testing for lung cancer susceptibility does not alter patient management; at the moment psysicians recommed smoking cessation for all patients. Therefore genetic tests cannot be recommended before intervention methods are available [330].

AIMS OF THE PRESENT STUDY

This study was undertaken to study the role of polymorphic *GST* and *EPHX1* genes in relation to tobacco-induced cancer. Special emphasis was directed to:

- study the frequency of *GSTM1* and *GSTT1* gene deletions and *GSTP1* polymorphism in cancer-free Finnish Caucasian population in different age groups;
- study the role of *GSTM1* and *GSTT1* gene deletions and *GSTP1* polymorphism as a susceptibility factor for cancer with and without smoking exposure;
- examine whether the genetic polymorphisms of *GSTM1*, *GSTM3*, *GSTP1* and *GSTT1* genes modulate individual susceptibility to smoking-related lung and oropharyngeal cancers, either separately or in combination;
- examine whether the genetic polymorphisms of the *EPHX1* gene modulate individual susceptibility to smoking-related lung cancer alone or in combination with *GSTM1*, *GSTT1* and *GSTP1* genotypes.

MATERIALS AND METHODS

Study populations

Finnish study populations

Lung cancer patiens

The Finnish lung cancer patients analysed in Papers I, IV and V were admitted to Helsinki University Central Hospital during a 8-year period (1988–1997) for surgical pulmonectomy or lobectomy for suspected, operable cancer. All consecutive patients from two of the three surgical units of the Department of Thoracic and Cardiovascular Surgery were included. All the patients were Finnish Caucasians.

Both lung tissue and blood samples were collected from the patients and they were personally interviewed to obtain detailed smoking and occupational histories. Their probability of being occupationally exposed to asbestos was classified according to work history [331]. From the surgical lung specimens of the lung cancer patients, the histological type of lung tumour was classified according to the WHO classification. SCC accounted for 43.0%, AC for 35.6%, all other types present 21.4% of the cases. Details of the population are shown in Table I in Paper I.

Population controls

The Finnish population controls analysed in Paper I consisted of 294 healthy working-aged blood donors from the Finnish Red Cross Transfusion Service (Paper I, Table I). All the controls were males. To protect the privacy of the blood donors, only sex, age, and smoker/non-smoker data were available for them.

Finnish population controls

The cohort study (Paper III) and the population controls used in Papers IV and V consisted of 2155 non-cancer individuals (1069 men, 1086 women) from a survey conducted by the Social Insurance Institution of Finland, Research and Development Centre. The study subjects were recruited to the health examination from the random list, which was based on a population register. Altogether 3252 subjects were contacted, 66% of whom (n=2155) attended the examination. The population controls were all Finnish Caucasians without previous or current malignant disease. They had been living either in the city of Turku or in some rural and urban communities in southwestern Finland. The sampling aimed at roughly equal numbers of subjects of both gender in each of five age strata (27, 37, 47, 57 and 67 years old). Forty-nine subjects had been diagnosed with some form of malignant

(cancerous) disease. One subject was excluded because of missing basic information. After these exclusions the final study population consisted of 2105 subjects (1051 men, 1054 women). Blood samples were collected into EDTA tubes.

The participants were asked about their health habits and health status, including their smoking history and diseases that had been diagnosed by a physician. Smokers were classified into current, occasional, and ex-smokers. Detailed information on smoking was obtained only from the current smokers, defined as those who had smoked within the year prior to the study. This approach was chosen because it was thought that detailed information on cumulative smoking would not be reliably obtained from ex-smokers or those who had smoked only occasionally. Selected characteristics of the study population are shown in Table I in Papers III and IV.

French oral and pharyngeal cancer study population

Oral and pharyngeal cancer patients

Oral and pharyngeal cancer cases were analysed in Paper II. The study was drawn from a case-control study performed in France from 1988 to 1992. The cases were recruited in 10 private or public hospitals, nine of which are located in Paris. They were Caucasians with histologically confirmed squamous primary cancers of the oral cavity and pharynx.

All cases were regular smokers, defined as people having smoked at least five cigarettes (or cigars or pipes) per day for at least five years. Blood samples were available for 121 patients (95%) with cancers of the oral cavity of pharynx. They were recruited by seven trained study interviewers. Detailed information on tobacco use, alcohol consumption and occupational exposures was recorded during a personal interview. The main characteristics of the study population are shown in Table I in Paper II.

Hospital controls

The French control group comprised of eligible Caucasian patients without previous or current malignant disease. They were admitted into the same hospital and recruited according to the age, sex and hospital distributions observed in cases. The main medical diagnoses in the control population were rheumatological (33%), infectious and parasitic (10%), respiratory (9%), cardiovascular (8%), and digestive (6%). Only regular smokers were included.

The controls were interviewed by the same interviewers as the cases. Detailed information on tobacco use, alcohol consumption and occupational exposure was recorded. Selected characteristics of the study population are shown in Table I in Paper II.

Methods

Genotyping analysis

For genotyping analyses, the DNA was isolated from peripheral blood samples collected into EDTA tubes and stored at -20°C until use. One hundred ng of DNA was used as a template in all PCR-RFLP analysis and 30 ng in TaqMan analysis. To ensure laboratory quality control in genotyping tests, two independent individuals interpreted the results. Any sample with ambiguous results was re-tested, and a random selection of 10% of all samples was re-tested. No discrepancies were discovered upon replicate testing within a method used.

Genotyping of GST genes

The *GSTM1* genotypes were determined in Paper I using a PCR-based method as described earlier [208, 209], except that β-globin primers were added to the amplification reaction to produce the internal control fragment. Similarly, the *GSTT1* genotype was analysed as previously described [151], but the vitamin D receptor specific primers were included to produce an internal control fragment [332].

In Papers II, III and V, a multiplex PCR-method was used for determination of the *GSTM1* and *GSTT1* genotypes. Briefly, *GSTM1* and *GSTT1*-specific primer pairs were used together with a third primer pair for β-globin in a multiplex polymerase chain reaction (PCR) analysis. The absence of the *GSTM1* and/or *GSTT1*-specific PCR-product indicated the corresponding null genotype whereas the β-globin specific fragment confirmed proper functioning of the reaction [333, 334].

In the *GSTM3* genotyping analysis (Paper I and II), a restriction enzyme digestion with *Mnl*I was performed subsequent to the PCR reaction; the presence of the digestion site revealed the *GSTM3*B* variant allele [200].

The *GSTP1* genotyping analysis (Paper I, II, III and V) was performed according to a protocol developed by Dr. Chirstian Malaveille at IARC as described in detail in Paper I. Briefly, the variant alleles containing a base substitution at the nucleotide 313 (*GSTP1*2* and *GSTP1*3*) resulting in Ile^{105}Val amino acid change were differentiated from the wild type allele (*GSTP1*1*) by *Sna*BI restriction enzyme digestion subsequent to a PCR amplification. Ala^{114}Val polymorphism present only in the *GSTP1*3* allele, was differentiated from the wild type with *Bst*UI restriction digestion.

Genotyping of EPHX1 gene

Genotyping of Tyr^{113}His and His^{139}Arg polymorphisms using PCR-RFLP assay (Papers IV and V)

The *EPHX* exon 3 *Tyr^{113}His* polymorphism was detected for lung cancer cases by RFLP assay using the method published previously identifying the presence of *Asp*I site [335]. The His^{139}Arg polymorphism at exon 4 was identified through the presence of the *RsaI* restriction site [336].

Genotyping of Tyr^{113}His polymorphism using TaqMan assay (Papers IV and V)

All the samples in Paper VI and V were analysed for Tyr^{113}His polymorphism using a TaqMan allelic discrimination assay. Following PCR amplification, end-point fluorescence was read with Applied Biosystems 7700 instrument and genotypes were assigned using Allelic Discrimination software (Applied Biosystems SDS Software v1.7) [337].

Sequencing

From eight samples (Paper IV) the *EPHX1* gene exon 3 area was sequenced by a standard manner to evaluate the role Lys^{119}Lys polymorphism in the analysis of Tyr^{113}His polymorphism (Table 7). Fordward primer 5´-CTT GAG CTC TGT CCT TCC CAT CCC-3´ and reverse primer 5´-CTC TGG CTG GCG TTT TGC-3´ were used for PCR amplification. PCR reactions were run in an agarose gel and the DNA was extracted from the gel using the Gel Extraction Kit (QIAGEN). Sequencing reactions were undertaken using Big Dye® v3.1 Terminator Cycle Sequencing Kit (Applied Biosystems), 1.6 pmol of forward primer and 5 ng of DNA. Dye-terminators were removed using DyeEx Spin Kit (QIAGEN) and the reactions were purified in a standard manner using EtOH. Sequences were read in an automatic sequencer ABI PRISM 377 (Applied Biosystems).

Statistical analysis

In Paper I, the odds ratios (OR) and 95% confidence intervals (CI) were calculated using the two-sided Mantel-Haenszel method, in paper II by unconditional logistic regression using Statistical Analysis Software (SAS, Cary, NC, version 6.11), and in papers III to V by Statistical Package for Social Sciences (SPSS, version 11.5 and 12.0). The interactive effects between genotypes and smoking exposure (Papers II-V) or between genotype and age (Paper III) were assessed by the likelihood ratio test to compare the goodness of fit of the model with the interaction term to that of the model including the main effect variables and the adjusting variables mentioned within each paper. All presented *p*-values are two-sided.

When calculating the pack-years (PYs) (20g/day for one year = 1 PY) the daily consumption of each type of tobacco was expressed in g/day (1 g for cigarette, 2 g for cigar, and 3

g for pipe) [7]. The predicted EPHX1 activity categories were in Paper IV based on the classification established by Benhamou *et al.* [297]. In Paper V a new categorization was used which was based on a recent functional analysis by Hosagrahara *et al.* [295]. Due to their small numbers, the *GSTM3 *A/*B* and **B/*B* genotypes, as well as the *GSTP1*1/*2* and **2/*2* genotypes, were combined in most analyses. Also *EPHX1 Tyr^{113}/His^{113}* and *His^{113}/His^{113}* genotypes and the *EPHX1 His^{139}/Arg^{139}* and *Arg^{139}/Arg^{139}* genotypes were combined in the analyses to increase statistical power.

In Paper II, the interactive effects between either any two *GST* genotypes or each *GST* genotype were studied. Sex, age (<50, 50–54, 55–59, 60–64, 65+), daily consumption of tobacco in g/day (≤ 20, 21–30, 31+), duration of smoking in years (≤25, 26–34, 35+), smoking status (former/current smoking), exlusive cigarette smokers (no/yes), daily consumption of alcohol in g/day (≤ 40, 41–80, 81–120, 121+) were used as adjusting variables in the analysis. The cut-off points were defined so that sufficient numbers of individuals were included in each subgroup. Those values for duration of smoking were defined according to approximate tertiles in the control population.

In paper III, chi-square test and logistic regression were used to compare the distributions of *GSTM1, GSTP1* and *GSTT1* genotypes between the age groups. When the concurrent effect of all three *GST* genes was examined, the combination of the most favorable genotypes of these genes was compared to all other genotypes. The effect of PYs was evaluated only for the two oldest age categories in which a meaningful portion of the subjects could be expected to have a cumulative smoking dose exceeding the cut off point of 40 PYs.

In Paper IV, risk estimates for the association between *EPHX1* polymorphisms and lung cancer were adjusted for age, sex, and smoking habits (never/ex/current smoker). Those analyses restricted to current smokers were additionally adjusted for PY.

A possible joint effect of *EPHX1* and *GST* genotypes in lung cancer proneness (Paper V) were examined calculating multivariate adjusted ORs for the combinations of one and two at-risk genotypes or diplotypes. The reference group consisted of individuals with the putatively most advantageous combinations of genotypes.

RESULTS

Characteristics of the study populations

Finnish study populations

The Finnish lung cancer studies included 230 lung cancer cases (Paper I, IV and V), 294 healthy working-aged blood donors to the Finnish Red Cross Blood Service (Paper I) and 2105 (Paper IV-V) cancer-free population controls living in southwestern Finland. Information from the control population studied in paper I was restricted to age (mean 41, SD 11), sex (all men) and smoker/non-smoker status (23%/77%) (Table 7).

The distribution of men and women was 79.1% and 20.9%, and 49.9% and 50.1%, in the cancer group and in the cancer-free population control group, respectively. The mean age among lung cancer cases was 61.4 (SD 10.2, range 18-81) and among controls 47 (SD 14.0, range 27-67). Current smoking was more common among cases (65.2%) than among controls (29.4%). PYs were calculated only for current smokers because the smoking habits of ex-smokers were not collected for controls. Among the cases, the mean PYs were 45.0 ± 20.3 and among controls 21.4 ± 19.4. Fourtythree percent of the lung cancer patients were diagnosed with SCC, 35.6% with AC, 5.7% with large cell carcinoma, 5.7% with small cell carcinoma, and 10% with other lung carcinomas.

More detailed characteristics of the Finnish cancer-free population controls are given in Paper III (Table I). Current smoking was more common among men (27.7%) than among women (18.4%). The prevalence of current smokers decreased with age in both men and women; in the youngest age group (27 years of age) 37.2% of men and 28.6% of women were current smokers, whereas in the oldest age group (67 years of age) the respective percentages (14.4% and 6.4%) were considerably lower. At the same time, 51.5% of 67 year old men were ex-smokers. This effect was not seen among women; only 8.3% of 67 year olds were ex-smokers. In the 67 year olds the current daily tobacco consumption was lower in both sexes than in the younger age groups. This also affected the calculated PYs, which were lower in the 67 year olds than in the 57 year olds.

French oral and pharyngeal cancer study population

Among the cancer patients 55% had oral cavity cancer, 41% had pharyngeal cancer and 3% had some unspecified or unclassifiable cancer of the oral cavity or pharynx. The mean age was similar in cases (54.4, SD 10.2) and controls (54.9, SD 11.1). Most of the subjects were male in both groups (93% of cases and 95% of controls). The smoking and alcohol related parametres were also very similar in both groups; parametres are defined in greater detail in table 1 in Paper II.

GSTs and cancer proneness

Distribution of *GSTM1*, *GSTP1* and *GSTT1* genotypes in the population controls

The overall prevalences of the *GSTM1* and *GSTT1* null genotypes in the study population were 47.5% and 12.4%, respectively (Table 7 and Paper III, Table 3). The distribution of *GSTM1* and *GSTT1* genotypes did not differ between men and women; the *GSTM1* gene was lacking from 47.7% of men and from 47.3% of women (p=0.844). The respective numbers for the *GSTT1* null genotype were 12.5% and 12.3% (p=0.888). There were no statistically significant overall age-related deviations in the distribution of the *GSTM1* and *GSTT1* genotypes, but when the sexes were considered separately, the *GSTM1* genotype distribution appeared to differ significantly between the different age groups in men (χ^2=14.20; df=4; p=0.007).

The frequencies of *GSTP1*A* and *GSTP1*B* alleles were 71.8% and 28.2%; the frequency of the *GSTP1*A/*A*, *GSTP1*A/*B* and *GSTP1*B/*B* genotypes was 51.9%, 39.8% and 8.3% respectively in the total study population (Table 7). These frequencies were in agreement with those predicted under Hardy-Weinberg equilibrium (p=0.44). No differences were seen in the distribution of *GSTP1* genotypes, either in men and women (p=0.491) nor were the frequencies different between the age groups in the whole study population (χ^2=8.20; df=8; p=0.414).

When the genotype frequencies were examined by smoking history, no differences were seen in the distributions of *GSTM1* and *GSTT1* genotypes in the different age groups. However, for the *GSTM1* null genotype, a significant interaction (p=0.003) was found between gender and age among never smokers. Since a *GSTM1* positive genotype has been shown to be over-represented in old smoking controls [338], we compared the prevalence of *GSTM1* positivity in the highest age group (67 years old) with that in the younger age group (27–57 years old). In this comparison, current smokers in the oldest age group tended more likely to exhibit the functional *GSTM1* gene (OR 1.34, 95%CI 0.71–2.56).

The *GSTP1* genotype frequencies were significantly different between the age groups in current smokers (χ^2=17.08; df=8; p=0.029). A significant interaction was found between gender and genotype (p=0.029). When stratified by gender, a significant deviation in the distribution was observed in women (χ^2=17.50; df=8; p=0.025). Current smoking men tended to be less likely (OR 0.57, 95%CI 0.31–1.03), whereas current smoking women were more likely (OR 1.70, 95%CI 0.97–2.97) to be homozygotes for the *GSTP1*B* allele when compared with never smokers.

In the oldest age group (67 years old) the *GSTP1*B* allele carriers were more prevalent among current-smokers, as compared with the never-smokers (OR 3.84, 95%CI 1.47–10.05). This finding appeared to be mainly confined to women (p for interaction between genotype and gender 0.017). However, due to the small numbers of subjects, this association could not be evaluated in detail.

We did not find any associations between tobacco smoking and the *GSTT1* genotype distributions in the different age groups (data not shown). However, current smoking women were more likely to lack the *GSTT1* gene compared to the never smokers (OR 0.55, 95%CI 0.35–0.87), whereas no such association was seen in men (OR 1.04, 95%CI 0.66–1.66) (data not shown).

When the combined genotype distributions were examined, no overall difference was observed for the *GSTM1/GSTT1* null genotype combination between the different age groups (χ^2=10.81; df=12; p=0.545) (data not shown). However, there was a significant interaction between gender and age among never smokers (p=0.045). When this was stratified by gender, the distribution was significantly different in men (χ^2=25.89; df=12; p=0.011). This finding was mainly attributable to never smokers (χ^2=21.71; df=12; p=0.041). Moreover, when the so-called *GSTM1*, *GSTT1* and *GSTP1* at-risk genotypes were combined, the distributions did not differ between the different age groups (χ^2=23.08; df=28; p=0.729), but again, a significant interaction was found between gender and age among never smokers (p=0.013).

Unfortunately, due to the small number of subjects in each category, we were not able to analyse the distribution of combined genotypes according to smoking history across the age groups.

When the data was stratified by PYs, according to our *a priori* hypothesis the putative protective *GST* genotypes were over-represented in subjects who had smoked over 40 PYs compared to lighter smokers. The difference was greater when two or three protective

Table 7. Characteristics of the control populations and distribution of *GST* polymorphism.

Study population	Size	Sex men/ women	Mean age	Smoking status (%)			GST				
							M1 null	T1 null	P1 *1/*1	P1 *1/*2	P1 *2/*2
	(n)	(%)	(SD)	never	ex	current	(%)	(%)	(%)	(%)	(%)
Finnish (Paper I)	294	100/-	41 (11)	78		22	46.6	13.3	52.6	38.2	9.2
Finnish (Paper III-V)	2105	50/50	47 (14)	47	24	29	47.5	12.4	51.9	39.8	8.3
French (Paper II)	150	95/5	55 (11)	-	33	67	52.3	15.7	50.0	37.2	12.8

genotypes were combined. This finding was more pronounced and statistically significant in subjects concurrently carrying the *GSTM1* and *GSTT1* positive genotypes (OR 2.50, 95%CI 1.12–5.58), *GSTT1* positive and *GSTP1*A/*A* genotypes (OR 2.44, 95%CI 1.09–5.45), or all the three putative protective *GST* genotypes (OR 2.80, 95%CI 1.10–7.12), when compared with all the other genotype combinations of the respective genes.

GST genotypes and lung cancer risk

The prevalence of the *GSTM1* null genotype was similar in lung cancer cases (48.1%) and population controls (46.6%)(OR 1.06; 95%CI 0.74–1.51) (Table 8). When the cases were considered according to cancer type, the *GSTM1* null genotype was slightly over-represented in the SCC patients (53.2%)(OR 1.30; 95%CI 0.82–2.07) and somewhat under-represented in AC patients (41.5%)(OR 0.81; 95%CI 0.49–1.33) compared to controls. However, the differencies were not statistically significant. Lung cancer risk associated with heterozygous and homozygous *GSTM3*B* carriers was 1.32 (95%CI 0.86–2.02) and 0.66 (0.17–2.61), respectively, when compared to carriers of the homozygous wild type *GSTM3* gene. The prevalence of the *GSTM3*B* allele carrying genotypes was quite similar in cases and controls (OR 1.25; 95%CI 0.83–1.89) when compared to the *GSTM3*A* carriers.

The prevalence of *GSTT1* null lung cancer cases (13.3%) was comparable to that found in the healthy controls (12.7%)(OR 0.96 95%CI 0.56–1.63). Similarly to the findings for *GSTM1*, a non-significant over-representation of the *GSTT1* null genotype was observed among patients with SCC (16.5%) (OR 1.29; 95%CI 0.67–2.47).

OR for the carriers of heterozygous and homozygous variant *GSTP1* genotypes was 1.10 (95%CI 0.76–1.59) and 0.96 (95%CI 0.50–1.85), respectively, when compared to homozygous *GSTP1*1* carriers. For carriers of at least one *GSTP1* variant allele it was 1.15 (95%CI 0.81–1.65).

When the joint effect of all *GST* genes was examined, those subjects simultaneously deficient in *GSTM1* and *GSTT1* were not an overall incresed lung cancer risk (OR 1.49; 95%CI 0.70–3.16) but showed a borderline increased risk for SCC (OR 2.35; 95%CI 0.98–5.65) when compared to *GSTM1* and *GSTT1* positive subjects. No other gene combinations exhibited any interactive effects (data not shown).

GST genotypes and oral and pharyngeal cancer risk

No association between *GSTM1* null genotypes and oral or pharyngeal cancers were observed (Table 9) (OR 0.8; 95%CI 0.4–1.4 and OR 1.3; 95%CI 0.6–2.6, respectively). Furthermore, neither of the variant *GSTM3* genotypes were also nor associated for these cancers; OR for *GSTM3*A/*A* genotypes compared to genotypes carrying *GSTM3*B* allele was 1.4 (95%CI 0.7–3.0) for oral cancer and 1.0 (95%CI 0.5–2.2) for pharyngeal cancer. Smoking and alcohol exposure did not modify the cancer risk.

Table 8. Associations between *GST* genotypes and lung cancer risk.

Gene	Genotype	Cancer type	Cases n (%)	Controls n (%)	OR	(95 % CI)
GSTM1	positive	All lung	108 (51.9)	157 (53.4)	1	
	null		100 (48.1)	137 (46.6)	1.06	(0.74-1.51)
	positive	SCC	44 (46.8)	157 (53.4)	1	
	null		50 (53.2)	137 (46.6)	1.30	(0.82-2.07)
	positive	AC	48 (58.5)	157 (53.4)	1	
	null		34 (41.5)	137 (46.6)	0.81	(0.49-1.33)
GSTM3	*A/*A	All lung	146 (72.6)	226 (76.9)	1	
	*A/*B		52 (25.9)	61 (20.7)	1.32	(0.86-2.02)
	*B/*B		3 (1.5)	7 (2.4)	0.66	(0.17-2.61)
	*B-carrying		55 (27.4)	68 (23.1)	1.25	(0.83-1.89)
GSTP1	*1/*1	All lung	101 (49.0)	154 (50.8)	1	
	*1/*2		88 (42.7)	122 (40.3)	1.10	(0.76-1.59)
	*2/*2		17 (8.3)	27 (8.9)	0.96	(0.50-1.85)
	*2-carrying		105 (51.0)	139 (45.9)	1.15	(0.81-1.65)
GSTT1	positive	All lung	178 (87.3)	255 (86.7)	1	
	null		26 (12.7)	39 (13.3)	0.96	(0.56-1.63)
	positive	SCC	76 (83.5)	255 (86.7)	1	
	null		15 (16.5)	39 (13.3)	1.29	(0.67-2.47)
GSTM1 and *GSTT1*	positive	All lung	96 (85.7)	134 (89.9)	1	
	null		16 (14.3)	15 (10.1)	1.49	(0.70-3.16)
GSTM1 and *GSTT1*	positive	SCC	38 (79.2)	134 (89.9)	1	
	null		10 (20.8)	15 (10.1)	2.35	(0.98-5.65)

The frequency of the *GSTT1* null genotype was higher in cancer patients (21.5%) than in controls (15.7%); OR for the oral and pharyngeal cancer was 2.0 (95%CI 1.0–4.0). In a separate analysis according to the cancer site, an increased risk was seen for oral cancer (OR 2.4; 95%CI 1.0–5.5) but not for pharyngeal cancer (OR 1.4; 95%CI 0.6–3.7). When examined by smoking exposure, the risk was increased only among smokers with a history of more than 30 years of smoking (OR 3.3; 95%CI 1.3–8.1) but not among smokers with a shorter history (OR 0.8; 95%CI 0.3–2.6). No interactions were observed between alcohol consumption and *GSTT1* genotypes (data not shown).

The frequency of the *GSTP1* (**1/*2* or **2/*2*) genotype was higher in cancer cases (58.7%) than in controls (50.0%). A borderline increased risk was seen for *GSTP1* (**1/*2* or **2/*2*) genotypes associated with oral and pharyngeal cancers (OR 1.6; 95%CI 1.0–2.8). There was an increased cancer risk (OR 2.0; 95%CI 1.0–3.9) in long time smokers (over 30 years) associated with *GSTP1* variant genotypes though this was not observed in smokers with a shorter history (OR 1.0; 95%CI 0.5–2.3). No interactions were observed between alcohol consumption and *GSTP1* (data not shown).

No significant interactive effects between different *GST* genotypes on the risk of cancers of the oral cavity and pharynx were observed (data not shown). Nor was any association found within the control group between genotypes and the main medical diagnoses (data not shown).

EPHX1 and cancer proneness

EPHX1 genotyping methodology

In an early stage of our studies we observed that the genotype results of the *EPHX1* $Tyr^{113}His$ locus obtained with the PCR-RFLP did not fully match with these from the TaqMan assay. To clarify the reason for these inconsistencies we sequenced eight samples that had previously been genotyped with the two methods. The results from the different analysis are shown in Table 10. Four of those samples detected in the RFLP analysis as homozygotes for the *EPHX1* His^{113} allele, were actually heterozygotes when sequenced and genotyped using the TaqMan method. These subjects were heterozygotes also for codon 119 polymorphism. One subject was a true homozygote for the *EPHX1* His^{113} allele and was also a homozygote for G at codon 119. In the other analysis, the results were similar in all analyses, *i.e.*, in RFLP, sequencing and TaqMan analysis.

When all lung cancer cases were analysed using the RFLP method (Paper IV), a total of 22 individuals were determined as homozygotes for the *EPHX1* His^{113} allele (9.7% of all cases). In the TaqMan assay 13 of these were confirmed to be "true" *EPHX1* His^{113} homozygotes (5.7% of all cases). The silent polymorphism of G to A at codon 119 ($Lys^{119}Lys$) was therefore shown to disturb the PCR analysis. The allele with A at codon 119 is not amplified in the presence of G allele, because the primers are more likely to bind to the sequence with G. Since A is in linkage with the Tyr^{113} (T), this leads to false genotype designation only in the case of His^{113}/His^{113}.

It appears that the substitution of G to A at codon 119 clearly disturbs the amplification of the *EPHX1* Tyr^{113} allele in heterozygotes also with the primers designed by Lancaster *et al.* [292, 339]. Thus it seems evident that the results obtained with the traditional PCR-RFLP assay for the *EPHX1* $Tyr^{113}His$ polymorphism need to be verified by an appropriate method taking into account the silent polymorphism in codon 119.

Table 9. The *EPHX1* exon 3 genotype results achieved by different analysis methods.

	PCR-RFLP	TaqMan	Sequencing	
			codon 113	codon 119
	Amino acid change		Nucleotide change	
Sample	Tyr^{113} to His^{113}		T to C	G to A
1	Tyr^{113}/Tyr^{113}	Tyr^{113}/Tyr^{113}	T	G
2	Tyr^{113}/Tyr^{113}	Tyr^{113}/Tyr^{113}	T	G and A
3	Tyr^{113}/His^{113}	Tyr^{113}/His^{113}	T and C	G
4	His^{113}/His^{113}	**Tyr^{113}**/His^{113}	**T** and C	G and **A**
5	His^{113}/His^{113}	**Tyr^{113}**/His^{113}	**T** and C	G and **A**
6	His^{113}/His^{113}	**Tyr^{113}**/His^{113}	**T** and C	G and **A**
7	His^{113}/His^{113}	**Tyr^{113}**/His^{113}	**T** and C	G and **A**
8	His^{113}/His^{113}	His^{113}/His^{113}	C	G

Distribution of *EPHX1* genotypes in the population controls

The *EPHX1* genotype frequencies in the control population were in agreement with those predicted from Hardy-Weinberg equilibrium (p=0.71 for both genotypes). The frequencies of *EPHX1* His^{113} and Arg^{139} alleles were 0.298 and 0.176, respectively. We did not observe any age-related changes in the *EPHX1* $Tyr^{113}His$ genotype distributions (p=0.988). In contrast, the distributions of *EPHX1* $His^{139}Arg$ genotypes in the different age groups were significantly different (p=0.034). However, there was no significant trend and when the rare homozygotes were combined with heterozygotes, this difference disappeared (p=0.127).

EPHX1 genotypes and lung cancer risk

The *EPHX1* His^{113} allele-carrying genotypes tended to pose a decreased lung cancer risk; the OR was 0.68 (95% CI, 0.49–0.94) when compared with the homozygous wild-type *EPHX1* Tyr^{113}/Tyr^{113} genotype. When considering only current smokers the cancer risk

associated with the *EPHX1* His^{113} allele-carrying genotypes was 0.59 (95% CI, 0.37–0.95) compared with the homozygous wild-type genotype. The results obtained for the exon 4 polymorphism conversely revealed a decreased risk for the putative fast *EPHX1* Arg^{139} allele-containing genotypes; the OR was 0.64 (95% CI, 0.43–0.93) for the variant *EPHX1* Arg^{139} allele-containing genotypes compared with the homozygous wild-type *EPHX1* His^{139}/His^{139} genotype.

When the *EPHX1* exon 3 and 4 genotype data was combined to predicted low, intermediate and high activity categories as proposed by Benhamou *et al.,* [297], a tendency for decreased risk was seen for both low and high activity phenotypes compared to the intermediate activity phenotype (OR, 0.75; 95% CI, 0.53–1.07 and OR, 0.65; 95% CI, 0.38–1.19, respectively). However, when the *EPHX1* genotypes were grouped to diplotypes based on the recent functional analysis which suggested that the highest activity was attributableto the wild type *EPHX1* $Tyr^{113}His^{139}$ haplotype [295], then the homozygous wildtype *EPHX1* $Tyr^{113}His^{139}$ haplotype was associated a doubled lung cancer risk (OR, 1.95; 95% CI, 1.22–3.12).

The age did not modify the association between variant genotypes and lung cancer risk (data not shown). Nor could we detect any interaction between smoking and the *EPHX1* genotypes.

Combined effect of GSTs and EPHX1 on lung cancer proneness

In Paper V, the *EPHX1* haplotypes were assessed from the genotype results and grouped to diplotypes. When grouped based on the recent functional analysis the homozygous wild type diplotype represent the highest activity group and conversely homozygous variant diplotypes are the lowest activity group [295].

When the combined genotype effects were examined (Table 11), the concurrent presence of the *EPHX1* wildtype diplotype and *GSTM1* null genotype possed an OR of 2.48 (95% CI, 1.29–4.77). The risk did not change when only current smokers were considered (OR 2.30; 95%CI, 0.94–5.62). However, individuals having smoked over 40 PYs were at a clearly elevated risk of lung cancer (OR, 8.89; 95% CI, 1.75–45.13), especially of SCC (OR, 13.33; 95% CI, 1.47–120.71).

The *EPHX1* diplotypes and *GSTT1* genotypes exhibited a significant interaction (p=0.022 for interaction term). The highest risk was seen for carriers of the homozygous wildtype *EPHX1* diplotype in combination with *GSTT1* positive genotype (OR, 2.47; 95% CI, 1.47–4.13). The risk was only moderately increased in current smokers (OR, 2.56; 95% CI, 1.25–5.27).

The *GSTP1* genotypes did not interact with *EPHX1* diplotypes, and when the joint effects were studied, no additional risk compared to *EPHX1* diplotypes alone were seen.

The age did not significantly affect the risk between variant genotypes and lung cancer. When age-related differences were examined in our control population, no deviations in genotype distributions between different age groups were seen either for *EPHX1* genotypes or diplotypes or for *GSTM1*, *GSTT1* and *GSTP1* genotypes (data not shown). No gene-environment interactions between combined genotypes and smoking related variables were found (data not shown).

DISCUSSION

Cancer is a global disease having an incidence of about 10 million new cases worldwide. Tobacco-related cancers comprise 30% of all malignant tumours of which lung cancer is globally the most common and a leading cause of cancer-related death. Lung cancer is largely due to exposure to the chemical carcinogens in tobacco smoke. Around 15% of all heavy smokers develop lung cancer. Cancer risk is either increased or decreased by genetic risk factors as well as by several other environmental factors [8].

Living beings are faced with a continuous attack of reactive chemical species by endogenous and exogenous agents. BRIs modify the chemical integrity of DNA and in turn change its informational content. Tobacco exposes the individual to a mixture of compounds, some of which are established carcinogens, and these are thought to be involved in cancer initiation as well as non-genotoxic carcinogens that facilitate carcinogenesis by stimulating cell division. Other environmental and nutritional factors may modulate responses to tobacco smoke via the genetic host factors and thus the risk of neoplasm formation. In this thesis the role of polymorphic *GST*s and *EPHX1* genes in individual susceptibility to lung cancer was explored in Finnish and French study populations.

GSTs

General

GSTs are detoxification enzymes involved in the metabolism of tobacco derived BRIs. They inactivate chemical carcinogens into less toxic or inactive metabolites, and are thought to have an essential role in chemical carcinogenesis by reducing their capacity to inflickt DNA damage [340]. DNA adduct formation is generally considered to be necessary for tumour formation [341]. There is evidence that PAH-induced DNA damage is one of the causes of lung cancer which derives from studies showing higher-levels of PAH-adducts in cancer cases than controls [57]. Metabolic modulation without reducing levels of chemical exposure has been shown to reduce the formation of DNA adducts and cancer risk [342].

GSTM1

Epoxides of PAHs like BPDE are substrates for GSTM1 [184], which is probably the most studied of the *GST* genes in disease susceptibility. Since, it is believed that deletion of the entire gene will completely eliminate the enzyme activity, this is thought to have significant effects on carcinogen metabolism. Consequently the *GSTM1* polymorphism has been extensively studied in lung [213], head and neck [223, 343], colon [344, 345], gastric [346], breast [347–349], prostate [350], and bladder cancers [351, 352]. Associa-

tions have been found in several tobacco-related cancers like lung, head and neck, and bladder cancers [213, 223, 343, 352]. In our study a tendency of increased risk was seen in lung SCC patients having *GSTM1* null genotype. This finding agrees with most studies, which show a modest effect of *GSTM1* null on lung cancer proneness [213]. However, no associations between *GSTM1* genotypes and oropharyngeal cancer risk were observed.

Although the *GSTM1* null genotype was only weakly associated with lung cancer risk in our study, multiple lines of evidence from molecular epidemiological studies do indicate that *GSTM1* is involved in cancer susceptibility. The *GSTM1* genotypes have been related with intermediate biomarkers of exposure; BPDE-DNA adducts were not found in *GSTM1* positive individuals exposed to PAHs, while 93% of the *GSTM1* null individuals showed detectable adducts [182]. Also early phenotyping studies on TSO, a specific substrate of GSTM1, indicated that fewer smoking related cancers occurred in the group exhibiting GSTM1 activity [353]. The *GSTM1* expression has also been shown to be lower in cancer cases than in controls [354].

Our novel study design employing a large control group in different age groups revealed some evidence of the role of *GSTM1* null genotype in cancer proneness. Our hypothesis was that the advantageous genotypes are over-represented in older controls compared to younger controls; other genotypes are suspected to have been eliminated from the study population because of tumour formation or death.

In this study setting, the overall prevalence of the *GSTM1* null (47.5%) genotype was in good agreement with previous observations in Finns [355] and other Caucasian populations [192]. However, when the age-related differences in genotype distribution was studied significant deviations were observed among men. The *GSTM1* null genotype appeared to be somewhat less prevalent among the oldest (67 years old) current male smokers (38.1%) as compared with the younger (27–57 years old) current smokers (47.3%) (p=0.260). Although failing to reach statistical significance, this finding agrees with some previous observations of lower prevalence of *GSTM1* null genotypes among older controls [338, 356, 357]. More *GSTM1* null individuals have also been observed in never smokers compared to ever smokers [238]. Since the prevalence of *GSTM1* genotypes may be significantly affected by gender, age and smoking status [358] it is very important to control for these factores in the study design.

GSTM3

The *GSTM3* gene has been less frequently studied in association studies. GSTM3 enzyme is detected in human lungs and the low expression levels of *GSTM3* have been suggested to result in an increased risk of lung malignancies. In our study, however, the *GSTM3* genotypes were not associated with lung cancer risk, a finding which is in agreement with other studies [201, 217, 218]. Furthermore, it was not associated with oral or pharyngeal cancer risk. There is limited previous evidence of a role for *GSTM3* genotypes in UAT

cancers. Decreased [203, 224, 225] and increased [226] risk of UAT cancers associated with *GSTM3*B/B* genotype have been observed.

Although it is possible that the influence of genotypes may only be exerted in certain regions of the upper aerodigestive tract, the function of the *GSTM3* polymorphism studied is not clear but a recent finding of promoter region polymorphism [199] affecting significantly on expression levels may explain the negative findings. The *GSTM3* gene is still a promising susceptibility gene in cancers of the aerodigestive tract and it is interesting to see the results from future studies taking into account other polymorphic sites.

GSTP1

GSTP1 is known to metabolize BPDE, a highly carcinogenic metabolite of B[a]P. In addition, the high expression level in lung has made it particularly interesting as a susceptible factor for cancers of this site. The association of the common polymorphism ($Ile^{105}Val$) and lung and head and neck cancers has been extensively studied. The other polymorphism ($Ala^{114}Val$) has only been studied to some extent. Functionally both alleles in the $Ile^{105}Val$ locus seem to have some advantages. Compared with the *GSTP1*1* (Ile^{105}) allele, the *GSTP1*2* (Val^{105}) allele exhibits a decreased activity towards CDNB but a greater activity towards PAH diolepoxides [229, 230, 232].

In our case-control study we did not see any effect in lung cancer risk for the *GSTP1* polymorphism, when the genes were studied separately. This is in agreement with most of the previous studies which have not detected any significant associations [201, 217, 218, 236–245]. However, a few studies have indicated that the *GSTP1*2* allele is a risk factor for lung cancer [210, 246, 247]. The most extensive study thus far on *GSTP1* and lung cancer proneness found an increased risk for current smokers with the risk being increased with increasing PYs [247]. This supports our finding in the control population, showing that *GSTP1* may interact with smoking. Also in our study of oropharyngeal cancers, the variant *GSTP1* genotypes carried an increased cancer risk and the effect was higher in long term smokers. However, meta- and pooled analysis on HNC did not find any increased risk in *GSTP1*2* carriers [223].

In our control study setting smoking appeared to have an effect on the distribution of the *GSTP1* genotypes; the *GSTP1* genotype distribution deviated significantly between the different age-groups in current smokers. Moreover, a significant gender-related effect was observed; current smoking men tended to be less likely and current smoking women more likely homozygotes for the *GSTP1*2* allele when compared with never smokers. There is one previous observation that current smoking controls are less likely to have the *GSTP1*2* allele than never smoking controls [238].

None of the *GSTP1* genotypes evaluated have shown have any major effect on the function of the protein, and no clear evidence for different expression levels have been observed [227]. However, a significant association between the *GSTP1*1/*1* genotypes and

increased overall *GSTP1* expression, and the *GSTP1*1/*2* genotypes and reduced overall *GSTP1* expression has been observed in clinical cancer [359]. It is also speculated that the heterozygote *GSTP1*1/*2* genotypes are associated with improved outcome because the enzyme expressed may provide a better balance between the effects of detoxification of carcinogens and the ability to metabolize chemoprevention agents [359].

As many genetic variants have been described to be present in the *GSTP1* gene region it may be possible to identify those SNPs which may affect the expression levels of *GSTP1*. As the gene is divided into the two linkage blocks it would be advantageous to study one hypothetical SNP in regulatory regions in addition to the $Ile^{105}Val$ polymorphism. Since the $Ala^{114}Val$ polymorphism has been shown to be in linkage disequilibrium with $Ile^{105}Val$ polymorphism, and the SNP has not been shown to affect functionality, it is not reasonable to include this SNP in candidate gene studies.

GSTT1

GSTT1 metabolizes monohalomethanes, ethylene oxide and to a lesser extent, the PAHs found in tobacco smoke. About 12% of the Finnish population lack the entire *GSTT1* gene and therefore do not exhibit the respective enzyme activity. The *GSTT1* gene was not associated with lung cancer risk in our study, but the *GSTT1* null genotype conferred an increased factor for oral cancer. In heavy smokers, a 3-fold increased risk for oral and pharyngeal cancers was seen. No differences were seen in the distribution of *GSTT1* genotypes in different age-groups in our control study setting. The role of *GSTT1* in cancer proneness thus remains unclear. In many studies it seems that the *GSTT1* null genotype is more protective than *GSTT1* positive genotype. This discrepancy cannot be explained solely by the role of *GSTT1* in the generation of toxic compounds [150].

EPHX1

EPHX1 null mice were found to be highly resistant to DMBA-induced carcinogenesis [276]. This points to an important role for EPHX1 in the activation of PAHs such as DMBA and B[a]P. In humans, wide interindividual differences in the EPHX1 activity have been observed. A 1.6-fold range of *in vivo* enzyme activities was observed in a healthy Caucasian population [360]. However, in *in vitro* experiments, the EPHX1 activity has been shown to exhibit an individual variation ranging from 2.5- to 63-fold, depending on substrate and the source and handling of liver tissue [361, 362]. Tissue differences in the expression of EPHX1 have also been detected. The enzyme is very abundant in liver and lung and much lower levels have been detected in lymphocytes. The genetic polymorphism detected thus far covers part of the detected variation in activity. Individuals with the Tyr^{113}/Arg^{139}

genotype exhibit the highest amount of enzyme [293], abut those with the Tyr^{113}/His^{139} combination possess the highest enzyme activity [295].

Most of the association studies done so far have determined the *EPHX1* $Tyr^{113}His$ genotype detected with a PCR-RFLP method [297, 301, 304, 305], which has been shown to be disturbed by an additional polymorphic nucleotide at codon 119 [291, 363]. This polymorphism leads to false genotype designation in the case of *EPHX1* His^{113}/His^{113}.

In our study, the lung cancer risk was significantly decreased in both *EPHX1* His^{113} allele-carrying and *EPHX1* Arg^{139} allele-carrying genotypes, when studied separately. The results were not affected by smoking status. The findings related to *EPHX1* $Tyr^{113}His$ polymorphism were in good agreement with a recent pooled analysis of 986 cases and 1633 controls [305]. However, the *EPHX1* $His^{139}Arg$ polymorphism has not yielded clear results in previous studies. In a recent meta- and pooled analysis of *EPHX1* $His^{139}Arg$ polymorphism, no effect was seen on lung cancer risk [305].

When exon 3 and 4 genotype data were combined to predict phenotype categories as established from the first functionality studies, no clear association with lung cancer was seen. This agrees with the results from the pooled analysis [305] although in some individual studies, the high activity phenotype has been suggested to be a risk factor for lung cancer [297, 301, 304]. The validity of previous interpretation, however, is questioned by the fact that most of these studies employed the defective RFLP-PCR method for the determination of the *EPHX1* $Tyr^{113}His$ genotypes. This may have caused some bias in the combined analysis.

Initial studies on the functionality of *EPHX1* genetic variation suggested increased activity to *EPHX1* Tyr^{113} and Arg^{139} alleles due to differences in the stability of the enzymes encoded [293]. The EPHX1 activity was reduced in lymphocytes of individuals carrying at least one His^{113} allele and somewhat higher for individuals possessing at least one Arg^{139} allele [296]. However, the most recent study attributes the highest activity to the wild type *EPHX1* $Tyr^{113}His^{139}$ haplotype [295].

In our study, we detected a decreased lung cancer risk for both *EPHX1* His^{113} and *EPHX1* Arg^{139} alleles. Since the His^{113} allele is thought to exhibit low enzyme activity and the Arg^{139} high enzyme activity, the results first appeared to be rather conflicting. However, in a recent functionality study, the Tyr^{113}/His^{139} combination exhibited the highest enzyme activity [295]. Diplotypes formed according to this study showed an increased lung cancer risk for homozygous *EPHX1* $Tyr^{113}His^{139}$ haplotypes. EPHX1 is involved in the activation of B[a]P to the highly carcinogenic BPDE and thus low enzyme activity can be thought to generate to a lower amount of reactive compounds and a lower cancer risk. However, it should be kept in mind that EPHX1 has a dual role in carcinogenesis; normally it is considered as a detoxification enzyme taking part in the hydration of reactive and toxic epoxides to their corresponding dihydrodiol products. Thus its role may depend on which environmental substrates the enzyme is metabolizing.

Focusing on pathways

Carcinogenesis is a multifactorial and multistage process that is dependent on a myriad of mechanisms and pathways that are under genetic control. At the moment an increasing number of genes are implicated as being relevant to carcinogenic outcome [135]. It will therefore be difficult to detect subtle differences in phenotype related to only a single polymorphism of a single gene [19].

In the current study in a separate analysis, only the *EPHX1* genotypes were significantly associated with lung cancer risk. In the joint analysis of the three *GST* genes, concurrent deficiency of *GSTM1* and *GSTT1* genes was associated with an increased risk of lung SCC in light smokers. No combined effect of *GSTM1*, *GSTP1* and *GSTT1* genotypes were detected. Joint analysis on *EPHX1* and *GST* genes focused on their role in heavy smokers and expecially in the development of SCC which emphasizes their role in the metabolism of PAHs; PAHs are thought to be crucial in the development of SCC cancer type of lung [364]. An eight-fold risk was seen for heavy smokers having a combination of the *GSTM1* null and *EPHX1* homozygous $Tyr^{113}His^{139}$. The *EPHX1* and *GSTT1* genotypes showed a significant interaction.

Since GSTs have overlapping substrate specificities, a deficiency of an individual GST isoenzyme may be compensated by other isoforms. Therefore, it is thought that simultaneous determination of all *GST* genotypes is a prerequisite for reliable interpretation of the role of the GST family in cancer development. However, recent studies have not found any clear evidence that the functional GSTs could compensate for a defective GST. Although the GST superfamilies exhibit some redundancy in their activities, the overlap in substrate specificity of individual isoenzymes may not be as extensive as originally was thought [365]. Information about tissue-spesific expression of GSTs is relatively scant. More information about GST distribution in man is therefore required to evaluate the types of disease processes that these enzymes may influence [204].

Experimental studies has shown that the lowest level of hydrophobic DNA adducts was found among patients with *GSTM1* positive and *GSTP1*1/*1* genotypes and the highest level was found in the group with combination of *GSTM1* null and *GSTP1*2/*2* genotype [210]. However, the reports of the association between *GSTM1* and *P1* and DNA adducts have also been mixed, with results varying by type of tissue, laboratory methods, and specific adducts measured [366]. Since tobacco smoke is a very complex mixture of chemicals, it is possible that when one tries to construct haplotype profiles, these alleles can both have advantages and disadvantages, depending on the composition of other genes in the profile. The different results of the published studies may thus be attributed to differences in the study populations and their exposure to environmental or dietary factors.

In the control population (Study III), the heaviest smokers were more likely to have the most protective combination of the *GST* genotypes. The risk increased somewhat but not statistically significantly when two or three genes were combined. This may be a

consequency of the low statistical power to detect the effect. It may also be linked to lack of overlapping substrate specificities or substantially different expression patterns of GSTs in different tissues.

An alternative hypothesis is that the GSH depletion to about 20–30% of total glutathione levels can impair the conjugation defence against the toxic actions of such compounds and become detrimental to cellular processes. Thus, the combined conjugation activities of all GSTs may lead to GSH depletion and in that way even be counterproductive. Instead of providing protection, the GSTs collectively may expose the cell to injurious effects such as oxidative DNA damage and associated mutagenic lesions [367].

Activity in phase I and II enzymes may affect the balance of produced BRIs and their effective detoxification. The BPDE-DNA adducts levels were modulated by *CYP1A1* and *GSTM1* polymorphism [182]. Another possible joint effect may derive from *CYPs* and *EPHX1* genes. It is probable that CYPs and EPHX1 cooperate via protein-protein interactions for fast metabolism [272, 273]. Certain CYP subfamilies have a quantifiable affinity for EPHX1. The CYP2C11 enzyme exhibits a higher activity for EPHX1 in rats than other the P450s, consistent with the ability of CYP2C11 to enhance the EPHX1-catalyzed hydrolysis of styrene oxide and B(a)P-oxide; B(a)P is activated in rats by CYP2C11 but not by CYP1A1 or CYP2B1.

As tobacco smoke influences a variety of biological processes (enzyme induction, oxidation, signal transduction) and contains multiple carcinogenic compounds, it is not sufficient to study only one pathway. Major carcinogens such as TSNAs and aromatic amines are metabolized by CYPs as well as the detoxification enzymes like NATs and UDPs. Multi-pathway interactions are also likely to occur between different carcinogenic pathways e.g. those involved in DNA repair and cell cycle control. The liability to suffer a disease is a complex interplay between the environment and genetic variation in several genes. At the moment there is an intense debate about gene-gene interactions. It is claimed that the interactions between genes and between genes and the environment are the most critical factors to explain why single gene studies produce such inconsistent results [368].

Validity

The results of studies on the effect of polymorphisms on lung and oropharyngeal cancer risks have been largely conflicting. The likelihood of a specific SNP being associated with altered cancer risk will depend on relevant environmental exposures. The lack of detailed information about the exposure may be one reason for the inconsistent results. Interactions between dietary and other environmental factors and SNPs in specific genes also may explain some of the interindividual variation in cancer experience and provide a mechanistic explanation for the lack of evidence of effect.

The careful measurement of strong confounders and effect modifiers as well as measurements of the exposure and disease variables and biomarkers are important in association studies. As many previous studies have had small study numbers and a lack of stringent study design, it is understandable that the results have been conflicting.

In the study design, all possible confounding factors and effect modifiers which are thought to modify the levels of enzymes, should be taken into account. For example, cruciferous vegetables are known to induce the expression of *GSTs* [369, 370]. GST induction may improve detoxification and excretion of potentially harmful compounds. Furthermore, the elevation of the activity of phase II detoxification enzymes may provide protection against neoplasia. Dietary factors are also likely to be critical because of their effect on DNA damage, mutation and repair. However, these is one hypothesis that *GSTM1* null individuals may experience greater exposure to dietary chemoprotective agents that are typically deactivated by *GSTM1*. In this case, the null genotype could be protective against lung cancer [204, 371–373].

More emphasis should be placed on the actual functional consequences of the genetic variation. For instance, the effects in the activity of different variants of *EPHX1* and *GSTP1* genotypes are relatively weak. Careful control of the factors influencing expression and protein levels should therefore be included in any future studies. It is thought that interindividual variation in activity, *e.g.*, in phenotype, is a result of the combination of induction, inhibitory, and genetic factors.

For statistical reasons, the gene-dosage effect is often ignored. Larger study designs could enable investigation of, the role of heterozygous and homozygous *GSTM1* positive genotypes. We genotyped the study subjects with a method which allowed only the identification of homozygous deficient *GSTM1* and *GSTT1* carriers. However, it is, known that the activity of these enzymes exhibit a trimodal phenotypic distribution, corresponding to the positive/positive, positive/null, null/null genotypes [374, 375].

An other possible source of bias is the differences in control groups. In many case-control studies, hospital controls are used. Sometimes this may carry an inherent source of bias. For instance, in one study, the control group consisted of hospital controls whose main diagnoses were chronic obstructive pulmonary disease (COPD). To illustrate this quandry, the *EPHX1* genotypes have been suggested to have opposite roles in the development of lung cancer and COPD [201].

Exposure assessment is another important issue to be considered. In many cases the lifetime exposure to tobacco exposure is imprecisely quantified. Although the duration and intensity of smoking have been shown to influence cancer risk, most molecular epidemiologic studies still classify tobacco exposure crudely as ´ever´ versus ´never´ or according to PYs smoked, and do not consider, *e.g.*, the important effects of cessation on the risk of tobacco-related cancers [32]. In our studies we examined only current smokers, or collected detailed data on smoking only from the current smokers, defined as those who had

smoked within the year prior to the study. From the rest of the subjects the information on smoking was limited to that needed to classify them as never-smokers, ex-smokers, or occasional smokers. This approach was chosen because it was thought that detailed information on cumulative smoking could not be reliably obtained from ex-smokers or those who had smoked only occasionally.

Finally, statistical methods should be further developed to be able to take account of all of the factors which are thought to affect cancer susceptibility [368, 376].

Future directions

DNA has a limited chemical stability and the maintenace of its integrity is a major goal of cells. Most cancers are known to be polygenic and multifactorial cancers are an output of the combination of two, usually dozens or perhaps even hundreds of genes and exogenous factors [135]. Carcinogenesis is a multifactorial and multistage process that is dependent on a myriad of mechanisms and pathways regulating absorption, metabolic activation and excretion, DNA repair, control of the mitotic cycle, hormonal stimulation of cell growth, inflammatory responses, and many other local or distant events that themselves are under genetic control [19]. For this reason it will be difficult to detect subtle differences in phenotype as the result of a single polymorphism of a single gene in a very complex pathway. It may necessary to evaluate the functional relevance of SNPs by concurrently genotyping a comprehensive list of SNPs in each pathway to test them globally as genetic risk factors for tobacco-related cancer and carefully consider the putative genetic and epigenetic pathways involved for the relevant agent and organ system [19].

To highlight the complexity of the topic, in one experiment two inbred strains of mice with different susceptibilities to lung cancer were interbred, and the phenotype and genotype was characterized in the progeny of the cross. In this experiment, at least 30 loci and many genetic interactions were associated with strain-specific cancer susceptibility [377]. This example demonstrates the tremendous genetic complexity associated with disease susceptibility.

Genomes are dynamic, fluctuating entities which have evolved by duplicating and by inducing variation when they duplicate. The recently completed determination of the whole human genome sequence provides a priceless research tool, which could ultimately enable analysis and understanding of human gene-interaction networks. Investigation of the role of hereditary factors and their interaction with environmental agents in the causation of human cancer will be of assistance in case-control studies aimed at identifying the specific genes that cause elevated cancer risk and help in the design of studies to examine their interaction with environmental risk factors. They will also investigate and develop the optimal epidemiological designs and statistical methods for analyzing the data generated from the aforementioned activities. The greatest challenges in human genetics, however, are

still ahead. Many genes, SNPs, CNPs and haplotypes have patterns of variation that differ considerably among populations. This emphasizes the need for rigor in the ascertainment of population origin when conducting association studies.

The implementation of the results to health benefits is also a major challenge for the future. Benefits would include being able to identify those smokers at the highest risk for developing cancer. These subgroups could be targeted for the most intensive smoking cessation interventions, could be enrolled into chemopreventive trials and might be suitable for more aggressive screening programmes that are not appropriate for the general population [26].

CONCLUSIONS

Most cancers show a complicated dependence on environmental factors. Individual differences in the ability to metabolize carcinogens modify the effects of environmental exposures and thereby can also affect the consequent cancer risk. In this study, the *GSTM1, GSTP1* and *GSTT1* polymorphisms were not associated with increased lung cancer risk, when studied alone. The *GSTP1* and *GSTT1* polymorphisms were associated with borderline increased oral and pharyngeal cancer risk, while no association with *GSTM1* was seen. Similarly the studied *GSTM3* polymorphism was not associated with lung or oral and pharyngeal cancer risk. However, since the recent findings of a functional SNP in the promoter region were not taken into account in this study, the *GSTM3* gene should still be considered as a potential candidate gene in future association studies.

The *EPHX1* His^{113} and Arg^{139} genotypes were associated with increased lung cancer risk and the risk was higher when the *EPHX1* and *GSTM1* genotypes were combined. These two genes showed an additive mode of interaction. Although the *EPHX1* and *GSTT1* genes showed a synergistic interaction, no significant combined effect on lung cancer susceptibility was observed. The above effects of genotypes in lung cancer susceptibility were seen to be mostly attributable to SCC. This is biologically plausible since the studied enzymes participate in metabolism of PAHs, which are thought to play an important role in development of SCC. However, due to the limited study size, these results remain speculative.

Our studies in the large control population stress that major emphasis should be put on the origin of the control population in association studies. The frequency of *GSTM1* genotypes was significantly different in males in the age groups studied. Moreover, the frequency of *GSTP1* genotypes differed in males and females and was affected by smoking status. Furthermore, when studying the combined effect of *GST* genes, overpresentation of *GST* protective genotypes was observed in cancer-free heavy smokers compared to lighter smokers in the oldest age groups studied.

Taken together, our results point to a modulating role of XME gene polymorphism on cancer proneness. However, as the knowledge on the variation on human genome is increasing exponentially, new methods to evaluate the role of these genes in cancer proneness need to be developed.

ACKNOWLEDGEMENTS

This study was carried out at the Department of Industrial Hygiene and Toxicology at the Finnish Institute of Occupational Health (FIOH). I wish to thank Professor Jorma Rantanen, the former Director General of the FIOH, Professor Harri Vainio, the Director General of the FIOH, and Professor Kai Savolainen the Head of the Department of Industrial Hygiene and Toxicology, for placing the excellent facilities and resources of the Institute at my disposal.

I wish to thank Professor Carl G. Gahmberg, the Head of Biochemistry Division, University of Helsinki.

I am grateful to my official referees, Professors Ingolf Cascorbi, and Hannu Raunio, for their careful review of the thesis and their valuable suggestions for improving the manuscript.

I wish to express my warmest gratitude to my supervisor Dr. Ari Hirvonen, for introducing me to the interesting world of molecular epidemiology. I am grateful for his support to pursue the PhD and to give independence in my studies.

I kindly thank my co-authors of the original publications included in this work. I am very grateful to Professor Olli Impivaara, Professor Jorma Järvisalo, Professor Kirsti Husgafvel-Pursiainen, Dr. Sisko Anttila, Dr. Nadejda Jourenkova-Mironova, and Dr. Simone Benhamou, for their expertise and fruitful collaboration.

I wish to thank the whole staff of the Department. I would like to express my sincere thanks to Dr. Hannu Norppa, for providing me with a place in his laboratory, in this ever changing world, and for the interesting scientific discussions and encouragement.

I wish to thank Maria Hirvonen and Tuula Suitiala for helping me in the laboratory. I also wish to thank all of the visitors and staff members in our lab who have worked under my guidance, Maria Pärepalo, Evgeny Suspitzin, Alexander Togo, Riina Pevgonen and Satu Hämäläinen.

I owe my warmest thanks to Katja Metsola for helping me with all the questions from statistics to dresscode. I am deeply grateful to Sirkku Saarikoski for giving me a scientific model to follow. I wish to thank Harriet Wikman, Pia Sillanpää, Iiris Heilimo, Päivi Rosenström, Reetta Holmila, Jarno Tuimala, and Hanna Sandvik, who together with my other collegues, contributed to the relaxed atmosphere.

I thank all my friends for taking my mind off the work. Special thanks to my parents and parents-in-law for all of the support and faith inme and especially for taking care of my children. Sincere thanks to Tuula Leino, Lena Ratazcak and the group of Lynx and Hedgehog in the kindergarten of Kuoppanummi for the warm athmosphere and loving care of my children.

Thanks for the love and patience of my loving husband Jaakko and our great children Eemil, Helmi and Eliel.

The financial support of Finnish Foundation for Cancer Research is gratefully acknowledged.

Helsinki, August 2005

REFERENCES

1. Collins, F.S., E.D. Green, A.E. Guttmacher, and M.S. Guyer, *A vision for the future of genomics research.* Nature, 2003. **422**(6934): p. 835–47.
2. WHO, *Tobacco or health: A global status report*. 1997, Geneva: World Health Organization. 495.
3. IARC, *Tobacco Smoke and Involuntary Smoking*. Monographs on the Evaluation of the Carcinogenic Risk of Chemicals to Humans. Vol. 83. 2004, Lyon: IARC Scientific Publications.
4. WHO, *The World Health Report 1999: Making a Difference*. 1999, Geneva: World Health Organization. 122.
5. WHO, *The World Health Report 2002: Reducing Risks, Promoting Healthy Life*. 2002, Geneva: World Health Organization.
6. Mackay, J.E., M, *The Tobacco Atlas*. 2002, Geneva: World Health Organization.
7. IARC, *Tobacco smoking*. Monographs on the Evaluation of the Carcinogenic Risk of Chemicals to Humans. Vol. 38. 1986, Lyon: IARC Scientific Publications. 270–303.
8. IARC, *World Cancer Report*, ed. B. Stewart and P. Kleihues. 2003, Lyon: IARC Press.
9. Takahashi, T., *Lung cancer: an ever increasing store of in-depth basic knowledge and the beginning of its clinical application.* Oncogene, 2002. **21**(45): p. 6868–9.
10. Doll, R. and A.B. Hill, *The mortality of doctors in relation to their smoking habits; a preliminary report.* Br Med J, 1954. **4877**: p. 1451–5.
11. IARC, *Cancer Incidence in Five Continents Vol. VIII*. IARC Scientific Publications, ed. D.M. Parkin, et al. Vol. No.155. 2002, Lyon: International Agency for Research on Cancer.
12. Peto, R., S. Darby, H. Deo, P. Silcocks, E. Whitley, and R. Doll, *Smoking, smoking cessation, and lung cancer in the UK since 1950: combination of national statistics with two case-control studies.* Bmj, 2000. **321**(7257): p. 323–9.
13. Darby, S., D. Hill, A. Auvinen, J.M. Barros-Dios, H. Baysson, F. Bochicchio, H. Deo, R. Falk, F. Forastiere, M. Hakama, I. Heid, L. Kreienbrock, M. Kreuzer, F. Lagarde, I. Makelainen, C. Muirhead, W. Oberaigner, G. Pershagen, A. Ruano-Ravina, E. Ruosteenoja, A.S. Rosario, M. Tirmarche, L. Tomasek, E. Whitley, H.E. Wichmann, and R. Doll, *Radon in homes and risk of lung cancer: collaborative analysis of individual data from 13 European case-control studies.* Bmj, 2004.
14. Lee, P.N., *Relation between exposure to asbestos and smoking jointly and the risk of lung cancer.* Occup Environ Med, 2001. **58**(3): p. 145–53.
15. Nelson, H.H. and K.T. Kelsey, *The molecular epidemiology of asbestos and tobacco in lung cancer.* Oncogene, 2002. **21**(48): p. 7284–8.
16. Nordlund, L.A., J.M. Carstensen, and G. Pershagen, *Are male and female smokers at equal risk of smoking-related cancer: evidence from a Swedish prospective study.* Scand J Public Health, 1999. **27**(1): p. 56–62.
17. Stellman, S.D., Y. Chen, J.E. Muscat, M.V. Djordjevic, J.P. Richie, Jr., P. Lazarus, S. Thompson, N. Altorki, M. Berwick, M.L. Citron, S. Harlap, T.B. Kaur, A.I. Neugut, S. Olson, J.M. Travaline, P. Witorsch, and Z.F. Zhang, *Lung cancer risk in white and black Americans.* Ann Epidemiol, 2003. **13**(4): p. 294–302.

18. Benowitz, N.L., E.J. Perez-Stable, B. Herrera, and P. Jacob, 3rd, *Slower metabolism and reduced intake of nicotine from cigarette smoking in Chinese-Americans.* J Natl Cancer Inst, 2002. **94**(2): p. 108–15.

19. Wu, X., H. Zhao, R. Suk, and D.C. Christiani, *Genetic susceptibility to tobacco-related cancer.* Oncogene, 2004. **23**(38): p. 6500–23.

20. Amos, C.I., N.E. Caporaso, and A. Weston, *Host factors in lung cancer risk: a review of interdisciplinary studies.* Cancer Epidemiol Biomarkers Prev, 1992. **1**(6): p. 505–13.

21. Schwartz, A.G., P. Yang, and G.M. Swanson, *Familial risk of lung cancer among nonsmokers and their relatives.* Am J Epidemiol, 1996. **144**(6): p. 554–62.

22. Travis, W.D., L.B. Travis, and S.S. Devesa, *Lung cancer.* Cancer, 1995. **75**(1 Suppl): p. 191–202.

23. Sobue, T., H. Tsukuma, A. Oshima, K. Genka, H. Tamori, N. Nishizawa, and S. Natsukawa, *Lung cancer incidence rates by histologic type in high- and low-risk areas; a population-based study in Osaka, Okinawa, and Saku Nagano, Japan.* J Epidemiol, 1999. **9**(3): p. 134–42.

24. Russo, A., P. Crosignani, S. Franceschi, and F. Berrino, *Changes in lung cancer histological types in Varese Cancer Registry, Italy 1976–1992.* Eur J Cancer, 1997. **33**(10): p. 1643–7.

25. Hecht, S.S., *Cigarette smoking and lung cancer: chemical mechanisms and approaches to prevention.* Lancet Oncol, 2002. **3**(8): p. 461–9.

26. Spitz, M.R., R.M. Santella, P. Boffetta, and P. Buffler, *Subgroup report: head and neck cancer.* IARC Sci Publ, 2004(157): p. 41–7.

27. Franceschi, S., F. Levi, C. La Vecchia, E. Conti, L. Dal Maso, L. Barzan, and R. Talamini, *Comparison of the effect of smoking and alcohol drinking between oral and pharyngeal cancer.* Int J Cancer, 1999. **83**(1): p. 1–4.

28. Franceschi, S., S. Barra, C. La Vecchia, E. Bidoli, E. Negri, and R. Talamini, *Risk factors for cancer of the tongue and the mouth. A case-control study from northern Italy.* Cancer, 1992. **70**(9): p. 2227–33.

29. McLaughlin, J.K., Z. Hrubec, W.J. Blot, and J.F. Fraumeni, Jr., *Smoking and cancer mortality among U.S. veterans: a 26-year follow-up.* Int J Cancer, 1995. **60**(2): p. 190–3.

30. Brennan, P. and P. Boffetta, *Mechanistic considerations in the molecular epidemiology of head and neck cancer.* IARC Sci Publ, 2004(157): p. 393–414.

31. Franceschi, S., R. Talamini, S. Barra, A.E. Baron, E. Negri, E. Bidoli, D. Serraino, and C. La Vecchia, *Smoking and drinking in relation to cancers of the oral cavity, pharynx, larynx, and esophagus in northern Italy.* Cancer Res, 1990. **50**(20): p. 6502–7.

32. Thun, M.J., S.J. Henley, and E.E. Calle, *Tobacco use and cancer: an epidemiologic perspective for geneticists.* Oncogene, 2002. **21**(48): p. 7307–25.

33. Hecht, S.S., *Tobacco carcinogens, their biomarkers and tobacco-induced cancer.* Nat Rev Cancer, 2003. **3**(10): p. 733–44.

34. Rodgman, A., C.J. Smith, and T.A. Perfetti, *The composition of cigarette smoke: a retrospective, with emphasis on polycyclic components.* Hum Exp Toxicol, 2000. **19**(10): p. 573–95.

35. IARC, *Polynuclear aromatic compounds, Part 1: Chemical, Environmental and Experimental Data*. IARC Monographs on the Evaluation of the Carcinogenic Risk of Chemicals to Humans. Vol. 32. 1983, Lyon.

36. Goldman, R., L. Enewold, E. Pellizzari, J.B. Beach, E.D. Bowman, S.S. Krishnan, and P.G. Shields, *Smoking increases carcinogenic polycyclic aromatic hydrocarbons in human lung tissue.* Cancer Res, 2001. **61**(17): p. 6367–71.

37. Lodovici, M., V. Akpan, L. Giovannini, F. Migliani, and P. Dolara, *Benzo[a]pyrene diol-epoxide DNA adducts and levels of polycyclic aromatic hydrocarbons in autoptic samples from human lungs.* Chem Biol Interact, 1998. **116**(3): p. 199–212.

38. Pfeifer, G.P., M.F. Denissenko, M. Olivier, N. Tretyakova, S.S. Hecht, and P. Hainaut, *Tobacco smoke carcinogens, DNA damage and p53 mutations in smoking-associated cancers.* Oncogene, 2002. **21**(48): p. 7435–51.

39. Gerde, P., B.A. Muggenburg, J.R. Thornton-Manning, J.L. Lewis, K.H. Pyon, and A.R. Dahl, *Benzo[a]pyrene at an environmentally relevant dose is slowly absorbed by, and extensively metabolized in, tracheal epithelium.* Carcinogenesis, 1997. **18**(9): p. 1825–32.

40. Cavalieri, E.L. and E.G. Rogan, *The approach to understanding aromatic hydrocarbon carcinogenesis. The central role of radical cations in metabolic activation.* Pharmacol Ther, 1992. **55**(2): p. 183–99.

41. Ross, J.A., G.B. Nelson, K.H. Wilson, J.R. Rabinowitz, A. Galati, G.D. Stoner, S. Nesnow, and M.J. Mass, *Adenomas induced by polycyclic aromatic hydrocarbons in strain A/J mouse lung correlate with time-integrated DNA adduct levels.* Cancer Res, 1995. **55**(5): p. 1039–44.

42. Hecht, S.S., *Biochemistry, biology, and carcinogenicity of tobacco-specific N- nitrosamines.* Chem Res Toxicol, 1998. **11**(6): p. 559–603.

43. Hecht, S.S., *DNA adduct formation from tobacco-specific N-nitrosamines.* Mutat Res, 1999. **424**(1–2): p. 127–42.

44. Peterson, L.A. and S.S. Hecht, *O6-methylguanine is a critical determinant of 4-(methylnitrosamino)-1- (3-pyridyl)-1-butanone tumorigenesis in A/J mouse lung.* Cancer Res, 1991. **51**(20): p. 5557–64.

45. Phillips, D.H., *Smoking-related DNA and protein adducts in human tissues.* Carcinogenesis, 2002. **23**(12): p. 1979–2004.

46. Hecht, S.S., *Tobacco smoke carcinogens and lung cancer.* J Natl Cancer Inst, 1999. **91**(14): p. 1194–210.

47. IARC, *Overall Evaluations of Carcinogenicity: An Updating of IARC Monographs Volumes 1 to 42.* 1987. **Suppl. 6**.

48. Castelao, J.E., J.M. Yuan, P.L. Skipper, S.R. Tannenbaum, M. Gago-Dominguez, J.S. Crowder, R.K. Ross, and M.C. Yu, *Gender- and smoking-related bladder cancer risk.* J Natl Cancer Inst, 2001. **93**(7): p. 538–45.

49. Pryor, W.A., K. Stone, L.Y. Zang, and E. Bermudez, *Fractionation of aqueous cigarette tar extracts: fractions that contain the tar radical cause DNA damage.* Chem Res Toxicol, 1998. **11**(5): p. 441–8.

50. Pryor, W.A., *Cigarette smoke radicals and the role of free radicals in chemical carcinogenicity.* Environ Health Perspect, 1997. **105 Suppl 4**: p. 875–82.

51. Arora, A., C.A. Willhite, and D.C. Liebler, *Interactions of beta-carotene and cigarette smoke in human bronchial epithelial cells.* Carcinogenesis, 2001. **22**(8): p. 1173–8.

52. Smith, C.J., T.A. Perfetti, M.J. Morton, A. Rodgman, R. Garg, C.D. Selassie, and C. Hansch, *The relative toxicity of substituted phenols reported in cigarette mainstream smoke.* Toxicol Sci, 2002. **69**(1): p. 265–78.

53. Lampe, J.W., S.B. Stepaniants, M. Mao, J.P. Radich, H. Dai, P.S. Linsley, S.H. Friend, and J.D. Potter, *Signatures of environmental exposures using peripheral leukocyte gene expression: tobacco smoke.* Cancer Epidemiol Biomarkers Prev, 2004. **13**(3): p. 445–53.

54. Snyder, R. and J.Y. Hong, *Metabolic and biochemical issues in the molecular epidemiology of cancer*, in *Mechanisms of Carcinogenesis Contributions of Molecular Epidemiology*, P. Buffler, et al., Editors. 2004, IARC Sci Publ. p. 51–69.

55. Guengerich, F.P. and T. Shimada, *Activation of procarcinogens by human cytochrome P450 enzymes.* Mutat Res, 1998. **400**(1–2): p. 201–13.

56. Lang, M. and O. Pelkonen, *Chapter 3. Metabolism of xenobiotics and chemical carcinogenesis*, in *Metabolic Polymorphisms and Susceptibility to Cancer*, P. Vineis, et al., Editors. 1999, IARC Sci Publ. p. 13–22.

57. Tang, D., D.H. Phillips, M. Stampfer, L.A. Mooney, Y. Hsu, S. Cho, W.Y. Tsai, J. Ma, K.J. Cole, M.N. She, and F.P. Perera, *Association between carcinogen-DNA adducts in white blood cells and lung cancer risk in the physicians health study.* Cancer Res, 2001. **61**(18): p. 6708–12.

58. Taningher, M., D. Malacarne, A. Izzotti, D. Ugolini, and S. Parodi, *Drug metabolism polymorphisms as modulators of cancer susceptibility.* Mutat Res, 1999. **436**(3): p. 227–61.

59. Piipari, R., K. Savela, T. Nurminen, J. Hukkanen, H. Raunio, J. Hakkola, T. Mantyla, P. Beaune, R.J. Edwards, A.R. Boobis, and S. Anttila, *Expression of CYP1A1, CYP1B1 and CYP3A, and polycyclic aromatic hydrocarbon-DNA adduct formation in bronchoalveolar macrophages of smokers and non-smokers.* Int J Cancer, 2000. **86**(5): p. 610–6.

60. Villard, P.H., E. Seree, B. Lacarelle, M.C. Therene-Fenoglio, Y. Barra, L. Attolini, B. Bruguerole, A. Durand, and J. Catalin, *Effect of cigarette smoke on hepatic and pulmonary cytochromes P450 in mouse: evidence for CYP2E1 induction in lung.* Biochem Biophys Res Commun, 1994. **202**(3): p. 1731–7.

61. Villard, P.H., E.M. Seree, J.L. Re, M. De Meo, Y. Barra, L. Attolini, G. Dumenil, J. Catalin, A. Durand, and B. Lacarelle, *Effects of tobacco smoke on the gene expression of the Cyp1a, Cyp2b, Cyp2e, and Cyp3a subfamilies in mouse liver and lung: relation to single strand breaks of DNA.* Toxicol Appl Pharmacol, 1998. **148**(2): p. 195–204.

62. Villard, P.H., R. Herber, E.M. Seree, L. Attolini, J. Magdalou, and B. Lacarelle, *Effect of cigarette smoke on UDP-glucuronosyltransferase activity and cytochrome P450 content in liver, lung and kidney microsomes in mice.* Pharmacol Toxicol, 1998. **82**(2): p. 74–9.

63. Anttila, S., A. Hirvonen, K. Husgafvel-Pursiainen, A. Karjalainen, T. Nurminen, and H. Vainio, *Combined effect of CYP1A1 inducibility and GSTM1 polymorphism on histological type of lung cancer.* Carcinogenesis, 1994. **15**(6): p. 1133–5.

64. Wardlaw, S.A., K.J. Nikula, D.A. Kracko, G.L. Finch, J.R. Thornton-Manning, and A.R. Dahl, *Effect of cigarette smoke on CYP1A1, CYP1A2 and CYP2B1/2 of nasal mucosae in F344 rats.* Carcinogenesis, 1998. **19**(4): p. 655–62.

65. Nebert, D.W., D.D. Petersen, and A.J. Fornace, Jr., *Cellular responses to oxidative stress: the [Ah] gene battery as a paradigm.* Environ Health Perspect, 1990. **88**: p. 13–25.

66. Hankinson, O., *The aryl hydrocarbon receptor complex.* Annu Rev Pharmacol Toxicol, 1995. **35**: p. 307–40.

67. Bauer, E., Z. Guo, Y.F. Ueng, L.C. Bell, D. Zeldin, and F.P. Guengerich, *Oxidation of benzo[a]pyrene by recombinant human cytochrome P450 enzymes.* Chem Res Toxicol, 1995. **8**(1): p. 136–42.

68. Shou, M., K.R. Korzekwa, C.L. Crespi, F.J. Gonzalez, and H.V. Gelboin, *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.* Mol Carcinog, 1994. **10**(3): p. 159–68.

69. Luch, A., *Nature and nurture – lessons from chemical carcinogenesis.* Nat Rev Cancer, 2005. **5**(2): p. 113–25.

70. Wiencke, J.K., S.W. Thurston, K.T. Kelsey, A. Varkonyi, J.C. Wain, E.J. Mark, and D.C. Christiani, *Early age at smoking initiation and tobacco carcinogen DNA damage in the lung.* J Natl Cancer Inst, 1999. **91**(7): p. 614–9.

71. Phillips, D.H., B. Schoket, A. Hewer, E. Bailey, S. Kostic, and I. Vincze, *Influence of cigarette smoking on the levels of DNA adducts in human bronchial epithelium and white blood cells.* Int J Cancer, 1990. **46**(4): p. 569–75.

72. Izzotti, A., G.A. Rossi, M. Bagnasco, and S. De Flora, *Benzo[a]pyrene diolepoxide-DNA adducts in alveolar macrophages of smokers.* Carcinogenesis, 1991. **12**(7): p. 1281–5.

73. Crawford, F.G., J. Mayer, R.M. Santella, T.B. Cooper, R. Ottman, W.Y. Tsai, G. Simon-Cereijido, M. Wang, D. Tang, and F.P. Perera, *Biomarkers of environmental tobacco smoke in preschool children and their mothers.* J Natl Cancer Inst, 1994. **86**(18): p. 1398–402.

74. Melikian, A.A., P. Sun, C. Pierpont, S. Coleman, and S.S. Hecht, *Gas chromatographic-mass spectrometric determination of benzo[a]pyrene and chrysene diol epoxide globin adducts in humans.* Cancer Epidemiol Biomarkers Prev, 1997. **6**(10): p. 833–9.

75. Pastorelli, R., M. Guanci, A. Cerri, E. Negri, C. La Vecchia, F. Fumagalli, M. Mezzetti, R. Cappelli, T. Panigalli, R. Fanelli, and L. Airoldi, *Impact of inherited polymorphisms in glutathione S-transferase M1, microsomal epoxide hydrolase, cytochrome P450 enzymes on DNA, and blood protein adducts of benzo(a)pyrene-diolepoxide.* Cancer Epidemiol Biomarkers Prev, 1998. **7**(8): p. 703–9.

76. Scherer, G., S. Frank, K. Riedel, I. Meger-Kossien, and T. Renner, *Biomonitoring of exposure to polycyclic aromatic hydrocarbons of nonoccupationally exposed persons.* Cancer Epidemiol Biomarkers Prev, 2000. **9**(4): p. 373–80.

77. Bartsch, H., M. Rojas, K. Alexandrov, and A. Risch, *Impact of adduct determination on the assessment of cancer susceptibility.* Recent Results Cancer Res, 1998. **154**: p. 86–96.

78. Kriek, E., F.J. Van Schooten, M.J. Hillebrand, F.E. Van Leeuwen, L. Den Engelse, A.J. De Looff, and A.P. Dijkmans, *DNA adducts as a measure of lung cancer risk in humans exposed to polycyclic aromatic hydrocarbons.* Environ Health Perspect, 1993. **99**: p. 71–5.

79. Vineis, P. and F. Perera, *DNA adducts as markers of exposure to carcinogens and risk of cancer.* Int J Cancer, 2000. **88**(3): p. 325–8.

80. Godschalk, R.W., E.J. Moonen, P.A. Schilderman, W.M. Broekmans, J.C. Kleinjans, and F.J. Van Schooten, *Exposure-route-dependent DNA adduct formation by polycyclic aromatic hydrocarbons.* Carcinogenesis, 2000. **21**(1): p. 87–92.

81. Tang, D., R.M. Santella, A.M. Blackwood, T.L. Young, J. Mayer, A. Jaretzki, S. Grantham, W.Y. Tsai, and F.P. Perera, *A molecular epidemiological case-control study of lung cancer.* Cancer Epidemiol Biomarkers Prev, 1995. **4**(4): p. 341–6.

82. Stone, J.G., N.J. Jones, A.D. McGregor, and R. Waters, *Development of a human biomonitoring assay using buccal mucosa: comparison of smoking-related DNA adducts in mucosa versus biopsies.* Cancer Res, 1995. **55**(6): p. 1267–70.

83. Romano, G., A. Sgambato, A. Boninsegna, G. Flamini, G. Curigliano, Q. Yang, V. La Gioia, C. Signorelli, A. Ferro, G. Capelli, R.M. Santella, and A. Cittadini, *Evaluation of polycyclic aromatic hydrocarbon-DNA adducts in exfoliated oral cells by an immunohistochemical assay.* Cancer Epidemiol Biomarkers Prev, 1999. **8**(1): p. 91–6.

84. Besarati Nia, A., H.W. Van Straaten, R.W. Godschalk, N. Van Zandwijk, A.J. Balm, J.C. Kleinjans, and F.J. Van Schooten, *Immunoperoxidase detection of polycyclic aromatic hydrocarbon-DNA adducts in mouth floor and buccal mucosa cells of smokers and nonsmokers.* Environ Mol Mutagen, 2000. **36**(2): p. 127–33.

85. Hsu, T.M., Y.J. Zhang, and R.M. Santella, *Immunoperoxidase quantitation of 4-aminobiphenyl- and polycyclic aromatic hydrocarbon-DNA adducts in exfoliated oral and urothelial cells of smokers and nonsmokers.* Cancer Epidemiol Biomarkers Prev, 1997. **6**(3): p. 193–9.

86. Latvala, J., S. Parkkila, J. Melkko, and O. Niemela, *Acetaldehyde adducts in blood and bone marrow of patients with ethanol-induced erythrocyte abnormalities.* Mol Med, 2001. **7**(6): p. 401–5.

87. Fang, J.L. and C.E. Vaca, *Detection of DNA adducts of acetaldehyde in peripheral white blood cells of alcohol abusers.* Carcinogenesis, 1997. **18**(4): p. 627–32.

88. Palli, D., G. Masala, M. Peluso, L. Gaspari, V. Krogh, A. Munnia, S. Panico, C. Saieva, R. Tumino, P. Vineis, and S. Garte, *The effects of diet on DNA bulky adduct levels are strongly modified by GSTM1 genotype: a study on 634 subjects.* Carcinogenesis, 2004. **25**(4): p. 577–584.

89. Peluso, M., L. Airoldi, C. Magagnotti, L. Fiorini, A. Munnia, A. Hautefeuille, C. Malaveille, and P. Vineis, *White blood cell DNA adducts and fruit and vegetable consumption in bladder cancer.* Carcinogenesis, 2000. **21**(2): p. 183–7.

90. Leanderson, P. and C. Tagesson, *Cigarette smoke-induced DNA-damage: role of hydroquinone and catechol in the formation of the oxidative DNA-adduct, 8-hydroxydeoxyguanosine.* Chem Biol Interact, 1990. **75**(1): p. 71–81.

91. Yang, Q., M. Hergenhahn, A. Weninger, and H. Bartsch, *Cigarette smoke induces direct DNA damage in the human B-lymphoid cell line Raji.* Carcinogenesis, 1999. **20**(9): p. 1769–75.

92. Liu, X., J. Lu, and S. Liu, *Synergistic induction of hydroxyl radical-induced DNA single-strand breaks by chromium(VI) compound and cigarette smoke solution.* Mutat Res, 1999. **440**(1): p. 109–17.

93. Hess, R.D. and G. Brandner, *DNA damage by filtered, tar- and aerosol-free cigarette smoke in rodent cells: a novel evaluation.* Toxicol Lett, 1996. **88**(1–3): p. 9–13.

94. Leanderson, P. and C. Tagesson, *Cigarette tar promotes neutrophil-induced DNA damage in cultured lung cells.* Environ Res, 1994. **64**(2): p. 103–11.

95. Kiyosawa, H., M. Suko, H. Okudaira, K. Murata, T. Miyamoto, M.H. Chung, H. Kasai, and S. Nishimura, *Cigarette smoking induces formation of 8-hydroxydeoxyguanosine, one of the oxidative DNA damages in human peripheral leukocytes.* Free Radic Res Commun, 1990. **11**(1–3): p. 23–7.

96. Weitberg, A.B. and D. Corvese, *Oxygen radicals potentiate the genetic toxicity of tobacco-specific nitrosamines.* Clin Genet, 1993. **43**(2): p. 88–91.

97. Asami, S., H. Manabe, J. Miyake, Y. Tsurudome, T. Hirano, R. Yamaguchi, H. Itoh, and H. Kasai, *Cigarette smoking induces an increase in oxidative DNA damage, 8-hydroxydeoxyguanosine, in a central site of the human lung.* Carcinogenesis, 1997. **18**(9): p. 1763–6.

98. Pignatelli, B., C.Q. Li, P. Boffetta, Q. Chen, W. Ahrens, F. Nyberg, A. Mukeria, I. Bruske-Hohlfeld, C. Fortes, V. Constantinescu, H. Ischiropoulos, and H. Ohshima, *Nitrated and oxidized plasma proteins in smokers and lung cancer patients.* Cancer Res, 2001. **61**(2): p. 778–84.

99. Spencer, J.P., A. Jenner, K. Chimel, O.I. Aruoma, C.E. Cross, R. Wu, and B. Halliwell, *DNA damage in human respiratory tract epithelial cells: damage by gas phase cigarette smoke apparently involves attack by reactive nitrogen species in addition to oxygen radicals.* FEBS Lett, 1995. **375**(3): p. 179–82.

100. Goode, E.L., C.M. Ulrich, and J.D. Potter, *Polymorphisms in DNA repair genes and associations with cancer risk.* Cancer Epidemiol Biomarkers Prev, 2002. **11**(12): p. 1513–30.

101. Norbury, C.J. and I.D. Hickson, *Cellular responses to DNA damage.* Annu Rev Pharmacol Toxicol, 2001. **41**: p. 367–401.

102. Hanawalt, P.C., *Controlling the efficiency of excision repair.* Mutat Res, 2001. **485**(1): p. 3–13.

103. Scicchitano, D.A. and P.C. Hanawalt, *Intragenomic repair heterogeneity of DNA damage.* Environ Health Perspect, 1992. **98**: p. 45–51.

104. Hainaut, P. and G.P. Pfeifer, *Patterns of p53 G-->T transversions in lung cancers reflect the primary mutagenic signature of DNA-damage by tobacco smoke.* Carcinogenesis, 2001. **22**(3): p. 367–74.

105. Hall, J., H. Bresil, F. Donato, C.P. Wild, N.A. Loktionova, O.I. Kazanova, I.P. Komyakov, V.G. Lemekhov, A.J. Likhachev, and R. Montesano, *Alkylation and oxidative-DNA damage repair activity in blood leukocytes of smokers and non-smokers.* Int J Cancer, 1993. **54**(5): p. 728–33.

106. Chen, A.T., J.A. Reidy, J.L. Annest, T.K. Welty, and H.G. Zhou, *Increased chromosome fragility as a consequence of blood folate levels, smoking status, and coffee consumption.* Environ Mol Mutagen, 1989. **13**(4): p. 319–24.

107. Kriek, E., M. Rojas, K. Alexandrov, and H. Bartsch, *Polycyclic aromatic hydrocarbon-DNA adducts in humans: relevance as biomarkers for exposure and cancer risk.* Mutat Res, 1998. **400**(1–2): p. 215–31.

108. Mollerup, S., D. Ryberg, A. Hewer, D.H. Phillips, and A. Haugen, *Sex differences in lung CYP1A1 expression and DNA adduct levels among lung cancer patients.* Cancer Res, 1999. **59**(14): p. 3317–20.

109. Denissenko, M.F., A. Pao, M. Tang, and G.P. Pfeifer, *Preferential formation of benzo[a]pyrene adducts at lung cancer mutational hotspots in P53.* Science, 1996. **274**(5286): p. 430–2.

110. Hussain, S.P., L.J. Hofseth, and C.C. Harris, *Tumor suppressor genes: at the crossroads of molecular carcinogenesis, molecular epidemiology and human risk assessment.* Lung Cancer, 2001. **34 Suppl 2**: p. S7–15.

111. Olivier, M., S.P. Hussain, C. Caron de Fromentel, P. Hainaut, and C.C. Harris, *TP53 mutation spectra and load: a tool for generating hypotheses on the etiology of cancer*, in *Mechanisms of Carcinogenesis Contributions of Molecular Epidemiology*, P. Buffler, et al., Editors. 2004, IARC Sci Publ. p. 247–70.

112. Hussain, S.P., P. Amstad, K. Raja, M. Sawyer, L. Hofseth, P.G. Shields, A. Hewer, D.H. Phillips, D. Ryberg, A. Haugen, and C.C. Harris, *Mutability of p53 hotspot codons to benzo(a)pyrene diol epoxide (BPDE) and the frequency of p53 mutations in nontumorous human lung.* Cancer Res, 2001. **61**(17): p. 6350–5.

113. Burkhart-Schultz, K.J., C.L. Thompson, and I.M. Jones, *Spectrum of somatic mutation at the hypoxanthine phosphoribosyltransferase (hprt) gene of healthy people.* Carcinogenesis, 1996. **17**(9): p. 1871–83.

114. Podlutsky, A., S.M. Hou, F. Nyberg, G. Pershagen, and B. Lambert, *Influence of smoking and donor age on the spectrum of in vivo mutation at the HPRT-locus in T lymphocytes of healthy adults.* Mutat Res, 1999. **431**(2): p. 325–39.

115. Hackman, P., S.M. Hou, F. Nyberg, G. Pershagen, and B. Lambert, *Mutational spectra at the hypoxanthine-guanine phosphoribosyltransferase (HPRT) locus in T-lymphocytes of nonsmoking and smoking lung cancer patients.* Mutat Res, 2000. **468**(1): p. 45–61.

116. Smith, L.E., M.F. Denissenko, W.P. Bennett, H. Li, S. Amin, M. Tang, and G.P. Pfeifer, *Targeting of lung cancer mutational hotspots by polycyclic aromatic hydrocarbons.* J Natl Cancer Inst, 2000. **92**(10): p. 803–11.

117. Denissenko, M.F., J.X. Chen, M.S. Tang, and G.P. Pfeifer, *Cytosine methylation determines hot spots of DNA damage in the human P53 gene.* Proc Natl Acad Sci U S A, 1997. **94**(8): p. 3893–8.

118. Denissenko, M.F., A. Pao, G.P. Pfeifer, and M. Tang, *Slow repair of bulky DNA adducts along the non-transcribed strand of the human p53 gene may explain the strand bias of transversion mutations in cancers.* Oncogene, 1998. **16**(10): p. 1241–7.

119. Sanchez-Cespedes, M., S.A. Ahrendt, S. Piantadosi, R. Rosell, M. Monzo, L. Wu, W.H. Westra, S.C. Yang, J. Jen, and D. Sidransky, *Chromosomal alterations in lung adenocarcinoma from smokers and non-smokers.* Cancer Res, 2001. **61**(4): p. 1309–13.

120. Slattery, M.L., K. Curtin, K. Anderson, K.N. Ma, L. Ballard, S. Edwards, D. Schaffer, J. Potter, M. Leppert, and W.S. Samowitz, *Associations between cigarette smoking, lifestyle factors, and microsatellite instability in colon tumors.* J Natl Cancer Inst, 2000. **92**(22): p. 1831–6.

121. Zhang, Z.F., X.M. Shu, C. Cordon-Cardo, I. Orlow, M.L. Lu, T.V. Millon, P.Q. Cao, C. Connolly-Jenks, G. Dalbagni, P. Lianes, L. Lacombe, V.E. Reuter, and H. Scher, *Cigarette smoking and chromosome 9 alterations in bladder cancer.* Cancer Epidemiol Biomarkers Prev, 1997. **6**(5): p. 321–6.

122. Curvall, M., T. Jansson, B. Pettersson, A. Hedin, and C.R. Enzell, *In vitro studies of biological effects of cigarette smoke condensate. I. Genotoxic and cytotoxic effects of neutral, semivolatile constituents.* Mutat Res, 1985. **157**(2–3): p. 169–80.

123. Salomaa, S., J. Tuominen, and E. Skytta, *Genotoxicity and PAC analysis of particulate and vapour phases of environmental tobacco smoke.* Mutat Res, 1988. **204**(2): p. 173–83.

124. Jansson, T., M. Curvall, A. Hedin, and C.R. Enzell, *In vitro studies of biological effects of cigarette smoke condensate. II. Induction of sister-chromatid exchanges in human lymphocytes by weakly acidic, semivolatile constituents.* Mutat Res, 1986. **169**(3): p. 129–39.

125. Asita, A.O., M. Matsui, T. Nohmi, A. Matsuoka, M. Hayashi, M. Ishidate, Jr., T. Sofuni, M. Koyano, and H. Matsushita, *Mutagenicity of wood smoke condensates in the Salmonella/microsome assay.* Mutat Res, 1991. **264**(1): p. 7–14.

126. Balansky, R.M., D.A. F, and S. De Flora, *Induction, persistence and modulation of cytogenetic alterations in cells of smoke-exposed mice.* Carcinogenesis, 1999. **20**(8): p. 1491–7.

127. Mohtashamipur, E., T. Steinforth, and K. Norpoth, *Comparative bone marrow clastogenicity of cigarette sidestream, mainstream and recombined smoke condensates in mice.* Mutagenesis, 1988. **3**(5): p. 419–22.

128. Balansky, R.M., P.M. Blagoeva, and Z.I. Mircheva, *The mutagenic and clastogenic activity of tobacco smoke.* Mutat Res, 1988. **208**(3–4): p. 237–41.

129. Herman, J.G. and S.B. Baylin, *Gene silencing in cancer in association with promoter hypermethylation.* N Engl J Med, 2003. **349**(21): p. 2042–54.

130. Mukai, T. and M. Sekiguchi, *Gene silencing in phenomena related to DNA repair.* Oncogene, 2002. **21**(58): p. 9033–42.

131. Knudson, A.G., Jr., *Mutation and cancer: statistical study of retinoblastoma.* Proc Natl Acad Sci U S A, 1971. **68**(4): p. 820–3.

132. Jones, P.A. and P.W. Laird, *Cancer epigenetics comes of age.* Nat Genet, 1999. **21**(2): p. 163–7.

133. Chen, S.Y., J.R. Lin, R. Darbha, P. Lin, T.Y. Liu, and Y.M. Chen, *Glycine N-Methyltransferase Tumor Susceptibility Gene in the Benzo(a)pyrene-Detoxification Pathway.* Cancer Res, 2004. **64**(10): p. 3617–23.

134. Hanahan, D. and R.A. Weinberg, *The hallmarks of cancer.* Cell, 2000. **100**(1): p. 57–70.

135. Ladiges, W., C. Kemp, J. Packenham, and J. Velazquez, *Human gene variation: from SNPs to phenotypes.* Mutat Res, 2004. **545**(1–2): p. 131–9.

136. Iafrate, A.J., L. Feuk, M.N. Rivera, M.L. Listewnik, P.K. Donahoe, Y. Qi, S.W. Scherer, and C. Lee, *Detection of large-scale variation in the human genome.* Nat Genet, 2004. **36**(9): p. 949–51.

137. *The ENCODE (ENCyclopedia Of DNA Elements) Project.* Science, 2004. **306**(5696): p. 636–40.

138. Rebbeck, T.R., M. Spitz, and X. Wu, *Assessing the function of genetic variants in candidate gene association studies.* Nat Rev Genet, 2004. **5**(8): p. 589–97.

139. Shields, P.G., *Molecular epidemiology of smoking and lung cancer.* Oncogene, 2002. **21**(45): p. 6870–6.

140. de la Chapelle, A., *Genetic predisposition to colorectal cancer.* Nature Reviews Cancer, 2004. **4**: p. 769–780.

141. Stubbins, M.J. and C.R. Wolf, *Chapter 21. Additional polymorphisms and cancer*, in *Metabolic Polymorphisms and Susceptibility to Cancer*, P. Vineis, et al., Editors. 1999, IARC Sci Publ. p. 271–302.

142. Su, T., Z. Bao, Q.Y. Zhang, T.J. Smith, J.Y. Hong, and X. Ding, *Human cytochrome P450 CYP2A13: predominant expression in the respiratory tract and its high efficiency metabolic activation of a tobacco-specific carcinogen, 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone.* Cancer Res, 2000. **60**(18): p. 5074–9.

143. Houlston, R.S., *Glutathione S-transferase M1 status and lung cancer risk: a meta- analysis.* Cancer Epidemiol Biomarkers Prev, 1999. **8**(8): p. 675–82.

144. Euskirchen, G. and M. Snyder, *A plethora of sites.* Nat Genet, 2004. **36**(4): p. 325–6.

145. *Glutathione S-transeferases: Structure, Function and Clinical Implications*, ed. N.P.E. Vermeulen, et al. 1996, London: Taylor & Francis Ltd. 261.

146. Ketterer, B., J.M. Harris, G. Talaska, D.J. Meyer, S.E. Pemble, J.B. Taylor, N.P. Lang, and F.F. Kadlubar, *The human glutathione S-transferase supergene family, its polymorphism, and its effects on susceptibility to lung cancer.* Environ Health Perspect, 1992. **98**: p. 87–94.

147. Mannervik, B., Y.C. Awasthi, P.G. Board, J.D. Hayes, C. Di Ilio, B. Ketterer, I. Listowsky, R. Morgenstern, M. Muramatsu, W.R. Pearson, and et al., *Nomenclature for human glutathione transferases.* Biochem J, 1992. **282**(Pt 1): p. 305–6.

148. Dekant, W. and S. Vamvakas, *Glutathione-dependent bioactivation of xenobiotics.* Xenobiotica, 1993. **23**(8): p. 873–87.

149. Dekant, W., *Chemical-induced nephrotoxicity mediated by glutathione S-conjugate formation.* Toxicol Lett, 2001. **124**(1–3): p. 21–36.

150. Thier, R., J.B. Taylor, S.E. Pemble, W.G. Humphreys, M. Persmark, B. Ketterer, and F.P. Guengerich, *Expression of mammalian glutathione S-transferase 5-5 in Salmonella typhimurium TA1535 leads to base-pair mutations upon exposure to dihalomethanes.* Proc Natl Acad Sci U S A, 1993. **90**(18): p. 8576–80.

151. Pemble, S., K.R. Schroeder, S.R. Spencer, D.J. Meyer, E. Hallier, H.M. Bolt, B. Ketterer, and J.B. Taylor, *Human glutathione S-transferase theta (GSTT1): cDNA cloning and the characterization of a genetic polymorphism.* Biochem J, 1994. **300**(Pt 1): p. 271–6.

152. Board, P.G., M. Coggan, G. Chelvanayagam, S. Easteal, L.S. Jermiin, G.K. Schulte, D.E. Danley, L.R. Hoth, M.C. Griffor, A.V. Kamath, M.H. Rosner, B.A. Chrunyk, D.E. Perregaux, C.A. Gabel, K.F. Geoghegan, and J. Pandit, *Identification, characterization, and crystal structure of the Omega class glutathione transferases.* J Biol Chem, 2000. **275**(32): p. 24798–806.

153. Strange, R.C., M.A. Spiteri, S. Ramachandran, and A.A. Fryer, *Glutathione-S-transferase family of enzymes.* Mutat Res, 2001. **482**(1–2): p. 21–6.

154. Hayes, J.D. and D.J. Pulford, *The glutathione S-transferase supergene family: regulation of GST and the contribution of the isoenzymes to cancer chemoprotection and drug resistance.* Crit Rev Biochem Mol Biol, 1995. **30**(6): p. 445–600.

155. Eaton, D.L. and T.K. Bammler, *Concise review of the glutathione S-transferases and their significance to toxicology.* Toxicol Sci, 1999. **49**(2): p. 156–64.

156. Hayes, J.D. and M. McMahon, *Molecular basis for the contribution of the antioxidant responsive element to cancer chemoprevention.* Cancer Lett, 2001. **174**(2): p. 103–13.

157. Sundberg, K., M. Widersten, A. Seidel, B. Mannervik, and B. Jernstrom, *Glutathione conjugation of bay- and fjord-region diol epoxides of polycyclic aromatic hydrocarbons by glutathione transferases M1-1 and P1-1.* Chem Res Toxicol, 1997. **10**(11): p. 1221–7.

158. Meyer, D.J., B. Coles, S.E. Pemble, K.S. Gilmore, G.M. Fraser, and B. Ketterer, *Theta, a new class of glutathione transferases purified from rat and man.* Biochem J, 1991. **274**(Pt 2): p. 409–14.

159. Landi, S., *Mammalian class theta GST and differential susceptibility to carcinogens: a review.* Mutat Res, 2000. **463**(3): p. 247–283.

160. Zhang, Y., R.H. Kolm, B. Mannervik, and P. Talalay, *Reversible conjugation of isothiocyanates with glutathione catalyzed by human glutathione transferases.* Biochem Biophys Res Commun, 1995. **206**(2): p. 748–55.

161. Szarka, C.E., G.R. Pfeiffer, S.T. Hum, L.C. Everley, A.M. Balshem, D.F. Moore, S. Litwin, E.B. Goosenberg, H. Frucht, P.F. Engstrom, and et al., *Glutathione S-transferase activity and glutathione S-transferase mu expression in subjects with risk for colorectal cancer.* Cancer Res, 1995. **55**(13): p. 2789–93.

162. Anttila, S., A. Hirvonen, H. Vainio, K. Husgafvel-Pursiainen, J.D. Hayes, and B. Ketterer, *Immunohistochemical localization of glutathione S-transferases in human lung.* Cancer Res, 1993. **53**(23): p. 5643–8.

163. Strange, R.C., C.G. Faulder, B.A. Davis, R. Hume, J.A. Brown, W. Cotton, and D.A. Hopkinson, *The human glutathione S-transferases: studies on the tissue distribution and genetic variation of the GST1, GST2 and GST3 isozymes.* Ann Hum Genet, 1984. **48 (Pt 1)**: p. 11–20.

164. van Lieshout, E.M. and W.H. Peters, *Age and gender dependent levels of glutathione and glutathione S-transferases in human lymphocytes.* Carcinogenesis, 1998. **19**(10): p. 1873–5.

165. Juronen, E., G. Tasa, M. Uuskula, M. Pooga, and A.V. Mikelsaar, *Purification, characterization and tissue distribution of human class theta glutathione S-transferase T1-1.* Biochem Mol Biol Int, 1996. **39**(1): p. 21–9.

166. Juronen, E., G. Tasa, M. Uuskula, M. Pooga, and A.V. Mikelsaar, *Production and characterization of monoclonal antibodies against class theta glutathione S-transferase T1-1.* Hybridoma, 1996. **15**(1): p. 77–82.

167. Smith, M.T., C.G. Evans, P. Doane-Setzer, V.M. Castro, M.K. Tahir, and B. Mannervik, *Denitrosation of 1,3-bis(2-chloroethyl)-1-nitrosourea by class mu glutathione transferases and its role in cellular resistance in rat brain tumor cells.* Cancer Res, 1989. **49**(10): p. 2621–5.

168. Listowsky, I., J.D. Rowe, Y.V. Patskovsky, T. Tchaikovskaya, N. Shintani, E. Novikova, and E. Nieves, *Human testicular glutathione S-transferases: insights into tissue-specific expression of the diverse subunit classes.* Chem Biol Interact, 1998. **111–112**: p. 103–12.

169. Ramos-Gomez, M., M.K. Kwak, P.M. Dolan, K. Itoh, M. Yamamoto, P. Talalay, and T.W. Kensler, *Sensitivity to carcinogenesis is increased and chemoprotective efficacy of enzyme inducers is lost in nrf2 transcription factor-deficient mice.* Proc Natl Acad Sci U S A, 2001. **98**(6): p. 3410–5.

170. Talalay, P., *Chemoprotection against cancer by induction of phase 2 enzymes.* Biofactors, 2000. **12**(1–4): p. 5–11.

171. Rinaldi, R., E. Eliasson, S. Swedmark, and R. Morgenstern, *Reactive intermediates and the dynamics of glutathione transferases.* Drug Metab Dispos, 2002. **30**(10): p. 1053–8.

172. van Lieshout, E.M., M.M. Bedaf, M. Pieter, C. Ekkel, W.A. Nijhoff, and W.H. Peters, *Effects of dietary anticarcinogens on rat gastrointestinal glutathione S-transferase theta 1-1 levels.* Carcinogenesis, 1998. **19**(11): p. 2055–7.

173. Seow, A., C.Y. Shi, F.L. Chung, D. Jiao, J.H. Hankin, H.P. Lee, G.A. Coetzee, and M.C. Yu, *Urinary total isothiocyanate (ITC) in a population-based sample of middle-aged and older Chinese in Singapore: relationship with dietary total ITC and glutathione S-transferase M1/T1/P1 genotypes.* Cancer Epidemiol Biomarkers Prev, 1998. **7**(9): p. 775–81.

174. Clapper, M.L. and C.E. Szarka, *Glutathione S-transferases--biomarkers of cancer risk and chemopreventive response.* Chem Biol Interact, 1998. **111–112**: p. 377–88.

175. Owuor, E.D. and A.N. Kong, *Antioxidants and oxidants regulated signal transduction pathways.* Biochem Pharmacol, 2002. **64**(5–6): p. 765–70.

176. Moffat, G.J., A.W. McLaren, and C.R. Wolf, *Sp1-mediated transcriptional activation of the human Pi class glutathione S-transferase promoter.* J Biol Chem, 1996. **271**(2): p. 1054–60.

177. Moffat, G.J., A.W. McLaren, and C.R. Wolf, *Functional characterization of the transcription silencer element located within the human Pi class glutathione S-transferase promoter.* J Biol Chem, 1996. **271**(34): p. 20740–7.

178. Moffat, G.J., A.W. McLaren, and C.R. Wolf, *Transcriptional and post-transcriptional mechanisms can regulate cell- specific expression of the human Pi-class glutathione S-transferase gene.* Biochem J, 1997. **324**(Pt 1): p. 91–5.

179. Jhaveri, M.S. and C.S. Morrow, *Methylation-mediated regulation of the glutathione S-transferase P1 gene in human breast cancer cells.* Gene, 1998. **210**(1): p. 1–7.

180. Eke, B.C., N. Vural, and M. Iscan, *Combined effects of ethanol and cigarette smoke on hepatic and pulmonary xenobiotic metabolizing enzymes in rats.* Chem Biol Interact, 1996. **102**(3): p. 155–67.

181. Sherratt, P.J., D.J. Pulford, D.J. Harrison, T. Green, and J.D. Hayes, *Evidence that human class Theta glutathione S-transferase T1-1 can catalyse the activation of dichloromethane, a liver and lung carcinogen in the mouse. Comparison of the tissue distribution of GST T1-1 with that of classes Alpha, Mu and Pi GST in human.* Biochem J, 1997. **326**(Pt 3): p. 837–46.

182. Rojas, M., I. Cascorbi, K. Alexandrov, E. Kriek, G. Auburtin, L. Mayer, A. Kopp-Schneider, I. Roots, and H. Bartsch, *Modulation of benzo[a]pyrene diolepoxide-DNA adduct levels in human white blood cells by CYP1A1, GSTM1 and GSTT1 polymorphism.* Carcinogenesis, 2000. **21**(1): p. 35–41.

183. Godschalk, R.W., J.W. Dallinga, H. Wikman, A. Risch, J.C. Kleinjans, H. Bartsch, and F.J. Van Schooten, *Modulation of DNA and protein adducts in smokers by genetic polymorphisms in GSTM1,GSTT1, NAT1 and NAT2.* Pharmacogenetics, 2001. **11**(5): p. 389–98.

184. Ketterer, B. and L.G. Christodoulides, *Enzymology of cytosolic glutathione S-transferases.* Adv Pharmacol, 1994. **27**: p. 37–69.

185. Kolm, R.H., U.H. Danielson, Y. Zhang, P. Talalay, and B. Mannervik, *Isothiocyanates as substrates for human glutathione transferases: structure-activity studies.* Biochem J, 1995. **311**(Pt 2): p. 453–9.

186. Comstock, K.E., M. Widersten, X.Y. Hao, W.D. Henner, and B. Mannervik, *A comparison of the enzymatic and physicochemical properties of human glutathione transferase M4-4 and three other human Mu class enzymes.* Arch Biochem Biophys, 1994. **311**(2): p. 487–95.

187. Seidegard, J., W.R. Vorachek, R.W. Pero, and W.R. Pearson, *Hereditary differences in the expression of the human glutathione transferase active on trans-stilbene oxide are due to a gene deletion.* Proc Natl Acad Sci U S A, 1988. **85**(19): p. 7293–7.

188. DeJong, J.L., C.M. Chang, J. Whang-Peng, T. Knutsen, and C.P. Tu, *The human liver glutathione S-transferase gene superfamily: expression and chromosome mapping of an Hb subunit cDNA.* Nucleic Acids Res, 1988. **16**(17): p. 8541–54.

189. Mannervik, B., P. Alin, C. Guthenberg, H. Jensson, M.K. Tahir, M. Warholm, and H. Jornvall, *Identification of three classes of cytosolic glutathione transferase common to several mammalian species: correlation between structural data and enzymatic properties.* Proc Natl Acad Sci U S A, 1985. **82**(21): p. 7202–6.

190. Tsuchida, S., T. Maki, and K. Sato, *Purification and characterization of glutathione transferases with an activity toward nitroglycerin from human aorta and heart. Multiplicity of the human class Mu forms.* J Biol Chem, 1990. **265**(13): p. 7150–7.

191. Campbell, E., Y. Takahashi, M. Abramovitz, M. Peretz, and I. Listowsky, *A distinct human testis and brain mu-class glutathione S-transferase. Molecular cloning and characterization of a form present even in individuals lacking hepatic type mu isoenzymes.* J Biol Chem, 1990. **265**(16): p. 9188–93.

192. Garte, S., L. Gaspari, A.K. Alexandrie, C. Ambrosone, H. Autrup, J.L. Autrup, H. Baranova, L. Bathum, S. Benhamou, P. Boffetta, C. Bouchardy, K. Breskvar, J. Brockmoller, I. Cascorbi, M.L. Clapper, C. Coutelle, A. Daly, M. Dell'Omo, V. Dolzan, C.M. Dresler, A. Fryer, A. Haugen, D.W. Hein, A. Hildesheim, A. Hirvonen, L.L. Hsieh, M. Ingelman-Sundberg, I. Kalina, D. Kang, M. Kihara, C.

Kiyohara, P. Kremers, P. Lazarus, L. Le Marchand, M.C. Lechner, E.M. van Lieshout, S. London, J.J. Manni, C.M. Maugard, S. Morita, V. Nazar-Stewart, K. Noda, Y. Oda, F.F. Parl, R. Pastorelli, I. Persson, W.H. Peters, A. Rannug, T. Rebbeck, A. Risch, L. Roelandt, M. Romkes, D. Ryberg, J. Salagovic, B. Schoket, J. Seidegard, P.G. Shields, E. Sim, D. Sinnet, R.C. Strange, I. Stucker, H. Sugimura, J. To-Figueras, P. Vineis, M.C. Yu, and E. Taioli, *Metabolic gene polymorphism frequencies in normal control populations.* Cancer Epidemiol Biomarkers Prev, 2001. **10**(12): p. 1239–48.

193. Dandara, C., J. Sayi, C.M. Masimirembwa, A. Magimba, S. Kaaya, K. De Sommers, J.R. Snyman, and J.A. Hasler, *Genetic polymorphism of cytochrome P450 1A1 (Cyp1A1) and glutathione transferases (M1, T1 and P1) among Africans.* Clin Chem Lab Med, 2002. **40**(9): p. 952–7.

194. Xu, S., Y. Wang, B. Roe, and W.R. Pearson, *Characterization of the human class Mu glutathione S-transferase gene cluster and the GSTM1 deletion.* J Biol Chem, 1998. **273**(6): p. 3517–27.

195. Roodi, N., W.D. Dupont, J.H. Moore, and F.F. Parl, *Association of homozygous wild-type glutathione S-transferase M1 genotype with increased breast cancer risk.* Cancer Res, 2004. **64**(4): p. 1233–6.

196. Widersten, M., W.R. Pearson, A. Engstrom, and B. Mannervik, *Heterologous expression of the allelic variant mu-class glutathione transferases mu and psi.* Biochem J, 1991. **276**(Pt 2): p. 519–24.

197. Tetlow, N., A. Robinson, T. Mantle, and P. Board, *Polymorphism of human mu class glutathione transferases.* Pharmacogenetics, 2004. **14**(6): p. 359–68.

198. McLellan, R.A., M. Oscarson, A.K. Alexandrie, J. Seidegard, D.A. Evans, A. Rannug, and M. Ingelman-Sundberg, *Characterization of a human glutathione S-transferase mu cluster containing a duplicated GSTM1 gene that causes ultrarapid enzyme activity.* Mol Pharmacol, 1997. **52**(6): p. 958–65.

199. Liu, X., M.R. Campbell, G.S. Pittman, E.C. Faulkner, M.A. Watson, and D.A. Bell, *Expression-based discovery of variation in the human glutathione S-transferase M3 promoter and functional analysis in a glioma cell line using allele-specific chromatin immunoprecipitation.* Cancer Res, 2005. **65**(1): p. 99–104.

200. Inskip, A., J. Elexperu-Camiruaga, N. Buxton, P.S. Dias, J. MacIntosh, D. Campbell, P.W. Jones, L. Yengi, J.A. Talbot, R.C. Strange, and et al., *Identification of polymorphism at the glutathione S-transferase, GSTM3 locus: evidence for linkage with GSTM1*A.* Biochem J, 1995. **312**(Pt 3): p. 713–6.

201. Risch, A., H. Wikman, S. Thiel, P. Schmezer, L. Edler, P. Drings, H. Dienemann, K. Kayser, V. Schulz, B. Spiegelhalder, and H. Bartsch, *Glutathione-S-transferase M1, M3, T1 and P1 polymorphisms and susceptibility to non-small-cell lung cancer subtypes and hamartomas.* Pharmacogenetics, 2001. **11**(9): p. 757–64.

202. Mitrunen, K., N. Jourenkova, V. Kataja, M. Eskelinen, V.M. Kosma, S. Benhamou, H. Vainio, M. Uusitupa, and A. Hirvonen, *Glutathione S-transferase M1, M3, P1, and T1 genetic polymorphism and susceptibility to breast cancer.* Cancer Epidemiol Biomarkers Prev, 2001. **10**(3): p. 229–36.

203. Jahnke V, M.C., Fryer A, Strange R., *Glutathione S-transferase and cytochrome P-450 polymorphism as risk factors for squamous cell carcinoma of the larynx.* Am J Surg, 1996. **172**(6): p. 671–673.

204. Hayes, J.D. and R.C. Strange, *Glutathione S-transferase polymorphisms and their biological consequences.* Pharmacology, 2000. **61**(3): p. 154–66.

205. Rebbeck, T.R., *Molecular epidemiology of the human glutathione S-transferase genotypes GSTM1 and GSTT1 in cancer susceptibility.* Cancer Epidemiol Biomarkers Prev, 1997. **6**(9): p. 733–43.

206. McWilliams, J.E., B.J. Sanderson, E.L. Harris, K.E. Richert-Boe, and W.D. Henner, *Glutathione S-transferase M1 (GSTM1) deficiency and lung cancer risk.* Cancer Epidemiol Biomarkers Prev, 1995. **4**(6): p. 589–94.

207. Kihara, M., K. Noda, and N. Okamoto, *Increased risk of lung cancer in Japanese smokers with class mu glutathione S-transferase gene deficiency.* Cancer Lett, 1993. **71**(1–3): p. 151–5.

208. Hirvonen, A., K. Husgafvel-Pursiainen, S. Anttila, and H. Vainio, *The GSTM1 null genotype as a potential risk modifier for squamous cell carcinoma of the lung.* Carcinogenesis, 1993. **14**(7): p. 1479–81.

209. Zhong, S., A.F. Howie, B. Ketterer, J. Taylor, J.D. Hayes, G.J. Beckett, C.G. Wathen, C.R. Wolf, and N.K. Spurr, *Glutathione S-transferase mu locus: use of genotyping and phenotyping assays to assess association with lung cancer susceptibility.* Carcinogenesis, 1991. **12**(9): p. 1533–7.

210. Ryberg, D., V. Skaug, A. Hewer, D.H. Phillips, L.W. Harries, C.R. Wolf, D. Ogreid, A. Ulvik, P. Vu, and A. Haugen, *Genotypes of glutathione transferase M1 and P1 and their significance for lung DNA adduct levels and cancer risk.* Carcinogenesis, 1997. **18**(7): p. 1285–9.

211. Pinarbasi, H., Y. Silig, O. Cetinkaya, Z. Seyfikli, and E. Pinarbasi, *Strong association between the GSTM1-null genotype and lung cancer in a Turkish population.* Cancer Genet Cytogenet, 2003. **146**(2): p. 125–9.

212. London, S.J., A.K. Daly, J. Cooper, W.C. Navidi, C.L. Carpenter, and J.R. Idle, *Polymorphism of glutathione S-transferase M1 and lung cancer risk among African-Americans and Caucasians in Los Angeles County, California.* J Natl Cancer Inst, 1995. **87**(16): p. 1246–53.

213. Benhamou, S., W.J. Lee, A.K. Alexandrie, P. Boffetta, C. Bouchardy, D. Butkiewicz, J. Brockmoller, M.L. Clapper, A. Daly, V. Dolzan, J. Ford, L. Gaspari, A. Haugen, A. Hirvonen, K. Husgafvel-Pursiainen, M. Ingelman-Sundberg, I. Kalina, M. Kihara, P. Kremers, L. Le Marchand, S.J. London, V. Nazar-Stewart, M. Onon-Kihara, A. Rannug, M. Romkes, D. Ryberg, J. Seidegard, P. Shields, R.C. Strange, I. Stucker, J. To-Figueras, P. Brennan, and E. Taioli, *Meta- and pooled analyses of the effects of glutathione S-transferase M1 polymorphisms and smoking on lung cancer risk.* Carcinogenesis, 2002. **23**(8): p. 1343–50.

214. Woodson, K., C. Stewart, M. Barrett, N.K. Bhat, J. Virtamo, P.R. Taylor, and D. Albanes, *Effect of vitamin intervention on the relationship between GSTM1, smoking, and lung cancer risk among male smokers.* Cancer Epidemiol Biomarkers Prev, 1999. **8**(11): p. 965–70.

215. Nazar-Stewart, V., A.G. Motulsky, D.L. Eaton, E. White, S.K. Hornung, Z.T. Leng, P. Stapleton, and N.S. Weiss, *The glutathione S-transferase mu polymorphism as a marker for susceptibility to lung carcinoma.* Cancer Res, 1993. **53**(10 Suppl): p. 2313–8.

216. Anttila, S., L. Luostarinen, A. Hirvonen, E. Elovaara, A. Karjalainen, T. Nurminen, J.D. Hayes, H. Vainio, and B. Ketterer, *Pulmonary expression of glutathione S-transferase M3 in lung cancer patients: association with GSTM1 polymorphism, smoking, and asbestos exposure.* Cancer Res, 1995. **55**(15): p. 3305–9.

217. Sorensen, M., H. Autrup, A. Tjonneland, K. Overvad, and O. Raaschou-Nielsen, *Glutathione S-transferase T1 null-genotype is associated with an increased risk of lung cancer.* Int J Cancer, 2004. **110**(2): p. 219–24.

218. Jourenkova-Mironova, N., H. Wikman, C. Bouchardy, A. Voho, P. Dayer, S. Benhamou, and A. Hirvonen, *Role of glutathione S-transferase GSTM1, GSTM3, GSTP1 and GSTT1 genotypes in modulating susceptibility to smoking-related lung cancer.* Pharmacogenetics, 1998. **8**(6): p. 495–502.

219. Hayashi, S., J. Watanabe, and K. Kawajiri, *High susceptibility to lung cancer analyzed in terms of combined genotypes of P450IA1 and Mu-class glutathione S-transferase genes.* Jpn J Cancer Res, 1992. **83**(8): p. 866–70.

220. Nakachi, K., K. Imai, S. Hayashi, and K. Kawajiri, *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 Res, 1993. **53**(13): p. 2994–9.

221. Alexandrie, A.K., M.I. Sundberg, J. Seidegard, G. Tornling, and A. Rannug, *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, 1994. **15**(9): p. 1785–90.

222. Hung, R.J., P. Boffetta, J. Brockmoller, D. Butkiewicz, I. Cascorbi, M.L. Clapper, S. Garte, A. Haugen, A. Hirvonen, S. Anttila, I. Kalina, L. Le Marchand, S.J. London, A. Rannug, M. Romkes, J. Salagovic, B. Schoket, L. Gaspari, and E. Taioli, *CYP1A1 and GSTM1 genetic polymorphisms and lung cancer risk in Caucasian non-smokers: a pooled analysis.* Carcinogenesis, 2003. **24**(5): p. 875–82.

223. Hashibe, M., P. Brennan, R.C. Strange, R. Bhisey, I. Cascorbi, P. Lazarus, M.B. Oude Ophuis, S. Benhamou, W.D. Foulkes, T. Katoh, C. Coutelle, M. Romkes, L. Gaspari, E. Taioli, and P. Boffetta, *Meta- and pooled analyses of GSTM1, GSTT1, GSTP1, and CYP1A1 genotypes and risk of head and neck cancer.* Cancer Epidemiol Biomarkers Prev, 2003. **12**(12): p. 1509–17.

224. Matthias, C., U. Bockmuhl, V. Jahnke, P.W. Jones, J.D. Hayes, J. Alldersea, J. Gilford, L. Bailey, J. Bath, S.F. Worrall, P. Hand, A.A. Fryer, and R.C. Strange, *Polymorphism in cytochrome P450 CYP2D6, CYP1A1, CYP2E1 and glutathione S-transferase, GSTM1, GSTM3, GSTT1 and susceptibility to tobacco-related cancers: studies in upper aerodigestive tract cancers.* Pharmacogenetics, 1998. **8**(2): p. 91–100.

225. Jahnke, V., R. Strange, C. Matthias, and A. Fryer, *Glutathione S-transferase and cytochrome P450 genotypes as risk factors for laryngeal carcinoma.* Eur Arch Otorhinolaryngol Suppl, 1997. **1**(9): p. S147–9.

226. Jourenkova-Mironova, N., A. Voho, C. Bouchardy, H. Wikman, P. Dayer, S. Benhamou, and A. Hirvonen, *Glutathione S-transferase GSTM3 and GSTP1 genotypes and larynx cancer risk.* Cancer Epidemiol Biomarkers Prev, 1999. **8**(2): p. 185–8.

227. Coles, B., M. Yang, N.P. Lang, and F.F. Kadlubar, *Expression of hGSTP1 alleles in human lung and catalytic activity of the native protein variants towards 1-chloro-2,4-dinitrobenzene, 4-vinylpyridine and (+)-anti benzo[a]pyrene-7,8-diol-9,10-oxide.* Cancer Lett, 2000. **156**(2): p. 167–75.

228. Ali-Osman, F., O. Akande, G. Antoun, J.X. Mao, and J. Buolamwini, *Molecular cloning, characterization, and expression in Escherichia coli of full-length cDNAs of three human glutathione S-transferase Pi gene variants. Evidence for differential catalytic activity of the encoded proteins.* J Biol Chem, 1997. **272**(15): p. 10004–12.

229. Sundberg, K., A.S. Johansson, G. Stenberg, M. Widersten, A. Seidel, B. Mannervik, and B. Jernstrom, *Differences in the catalytic efficiencies of allelic variants of glutathione transferase P1-1 towards carcinogenic diol epoxides of polycyclic aromatic hydrocarbons.* Carcinogenesis, 1998. **19**(3): p. 433–6.

230. Sundberg, K., A. Seidel, B. Mannervik, and B. Jernstrom, *Detoxication of carcinogenic fjord-region diol epoxides of polycyclic aromatic hydrocarbons by glutathione transferase P1-1 variants and glutathione.* FEBS Lett, 1998. **438**(3): p. 206–10.

231. Hu, X., X. Ji, S.K. Srivastava, H. Xia, S. Awasthi, B. Nanduri, Y.C. Awasthi, P. Zimniak, and S.V. Singh, *Mechanism of differential catalytic efficiency of two polymorphic forms of human glutathione S-transferase*

P1-1 in the glutathione conjugation of carcinogenic diol epoxide of chrysene. Arch Biochem Biophys, 1997. **345**(1): p. 32–8.

232. Hu, X., H. Xia, S.K. Srivastava, C. Herzog, Y.C. Awasthi, X. Ji, P. Zimniak, and S.V. Singh, *Activity of four allelic forms of glutathione S-transferase hGSTP1-1 for diol epoxides of polycyclic aromatic hydrocarbons.* Biochem Biophys Res Commun, 1997. **238**(2): p. 397–402.

233. Watson, M.A., R.K. Stewart, G.B. Smith, T.E. Massey, and D.A. Bell, *Human glutathione S-transferase P1 polymorphisms: relationship to lung tissue enzyme activity and population frequency distribution.* Carcinogenesis, 1998. **19**(2): p. 275–80.

234. Zimniak, P., B. Nanduri, S. Pikula, J. Bandorowicz-Pikula, S.S. Singhal, S.K. Srivastava, S. Awasthi, and Y.C. Awasthi, *Naturally occurring human glutathione S-transferase GSTP1-1 isoforms with isoleucine and valine in position 104 differ in enzymic properties.* Eur J Biochem, 1994. **224**(3): p. 893–9.

235. Whyatt, R.M., F.P. Perera, W. Jedrychowski, R.M. Santella, S. Garte, and D.A. Bell, *Association between polycyclic aromatic hydrocarbon-DNA adduct levels in maternal and newborn white blood cells and glutathione S-transferase P1 and CYP1A1 polymorphisms.* Cancer Epidemiol Biomarkers Prev, 2000. **9**(2): p. 207–12.

236. Harris, M.J., M. Coggan, L. Langton, S.R. Wilson, and P.G. Board, *Polymorphism of the Pi class glutathione S-transferase in normal populations and cancer patients.* Pharmacogenetics, 1998. **8**(1): p. 27–31.

237. To-Figueras, J., M. Gene, J. Gomez-Catalan, E. Pique, N. Borrego, J.L. Carrasco, J. Ramon, and J. Corbella, *Genetic polymorphism of glutathione S-transferase P1 gene and lung cancer risk.* Cancer Causes Control, 1999. **10**(1): p. 65–70.

238. Kihara, M. and K. Noda, *Lung cancer risk of the GSTM1 null genotype is enhanced in the presence of the GSTP1 mutated genotype in male Japanese smokers.* Cancer Lett, 1999. **137**(1): p. 53–60.

239. Lewis, S.J., N.M. Cherry, R.M. Niven, P.V. Barber, and A.C. Povey, *GSTM1, GSTT1 and GSTP1 polymorphisms and lung cancer risk.* Cancer Lett, 2002. **180**(2): p. 165–71.

240. Wang, Y., M.R. Spitz, M.B. Schabath, F. Ali-Osman, H. Mata, and X. Wu, *Association between glutathione S-transferase p1 polymorphisms and lung cancer risk in Caucasians: a case-control study.* Lung Cancer, 2003. **40**(1): p. 25–32.

241. Wang, J., Y. Deng, J. Cheng, J. Ding, and S. Tokudome, *GST genetic polymorphisms and lung adenocarcinoma susceptibility in a Chinese population.* Cancer Lett, 2003. **201**(2): p. 185–93.

242. Lin, P., Y.M. Hsueh, J.L. Ko, Y.F. Liang, K.J. Tsai, and C.Y. Chen, *Analysis of NQO1, GSTP1, and MnSOD genetic polymorphisms on lung cancer risk in Taiwan.* Lung Cancer, 2003. **40**(2): p. 123–9.

243. Nazar-Stewart, V., T.L. Vaughan, P. Stapleton, J. Van Loo, B. Nicol-Blades, and D.L. Eaton, *A population-based study of glutathione S-transferase M1, T1 and P1 genotypes and risk for lung cancer.* Lung Cancer, 2003. **40**(3): p. 247–58.

244. Schneider, J., U. Bernges, M. Philipp, and H.J. Woitowitz, *GSTM1, GSTT1, and GSTP1 polymorphism and lung cancer risk in relation to tobacco smoking.* Cancer Lett, 2004. **208**(1): p. 65–74.

245. Kiyohara, C., K.I. Yamamura, Y. Nakanishi, K. Takayama, and N. Hara, *Polymorphism in GSTM1, GSTT1, and GSTP1 and Susceptibility to Lung Cancer in a Japanese Population.* Asian Pac J Cancer Prev, 2000. **1**(4): p. 293–298.

246. Stucker, I., A. Hirvonen, I. de Waziers, A. Cabelguenne, K. Mitrunen, S. Cenee, E. Koum-Besson, D. Hemon, P. Beaune, and M.A. Loriot, *Genetic polymorphisms of glutathione S-transferases as modulators of lung cancer susceptibility.* Carcinogenesis, 2002. **23**(9): p. 1475–81.

247. Miller, D.P., D. Neuberg, I. de Vivo, J.C. Wain, T.J. Lynch, L. Su, and D.C. Christiani, *Smoking and the risk of lung cancer: susceptibility with GSTP1 polymorphisms.* Epidemiology, 2003. **14**(5): p. 545–51.

248. Miller, D.P., G. Liu, I. De Vivo, T.J. Lynch, J.C. Wain, L. Su, and D.C. Christiani, *Combinations of the variant genotypes of GSTP1, GSTM1, and p53 are associated with an increased lung cancer risk.* Cancer Res, 2002. **62**(10): p. 2819–23.

249. Reszka, E., W. Wasowicz, K. Rydzynski, N. Szeszenia-Dabrowska, and W. Szymczak, *Glutathione S-transferase M1 and P1 metabolic polymorphism and lung cancer predisposition.* Neoplasma, 2003. **50**(5): p. 357–62.

250. Coggan, M., L. Whitbread, A. Whittington, and P. Board, *Structure and organization of the human theta-class glutathione S-transferase and D-dopachrome tautomerase gene complex.* Biochem J, 1998. **334 (Pt 3)**: p. 617–23.

251. Webb, G., V. Vaska, M. Coggan, and P. Board, *Chromosomal localization of the gene for the human theta class glutathione transferase (GSTT1).* Genomics, 1996. **33**(1): p. 121–3.

252. Board, P., M. Harris, J. Flanagan, L. Langton, and M. Coggan, *Genetic heterogeneity of the structure and function of GSTT2 and GSTP1.* Chem Biol Interact, 1998. **111–112**: p. 83–9.

253. Sprenger, R., R. Schlagenhaufer, R. Kerb, C. Bruhn, J. Brockmoller, I. Roots, and U. Brinkmann, *Characterization of the glutathione S-transferase GSTT1 deletion: discrimination of all genotypes by polymerase chain reaction indicates a trimodular genotype-phenotype correlation.* Pharmacogenetics, 2000. **10**(6): p. 557–65.

254. Nelson, H.H., J.K. Wiencke, D.C. Christiani, T.J. Cheng, Z.F. Zuo, B.S. Schwartz, B.K. Lee, M.R. Spitz, M. Wang, X. Xu, and et al., *Ethnic differences in the prevalence of the homozygous deleted genotype of glutathione S-transferase theta.* Carcinogenesis, 1995. **16**(5): p. 1243–5.

255. Shen, J., G. Lin, W. Yuan, J. Tan, H.M. Bolt, and R. Thier, *Glutathione transferase T1 and M1 genotype polymorphism in the normal population of Shanghai.* Arch Toxicol, 1998. **72**(7): p. 456–8.

256. Bartsch, H., M. Rojas, U. Nair, J. Nair, and K. Alexandrov, *Genetic cancer susceptibility and DNA adducts: studies in smokers, tobacco chewers, and coke oven workers.* Cancer Detect Prev, 1999. **23**(6): p. 445–53.

257. To-Figueras, J., M. Gene, J. Gomez-Catalan, M.C. Galan, M. Fuentes, J.M. Ramon, M. Rodamilans, E. Huguet, and J. Corbella, *Glutathione S-transferase M1 (GSTM1) and T1 (GSTT1) polymorphisms and lung cancer risk among Northwestern Mediterraneans.* Carcinogenesis, 1997. **18**(8): p. 1529–33.

258. Malats, N., A.M. Camus-Radon, F. Nyberg, W. Ahrens, V. Constantinescu, A. Mukeria, S. Benhamou, H. Batura-Gabryel, I. Bruske-Hohlfeld, L. Simonato, A. Menezes, S. Lea, M. Lang, and P. Boffetta, *Lung cancer risk in nonsmokers and GSTM1 and GSTT1 genetic polymorphism.* Cancer Epidemiol Biomarkers Prev, 2000. **9**(8): p. 827–33.

259. Kelsey, K.T., M.R. Spitz, Z.F. Zuo, and J.K. Wiencke, *Polymorphisms in the glutathione S-transferase class mu and theta genes interact and increase susceptibility to lung cancer in minority populations.* Cancer Causes Control, 1997. **8**(4): p. 554–9.

260. el-Zein, R., J.B. Zwischenberger, T.G. Wood, S.Z. Abdel-Rahman, C. Brekelbaum, and W.W. Au, *Combined genetic polymorphism and risk for development of lung cancer.* Mutat Res, 1997. **381**(2): p. 189–200.

261. Jourenkova, N., M. Reinikanen, C. Bouchardy, K. Husgafvel-Pursiainen, P. Dayer, S. Benhamou, and A. Hirvonen, *Effects of glutathione S-transferases GSTM1 and GSTT1 genotypes on lung cancer risk in smokers.* Pharmacogenetics, 1997. **7**(6): p. 515–8.

262. Alexandrie, A.K., F. Nyberg, M. Warholm, and A. Rannug, *Influence of CYP1A1, GSTM1, GSTT1, and NQO1 genotypes and cumulative smoking dose on lung cancer risk in a Swedish population.* Cancer Epidemiol Biomarkers Prev, 2004. **13**(6): p. 908–14.

263. London, S.J., J.M. Yuan, F.L. Chung, Y.T. Gao, G.A. Coetzee, R.K. Ross, and M.C. Yu, *Isothiocyanates, glutathione S-transferase M1 and T1 polymorphisms, and lung-cancer risk: a prospective study of men in Shanghai, China [In Process Citation].* Lancet, 2000. **356**(9231): p. 724–9.

264. Spitz, M.R., C.M. Duphorne, M.A. Detry, P.C. Pillow, C.I. Amos, L. Lei, M. de Andrade, X. Gu, W.K. Hong, and X. Wu, *Dietary intake of isothiocyanates: evidence of a joint effect with glutathione S-transferase polymorphisms in lung cancer risk [In Process Citation].* Cancer Epidemiol Biomarkers Prev, 2000. **9**(10): p. 1017–20.

265. Deakin, M., J. Elder, C. Hendrickse, D. Peckham, D. Baldwin, C. Pantin, N. Wild, P. Leopard, D.A. Bell, P. Jones, H. Duncan, K. Brannigan, J. Alldersea, A.A. Fryer, and R.C. Strange, *Glutathione S-transferase GSTT1 genotypes and susceptibility to cancer: studies of interactions with GSTM1 in lung, oral, gastric and colorectal cancers.* Carcinogenesis, 1996. **17**(4): p. 881–4.

266. Fretland, A.J. and C.J. Omiecinski, *Epoxide hydrolases: biochemistry and molecular biology.* Chem Biol Interact, 2000. **129**(1–2): p. 41–59.

267. Coller, J.K., P. Fritz, U.M. Zanger, I. Siegle, M. Eichelbaum, H.K. Kroemer, and T.E. Murdter, *Distribution of microsomal epoxide hydrolase in humans: an immunohistochemical study in normal tissues, and benign and malignant tumours.* Histochem J, 2001. **33**(6): p. 329–36.

268. Beetham, J.K., D. Grant, M. Arand, J. Garbarino, T. Kiyosue, F. Pinot, F. Oesch, W.R. Belknap, K. Shinozaki, and B.D. Hammock, *Gene evolution of epoxide hydrolases and recommended nomenclature.* DNA Cell Biol, 1995. **14**(1): p. 61–71.

269. Sakaguchi, M., K. Mihara, and R. Sato, *A short amino-terminal segment of microsomal cytochrome P-450 functions both as an insertion signal and as a stop-transfer sequence.* Embo J, 1987. **6**(8): p. 2425–31.

270. Sato, T., M. Sakaguchi, K. Mihara, and T. Omura, *The amino-terminal structures that determine topological orientation of cytochrome P-450 in microsomal membrane.* Embo J, 1990. **9**(8): p. 2391–7.

271. Shimozawa, O., M. Sakaguchi, H. Ogawa, N. Harada, K. Mihara, and T. Omura, *Core glycosylation of cytochrome P-450(arom). Evidence for localization of N terminus of microsomal cytochrome P-450 in the lumen.* J Biol Chem, 1993. **268**(28): p. 21399–402.

272. Taura Ki, K., H. Yamada, E. Naito, N. Ariyoshi, M.A. Mori Ma, and K. Oguri, *Activation of microsomal epoxide hydrolase by interaction with cytochromes P450: kinetic analysis of the association and substrate-specific activation of epoxide hydrolase function.* Arch Biochem Biophys, 2002. **402**(2): p. 275–80.

273. Taura, K.I., H. Yamada, Y. Hagino, Y. Ishii, M.A. Mori, and K. Oguri, *Interaction between cytochrome P450 and other drug-metabolizing enzymes: evidence for an association of CYP1A1 with microsomal*

epoxide hydrolase and UDP-glucuronosyltransferase. Biochem Biophys Res Commun, 2000. **273**(3): p. 1048–52.

274. Shou, M., F.J. Gonzalez, and H.V. Gelboin, *Stereoselective epoxidation and hydration at the K-region of polycyclic aromatic hydrocarbons by cDNA-expressed cytochromes P450 1A1, 1A2, and epoxide hydrolase.* Biochemistry, 1996. **35**(49): p. 15807–13.

275. Phillips, D.H. and P.L. Grover, *Polycyclic hydrocarbon activation: bay regions and beyond.* Drug Metab Rev, 1994. **26**(1–2): p. 443–67.

276. Miyata, M., G. Kudo, Y.H. Lee, T.J. Yang, H.V. Gelboin, P. Fernandez-Salguero, S. Kimura, and F.J. Gonzalez, *Targeted disruption of the microsomal epoxide hydrolase gene. Microsomal epoxide hydrolase is required for the carcinogenic activity of 7,12-dimethylbenz[a]anthracene.* J Biol Chem, 1999. **274**(34): p. 23963–8.

277. Omiecinski, C.J., L. Aicher, R. Holubkov, and H. Checkoway, *Human peripheral lymphocytes as indicators of microsomal epoxide hydrolase activity in liver and lung.* Pharmacogenetics, 1993. **3**(3): p. 150–8.

278. Devereux, T.R., T.E. Massey, M.R. Van Scott, J. Yankaskas, and J.R. Fouts, *Xenobiotic metabolism in human alveolar type II cells isolated by centrifugal elutriation and density gradient centrifugation.* Cancer Res, 1986. **46**(10): p. 5438–43.

279. Liang, S.H., C. Hassett, and C.J. Omiecinski, *Alternative promoters determine tissue-specific expression profiles of the human microsomal epoxide hydrolase gene (EPHX1).* Mol Pharmacol, 2005. **67**(1): p. 220–30.

280. Zhu, Q.S., B. Qian, and D. Levy, *Regulation of human microsomal epoxide hydrolase gene (EPHX1) expression by the transcription factor GATA-4.* Biochim Biophys Acta, 2004. **1676**(3): p. 251–60.

281. Hassett, C., E.M. Laurenzana, J.S. Sidhu, and C.J. Omiecinski, *Effects of chemical inducers on human microsomal epoxide hydrolase in primary hepatocyte cultures.* Biochem Pharmacol, 1998. **55**(7): p. 1059–69.

282. Zhu, Q.S., B. Qian, and D. Levy, *CCAAT/enhancer-binding protein alpha (C/EBPalpha) activates transcription of the human microsomal epoxide hydrolase gene (EPHX1) through the interaction with DNA-bound NF-Y.* J Biol Chem, 2004. **279**(29): p. 29902–10.

283. Nebert, D.W. and F.J. Gonzalez, *P450 genes: structure, evolution, and regulation.* Annu Rev Biochem, 1987. **56**: p. 945–93.

284. Hartsfield, J.K., Jr., M.J. Sutcliffe, E.T. Everett, C. Hassett, C.J. Omiecinski, and J.A. Saari, *Assignment1 of microsomal epoxide hydrolase (EPHX1) to human chromosome 1q42.1 by in situ hybridization.* Cytogenet Cell Genet, 1998. **83**(1–2): p. 44–5.

285. Hassett, C., K.B. Robinson, N.B. Beck, and C.J. Omiecinski, *The human microsomal epoxide hydrolase gene (EPHX1): complete nucleotide sequence and structural characterization.* Genomics, 1994. **23**(2): p. 433–42.

286. Skoda, R.C., A. Demierre, O.W. McBride, F.J. Gonzalez, and U.A. Meyer, *Human microsomal xenobiotic epoxide hydrolase. Complementary DNA sequence, complementary DNA-directed expression in COS-1 cells, and chromosomal localization.* J Biol Chem, 1988. **263**(3): p. 1549–54.

287. Gaedigk, A., J.S. Leeder, and D.M. Grant, *Tissue-specific expression and alternative splicing of human microsomal epoxide hydrolase.* DNA Cell Biol, 1997. **16**(11): p. 1257–66.

288. London, S.J., J. Smart, and A.K. Daly, *Lung cancer risk in relation to genetic polymorphisms of microsomal epoxide hydrolase among African-Americans and Caucasians in Los Angeles County.* Lung Cancer, 2000. **28**(2): p. 147–55.

289. To-Figueras, J., M. Gene, J. Gomez-Catalan, E. Pique, N. Borrego, and J. Corbella, *Lung cancer susceptibility in relation to combined polymorphisms of microsomal epoxide hydrolase and glutathione S-transferase P1.* Cancer Lett, 2001. **173**(2): p. 155–62.

290. Gsur, A., T. Zidek, K. Schnattinger, E. Feik, G. Haidinger, P. Hollaus, A. Mohn-Staudner, C. Armbruster, S. Madersbacher, G. Schatzl, K. Trieb, C. Vutuc, and M. Micksche, *Association of microsomal epoxide hydrolase polymorphisms and lung cancer risk.* Br J Cancer, 2003. **89**(4): p. 702–6.

291. Baxter, S.W., D.Y. Choong, and I.G. Campbell, *Microsomal epoxide hydrolase polymorphism and susceptibility to ovarian cancer.* Cancer Lett, 2002. **177**(1): p. 75–81.

292. Yoshikawa, M., K. Hiyama, S. Ishioka, H. Maeda, A. Maeda, and M. Yamakido, *Microsomal epoxide hydrolase genotypes and chronic obstructive pulmonary disease in Japanese.* Int J Mol Med, 2000. **5**(1): p. 49–53.

293. Hassett, C., L. Aicher, J.S. Sidhu, and C.J. Omiecinski, *Human microsomal epoxide hydrolase: genetic polymorphism and functional expression in vitro of amino acid variants.* Hum Mol Genet, 1994. **3**(3): p. 421–8.

294. Hassett, C., J. Lin, C.L. Carty, E.M. Laurenzana, and C.J. Omiecinski, *Human hepatic microsomal epoxide hydrolase: comparative analysis of polymorphic expression.* Arch Biochem Biophys, 1997. **337**(2): p. 275–83.

295. Hosagrahara, V.P., A.E. Rettie, C. Hassett, and C.J. Omiecinski, *Functional analysis of human microsomal epoxide hydrolase genetic variants.* Chem Biol Interact, 2004. **150**(2): p. 149–59.

296. Haufroid, V., M. Jakubowski, B. Janasik, D. Ligocka, J.P. Buchet, E. Bergamaschi, P. Manini, A. Mutti, S. Ghittori, M. Arand, N. Hangen, F. Oesch, A. Hirvonen, and D. Lison, *Interest of genotyping and phenotyping of drug-metabolizing enzymes for the interpretation of biological monitoring of exposure to styrene.* Pharmacogenetics, 2002. **12**(9): p. 691–702.

297. Benhamou, S., M. Reinikainen, C. Bouchardy, P. Dayer, and A. Hirvonen, *Association between lung cancer and microsomal epoxide hydrolase genotypes.* Cancer Res, 1998. **58**(23): p. 5291–3.

298. Smith, C.A. and D.J. Harrison, *Association between polymorphism in gene for microsomal epoxide hydrolase and susceptibility to emphysema.* Lancet, 1997. **350**(9078): p. 630–3.

299. Yin, L., Y. Pu, T.Y. Liu, Y.H. Tung, K.W. Chen, and P. Lin, *Genetic polymorphisms of NAD(P)H quinone oxidoreductase, CYP1A1 and microsomal epoxide hydrolase and lung cancer risk in Nanjing, China.* Lung Cancer, 2001. **33**(2–3): p. 133–41.

300. Wu, X., K. Gwyn, C.I. Amos, N. Makan, W.K. Hong, and M.R. Spitz, *The association of microsomal epoxide hydrolase polymorphisms and lung cancer risk in African-Americans and Mexican-Americans.* Carcinogenesis, 2001. **22**(6): p. 923–8.

301. Zhou, W., S.W. Thurston, G. Liu, L.L. Xu, D.P. Miller, J.C. Wain, T.J. Lynch, L. Su, and D.C. Christiani, *The interaction between microsomal epoxide hydrolase polymorphisms and cumulative cigarette smoking in different histological subtypes of lung cancer.* Cancer Epidemiol Biomarkers Prev, 2001. **10**(5): p. 461–6.

302. Zhou, W., G. Liu, S.W. Thurston, L.L. Xu, D.P. Miller, J.C. Wain, T.J. Lynch, L. Su, and D.C. Christiani, *Genetic polymorphisms in N-acetyltransferase-2 and microsomal epoxide hydrolase, cumulative cigarette smoking, and lung cancer.* Cancer Epidemiol Biomarkers Prev, 2002. **11**(1): p. 15–21.

303. Zhao, H., M.R. Spitz, K.M. Gwyn, and X. Wu, *Microsomal epoxide hydrolase polymorphisms and lung cancer risk in non-Hispanic whites.* Mol Carcinog, 2002. **33**(2): p. 99–104.

304. Cajas-Salazar, N., W.W. Au, J.B. Zwischenberger, C.H. Sierra-Torres, S.A. Salama, S.K. Alpard, and S.K. Tyring, *Effect of epoxide hydrolase polymorphisms on chromosome aberrations and risk for lung cancer.* Cancer Genet Cytogenet, 2003. **145**(2): p. 97–102.

305. Lee, W.J., P. Brennan, P. Boffetta, S.J. London, S. Benhamou, A. Rannug, J. To-Figueras, M. Ingelman-Sundberg, P. Shields, L. Gaspari, and E. Taioli, *Microsomal epoxide hydrolase polymorphisms and lung cancer risk: a quantitative review.* Biomarkers, 2002. **7**(3): p. 230–41.

306. Park, J.Y., S.P. Schantz, and P. Lazarus, *Epoxide hydrolase genotype and orolaryngeal cancer risk: interaction with GSTM1 genotype.* Oral Oncol, 2003. **39**(5): p. 483–90.

307. Jourenkova-Mironova, N., K. Mitrunen, C. Bouchardy, P. Dayer, S. Benhamou, and A. Hirvonen, *High-activity microsomal epoxide hydrolase genotypes and the risk of oral, pharynx, and larynx cancers.* Cancer Res, 2000. **60**(3): p. 534–6.

308. Wenghoefer, M., B. Pesch, V. Harth, P. Broede, S. Fronhoffs, O. Landt, T. Bruning, J. Abel, H.M. Bolt, C. Herberhold, H. Vetter, and Y.D. Ko, *Association between head and neck cancer and microsomal epoxide hydrolase genotypes.* Arch Toxicol, 2003. **77**(1): p. 37–41.

309. To-Figueras, J., M. Gene, J. Gomez-Catalan, E. Pique, N. Borrego, M. Caballero, F. Cruellas, A. Raya, M. Dicenta, and J. Corbella, *Microsomal epoxide hydrolase and glutathione S-transferase polymorphisms in relation to laryngeal carcinoma risk.* Cancer Lett, 2002. **187**(1–2): p. 95–101.

310. Persson, I., I. Johansson, Y.C. Lou, Q.Y. Yue, L.S. Duan, L. Bertilsson, and M. Ingelman-Sundberg, *Genetic polymorphism of xenobiotic metabolizing enzymes among Chinese lung cancer patients.* Int J Cancer, 1999. **81**(3): p. 325–9.

311. Lin, P., S.L. Wang, H.J. Wang, K.W. Chen, H.S. Lee, K.J. Tsai, C.Y. Chen, and H. Lee, *Association of CYP1A1 and microsomal epoxide hydrolase polymorphisms with lung squamous cell carcinoma.* Br J Cancer, 2000. **82**(4): p. 852–7.

312. Rebbeck, T.R., C.B. Ambrosone, D.A. Bell, S.J. Chanock, R.B. Hayes, F.F. Kadlubar, and D.C. Thomas, *SNPs, haplotypes, and cancer: applications in molecular epidemiology.* Cancer Epidemiol Biomarkers Prev, 2004. **13**(5): p. 681–7.

313. Goldstein, D.B. and M.E. Weale, *Population genomics: linkage disequilibrium holds the key.* Curr Biol, 2001. **11**(14): p. R576–9.

314. Deitz, A.C., N. Rothman, T.R. Rebbeck, R.B. Hayes, W.H. Chow, W. Zheng, D.W. Hein, and M. Garcia-Closas, *Impact of misclassification in genotype-exposure interaction studies: example of N-acetyltransferase 2 (NAT2), smoking, and bladder cancer.* Cancer Epidemiol Biomarkers Prev, 2004. **13**(9): p. 1543–6.

315. Marchini, J., L.R. Cardon, M.S. Phillips, and P. Donnelly, *The effects of human population structure on large genetic association studies.* Nat Genet, 2004. **36**(5): p. 512–7.

316. Ng, P.C. and S. Henikoff, *Predicting deleterious amino acid substitutions.* Genome Res, 2001. **11**(5): p. 863–74.

317. Fleming, M.A., J.D. Potter, C.J. Ramirez, G.K. Ostrander, and E.A. Ostrander, *Understanding missense mutations in the BRCA1 gene: an evolutionary approach.* Proc Natl Acad Sci U S A, 2003. **100**(3): p. 1151–6.

318. Ramensky, V., P. Bork, and S. Sunyaev, *Human non-synonymous SNPs: server and survey.* Nucleic Acids Res, 2002. **30**(17): p. 3894–900.

319. Stoilov, P., E. Meshorer, M. Gencheva, D. Glick, H. Soreq, and S. Stamm, *Defects in pre-mRNA processing as causes of and predisposition to diseases.* DNA Cell Biol, 2002. **21**(11): p. 803–18.

320. Knight, J.C., *Functional implications of genetic variation in non-coding DNA for disease susceptibility and gene regulation.* Clin Sci (Lond), 2003. **104**(5): p. 493–501.

321. Botstein, D. and N. Risch, *Discovering genotypes underlying human phenotypes: past successes for mendelian disease, future approaches for complex disease.* Nat Genet, 2003. **33 Suppl**: p. 228–37.

322. Salisbury, B.A., M. Pungliya, J.Y. Choi, R. Jiang, X.J. Sun, and J.C. Stephens, *SNP and haplotype variation in the human genome.* Mutat Res, 2003. **526**(1–2): p. 53–61.

323. *The International HapMap Project.* Nature, 2003. **426**(6968): p. 789–796.

324. Pharoah, P.D., A.M. Dunning, B.A. Ponder, and D.F. Easton, *Association studies for finding cancer-susceptibility genetic variants.* Nat Rev Cancer, 2004. **4**(11): p. 850–60.

325. Collins, F.S., M. Morgan, and A. Patrinos, *The Human Genome Project: lessons from large-scale biology.* Science, 2003. **300**(5617): p. 286–90.

326. Vahakangas, K., *Ethical aspects of molecular epidemiology of cancer.* Carcinogenesis, 2004. **25**(4): p. 465–71.

327. Burgermeister, J., *Switzerland has opened door to genetic discrimination, say ethicists.* Bmj, 2004. **329**(7457): p. 70.

328. Burgermeister, J., *Germany may allow employers to introduce gene tests for job seekers.* Bmj, 2004. **329**(7472): p. 940.

329. Mayor, S., *Report warns of danger of genetic discrimination in the workplace.* Bmj, 2003. **327**(7417): p. 702.

330. Marcy, T.W., M. Stefanek, and K.M. Thompson, *Genetic testing for lung cancer risk: if physicians can do it, should they?* J Gen Intern Med, 2002. **17**(12): p. 946–51.

331. Tuomi, T., M.S. Huuskonen, L. Tammilehto, E. Vanhala, and M. Virtamo, *Occupational exposure to asbestos as evaluated from work histories and analysis of lung tissues from patients with mesothelioma.* Br J Ind Med, 1991. **48**(1): p. 48–52.

332. Morrison, N.A., J.C. Qi, A. Tokita, P.J. Kelly, L. Crofts, T.V. Nguyen, P.N. Sambrook, and J.A. Eisman, *Prediction of bone density from vitamin D receptor alleles.* Nature, 1994. **367**(6460): p. 284–7.

333. Chen, H., D.P. Sandler, J.A. Taylor, D.L. Shore, E. Liu, C.D. Bloomfield, and D.A. Bell, *Increased risk for myelodysplastic syndromes in individuals with glutathione transferase theta 1 (GSTT1) gene defect.* Lancet, 1996. **347**(8997): p. 295–7.

334. Hirvonen, A., S.T. Saarikoski, K. Linnainmaa, K. Koskinen, K. Husgafvel-Pursiainen, K. Mattson, and H. Vainio, *Glutathione S-transferase and N-acetyltransferase genotypes and asbestos-associated pulmonary disorders.* J Natl Cancer Inst, 1996. **88**(24): p. 1853–6.

335. Hassett, C., L. Aicher, J.S. Sidhu, and C.J. Omiecinski, *Human microsomal epoxide hydrolase: genetic polymorphism and functional expression in vitro of amino acid variants.* Hum. Mol. Genet., 1994. **3**: p. 421–428.

336. Lancaster, J.M., H.A. Browniee, D.A. Bell, A. Futreal, J.R. Marks, A. Berchuck, R.W. Wiseman, and J.A. Taylor, *Microsomal epoxide hydrolase polymorphism as a risk factor for ovarian cancer.* Mol. Carcinog., 1996. **17**: p. 160–162.

337. Spurdle, A.B., D.M. Purdie, M.W. Penelope, X. Chen, A. Green, and G. Chenevix-Trench, *The Microsomal Epoxide Hydrolase Tyr113His Polymorphism: Association With Risk of Ovarian Cancer.* Molecular Carcinogenesis, 2001. **30**: p. 71–78.

338. Belogubova, E.V., A.V. Togo, T.V. Kondratieva, V.G. Lemehov, K.P. Hanson, and E.N. Imyanitov, *GSTM1 genotypes in elderly tumour-free smokers and non-smokers.* Lung Cancer, 2000. **29**(3): p. 189–95.

339. Lancaster, J.M., H.A. Brownlee, D.A. Bell, P.A. Futreal, J.R. Marks, A. Berchuck, R.W. Wiseman, and J.A. Taylor, *Microsomal epoxide hydrolase polymorphism as a risk factor for ovarian cancer.* Mol Carcinog, 1996. **17**(3): p. 160–2.

340. Strange, R.C., P.W. Jones, and A.A. Fryer, *Glutathione S-transferase: genetics and role in toxicology.* Toxicol Lett, 2000. **112–113**: p. 357–63.

341. Poirier, M.C., *Chemical-induced DNA damage and human cancer risk.* Nat Rev Cancer, 2004. **4**(8): p. 630–7.

342. Kensler, T.W., P.A. Egner, N.E. Davidson, B.D. Roebuck, A. Pikul, and J.D. Groopman, *Modulation of aflatoxin metabolism, aflatoxin-N7-guanine formation, and hepatic tumorigenesis in rats fed ethoxyquin: role of induction of glutathione S-transferases.* Cancer Res, 1986. **46**(8): p. 3924–31.

343. Ye, Z., H. Song, and Y. Guo, *Glutathione S-transferase M1, T1 status and the risk of head and neck cancer: a meta-analysis.* J Med Genet, 2004. **41**(5): p. 360–5.

344. Chen, K., Q.T. Jiang, and H.Q. He, *Relationship between metabolic enzyme polymorphism and colorectal cancer.* World J Gastroenterol, 2005. **11**(3): p. 331–5.

345. Ye, Z. and J.M. Parry, *A meta-analysis of 20 case-control studies of the glutathione S-transferase M1 (GSTM1) status and colorectal cancer risk.* Med Sci Monit, 2003. **9**(10): p. SR83–91.

346. La Torre, G., S. Boccia, and G. Ricciardi, *Glutathione S-transferase M1 status and gastric cancer risk: a meta-analysis.* Cancer Lett, 2005. **217**(1): p. 53–60.

347. Vogl, F.D., E. Taioli, C. Maugard, W. Zheng, L.F. Pinto, C. Ambrosone, F.F. Parl, V. Nedelcheva-Kristensen, T.R. Rebbeck, P. Brennan, and P. Boffetta, *Glutathione S-transferases M1, T1, and P1 and breast cancer: a pooled analysis.* Cancer Epidemiol Biomarkers Prev, 2004. **13**(9): p. 1473–9.

348. Egan, K.M., Q. Cai, X.O. Shu, F. Jin, T.L. Zhu, Q. Dai, Y.T. Gao, and W. Zheng, *Genetic polymorphisms in GSTM1, GSTP1, and GSTT1 and the risk for breast cancer: results from the Shanghai Breast Cancer Study and meta-analysis.* Cancer Epidemiol Biomarkers Prev, 2004. **13**(2): p. 197–204.

349. Sull, J.W., H. Ohrr, D.R. Kang, and C.M. Nam, *Glutathione S-transferase M1 status and breast cancer risk: a meta-analysis.* Yonsei Med J, 2004. **45**(4): p. 683–9.

350. Ntais, C., A. Polycarpou, and J.P. Ioannidis, *Association of GSTM1, GSTT1, and GSTP1 gene polymorphisms with the risk of prostate cancer: a meta-analysis.* Cancer Epidemiol Biomarkers Prev, 2005. **14**(1): p. 176–81.

351. Johns, L.E. and R.S. Houlston, *Glutathione S-transferase mu1 (GSTM1) status and bladder cancer risk: a meta-analysis.* Mutagenesis, 2000. **15**(5): p. 399–404.

352. Engel, L.S., E. Taioli, R. Pfeiffer, M. Garcia-Closas, P.M. Marcus, Q. Lan, P. Boffetta, P. Vineis, H. Autrup, D.A. Bell, R.A. Branch, J. Brockmoller, A.K. Daly, S.R. Heckbert, I. Kalina, D. Kang, T. Katoh, A. Lafuente, H.J. Lin, M. Romkes, J.A. Taylor, and N. Rothman, *Pooled analysis and meta-analysis of glutathione S-transferase M1 and bladder cancer: a HuGE review.* Am J Epidemiol, 2002. **156**(2): p. 95–109.

353. Heckbert, S.R., N.S. Weiss, S.K. Hornung, D.L. Eaton, and A.G. Motulsky, *Glutathione S-transferase and epoxide hydrolase activity in human leukocytes in relation to risk of lung cancer and other smoking-related cancers.* J Natl Cancer Inst, 1992. **84**(6): p. 414–22.

354. Seidegard, J., R.W. Pero, M.M. Markowitz, G. Roush, D.G. Miller, and E.J. Beattie, *Isoenzyme(s) of glutathione transferase (class Mu) as a marker for the susceptibility to lung cancer: a follow up study.* Carcinogenesis, 1990. **11**(1): p. 33–6.

355. Saarikoski, S.T., A. Voho, M. Reinikainen, S. Anttila, A. Karjalainen, C. Malaveille, H. Vainio, K. Husgafvel-Pursiainen, and A. Hirvonen, *Combined effect of polymorphic GST genes on individual susceptibility to lung cancer.* Int J Cancer, 1998. **77**(4): p. 516–21.

356. Nyberg, F., S.M. Hou, K. Hemminki, B. Lambert, and G. Pershagen, *Glutathione S-transferase mu1 and N-acetyltransferase 2 genetic polymorphisms and exposure to tobacco smoke in nonsmoking and smoking lung cancer patients and population controls.* Cancer Epidemiol Biomarkers Prev, 1998. **7**(10): p. 875–83.

357. Ford, J.G., Y. Li, M.M. O'Sullivan, R. Demopoulos, S. Garte, E. Taioli, and P.W. Brandt-Rauf, *Glutathione S-transferase M1 polymorphism and lung cancer risk in african-americans.* Carcinogenesis, 2000. **21**(11): p. 1971–5.

358. Kihara, M. and K. Noda, *Distribution of GSTM1 null genotype in relation to gender, age and smoking status in Japanese lung cancer patients.* Pharmacogenetics, 1995. **5**(Spec No): p. S74-9.

359. Howells, R.E., K.K. Dhar, P.R. Hoban, P.W. Jones, A.A. Fryer, C.W. Redman, and R.C. Strange, *Association between glutathione-S-transferase GSTP1 genotypes, GSTP1 over-expression, and outcome in epithelial ovarian cancer.* Int J Gynecol Cancer, 2004. **14**(2): p. 242–50.

360. Kroetz, D.L., B.M. Kerr, L.V. McFarland, P. Loiseau, A.J. Wilensky, and R.H. Levy, *Measurement of in vivo microsomal epoxide hydrolase activity in white subjects.* Clin Pharmacol Ther, 1993. **53**(3): p. 306–15.

361. Mertes, I., R. Fleischmann, H.R. Glatt, and F. Oesch, *Interindividual variations in the activities of cytosolic and microsomal epoxide hydrolase in human liver.* Carcinogenesis, 1985. **6**(2): p. 219–23.

362. Seidegard, J. and G. Ekstrom, *The role of human glutathione transferases and epoxide hydrolases in the metabolism of xenobiotics.* Environ Health Perspect, 1997. **105 Suppl 4**: p. 791–9.

363. Keicho, N., M. Emi, M. Kajita, I. Matsushita, K. Nakata, A. Azuma, N. Ohishi, and S. Kudoh, *Overestimated frequency of a possible emphysema-susceptibility allele when microsomal epoxide hydrolase is genotyped by the conventional polymerase chain reaction-based method.* J Hum Genet, 2001. **46**(2): p. 96–8.

364. Brooks, D.R., J.H. Austin, R.T. Heelan, M.S. Ginsberg, V. Shin, S.H. Olson, J.E. Muscat, and S.D. Stellman, *Influence of type of cigarette on peripheral versus central lung cancer.* Cancer Epidemiol Biomarkers Prev, 2005. **14**(3): p. 576–81.

365. Seidel, A., T. Friedberg, B. Lollmann, A. Schwierzok, M. Funk, H. Frank, R. Holler, F. Oesch, and H. Glatt, *Detoxification of optically active bay- and fjord-region polycyclic aromatic hydrocarbon dihydrodiol epoxides by human glutathione transferase P1-1 expressed in Chinese hamster V79 cells.* Carcinogenesis, 1998. **19**(11): p. 1975–81.

366. Pavanello, S. and E. Clonfero, *Biological indicators of genotoxic risk and metabolic polymorphisms.* Mutat Res, 2000. **463**(3): p. 285–308.

367. Parl, F.F., *Glutathione S-transferase genotypes and cancer risk.* Cancer Lett, 2005. **221**(2): p. 123–9.

368. Moore, J.H. and S.M. Williams, *New strategies for identifying gene-gene interactions in hypertension.* Ann Med, 2002. **34**(2): p. 88–95.

369. Boone, C.W., G.J. Kelloff, and W.E. Malone, *Identification of candidate cancer chemopreventive agents and their evaluation in animal models and human clinical trials: a review.* Cancer Res, 1990. **50**(1): p. 2–9.

370. Steinkellner, H., S. Rabot, C. Freywald, E. Nobis, G. Scharf, M. Chabicovsky, S. Knasmuller, and F. Kassie, *Effects of cruciferous vegetables and their constituents on drug metabolizing enzymes involved in the bioactivation of DNA-reactive dietary carcinogens.* Mutat Res, 2001. **480–481**: p. 285–97.

371. Lampe, J.W., C. Chen, S. Li, J. Prunty, M.T. Grate, D.E. Meehan, K.V. Barale, D.A. Dightman, Z. Feng, and J.D. Potter, *Modulation of human glutathione S-transferases by botanically defined vegetable diets.* Cancer Epidemiol Biomarkers Prev, 2000. **9**(8): p. 787–93.

372. Lin, H.J., N.M. Probst-Hensch, A.D. Louie, I.H. Kau, J.S. Witte, S.A. Ingles, H.D. Frankl, E.R. Lee, and R.W. Haile, *Glutathione transferase null genotype, broccoli, and lower prevalence of colorectal adenomas.* Cancer Epidemiol Biomarkers Prev, 1998. **7**(8): p. 647–52.

373. Nijhoff, W.A., M.J. Grubben, F.M. Nagengast, J.B. Jansen, H. Verhagen, G. van Poppel, and W.H. Peters, *Effects of consumption of Brussels sprouts on intestinal and lymphocytic glutathione S-transferases in humans.* Carcinogenesis, 1995. **16**(9): p. 2125–8.

374. Brockmoller, J., D. Gross, R. Kerb, N. Drakoulis, and I. Roots, *Correlation between trans-stilbene oxide-glutathione conjugation activity and the deletion mutation in the glutathione S-transferase class mu gene detected by polymerase chain reaction.* Biochem Pharmacol, 1992. **43**(3): p. 647–50.

375. Seidegard, J., J.W. DePierre, and R.W. Pero, *Hereditary interindividual differences in the glutathione transferase activity towards trans-stilbene oxide in resting human mononuclear leukocytes are due to a particular isozyme(s).* Carcinogenesis, 1985. **6**(8): p. 1211–6.

376. Ritchie, M.D., L.W. Hahn, and J.H. Moore, *Power of multifactor dimensionality reduction for detecting gene-gene interactions in the presence of genotyping error, missing data, phenocopy, and genetic heterogeneity.* Genet Epidemiol, 2003. **24**(2): p. 150–7.

377. Tripodis, N., A.A. Hart, R.J. Fijneman, and P. Demant, *Complexity of lung cancer modifiers: mapping of thirty genes and twenty-five interactions in half of the mouse genome.* J Natl Cancer Inst, 2001. **93**(19): p. 1484–91.

When I know your genes I can say how you should live.
Is this reality in the future?

Humans are known to differ from each other in their responses to environmental exposures. During recent years these differences have been attributed to interindividual variations in the genes encoding for enzymes involved in the metabolism of exogenous agents. The individual variations in metabolic capacity have been suggested to be an important modifier of individual susceptibility to environmentally induced cancers.

This study discusses current knowledge of the interaction between tobacco smoke and genetic factors in lung and oropharyngeal cancer proneness.

Orders:
Finnish Institute of Occupational Health
Topeliuksenkatu 41 a A
FIN-00250 Helsinki
Finland

Fax +358-9-477 5071
kirjakauppa@ttl.fi
www.ttl.fi

ISBN (in print) 951-802-637-8
ISBN (PDF) 951-802-638-6
ISSN 1237-6183

Cover picture Jaakko Haaparanta

ISBN 951-802-637-8