

Genetic polymorphisms of phase I and phase II metabolic enzymes as modulators of lung cancer susceptibility

P. Mota · H. C. Silva · M. J. Soares · A. Pego ·
M. Loureiro · C. Robalo Cordeiro · F. J. Regateiro

Received: 9 August 2014 / Accepted: 21 October 2014 / Published online: 12 November 2014
© Springer-Verlag Berlin Heidelberg 2014

Abstract

Objectives Tobacco exposure remains the main etiologic factor for lung cancer (LC). Interactions between environment and individual genetic profile are particularly important for this disease. The aim of this study was to evaluate the contribution of *CYP1A1*2A*, *CYP1A1*2C*, *CYP2D6*4*, *GSTP1*, *GSTM1*, *GSTT1* and *NAT2* polymorphisms for the susceptibility to LC in a Portuguese population considering their demographic and clinical characteristics.

Materials and methods A total of 200 LC and 247 controls subjects from the Centre of Portugal were studied. Clinical and demographic characteristics were collected from clinical files and by individual questionnaires.

Polymorphisms of *CYP1A1*2A*, *CYP1A1*2C*, *CYP2D6*4*, *GSTP1*, *GSTM1*, *GSTT1* and *NAT2* were genotyped using PCR–RFLP, PCR multiplex, ARMS and real time.

Results Gender, family history of cancer, smoke cessation and alcohol consumption were independent risk factors ($p < 0.05$). Associations found between phases I and II genes and LC population reveal a sex dependent distribution. Logistic regression analysis demonstrates that enhanced activation by CYPs, associated by reduced or loss of function of phase II enzymes, can lead to a greater risk. *GSTP1* and *NAT2* polymorphisms studied have a significant contribution for the histological tumour types and the presence of metastases, at time of diagnosis, respectively, when males with smoking habits were considered.

Conclusion Multiple interactions between environment and individual characteristics are clearly associated to this disease. Variants of the detoxification genes may act synergistically contributing to this disease and modifying the risk posed by smoking and sex. The *GSTT1*0* and *GSTP1 (Ile462Val)* might contribute to the malignant phenotype through different mechanisms.

P. Mota (✉)

Unit of Medical Genetics, Faculty of Medicine, University of Coimbra, Coimbra, Portugal
e-mail: motapaula8@gmail.com

H. C. Silva · F. J. Regateiro

Unit of Medical Genetics, Faculty of Medicine, University of Coimbra and CIMAGO (Centre of Investigation of Cancer, Genetics and Environment), Coimbra, Portugal

M. J. Soares

Unit of Psychology, Faculty of Medicine, University of Coimbra, Coimbra, Portugal

A. Pego · M. Loureiro

Unit of Pneumology, Hospital of the University of Coimbra and Centre of Pneumology, Coimbra, Portugal

C. R. Cordeiro

Unit of Pneumology, Hospital of the University of Coimbra, Coimbra, Portugal

C. R. Cordeiro

Faculty of Medicine, University of Coimbra and Centre of Pneumology, Coimbra, Portugal

Keywords Lung cancer · Polymorphisms · Alcohol consumption · Smoking habits · Sex

Introduction

Lung cancer (LC) was the fourth most common cancer (410,000) and the most common cause of death (353,000) in the European Union (EU) in 2012 (Ferlay et al. 2013). Although Portugal has one of the lowest incidence rates in the EU, they have increased dramatically in the last 3 years, with values considered to be disturbing for women and the leading cause of death within males (Ferlay et al. 2013).

For the World Health Organization (WHO), the demographic evolution and the exposure to risk factors will determine the increased incidence of malignancies in the coming years.

The most important risk factor for LC is tobacco smoking (Hecht 2002; Tyczynski et al. 2003). Nevertheless, the response of the organism to carcinogen exposure depends on its metabolizing capacities modulated by drug metabolizing enzymes (DME). Interindividual variability found in these enzymes activities influences the biological responses to xenobiotic exposure, cancer susceptibility and eventual adverse drug reactions (Božina et al. 2009).

Cytochrome P450 (CYP) enzymes can metabolically activate carcinogenic compounds to form genotoxic electrophilic intermediates. Among these, CYP1A1 is responsible for the metabolic activation of aryl hydrocarbon carcinogens, such as benzo(a)pyrene, aromatic amines and polycyclic aromatic hydrocarbons (PAH), while CYP2D6 is a potential susceptibility factor for certain environmental diseases (Božina et al. 2009; Preissner et al. 2013). The *CYP1A1**2A (6235T>C; Msp1 polymorphism) and *CYP1A1**2C (4889A>G; *Ile462Val*) alleles have been associated with an increased activity towards PAH and aromatic amines which results in an increased amount of highly reactive genotoxic metabolites binding to DNA, producing DNA adducts capable of inducing mutation and initiating carcinogenesis (Božina et al. 2009; Kawajiri et al. 1990). Significant associations between these activities and LC risk were demonstrated in Asian studies, where genotype frequencies of these *CYP1A1* alleles were established between 30 and 40 % (Kawajiri et al. 1993). In Caucasians, *CYP2D6* is highly polymorphic, with *CYP2D6**2A (32.4 %), *CYP2D6**4 (20.7 %) and *CYP2C9**2 (16.0 %) responsible for a decrease in metabolism (Preissner et al. 2013). *CYP2D6**4 allele (1846G>A) is the most frequent and is responsible for 70–90 % of all poor metabolizers (PM; Preissner et al. 2013).

Glutathione S-transferases (GSTs) are phase II xenobiotic metabolizing enzymes, located in cytosol and involved in both biotransformation of electrophilic compounds and oxidative stress (Board and Menon 2013; Hayes and Strange 2000; Holley et al. 2007). Polymorphisms have been identified for *GSTM1*, *GSTP1* and *GSTT1* genes leading to possible impaired activity for the elimination of carcinogenic compounds and increased cancers' risk (Hayes and Strange 2000). *GSTM1* is referred to be involved in the detoxification of tobacco-related carcinogens, such as epoxides and hydroxylated metabolites of benzo(a)pyrene, whereas *GSTT1* is involved in the biotransformation of several low molecular weight substrates such as dichlorometano, methyl bromide and also ethylene oxides and butadiene, both constituents of tobacco smoke (Hayes and Strange 2000). In addition to its role in detoxication

processes, *GSTP1* can catalyse bioactivation and appears to be involved in response to cellular stress (exogenous and/or endogenous oxidative stress) and cell cycle control (Holley et al. 2007).

N-acetyltransferase (NAT) are phase II enzymes that catalyse acetylation of aromatic amines and hydrazines, including carcinogenic compounds and therapeutic drugs, and act as modulators of drug activity and carcinogen detoxifying agents (Gross et al. 1999). Human *NAT2* gene is highly polymorphic resulting in a wide range of N-acetylation activity in humans, from fast to slow acetylators. Several studies point to an association between slow acetylator phenotype and urinary bladder cancer (Gross et al. 1999; Borlak and Reamon-Buettner 2006). However, a role for *NAT2* acetylation status and LC is still unclear. The exposure of lungs to environmental carcinogens and cigarette smoke, as well as the influence of ethnic and genetic differences may contribute to the controversial results (Borlak and Reamon-Buettner 2006).

The main goal of this study was to evaluate the contribution of each polymorphism for the LC susceptibility when demographic features were considered and a possible association with the clinical features, at time of diagnosis. Additionally, genotype combinations were performed in order to identify a prevalent one within LC population which might modulate the risk for the disease.

Populations and methods

Populations

This case–control study was conducted in 200 LC patients attending the Pneumology Department of Coimbra University Hospitals and 247 cancer-free controls. All subjects were Caucasians from the Centre Region of Portugal. Controls were randomly selected from individuals admitted to the same department, without diagnostic of cancer.

The study was performed with the approval of the Ethical Committee of Coimbra University Hospitals and Faculty of Medicine. For all patients, LC diagnosis was confirmed by histological studies of biopsies or surgical pieces. For both populations, a questionnaire was answered, containing the following data, age, sex, family history, smoking habits, alcohol consumption and occupational exposure. Family history indicates the presence of first relatives with any type of sporadic cancer. Hereditary cancer syndromes were excluded.

Smoking status was categorized as never smokers, former smokers (stopped smoking until 1 year before diagnosis) and current smokers. Smoking intensity (pack-years) was defined as packs units smoked per day plus years as a smoker. Smoking cessation was categorized

considering that the lung epithelium takes about 15 years to be renewed in order to be considered normal functioning. Time consumption was also evaluated. Alcohol consumption was defined as usual extra-meals alcohol consumption. Occupational exposure was considered taking into account the definition in IARC monograph of 2013 [<http://monographs.iarc.fr/ENG/Classification/>]. Clinical variables (location, histological type, disease stage and survival) were collected from clinical files.

Genotyping

For genetic analysis, 5 mL of peripheral blood was collected from each subject in a tube with EDTA. Genomic DNA was isolated from blood samples using standard procedures.

*CYP*s genotypes were identified by PCR–RFLP assay. Allelic variants *CYP1A1**2A (rs4646903) and *CYP1A1**2C (rs1048943) were identified as described by Borlak and Reamon-Buettner 2006 and Hayashi et al. 1991, respectively (Hayashi et al. 1991; Nakachi et al. 1991). Identification of *CYP2D6**4 allele (rs3892097) was assessed according to the method described by (Nakachi et al. 1991; Gough et al. 1990).

GSTP1 Ile105Val (rs1695) was performed by real time PCR using TaqMan® Drug Metabolism SNP Genotyping Assay (Applied Biosystems; ID: C_3237198_20; Pool-Zobel et al. 2005). *GSTT1**0 (rs2266636) and *GSTM1**0 (rs737497) polymorphisms were determined by PCR multiplex as previously described (Saxena et al. 2012). Individuals' homozygous for the null allele was considered “null”, and other genotypes were described as “present” (+).

Identification of *NAT2* genotype associated with transitions 481C > T (rs1799929), 803A > G (rs1208), 590G > A (rs1799930) and 857G > A (rs1799931) were identified by PCR–RFLP, according to, Smith et al. (1997) allowing phenotypic classification of each individual as slow or fast acetylator, as described (Smith et al. 1997).

Samples were randomly genotyped in duplicate, and genotyping was repeated when discrepancies were found.

Statistical analysis

Statistical analysis was performed using the IBM SPSS STATISTICAL 20. A two-sided $p \leq 0.05$ was considered statistically significant in the analysis. A Chi-square test was used to determine whether the genotypes were in Hardy–Weinberg equilibrium. The Chi-square independence test was applied to characterize the distribution of all demographic variables considered, as well as the genetic polymorphisms between patients and controls. The Chi-square independence test or Fisher's exact test was used for evaluation of possible associations between

clinical variables and genotypic frequencies distributions. Odds ratio (OR) and 95 % confidence intervals (CI) were calculated whenever statistical significance for Chi-square test ($p < 0.05$) was observed. Logistic regression, to adjust for confounding factors such as the demographic variables, was applied whenever Chi-square test was statistically significant. Concatenation analysis was used to evaluate prevalent combination genotypes within populations.

Heterozygous genotypes and rare allele variants were associated whenever the value of their frequencies violated the prerequisites of the statistical tests applied.

Results

Population analysis shows that mean age (\pm SD) of patients ($N = 200$) and controls ($N = 247$) were 63.06 (\pm 10.17) and 70.92 (\pm 11.24) years, respectively, with an age extent of [36;87] for patients and of [40;97] for controls. The sex ratio (male:female) was 2.92 for patients and 1.71 for controls. Patients were mainly males, 74.5 % (OR = 1.7), with a higher prevalence between 67 and 76 years old, with smoking habits, 65.5 % (OR = 2.9), a period of smoking cessation of 1 year or less at time of diagnosis 69.3 %, alcohol consumption 65.9 % (OR = 1.90) and a significant number of individuals, 31.2 %, with first-degree relatives with cancer (OR = 5.34). Patients showed no prevalence for risky occupations. Also, the controls were predominantly men, 63.2 %, with a higher occurrence between 77 and 97 years old, mainly without smoking habits, 60.4 %, and a period of smoking cessation between 1 and 15 years at time of diagnosis, 42.2 %. They were alcohol consumers, 50.4 %, with risky occupations, 51.8 %, and a reduced number of individuals, 7.2 %, with first-degree relatives with any type of cancer. Population was stratified by smoking habits, showing that patients were mainly alcohol consumers, 82.6 %, with an increased OR (OR = 4.5) when compared with values found for smoking habits (OR = 2.9) and alcohol consumption (OR = 1.9) individually. Time consumption and pack-years revealed no statistically significant association between populations (Table 1). Populations were also evaluated taking into account their sex and smoking habits, simultaneously (Table 2). Patients with smoking habits were mainly men, 61.5 %, with a statistically significant association (OR = 3.74), while patients never smokers are mostly women, 21.5 %.

Logistic regression analysis, evaluating the demographic variables that contribute most for the “lung cancer patient profile” (LCPP) show that smoking cessation for 1 year or less at time of diagnosis and the presence of first-degree relatives with cancer assigns a higher susceptibility for LC (Table 3; smoking cessation: OR = 10.55; family history:

Table 1 Demographic characteristic for patients and controls

Characteristics	Patients Number of cases (%)	Control Number of cases (%)	<i>p</i> values OR 95 % CI
Age ranges (years)			
[36;56[54 (27.0)	26 (10.5)	
[57;66[65 (32.5)	60 (24.3)	<0.001
[67;76[68 (34.0)	79 (32.0)	–
[77;97[13 (6.5)	82 (33.2)	
$\bar{x} \pm \sigma$	63.06 \pm 10.17	70.92 \pm 11.24	
SA (Samp. amplitude)	[36;87]	[40;97]	
Sex			0.010
Male	149 (74.5)	156 (63.2)	1.704
Female	51 (25.5)	91 (36.8)	[1.13;5.67]
Occupational exposure			<0.001
No	114 (70.4)	113 (51.8)	2.21
Yes	48 (29.6)	105 (48.2)	[1.44;3.39]
Family history			<0.001
Yes	58 (31.2)	19 (7.2)	5.34
No	128 (68.8)	224 (92.2)	[3.05;9.37]
Smoking habits			<0.001
Yes	131 (65.5)	97 (39.6)	2.897
No	69 (34.5)	148 (60.4)	[1.97;4.27]
Pack-years (pack unit \times years)			
<5	5 (4.0)	3 (3.2)	
[5;20[7 (5.5)	6 (6.2)	0.123
[20;50[44 (34.9)	48 (50.0)	
>50	70 (55.6)	39 (40.6)	
Time of consumption (years)			
<30	29 (23.4)	19 (20.7)	
[30;40[39 (31.5)	26 (28.3)	0.796
[40;50[35 (28.2)	27 (29.3)	
>50	21 (16.9)	20 (21.7)	
Tobacco cessation (years)			
<1	79 (69.3)	31 (28.4)	
[1;15]	22 (19.3)	46 (42.2)	<0.001
>15	13 (11.4)	32 (29.4)	
Alcohol consumption			0.002
Yes	120 (65.9)	118 (50.4)	1.903
No	62 (34.1)	116 (49.6)	[1.21;2.84]
Smokers			<0.001
Alcohol consumption			
Yes	95 (82.6)	47 (51.1)	4.5
No	20 (17.4)	45 (48.9)	[2.42;8.56]
Never smokers			
Alcohol consumption			
Yes	25 (37.3)	70 (50.0)	
No	42 (62.7)	70 (50.0)	0.087

Association analysis by Chi-square test; risk evaluation by odds ratio

OR = 12.21). Being a male and alcohol consumer also contributes to the LCPP, although with a lower risk, 3.8 and 3.15, respectively (Table 3).

Genotypic frequencies distribution of all genes studied is shown in Table 4. No associations were found between *CYP1A1*2A*, *CYP1A1*2C* and *GSTM1* genotype

Table 2 Frequencies for the populations of patients and controls when sex and smoking habits were considered

Sex/smoking habits	Smokers (N %) (current + former)	Never smokers (N %)	p values OR; 95 % CI
Male			
Patients (N = 209)	123 (82.6)	26 (17.4)	<0.001 3.74 [2.2;6.35]
Control (N = 94)	86 (55.8)	68 (44.2)	
Female			
Patients (N = 19)	8 (15.7)	43 (84.3)	0.546
Control (N = 123)	11 (12.1)	80 (87.9)	

Table 3 Logistic regression study for the evaluation of the demographic characteristics who contribute most for the lung cancer patient profile (LCPP)

Characteristics	Statistical parameters		
	p value	OR	95 % CI
Sex (male)	0.041	3.80	[1.05;13.69]
Family history (yes)	0.000	12.21	[3.81;39.14]
Alcohol consumption (yes)	0.003	3.15	[1.48;6.71]
Smoking cessation (>15 years)	0.000		
<1 year	0.000	10.55	[3.83;29.09]
[1;15] years	0.162		
Constant	0.000	0.027	

frequencies and LC patients, not even when population were stratified by sex. *CYP2D6*4* allele ($p = 0.017$) and *GSTP1 (Ile105Val)* ($p = 0.016$) reveal a statistically significant association when females were considered, while *GSTT1* null allele reveals a significant association for population as a whole (OR = 1.6). When males were considered, a statistically significant association was found for *GSTT1* null allele (OR = 1.8) and for NAT2 “slow acetylator” phenotype (OR = 1.8). The same statistical analysis was performed considering non smokers, confirming the association found for *CYP2D6*4* allele ($p = 0.043$) and *GSTP1 (Ile105Val)* ($p = 0.018$) in the female subpopulation (Table 4).

Since sex and alcohol consumption were characteristics that revealed a significant contribution for LCPP (Table 3), a logistic regression study was conducted to evaluate the contribution of those in association with genes of phases I and II (Table 5). Family history and smoking cessation were excluded because of the scarce number on some of the categories of these variables within control population.

The genotypes of each phase were combined by concatenation. Two prevalent *CYPs* combinations, *CYP1A1*2A (T/T)/CYP1A1*2C (A/A)/CYP2D6*4 (G/A)* (Patients: 19.7 %/Controls: 22.3 %) and *CYP1A1*2A (T/T)/CYP1A1*2C (A/A)/CYP2D6*4 (G/G)* (Patients: 51.1 %/Controls: 40.1 %) and three prevalent GSTs combinations, *GSTP1 (A/G)/GSTT1(-)/GSTMI(+)* (Patients:18.5 %/Controls: 7.8 %), *GSTP1(A/G)/GSTT1 (+)/GSTMI(+)*

(Patients:16.6 %/Controls: 27.2 %) and *GSTP1(A/A)/GSTT1(-)/GSTMI(+)* (Patients: 10.2 %/Controls: 6.7 %), were highlighted once their values of adjusted residuals contribute most for Chi-square test. When associated with sex and alcohol consumption, “slow acetylator” phenotype (OR = 4.0), *GSTP1 (A/G)/GSTT1 (-)/GSTMI (+)* (OR = 6.89), and *GSTP1 (A/A)/GSTT1 (-)/GSTMI (+)* (OR = 6.13) reveal a statistically significant contribution for LCPP (Table 5).

Considering the clinical characteristics, histology, stage and survival were also evaluated. Three prevalent histological types were found: adenocarcinoma, 42.4 %, epidermoid carcinoma, 26.2 % and small cell carcinoma, 22.5 %. The adenocarcinoma was the major histological type when patients were evaluated as a whole, as well as when sex was considered (male: 34.5 %; female: 65.3 %). At the time of diagnosis, patients have a tumour location mainly at the upper lobe (65.0 %) and were in stage IV (75.4 %). Survival was below 2 years for 63.4 % (Table 6).

For male subpopulation of current and former smokers, a statistically significant association was found between *GSTP1* and the tumour histology ($p = 0.042$) and between NAT2 frequencies and the presence or absence of metastases, at time of diagnosis, ($p = 0.009$; Table 7). The “fast acetylator” phenotype is prevalent in individuals without metastasis, at the time of diagnosis (24.2 %).

Discussion

Smoke is an etiologic factor for LC, with a clear dose–response relationship between LC risk and the number of cigarettes smoked per day, the degree of inhalation and the age of smoking initiation (Hecht 2002; Nakachi et al. 1991). The type of cigarette smoked, the time of consumption and cessation are also important for LC development (Hecht 2002; Bertram and Rogers 1981; Parsons et al. 2010). Environmental factors have small contribution for LC cases; nevertheless, occupational exposure to asbestos, nickel and arsenic and gases like radon may act synergistically with smoking (Hecht 2002; Tyczynski et al. 2003; Parsons et al. 2010).

Table 4 Distribution of the genotypic frequencies (phenotypic frequencies for *NAT2*) of phase I and phase II genes among patients and controls and when stratified by sex

Genes	Patient/control <i>p</i> value OR; 95 % CI	Males			Females		
		Patient <i>N</i> (%)	Control <i>N</i> (%)	<i>p</i> value OR; 95 % CI	Patient <i>N</i> (%)	Control <i>N</i> (%)	<i>p</i> value OR; 95 % CI
<i>CYP1A1</i> *2A							
T/T		117 (81.8)	108 (77.1)		37 (75.5)	56 (67.5)	
T/C	0.277	24 (16.8)	31 (22.1)	0.460	11 (22.4)	23 (27.7)	0.537
C/C		2 (1.4)	1 (0.7)		1 (2.0)	4 (4.8)	
<i>CYP1A1</i> *2C							
A/A		115 (80.4)	102 (74.5)		37 (77.1)	64 (78.0)	
A/G	0.455	27 (18.9)	35 (25.5)	0.261	10 (20.8)	17 (20.7)	0.927
G/G		1 (0.7)	0 (0.0)		1 (2.1)	1 (1.2)	
<i>CYP2D6</i> *4							
G/G		108 (75.5)	104 (73.2)		31 (64.6)	48 (57.1)	
G/A	0.104	33 (22.9)	35 (24.6)	0.942	11 (22.9)	34 (40.5)	0.017
A/A		3 (2.1)	3 (2.1)		6 (12.5)	2 (2.4)	
<i>GSTP1</i> (Ile105Val)							
A/A		59 (40.1)	56 (39.4)		12 (23.5)	23 (28.5)	
A/G	0.358	69 (46.9)	66 (46.5)	0.959	25 (49.0)	51 (63.0)	0.016
G/G		19 (12.9)	20 (14.1)		14 (27.5)	7 (8.6)	
<i>GSTT1</i>							
Present (+)	0.030	91 (66.4)	113 (78.5)	0.024	31 (67.4)	61 (72.6)	0.531
Null (–)	1.6 [1.04;2.49]	46 (33.6)	31 (21.5)	1.8 [1.08;3.13]	15 (32.6)	23 (27.4)	
<i>GSTM1</i>							
Present (+)	0.500	79 (57.7)	78 (54.2)	0.555	29 (63.0)	49 (58.3)	0.600
Null (–)		58 (42.3)	66 (45.8)		17 (37.0)	35 (41.7)	
Phenotype <i>NAT2</i>							
Rapid acetylator	0.067	23 (16.2)	38 (26.0)	0.041	13 (27.7)	23 (27.7)	0.985
Slow acetylator		119 (83.8)	108 (74.0)	1.8 [1.02;3.25]	34 (72.3)	60 (72.3)	
Never smokers							
<i>CYP2D6</i> *4							
G/G		19 (28.4)	48 (71.6)		25 (37.3)	42 (62.7)	
G/A	0.099	5 (29.4)	12 (70.6)	0.980	10 (24.4)	31 (75.6)	0.043
A/A		1 (33.3)	2 (66.7)		5 (71.4)	2 (28.6)	
<i>GSTP1</i> (Ile105Val)							
A/A		9 (29.0)	22 (71.0)		10 (31.2)	22 (68.8)	
A/G	0.260	13 (30.2)	30 (69.8)	0.742	21 (31.8)	45 (68.2)	0.018
G/G		3 (20.0)	12 (80.0)		12 (66.7)	6 (33.3)	

Association studies by Chi-square test

In this case–control study, patterns of smoking like pack-years, time of consumption and smoking cessation were evaluated. Occupational exposures and alcohol consumption were also considered as factors that could contribute synergistically for a higher susceptibility for LC.

Although patients had prevalent smoking habits (65.5 %), the pack-years ($p = 0.123$) and time consumption ($p = 0.796$) patterns were similar to those found for controls. However, when populations were analysed for values

upper than 20 pack-years, a statistically significant association was found ($p = 0.019$). After smoking cessation, risk of LC declines steadily and may never reach those who never smoke, even after decades of cessation (Bertram and Rogers 1981; Parsons et al. 2010; Wu and Sin 2011). The risk remains among 30–50 %, after 10 years of abstinence, when compared with current smokers (Parsons et al. 2010; Wu and Sin 2011). Populations were evaluated for three distinct periods of smoking cessation, and a association

Table 5 Logistic regression study for evaluation of the contribution of the phase I and phase II genotypes combinations for lung cancer patient profile (LCP), when sex and alcohol consumption, at the time diagnosis, were considered

Characteristics	Statistical parameters		
	<i>p</i> value	OR	95 % CI
Sex (female)			
Male	0.307	–	–
Alcohol consumption (no)			
Yes	0.085	–	–
Genotypes <i>GST</i>			
(<i>GSTP1(A/G)/GSTT1(+)/GSTMI (+)</i>)	0.008		
<i>GSTP1(A/G)/GSTT1(-)/GSTMI(+)</i>	0.006	6.889	[1.75;27.12]
<i>GSTP1(A/A)/GSTT1(-)/GSTMI(+)</i>	0.022	6.132	[1.30;28.98]
Genotypes <i>CYP</i>			
(<i>IAI*2A(T/T)/IAI*2C(A/A)/2D6(G/A)</i>)			
<i>IAI(T/T)/IAI*2C(A/A)/C2D6(G/G)</i>	0.736	–	–
Phenotype <i>NAT2</i> (“rapid”)			
Slow	0.033	4.007	[1.12;14.48]
Constant	0.001	0.047	

was found revealing a different behaviour between patients, with a smoking cessation period <1 year before diagnosis (69.3 %), and controls, with a smoking cessation period between 1 and 15 years (42.2 %).

The relationship between alcohol consumption and LC risk remains controversial (Benedetti et al. 2006; Freudenheim et al. 2005). This association is often assumed as a result of residual confounding factors such as tobacco (Benedetti et al. 2006). Results show an independent risk

for each one of these characteristics ($OR_{\text{smoking habits}} = 2.9/ OR_{\text{alcohol consumption}} = 1.9$) and an increased OR when both variables were considered ($OR = 4.5$), suggesting a synergistic effect. The biological mechanism responsible for the harmful effect of alcohol is still unknown, but the acetaldehyde, oxidative metabolite of a primary alcohol, has been proposed, since it seems to be carcinogenic in animal studies. Also, the oxidative effect of alcohol might be responsible for such behaviour (Freudenheim et al. 2005).

The modulation of this disease by sex-associated factors has been referred (North and Christiani 2013). The higher frequencies of women in subpopulations of never smokers with LC are confirmed, not only in this but in several other studies ($LC_{\text{never smokers}}$: males 13.0 %; females 21.5 %/ LC_{smokers} (former and current): males 61.5 %; females 4.0 %; Pallis and Syrigos 2013; Cohen et al. 2007; Siegfried 2001). It is not clear whether this female incidence is due to inherent susceptibility or to a greater contribution of risk factors, such as the presence of estrogens receptors (ER), α and β , which have a higher expression in cancerous lung tissue (Cohen et al. 2007; Siegfried 2001). The activation of ER results in an increased expression of genes involved in cellular proliferation and tumour growth (Siegfried 2001). Sex demonstrates a statistically significant association ($OR = 1.7$) with LC patients (males: 74.5 %; Table 1) not only when smoking habits are considered, but also when genotype frequencies distribution was evaluated (Tables 2 and 4).

The activities of CYP450 enzymes towards carcinogenic compounds of tobacco, as benzo(a)pyrene and procarcinogens such as arylarenes, nitroarenes and arylamines, and their involvement in DNA adduct formation justify their evaluation in LC studies (Zanger and Schwab 2013;

Table 6 Clinical characteristics of patients at the time of diagnosis, and when stratified by sex

Clinical characteristics	Patients <i>N</i> (%)	Male <i>N</i> (%)	Female <i>N</i> (%)
Histology			
Small cell carcinoma	43 (22.5)	34 (23.9)	9 (18.4)
Epidermoid carcinoma	50 (26.2)	46 (32.4)	4 (8.2)
Adenocarcinoma	81 (42.4)	49 (34.5)	32 (65.3)
Pleomorphic carcinoma	13 (6.8)	10 (7.0)	3 (6.1)
Others (neuroendocrine carc., adenosquamous carc.)	4 (2.1)	3 (2.1)	1 (2.0)
Tumour localization			
Upper lobe, right and left lung	80 (65.0)	59 (65.6)	21 (63.6)
Lower lobe, right and left lung	43 (35.0)	31 (34.4)	12 (36.4)
Stage			
Up to IIIA	17 (13.5)	12 (12.2)	5 (17.9)
IIIB	14 (11.1)	9 (9.2)	5 (17.9)
IV	95 (75.4)	77 (78.6)	18 (64.3)
Survival			
<2 years	52 (63.4)	38 (62.3)	14 (66.7)
>2 years	30 (36.6)	23 (37.7)	7 (33.3)

Table 7 Genotype and phenotype frequencies distribution among clinical variables, histology and metastases, when males with smoking habits, current and former smokers, were considered

Males (current and former smokers)	<i>GSTP1</i>			<i>p</i> value*
	A/A <i>N</i> (%)	A/G <i>N</i> (%)	G/G <i>N</i> (%)	
Histology (<i>N</i> = 102)				
Small cell carcinoma	18 (40.0)	12 (26.7)	15 (33.3)	
Epidermoid carcinoma	7 (15.6)	22 (48.9)	16 (35.5)	0.042
Adenocarcinoma	2 (16.7)	7 (58.3)	3 (25.0)	
Metastases (<i>N</i> = 114)	NAT 2 phenotype			
	“Fast acetylator”	“Slow acetylator”		
Yes <i>N</i> (%)	3 (5.8)	49 (94.2)		0.009
No <i>N</i> (%)	15 (24.2)	47 (75.8)		

* Fisher’s exact test

Christensen et al. 1997; Wright et al. 2010; Chen et al. 2011). Sex influences CYP450 expression and function (Siegfried 2001). Studies indicate that women metabolize drugs more quickly than men with implications, not only from a clinical perspective, but also when response to environmental exposure is evaluated (Siegfried 2001). The lack of associations between *CYP1A1* alleles and LC ($P_{CYP1A1*2A} = 0.277/P_{CYP1A1*2C} = 0.455$) might reflect the low incidence of both polymorphisms in Portuguese population (*CYP1A1*2A* (C/C): Patients 1.6%/Controls 2.2%; *CYP1A1*2C* (G/G): Patients 1.0%/Controls 0.5%) which are in agreement with references of other Caucasian populations, 0–4%, respectively (Wright et al. 2010; Kawajiri et al. 1990). The positive associations found in Asian populations were probably due to a higher genotype frequencies for both alleles (Sreejal et al. 2005). In accordance, a global meta-analysis with *Msp1 CYP1A1* polymorphism revealed an association with increased susceptibility for LC. However, when stratified by ethnicities, only East Asian populations confirmed the results ($p < 0.05$; Chen et al. 2011).

The association between *CYP2D6*4* and LC female population (Table 5) might be difficult to explain, since this gene is usually implicated in clinical drug metabolism (Rostami-Hodjegan et al. 1998). However, the involvement of this polymorphism with at least a *CYP1A1 Msp1* allelic variant might modulate the LC development (Christensen et al. 1997; Rostami-Hodjegan et al. 1998).

The *GSTM1*, *GSTT1* and *GSTP1* genes have been referred as cancer modulators, given their ability to regulate the conjugation of carcinogenic compounds, converting them into hydrophilic metabolites for excretion and protecting cells from oxidative stress (Hayes and Strange 2000; Josephy 2010).

No statistically significant differences were found between populations for the distribution of *GSTP1* genotype frequencies (Table 4). Nevertheless, when female subpopulation was considered ($p = 0.016$), it was found

a higher Val/Val genotype frequency in patients (27.5%) than in controls (8.6%). *GSTP1* is the most abundant isoform in the lungs and is important in the detoxification of inhaled carcinogens (Hayes and Strange 2000). The 105Val variant has been associated with a sevenfold higher activity for diol epoxides of PAH, while 105Ile variant has shown a threefold less effective for 1-chloro-2,4-dinitrobenzene (CDNB) (Hayes and Strange 2000). *GSTP1* also appears to be involved in response to cellular stress (exogenous and/or endogenous oxidative stress) and control of the cell cycle (Holley et al. 2007).

The allelic frequencies found for *GSTT1* null (Patients: 33.3%; Controls: 23.7%) were higher than the values described for Caucasians, 10–18%, but close to those described for African-Americans, 22–29% (Carlsten et al. 2008; Gao et al. 2002). The statistically significant association found between *GSTT1* null and the LC population ($p = 0.030$), even when males ($p = 0.024$) were considered (Table 4), could reflect the prevalence of smoking habits in this subpopulation, since this enzyme participates in detoxification of monohalomethanes and reactive diol epoxides (Josephy 2010). *GSTM1* catalyses the conjugation of the tripeptide GSH to PAH diol epoxide (Carlsten et al. 2008). Nevertheless, no association was found between the *GSTM1* null frequencies (Patients: 41.0%/Controls: 44.3%) and LC, even when population was stratified by sex (Table 4). Although the frequencies were according to those found for Caucasians, 38–67%, the absence of association is in disagreement with former studies, since several meta-analysis and pooled analysis have been consistent in finding an increased risk for individuals with *GSTM1* null allele (Board and Menon 2013; Sweeney et al. 2003; Carlsten et al. 2008).

NAT2 is a highly polymorphic gene (Gross et al. 1999). Controls have a higher frequency of “fast acetylators” (26.6%) when compared with patients (19.0%). A 1.7 “fast acetylator” ratio (control/patients) is observed for population as a whole or when sex is considered. Male

patients appear to be at higher risk (OR = 1.8), which might be associated to the prevalence of smoking habits and alcohol consumption in this subpopulation (Table 1). As referred by others, “fast acetylators” might have a protective effect, when risky habits are considered (Table 1; Gross et al. 1999; Borlak and Reamon-Buettner 2006).

The hypothesis that enhanced activation by phase I enzymes associated by reduced or loss of phase II function can lead to greater risk was tested in the present study by logistic regression (Table 5). Two prevalent *CYPs* combinations and three prevalent *GSTs* combinations, *GSTP1(A/G)/GSTT1(-)/GSTMI(+)* (patients:18.5 %/controls: 7.8 %), *GSTP1(A/G)/GSTT1(+)/GSTMI(+)* (patients:16.6 %/controls:27.2 %) and *GSTP1(A/A)/GSTT1(-)/GSTMI(+)* (patients: 10.2 %/controls:6.7 %), were found and used to evaluate their contribution as modulators of LC susceptibility. A statistically significant association was found for the *GSTs* combinations and “slow acetylator” phenotype when sex, alcohol consumption and *CYPs* combinations were considered. No association was found for *CYPs* combinations. Although the *CYP1A1* genotypes present in both combinations have been referenced as having a lower activity than *CYP1A1*2A* and *CYP1A1*2C*, they modulate the contribution of phase II enzymes, since their activity contribute to an increased DNA adducts formation, when conjugation and acetylation were deficient. Genotypes combinations with *GSTT1 (-)* were always prevalent in patients. The higher values found for the upper limits of the confidence intervals (95 %) may reflect the scarce number of individuals involved in this study. However, the presence of lower limits higher than one should ensure the statistical significance of the results.

Considering smoking habits as a confounding factor, association studies were performed in a never smoker’s subpopulation, stratified into females and males (Table 4; Pallis and Syrigos 2013; Cohen et al. 2007). No statistically significant values were found for males, confirming the smoking habits as a confounding factor for this subpopulation. For female never smokers, the statistically significant values found for *CYP2D6*4* and *GSTP1* suggest an important role of both genes as modulators of susceptibility to LC. This sex related pattern of differentiation might represent a synergistic effect between different environmental exposures, hormone modulation and different gene metabolizers’ activities, beyond their roles on detoxification processes. The involvement of *GSTP1* in the regulation of stress signalling and resistance to apoptosis through a JNK-dependent mechanism should be considered, since their involvement in inflammatory processes might contribute to carcinogenesis (Ritchie et al. 2007; Seyfried and Shelton 2010).

Smoking habits and sex are also important, when tumour histology and the presence of metastases at time of diagnosis were evaluated (Table 7). An association was found between *GSTP1* genotype frequencies and the prevalent histological types (0.042). This result might be associated to the fact that *GSTP1* is responsible for more than 90 % of the *GST* activity within the adult human lung epithelial cell population, suggesting an important role in the detoxification of this organ (Holley et al. 2007; Sweeney et al. 2003). The incidence of cancer can be significantly reduced, avoiding exposure to agents or conditions that provoke tissue inflammation such as smoking, alcohol, carcinogenic chemicals and ionizing radiation (Seyfried and Shelton 2010). *GSTP1* as well as *NAT2* are pieces of this complex mechanism, modulating the individual response to each agent, contributing to their elimination or harmful effect. The possible protective role of the fast acetylator phenotype associated with the absence of metastases, at diagnosis, should be explored.

The prevalence of women in subpopulation of never smokers with LC suggests a different mechanism for its development, in women. Therefore, differences in exposure to risk factors and host characteristics must be evaluated (Gao et al. 2002). An unequivocal contribution of *CYP2D6* and *GSTP1* in female never smokers should be considered, while for subpopulations of current smokers, *GSTT1* appears to have a higher contribution. This study attempts to address the contribution of *CYP450*, *GSTs* and *NAT2* activities in lungs and their metabolic balance, which may be a determinant host factor underlying LC.

Conflict of interest The authors declare no conflict of interest.

References

- Benedetti A, Parent ME, Siemietycki J (2006) Consumption of alcoholic beverages and risk of lung cancer: results from two case-control studies in Montreal, Canada. *Cancer Causes Control* 17:469–480
- Bertram J, Rogers A (1981) Recovery of bronchial epithelium on stopping smoking. *Br Med J* 283:1567–1569
- Board P, Menon D (2013) Glutathione transferases, regulators of cellular metabolism and physiology. *Biochim Biophys Acta* 1830:3267–3288
- Borlak J, Reamon-Buettner S (2006) N-Acetyltransferases 2 (*NAT2*) gene polymorphisms in colon and lung cancer patients. *BMC Med Genet* 7:58. doi:10.1186/1471-2350-7-58
- Božina N, Bradamante V, Lovrić M (2009) Genetic polymorphism of metabolic enzymes P450 (*CYP*) as a susceptibility factor for drug response, toxicity, and cancer risk. *Arh Hig Rada Toksikol* 60:217–242
- Carlsten C, Saggio GS, Frodshem AJ, Burke W, Higgins JP (2008) Glutathione S-Transferase M1 (*GSTM1*) polymorphisms and lung cancer: a literature-based systematic HuGE review and meta-analysis. *Am J Epidemiol* 167:759–774

- Chen Z, Li Z, Niu X, Yu Y, Lu S (2011) The effect of CYP1A1 polymorphism on the risk of lung cancer: a global meta-analysis based on 71 case-control studies. *Mutagenesis* 26:437–446
- Christensen PM, Gotzsche PC, Broesen K (1997) The sparteine/debrisoquine (CYP2D6) oxidation polymorphism and the risk of lung cancer: a meta-analysis. *Eur J Clin Pharmacol* 51:389–393
- Cohen S, Paré P, Man S, Sin D (2007) The growing burden of chronic obstructive pulmonary disease and lung cancer in women. *Am J Respir Crit Care Med* 176:113–120
- Ferlay J, Steliarova-Foucher E, Lortet-Tieulent J, Rosso S, Coebergh JWV, Comber H, Forman D, Bray F (2013) Cancer incidence and mortality patterns in Europe: estimates for 40 countries in 2012. *Eur J Cancer* 49:1374–1403
- Freudenheim JL, Ritz J, Smith-Warner SA, Albanes D, Bandera EV, van den Brandt P, Colditz G, Feskanich D, Goldbohm RA, Har-nack L, Miller A, Rimm E, Rohan T, Sellers T, Virtamo J, Willett W, Hunter D (2005) Alcohol consumption and risk of lung cancer: a pooled analysis of cohort studies. *Am J Clin Nutr* 82:657–667
- Gao CM, Takezaki T, Wu JZ, Li ZY, Liu YT, Li SP et al (2002) Glutathione S-Transferase M1 (GSTM1) and GSTT1 genotype, smoking, consumption of alcohol and tea and risk of esophageal and stomach cancers: a case control study of a high incidence area in Jiangsu Province, China. *Cancer Lett* 188:95–102
- Gough AC, Miles JS, Spurr NK, Moss JE, Gaedigk A, Eichelbeum M, Wolf CR (1990) Identification of the primary gene defect at the cytochrome P450 CYP2D locus. Reprinted from *Nature, Macmillan Magazines Ltd* 347:773–776
- Gross M, Kruisselbrink T, Anderson K, Lang N, McGovern P, Delongchamp R, Kadlubar F (1999) Distribution and concordance of N-acetyltransferase genotype and phenotype in an American population. *Cancer Epidemiol Biomark Prev* 8:683–692
- Hayashi S, Watanabe J, Nakachi K, Kawajiri K (1991) Genetic linkage of lung cancer –associated Msp1 polymorphism with aminoacid replacement in the heme binding region of the human cytochrome P450 1A1 gene. *J Biochem* 110:407–411
- Hayes J, Strange R (2000a) Glutathione S-transferase polymorphisms and their biological consequences. *Pharmacology* 61:154–166
- Hayes J, Strange R (2000b) Glutathione S-transferase polymorphisms and their biological consequences. *Pharmacology* 61:154–166
- Hecht SS (2002) Cigarette smoking and lung cancer: chemical mechanisms and approaches to prevention. *Lancet oncol* 3:461–469
- Holley S, Fryer A, Haycock J, Grubb A, Strange R, Hoban P (2007) Differential effects of glutathione S-transferase pi (GSTP1) haplotypes on cell proliferation and apoptosis. *Carcinogenesis* 28(11):2268–2273
- Joseph P (2010) Genetic variations in human glutathione transferase enzymes: significance for pharmacology and toxicology. *Hum Genomics Proteomics* 2010:876940
- Kawajiri K, Nakachi K, Imai K et al (1990a) Identification of genetically high risk individuals to lung cancer by DNA polymorphisms of the cytochrome P450 1A1 gene. *FEBS Lett* 263:131–133
- Kawajiri K, Nakachi K, Imai K et al (1990b) Identification of genetically high risk individuals to lung cancer by DNA polymorphisms of the cytochrome P4501A1 gene. *FEBS Lett* 263:131–133
- Kawajiri K, Nakachi K, Imai K, Watanabe J, Hayashi S (1993) The CYP1A1 gene and cancer susceptibility. *Crit Rev Oncol Hematol* 14:77–87
- 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* 51:5177–5180
- North C, Christiani D (2013) Women and lung cancer: what is new? *Semin Thorac Surg* 25:87–94
- Pallis A, Syrigos K (2013) Lung cancer in never smokers: disease characteristics and risk factors. *Crit Rev Oncol/Hematol* 88:494–503
- Parsons A, Daley A, Begh R, Aveyard P (2010) Influence of smoking cessation after diagnosis of early stage lung cancer on prognosis. Systematic review of observational studies with meta-analysis. *BMJ* 340:b5569. doi:10.1136/bmj.b5569
- Pool-Zobel B, Veeriah S, Böhmer FD (2005) Modulation of xenobiotic metabolizing enzymes by anticarcinogens—focus on glutathione S-transferases and their role as targets of dietary chemoprevention in colorectal carcinogenesis. *Mutat Res* 591(1–2):74–92
- Preissner SC, Hoffmann MF, Preissner R, Dunkel M, Gewiess A, Preissner S (2013) Polymorphic cytochrome P450 enzymes (Cyps) and their role in personalized therapy. *PLoS ONE* 8(12):e82562. doi:10.1371/journal.pone.0082562
- Ritchie K, Henderson C, Wang X, Vassieva O, Carrie D et al (2007) Glutathione transferase plays a critical role in the development of lung carcinogenesis following exposure to tobacco-related carcinogens and urethane. *Cancer Res* 67:9248–9257
- Rostami-Hodjegan A, Lennard MS, Woods HF, Tucker GT (1998) Meta-analysis of studies of the CYP2D6 polymorphism in relation to lung cancer and Parkinson's disease. *Pharmacogenetics* 8:227–238
- Saxena A, Dhillon VS, Shahid M, Khalil HS et al (2012) *GSTP1* methylation and polymorphism increase the risk of breast cancer and the effects of diet and lifestyle in breast cancer patients. *Exp Ther Med* 4(6):1097–1103 PMID: PMC3494109
- Seyfried T, Shelton L (2010) Cancer as a metabolic disease. *Nutr Metab* 7:7
- Siegfried JM (2001) Women and lung cancer: does oestrogen play a role? *Lancet Oncol* 2:506–513
- Smith C, Wadelius M, Gough A, Harrison D, Wolf C, Rane A (1997) A simplified assay for the arylamine N-acetyltransferase 2 polymorphism validated by phenotyping with isoniazid. *J Med Genet* 34:758–760
- Sreejal S, Yamola V, Hariharan S et al (2005) Possible risk modification by CYP1A1, GSTM1 and GSTT1 gene polymorphisms in lung cancer susceptibility in a South Indian population. *J Hum Genet* 50:618–627
- Sweeney C, Nazar-Stewart V, Stapleton P, Eaton D, Vaughan T (2003) Glutathione S-transferase M1, T1 and P1 polymorphisms and survival among lung cancer patients. *Cancer Epidemiol Biomark Prev* 12:527–533
- Tyczynski J, Bray F, Parkin DM (2003) Lung cancer in Europe in 2000: epidemiology, prevention and early detection. *Lancet oncol* 4:45–55
- Wright CM, Larsen JE, Colosimo ML, Barr JJ, Chen L, McLachlan RE, Yang IA, Bowman RV, Fong KM (2010) Genetic association study of CYP1A1 polymorphisms identifies risk haplotypes in non small cell lung cancer. *Eur Respir J* 35:152–159
- Wu J, Sin D (2011) Improved patient outcome with smoking cessation: when is it too late? *Int J COPD* 6:259–267
- Zanger U, Schwab M (2013) Cytochrome P450 enzymes in drug metabolism: regulation of gene expression, enzyme activities and impact of genetic variation. *Pharmacol Ther* 138:103–141