Cancer Causes and Control (2005) 16:579–585
DOI 10.1007/s10552-004-7842-3

Risk of non-small cell lung cancer and the cytochrome P4501A1 Ile462Val polymorphism

Jill Everland Larsen[1,2,*], Maree Louise Colosimo[1], Ian Anthony Yang[1,2], Rayleen Bowman[2], Paul Victor Zimmerman[1,2], Kwun Meng Fong[1,2]
[1]*Department of Thoracic Medicine, The Prince Charles Hospital, Brisbane, Australia;* [2]*School of Medicine, University of Queensland, Brisbane, Australia*

Received 10 August 2004; accepted in revised form 20 December 2004

Key words: cytochrome P450 CYP1A1, carcinoma, non-small-cell lung, polymorphism (genetics), genetic predisposition to disease.

Abstract

Objective: The Ile462Val substitution in the cytochrome P450 1A1 gene (*CYP1A1*) results in increased enzymatic activity. Preliminary data suggesting a link between this polymorphism and lung cancer risk in Caucasians are inconsistent, reflecting small sample sizes and the relatively low frequency of the variant.
Methods: The data set consisted of 1050 primary non-small cell lung cancer cases and 581 controls, a large homogenous population designed specifically to address previous inconsistencies. Patients were genotyped using a PCR-RFLP technique.
Results: Carriers of the valine allele, *CYP1A1*2C*, (Ile/Val or Val/Val genotypes) were significantly over-represented in non-small cell lung cancer compared to controls (OR = 1.9; 95% CI = 1.2–2.9; $p = 0.005$) when adjusted for confounders, particularly in women (OR = 4.6; 95% CI = 1.7–12.4; $p = 0.003$). The valine variant was statistically significantly over-represented in cases of lung cancer younger than the median age (64 years) (OR = 2.5; 95% CI = 1.3–4.8; $p = 0.005$) and cases with less than the median cumulative tobacco-smoke exposure (46 pack-years) (OR = 2.4; 95% CI = 1.3–4.7; $p = 0.007$).
Conclusions: These new data establish an association between the *CYP1A1* Ile462Val polymorphism and the risk of developing non-small cell lung cancer, especially among women.

Introduction

Lung cancer is attributable to cigarette smoking yet less than 20% of life-long smokers will develop lung cancer suggesting the possibility of genetic predisposition [1]. Enzymes encoded by polymorphic genes that activate or detoxify harmful chemicals may account for wide inter-individual differences in sensitivity to cancer-inducing or cancer-promoting compounds [2].

Phase I enzymes (mainly cytochrome P450) metabolically activate carcinogens such as polycyclic aromatic hydrocarbons (PAHs) and *N*-nitrosamines to reactive intermediates [3]. These intermediates are capable of binding covalently to DNA to form DNA adducts, thereby potentially initiating the carcinogenic process. Increased activity of benzo[*a*]pyrene (B[*a*]P) hydroxylase, encoded by cytochrome P4501A1 (*CYP1A1*) has been associated with lung cancer risk [4–6]. Hayashi *et al*. [7] first described a base substitution of adenine to guanine at position 2455 in the heme-binding region of exon 7, that results in an isoleucine to valine amino acid substitution at codon 462 (Ile462Val). The valine allele, designated *CYP1A1*2C* following the initial trivial designation of *CYP1A1 m2*, correlates with increased enzymatic activity [8–10], thought to lead to greater carcinogen susceptibility and consequently higher risk of tobacco smoke-related lung cancer.

An over-representation of the valine allele among lung cancer cases has been reported in numerous studies in

* Author for correspondence: Ms Jill Larsen, Thoracic Research Laboratory, Level 1 Clinical Sciences Building, The Prince Charles Hospital, Rode Road, Chermside, Queensland 4032 Australia. Ph.: +61-7-3350-8437; Fax: +61-7-3350-8957; E-mail: jill_e_larsen@health.qld.gov.au

Asian and Caucasian populations [11–15]. Whilst results from Asian populations are fairly consistent, studies in Caucasian populations are more controversial, most likely caused by ethnic differences in allele frequency. Although relatively frequent in Asian populations [0.20–0.25], the valine allele is quite rare in Caucasian populations (0.05). Thus, smaller epidemiological studies may have lacked the statistical power to accurately define the relationship between the *CYP1A1* Ile462Val polymorphism and lung cancer. In addition, conflicting results may arise from population heterogeneity, as well as failure to control for the potential confounding effect of tobacco smoke, the main determinant of lung cancer.

To address this, a recent pooled analysis of original data sets combined the results of eleven studies creating a collective Caucasian population of 1153 lung cancer cases and 1449 controls [16]. The authors found individuals with the exon 7 variant at an increased risk of lung cancer, particularly squamous cell carcinoma (SCC). Concurrently, we have conducted an adequately powered association study in a homogenous population of non-small cell lung carcinomas (NSCLCs) (of similar sample size to the total of the eleven cumulative studies) to investigate the role of *CYP1A1* Ile462Val in modifying lung cancer risk. Our aim was to conclusively define the role of this polymorphism in the risk of lung cancer in a single, large homogenous population of NSCLCs.

Materials and methods

Cases consisted of patients with cytologically or histologically confirmed primary NSCLC (n = 1050) treated at The Prince Charles Hospital from 1980 to 2003. Controls consisted of patients with chronic obstructive pulmonary disease (COPD) but without lung cancer (n = 358), treated at the same hospital from 1998 to 2003, or healthy smokers attending a smoking cessation clinic held at the hospital from 2000 to 2003 (n = 223). These two patient groups were selected as controls as they possessed a median tobacco-smoke exposure approximately equivalent to cases, this factor being the greatest risk factor for lung cancer. Thus, in smokers of equivalent tobacco-smoke exposure, the effect of the polymorphism could be compared between those with and without lung cancer. Approximately 95% of eligible subjects consented to participate in the study and the population from which our study group was drawn was more than 99% Caucasian.

The study and use of archived NSCLC paraffin blocks (n = 259) was approved by the Ethics Committee of The Prince Charles Hospital, and patients (both cases and controls) gave informed written consent for use of resected fresh lung tissue or blood (n = 1372). Demographic characteristics of cases and controls (Table 1) were collected by the research nurse or treating physician, and data were checked against patient records and the institutional lung cancer database.

DNA from control subjects was extracted from peripheral blood. DNA from subjects diagnosed with NSCLC was extracted from peripheral blood or resected normal lung tissue, either fresh-frozen or paraffin embedded [17—19]. In 166 cases, DNA was extracted from more than one source (94 cases with paraffin-extracted and fresh tissue-extracted DNA, 72 cases with blood lymphocyte-extracted and fresh tissue-extracted DNA). In such cases, all available DNA samples were genotyped and the results were found to be identical, reinforcing the reproducibility of genotyping methodology.

PCR-based restriction fragment length polymorphism (RFLP) methods were used to analyze *CYP1A1* genotypes at the Ile462Val site as described [20]. About 10% of samples were randomly selected and retested for consistency. Genotyping was performed blinded to case/control status. Odds ratios (OR) and 95% confidence intervals (CI) were calculated using logistic regression analysis where log odds of lung cancer was adjusted for smoking (pack-years) and age (as continuous variables) and sex (as a categorical variable). In order to detect important differences in population subgroups, stratification by subgroup analysis of clinically relevant factors was performed. All tests were two-sided and a p value of 0.05 or less was considered significant. Formal tests of interaction were tested for significance by fit models comparing −2 log (likelihood) in models that included main effects (*CYP1A1* genotype and the potential high-risk variable) with the interaction term. Interactions were tested for significance by fit models comparing −2 log (likelihood) in models that included main effects (gene and smoking) with and without the interaction term. Data were analyzed using the SPSS for Windows Version 11.5 (SPSS Inc., Chicago, Illinois, USA). The population attributable risk (PAR%) for NSCLC of the *CYP1A1* valine allele, was estimated as follows: $\mathrm{PAR\%} = 1-(1/((p^2 \times g_2)+(2pq \times g_1)+(q^2))$, where p is the frequency of the variant in the general population, $q = 1-p$, and g_1 and g_2 are the estimated OR of the association between the heterozygous genotype and homozygous variant genotype respectively and the presence of NSCLC [21].

Results

Initially, to confirm the validity of combining the two subgroups of controls, the *CYP1A1* valine allele

Table 1. Characteristics of the study subjects

	NSCLC	Controls (All)[a]	Controls	
			COPD	Smokers
N	1050	581	358	223
Age (years)[b]				
Mean Age (SD)	63.3 (±9.4)	62.7 (±12.9)	68.5 (±9.0)	53.4 (±12.9)
Range	29–92	18–87	43–87	18–85
Sex[b,c]				
Male	760 (72.4)	396 (68.2)	232 (64.8)	164 (73.5)
Female	290 (27.6)	185 (31.8)	126 (35.2)	59 (26.5)
Pack-Years[b,d]				
Mean (SD)	52.1 (±37.3)	52.0 (±33.0)	58.0 (±34.9)	42.4 (±27.1)
Range	0–400	0.8–246.2	0.8–220.0	1.6–246.2
Status[c,e]				
Never	58 (5.5)	0 (0.0)	0 (0.0)	0 (0.0)
Former	150 (13.7)	298 (47.5)	248 (65.3)	50 (20.2)
Current	881 (80.5)	329 (52.5)	132 (34.7)	197 (79.8)
Data Missing[c]	5 (0.5)	0 (0.0)	0 (0.0)	0 (0.0)
Histologic Type[c]				
Adenocarcinoma	479 (45.6)			
Squamous cell carcinoma	479 (45.6)			
Others[f]	92 (8.8)			
TNM Staging[c]				
I	537 (51.1)			
II	237 (22.6)			
III	197 (18.8)			
IV	38 (3.6)			
Unknown	41 (3.9)			
CYP1A1 Ile462Val genotype				
Ile/Ile	958 (91.2)	552 (95.0)	341 (95.3)	211 (94.6)
Ile/Val	84 (8.0)	27 (4.7)	16 (4.5)	11 (4.9)
Val/Val	8 (0.8)	2 (0.3)	1 (0.3)	1 (0.4)

[a] As there was no significant difference in CYP1A1 genotype frequency between COPD patients and healthy smokers ($p > 0.05$), the two control populations were combined to form the control group.

[b] There was no significant difference in pack-years, age, or sex between cases and controls ($p = 0.87$, $p = 0.05$, and $p = 0.07$, respectively), χ^2 or t test. All statistical tests were two-sided.

[c] Presented as number (%).

[d] Pack-years (a measure of cumulative smoking exposure) was defined as the average number of packs (20 cigarettes/pack) of cigarettes smoked per day multiplied by the number of years of smoking.

[e] Never = less than 100 cigarettes in lifetime; Former = smoking cessation for one year or more at time of sample collection; Current = current smoker or smoking cessation less than one year at time of sample collection.

[f]Comprising of large cell carcinoma (n = 36), adenosquamous carcinoma (n = 32), unspecified NSCLC (n = 21).

frequency was compared between COPD patients and healthy smokers. There was no significant difference in the variant frequency between these two subgroups of controls ($p > 0.05$, data not shown) thus the two subgroups were combined to form one control group. The genotype frequencies are listed in Table 1. Genotype frequencies in the case and control population were in Hardy–Weinberg equilibrium and, similar to other Caucasian populations, the combined frequency of *CYP1A1* Ile/Val and Val/Val genotypes in the control group was 5% [22]. The very rare Val/Val homozygotes were grouped with the Ile/Val heterozygotes for statistical analysis (Table 2). The study had more than 80% statistical power to detect an odds ratio of 1.8 or more for the variant allele with an $@ = 0.05$.

The *CYP1A1* Ile/Val-Val/Val genotype was associated with a significantly increased risk of NSCLC when compared to controls (OR = 1.85; 95% CI = 1.20–2.85; $p = 0.005$, logistic regression) after adjustment for the confounding effects of smoking exposure, age, and sex. The population attributable risk percent of the heterozygous and homozygous valine variant genotypes of *CYP1A1* Ile462Val for the development of NSCLC was calculated as 4%.

Table 2. Association of *CYP1A1* Ile462Val polymorphism and lung cancer risk

	Ile462Val Genotype[a]	Case[b]	Control[b]	OR (95%CI)
Overall				
Crude[c]	Ile/Ile	958 (91.2)	552 (95.0)	1.0
	Ile/Val, Val/Val	92 (8.8)	29 (5.0)	1.83 (1.19–2.81)
Adjusted[d]	Ile/Ile	958 (91.2)	552 (95.0)	1.0
	Ile/Val, Val/Val	92 (8.8)	29 (5.0)	1.85 (1.20–2.85)
Sex[d]				
Male	Ile/Ile	695 (91.4)	372 (93.9)	1.0
	Ile/Val, Val/Val	65 (8.6)	24 (6.1)	1.47 (0.91–2.39)
Female	Ile/Ile	263 (90.7)	180 (97.3)	1.0
	Ile/Val, Val/Val	27 (9.3)	5 (2.7)	4.58 (1.69–12.42)
Pack-Years[e,f]				
< median (46)	Ile/Ile	466 (90.8)	288 (96.0)	
	Ile/Val-Val/Val	47 (9.2)	12 (4.0)	2.44 (1.27–4.68)
≥ median (46)	Ile/Ile	492 (91.6)	264 (94.0)	
	Ile/Val-Val/Val	45 (8.4)	17 (6.0)	1.46 (0.80–2.56)
Age[g,h]				
< median (63)	Ile/Ile	430 (89.2)	254 (95.1)	1.0
	Ile/Val, Val/Val	52 (10.8)	13 (4.9)	2.52 (1.33–4.79)
≥ median (63)	Ile/Ile	528 (93.0)	298 (94.9)	1.0
	Ile/Val, Val/Val	40 (7.0)	16 (5.1)	1.27 (0.69–2.32)
AC[d]	Ile/Ile	426 (88.9)	551 (95.0)	1.0
	Ile/Val, Val/Val	53 (11.1)	29 (5.0)	2.38 (1.49–3.82)
SCC[d]	Ile/Ile	450 (93.9)	551 (95.0)	1.0
	Ile/Val, Val/Val	29 (6.1)	29 (5.0)	1.22 (0.71–2.08)

[a] Due to the low frequency of Val/Val genotypes in *CYP1A1*, risk estimates associated by Ile/Ile *versus* Ile/Val + Val/Val.
[b] Presented as number (%).
[c] p Values and OR not adjusted for confounding factors.
[d] p Values and OR adjusted for pack years, age and sex.
[e] p Values and OR adjusted for age and sex.
[f] Cumulative tobacco-smoke exposure stratified around the median of 46 pack-years.
[g] Age stratified around the median age of 63 years.
[h] p values and OR adjusted for pack years and sex.

Stratification for the two main cell-types of lung cancer revealed a greater increase in risk for adenocarcinoma (AC) (OR = 2.38; 95% CI = 1.49–3.82; $p < 0.001$) than for SCC (OR = 1.22; 95% CI = 0.71–2.08; $p = 0.470$) (Table 2).

Due to reports of gender differences in lung cancer susceptibility, the *CYP1A1* Ile462Val genotype was stratified around sex. In nested analyses, the risk of NSCLC in subjects heterozygous or homozygous for the exon 7 polymorphism was statistically significant in females (OR = 4.58; 95% CI = 1.69–12.42; $p = 0.003$) but not in males (OR = 1.47; 95% CI = 0.97–2.39; $p = 0.118$). A formal test of interaction however, between *CYP1A1* Ile462Val and sex was not significant ($p > 0.05$, data not shown).

Cumulative tobacco-smoke exposure and age are known risk factors for lung cancer. Thus, cases and controls were stratified to determine if the risk effect of *CYP1A1* Ile462Val was modified by pack-years smoked or age. To investigate the interaction of *CYP1A1* Ile462Val genotype and smoking exposure, study samples were stratified around the median of 46 pack-years. The *CYP1A1* Ile/Val-Val/Val genotype was associated with a greater risk of NSCLC in subjects with a cumulative tobacco-smoke exposure less than the median (OR = 2.44; 95% CI = 1.27–4.68; $p = 0.007$) while in subjects with more than the median tobacco-smoke exposure, there was no significant association between the *CYP1A1* Ile/Val-Val/Val genotype and lung cancer risk (OR = 1.46; 95% CI = 0.80–2.56, $p = 0.222$) (Table 2). A formal test of interaction between *CYP1A1* Ile462Val and tobacco-smoke exposure (stratified around the median) was not significant ($p > 0.05$, data not shown).

Stratification around the median age of 63 years found *CYP1A1* Ile/Val-Val/Val individuals younger than the median age at a significantly higher risk of lung cancer (OR = 2.52; 95% CI = 1.33–4.79; $p = 0.005$) with no effect observed in *CYP1A1* Ile/Val-Val/Val individuals older than the median age (OR = 1.27; 95%

CI = 0.69–2.32; $p = 0.445$) (Table 2). A formal test of interaction between CYP1A1 Ile462Val and age was not significant ($p > 0.05$, data not shown).

Additional analyses of combined effects of risk factors were not performed due to small sample sizes.

Discussion

To resolve the uncertainty of increased lung cancer risk with the *CYP1A1* Ile462Val in Caucasians, we studied the largest ethnically homogenous population to date and found the *CYP1A1*2C* allele (heterozygous and variant homozygous genotypes) was associated with an increased risk of NSCLC (OR = 1.85) in tobacco smokers. Results from this single study substantiate a recent pooled analysis of similar size based on a heterogeneous population [16], in addition to being biologically consistent with the higher enzymatic activity reported for the valine variant [8–10].

Relative to histological subtypes, a significant effect for the *CYP1A1*2C* allele was observed in ACs (OR = 2.38) but not in SCCs (OR = 1.22). A Japanese study also reported a risk effect within ACs, particularly poorly differentiated carcinomas [11]. This difference, together with data linking tobacco-specific nitrosamines to lung ACs and tobacco smoke PAHs to SCCs, adds emphasis to the growing biological and epidemiological evidence that histological subtypes of NSCLC are distinct etiological entities that should be analyzed separately [23–25].

Previous studies have reported a stronger effect of the *CYP1A1*2C* allele in females as observed in our study with the odds ratio in females more than three times greater than in males (OR = 4.58 and OR = 1.47, respectively) [15, 16]. Recent epidemiological and biochemical studies have suggested increased susceptibility to tobacco carcinogens in women compared to men [26–28]. *CYP1A1* mRNA expression in the lung has been observed to be more than two-fold higher in female smokers compared with male smokers [29], a difference possibly due to the effect of circulation estrogens, which have been shown to induce expression of PAH-metabolizing enzymes, such as CYP1A1, thereby increasing metabolic activation of carcinogens [15, 30].

The *CYP1A1*2C* allele also conferred a heightened risk for smokers with lower cumulative tobacco-smoke exposure (less than the median of 46 pack-years), and younger individuals (less than the median age of 63 years). A possible explanation is that genetic susceptibility factors would be most apparent in those with modest conventional risk factors (age and smoking). Others have also reported that variants within the *CYP1A1* gene were associated with a higher risk of lung cancer in individuals with less smoking exposure [11, 31–33] as well as a younger age [34].

To date, four polymorphic variants have been described in the CYP1A1 gene: the exon 7 A2455G variant (*CYP1A1*2C*) investigated in this report, a T3801C variant in the 3'end (*CYP1A1*2A*), a C2453A substitution (*CYP1A1*4*) also located in exon 7, and an African-American specific T3205C variant (*CYP1A1*3*) located in intron 7. The former two variants have been shown to be in close linkage disequilibrium and lead to a more inducible form of CYP1A1 with increased enzyme activity [10, 35]. Considering the known linkage effect, and the comparable genotype effect observed in the *CYP1A1*2C* polymorphism in our study population and CYP1A1*2A in another large study of Caucasians [36], we now intend to investigate the combined effect of these two polymorphisms, in addition to other known CYP1A1 polymorphisms, in our study population through haplotyping. Haplotyping analyses facilitate the fine-scale mapping of susceptibility genes through linkage disequilibrium analysis of the surrounding markers [37]. This allows characterization of the effects of genetic variation across the entire gene to subsequently allow associations to be made between different haplotypes and the disease. A recent haplotype study found the variant frequency to differ significantly between prostate cancer when compared to controls [38] but, to date, the only haplotyping study performed in lung cancer examined the promoter region only and consequently did not include the exon 7 variant A2455G investigated in this report [39].

Potential limitations of this study should be addressed. The design of this study may generate a possible vulnerability to a misclassification bias from the control group – at risk individuals who may develop lung cancer in time. However this bias would be expected to reduce rather than increase the strength of association. A second polymorphism in exon 7, a base substitution of cytosine to adenine at position 2453 leading to the Thr461Asn polymorphism (*CYP1A1*4*) [13], may result in genotype misclassification in studies using genotyping methods that cannot distinguish between base changes at positions 2455 (Ile462Val) and 2453 (Thr461Asn). The RFLP method used in this study however, selectively digests in the presence of the guanine nucleotide (the valine variant) irrespective of the nucleotide at position 2453. Thus, we can be confident that polymorphic effects observed in this analysis are specifically associated with *CYP1A1* Ile462Val.

In conclusion, this large study provides evidence for *CYP1A1* as a susceptibility gene for tobacco smokers in lung cancer, confirming a recent pooled analysis of past

studies of inadequate sample size [16]. Whilst genetic epidemiological evidence should be interpreted cautiously and in the light of previous results and *in vitro* data, these data now consistently implicate *CYP1A1* Ile462Val in modifying lung cancer risk. Whilst the increased risk is modest, the potential effect on a population basis is highly relevant with a calculated population-attributable risk of 4%, which indicates the proportion of NSCLC in smokers attributable to the *CYP1A1* valine allele, *CYP1A1*2C*. The effect of the exon 7 *CYP1A1* polymorphism in lung cancer risk – especially in women, and younger and moderate smokers – may assist in risk stratification for early lung cancer detection and prevention efforts such as low dose CT screening and chemoprevention. Further research using adequately powered studies is needed to test the hypothesis that risk is modified by gender, age, and tobacco-smoke exposure. In addition, it would be beneficial to investigate the role of the other polymorphic variants present within the CYP1A1 gene and explore possible linkage effects between the variants.

Acknowledgements

We thank the physicians, surgeons, and pathologists at The Prince Charles Hospital; the patients and donors who participated in this study; Dr David Duffy and Dr Harry Bartlett for assistance in statistical analysis; Ainsley M Tunnicliffe and Jessie M Kelly for their help in DNA extractions; Linda H Passmore for recruiting subjects and confirming clinical data. This work was supported by The Prince Charles Hospital Foundation and The Queensland Cancer Fund.

References

1. Shopland DR, Eyre HJ, Pechacek TF (1991) Smoking-attributable cancer mortality in 1991: is lung cancer now the leading cause of death among smokers in the United States? *J Natl Cancer Inst* **83**(16): 1142–1148.
2. Nebert DW (1980) *Human Genetic Variation in the Enzymes of Detoxification*. New York: Academic Press.
3. Bouchardy C, Benhamou S, Jourenkova N, Dayer P, Hirvonen A (2001) Metabolic genetic polymorphisms and susceptibility to lung cancer. *Lung Cancer* **32**(2): 109–112.
4. Kouri RE, McKinney CE, Slomiany DJ, Snodgrass DR, Wray NP, McLemore TL (1982) Positive correlation between high aryl hydrocarbon hydroxylase activity and primary lung cancer as analyzed in cryopreserved lymphocytes. *Cancer Res* **42**(12): 5030–5037.
5. Jacquet M, Lambert V, Baudoux E, Muller M, Kremers P, Gielen J (1996) Correlation between P450 CYP1A1 inducibility, MspI genotype and lung cancer incidence. *Eur J Cancer* **32A**(10): 1701–1706.
6. Stucker I, Jacquet M, de Waziers I, *et al.* (2000) Relation between inducibility of CYP1A1, GSTM1 and lung cancer in a French population. *Pharmacogenetics* **10**(7): 617–627.
7. Hayashi S, Watanabe J, Nakachi K, Kawajiri K (1991) Genetic linkage of lung cancer-associated MspI polymorphisms with amino acid replacement in the heme binding region of the human cytochrome P450IA1 gene. *J Biochem (Tokyo)* **110**(3): 407–411.
8. Ingelman-Sundberg M, Johansson I, Persson I, *et al.* (1992) Genetic polymorphism of cytochromes P450: interethnic differences and relationship to incidence of lung cancer. *Pharmacogenetics* **2**(6): 264–271.
9. Kawajiri K, Nakachi K, Imai K, Watanabe J, Hayashi S (1993) The CYP1A1 gene and cancer susceptibility. *Crit Rev Oncol Hematol* **14**(1): 77–87.
10. Crofts F, Taioli E, Trachman J, *et al.* (1994) Functional significance of different human CYP1A1 genotypes. *Carcinogenesis* **15**(12): 2961–2963.
11. Nakachi K, Hayashi S, Kawajiri K, Imai K (1995) Association of cigarette smoking and CYP1A1 polymorphisms with adenocarcinoma of the lung by grades of differentiation. *Carcinogenesis* **16**(9): 2209–2213.
12. Kawajiri K, Eguchi H, Nakachi K, Sekiya T, Yamamoto M (1996) Association of CYP1A1 germ line polymorphisms with mutations of the p53 gene in lung cancer. *Cancer Res* **56**(1): 72–76.
13. Cascorbi I, Brockmoller J, Roots I (1996) A C4887A polymorphism in exon 7 of human CYP1A1: population frequency, mutation linkages, and impact on lung cancer susceptibility. *Cancer Res* **56**(21): 4965–4969.
14. Sugimura H, Wakai K, Genka K, *et al.* (1998) Association of Ile462Val (Exon 7) polymorphism of cytochrome P450 IA1 with lung cancer in the Asian population: further evidence from a case-control study in Okinawa. *Cancer Epidemiol Biomarkers Prev* **7**(5): 413–417.
15. Dresler CM, Fratelli C, Babb J, Everley L, Evans AA, Clapper ML (2000) Gender differences in genetic susceptibility for lung cancer. *Lung Cancer* **30**(3): 153–160.
16. Le Marchand L, Guo C, Benhamou S, *et al.* (2003) Pooled analysis of the CYP1A1 exon 7 polymorphism and lung cancer (United States). *Cancer Causes Control* **14**(4): 339–346.
17. Miller SA, Dykes DD, Polesky HF (1988) A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res* **16**(3): 1215.
18. Levi S, Urbano-Ispizua A, Gill R, *et al.* (1991) Multiple K-ras codon 12 mutations in cholangiocarcinomas demonstrated with a sensitive polymerase chain reaction technique. *Cancer Res* **51**(13): 3497–3502.
19. Fong KM, Kida Y, Zimmerman PV, Ikenaga M, Smith PJ (1995) Loss of heterozygosity frequently affects chromosome 17q in non-small cell lung cancer. *Cancer Res* **55**(19): 4268–4272.
20. Blomeke B, Bennett WP, Harris CC, Shields PG (1997) Serum, plasma and paraffin-embedded tissues as sources of DNA for studying cancer susceptibility genes. *Carcinogenesis* **18**(6): 1271–1275.
21. Rothman KJ (2002) *Epidemiology: An Introduction*. New York, N.Y.: Oxford University Press Inc; 2002.
22. Smith GB, Harper PA, Wong JM, *et al.* (2001) Human lung microsomal cytochrome P4501A1 (CYP1A1) activities: impact of smoking status and CYP1A1, aryl hydrocarbon receptor, and glutathione S-transferase M1 genetic polymorphisms. *Cancer Epidemiol Biomarkers Prev* **10**(8): 839–853.

23. Devesa SS, Shaw GL, Blot WJ (1991) Changing patterns of lung cancer incidence by histological type. *Cancer Epidemiol Biomarkers Prev* **1**(1): 29–34.
24. Yokoyama S, Yamakawa K, Tsuchiya E, Murata M, Sakiyama S, Nakamura Y (1992) Deletion mapping on the short arm of chromosome 3 in squamous cell carcinoma and adenocarcinoma of the lung. *Cancer Res* **52**(4): 873–877.
25. Sato S, Nakamura Y, Tsuchiya E (1994) Difference of allelotype between squamous cell carcinoma and adenocarcinoma of the lung. *Cancer Res* **54**(21): 5652–5655.
26. Engeland A, Andersen A, Haldorsen T, Tretli S (1996) Smoking habits and risk of cancers other than lung cancer: 28 years' follow-up of 26,000 Norwegian men and women. *Cancer Causes Control* **7**(5): 497–506.
27. Zang EA, Wynder EL (1996) Differences in lung cancer risk between men and women: examination of the evidence. *J Natl Cancer Inst* **88**(3–4): 183–192.
28. Prescott E, Osler M, Hein HO*, et al.* (1998) Gender and smoking-related risk of lung cancer. The Copenhagen Center for Prospective Population Studies. *Epidemiology* **9**(1): 79–83.
29. Mollerup S, Ryberg D, Hewer A, Phillips DH, Haugen A (1999) Sex differences in lung CYP1A1 expression and DNA adduct levels among lung cancer patients. Cancer Res **59**(14): 3317–3320.
30. Siegfried JM (2001) Women and lung cancer: does oestrogen play a role? *Lancet Oncol* **2**(8): 506–513.
31. Nakachi K, Imai K, Hayashi S, Watanabe J, Kawajiri K (1991) Genetic susceptibility to squamous cell carcinoma of the lung in relation to cigarette smoking dose. *Cancer Res* **51**(19): 5177–5180.
32. Ishibe N, Wiencke JK, Zuo ZF, McMillan A, Spitz M, Kelsey KT (1997) Susceptibility to lung cancer in light smokers associated with CYP1A1 polymorphisms in Mexican- and African-Americans. *Cancer Epidemiol Biomarkers Prev* **6**(12): 1075–1080.
33. Chen S, Xue K, Xu L, Ma G, Wu J (2001) Polymorphisms of the CYP1A1 and GSTM1 genes in relation to individual susceptibility to lung carcinoma in Chinese population. *Mutat Res* **458**(1–2): 41–47.
34. Taioli E, Gaspari L, Benhamou S*, et al.* (2003) Polymorphisms in CYP1A1, GSTM1, GSTT1 and lung cancer below the age of 45 years. *Int J Epidemiol* **32**(1): 60–63.
35. Nerurkar PV, Okinaka L, Aoki C*, et al.* (2000) CYP1A1, GSTM1, and GSTP1 genetic polymorphisms and urinary 1-hydroxypyrene excretion in non-occupationally exposed individuals. *Cancer Epidemiol Biomarkers Prev* **9**(10): 1119–1122.
36. Vineis P, Veglia F, Benhamou S*, et al.* (2003) CYP1A1 T3801 C polymorphism and lung cancer: a pooled analysis of 2451 cases and 3358 controls. *Int J Cancer* **104**(5): 650–657.
37. Rannala B, Slatkin M (1998) Likelihood analysis of disequilibrium mapping, and related problems. *Am J Hum Genet* **62**(2): 459–473.
38. Chang BL, Zheng SL, Isaacs SD*, et al.* (2003) Polymorphisms in the CYP1A1 gene are associated with prostate cancer risk. *Int J Cancer* **106**(3): 375–378.
39. Han W, Pentecost BT, Spivack SD (2003) Functional evaluation of novel single nucleotide polymorphisms and haplotypes in the promoter regions of CYP1B1 and CYP1A1 genes. *Mol Carcinog* **37**(3): 158–169.