

GSTM1 null genotype in COPD and lung cancer: evidence of a modifier or confounding effect?

Robert P Young^{1,2}
Raewyn J Hopkins¹
Bryan A Hay¹
Gregory D Gamble¹

¹Schools of Biological Science and Health Sciences, University of Auckland, ²Department of Medicine, Auckland City Hospital, Auckland, New Zealand

Background: Studies over the past two decades have reported associations between *GSTM1* (*glutathione S-transferase mu 1*) null genotype and chronic obstructive pulmonary disease (COPD) or lung cancer. However, a modifier or confounding effect from COPD mediating the *GSTM1* association with lung cancer has not been previously explored.

Aim and methods: This variant was examined in a case-control study of current or former smokers with COPD (n = 669), lung cancer (n = 454), or normal lung function (n = 488). Sex, age, and smoking history were comparable between groups.

Results: The *GSTM1* null genotype was found to be more frequent in smokers with COPD alone (odds ratio [OR] 1.30, 95% confidence interval [CI] 1.02–1.66, $P = 0.031$) and lung cancer (OR 1.26, 95% CI 0.96–1.65, $P = 0.083$) than in matched smokers with normal lung function (62%, 61%, and 56%, respectively). However, when smokers with lung cancer were subgrouped according to the presence of COPD, then the association with all COPD subjects (OR 1.34, 95% CI 1.07–1.70, $P = 0.010$) and with COPD and lung cancer (OR 1.50, 95% CI 1.06–2.12, $P = 0.018$) continued to be significant while that with lung cancer only was reduced (OR 1.11, 95% CI 0.78–1.56, $P = 0.55$). These associations were independent of age, sex, height, lung function, and smoking history.

Conclusion: Findings suggest that COPD is an important subphenotype of lung cancer and may underlie previously reported associations with the *GSTM1* null genotype.

Keywords: lung cancer, chronic obstructive pulmonary disease, *GSTM1*, association study, polymorphism, copy number variant

Introduction

Cigarette smoking is the single most important environmental exposure leading to chronic obstructive pulmonary disease (COPD) and lung cancer, accounting for 85%–90% of cases. Although nearly 50% of deaths from smoking can be attributed to these two pulmonary complications, only 10%–20% of smokers are affected, suggesting genetic predisposition may play an important part.^{1,2} Recently, the authors and others have used spirometry to show that COPD is present in 50%–70% of lung cancer cases.^{3–5} This is compared with a COPD prevalence of 15%–20% among smokers randomly recruited from the community.^{3,6,7} In addition, the presence of COPD is associated with a four- to sixfold increased risk of lung cancer compared with community-recruited smokers³ or smoking controls with normal lung function.⁶ Given this significant overlap (Figure 1), and that the heritability (genetic contribution) of COPD (estimated to be 40%–77%) is nearly three times that of lung cancer (estimated to be 15%–25%),^{8,9} the authors propose that genes conferring a propensity to COPD may

Correspondence: Robert P Young
Department of Medicine, Auckland
Hospital, Private Bag 92019, Auckland,
New Zealand
Tel +64 9 307 4949 ext 4016
Fax +64 9 623 6456
Email roberty@adhb.govt.nz

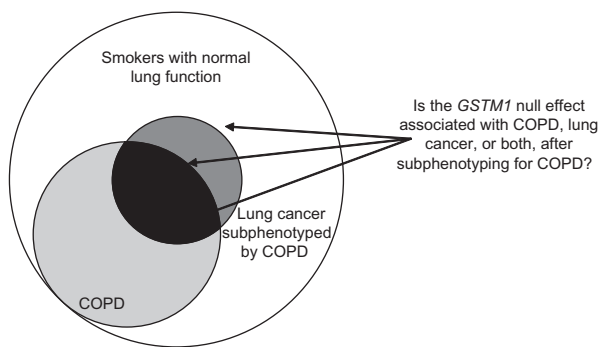


Figure 1 Overlapping relationship between chronic obstructive pulmonary disease (COPD) and lung cancer in smokers and the susceptibility effect conferred by the *GSTM1* (glutathione S-transferase mu 1) null genotype.

also be relevant in lung cancer susceptibility.¹⁰ Specifically among smokers, some loci which determine susceptibility to COPD may also be important in the susceptibility to lung cancer. Support for this hypothesis comes from recently published genome-wide association studies of COPD, lung function (forced expiratory volume in 1 second [FEV₁]), and lung cancer that report overlapping loci.^{11–16} Further support for this overlap hypothesis comes from studies where lung cancer cases have been subgrouped according to COPD status (based on spirometry; Figure 1), and show that COPD-related single nucleotide polymorphisms (SNPs) are also associated with lung cancer, independent of the presence of COPD.^{17,18} Using this approach, some SNP associations previously reported for lung cancer have been shown to be more closely related to COPD.¹⁹ This raises the possibility that previously reported genetic associations, or indeed genomic (ie, gene expression) associations, for lung cancer may in fact represent associations with COPD through a modifier or confounding effect.

In the past two decades there has been considerable interest in the phase I and II enzymes, which include glutathione S-transferase (GST) enzymes, in the metabolism and inactivation of smoking-related compounds including carcinogens.²⁰ The highly conserved phase II metabolic enzymes of the GST family are expressed in the airway epithelium and alveolar macrophages.²⁰ The gene for the isoenzyme GST Mu 1 (*GSTM1*) is located in a region of the genome susceptible to copy number variation, such that about 38%–67% of those with European ancestry are homozygote for the *GSTM1* gene deletion (null genotype or *GSTM1**0). In a recently reported study it was shown that for those who had homozygous gene deletion, the expression of *GSTM1* in the small airways epithelium and alveolar macrophages is approximately one-third of that in those with one or more copies.²¹ This shows that subjects who had homozygous gene

deletion (*GSTM1**0) have significantly reduced expression of this enzyme, compared with those carrying at least one copy of the gene, in the tissues thought to be directly involved in the development of lung cancer and COPD. This finding provides compelling evidence of a functional role for the *GSTM1* null genetic variant and might explain why this genotype has been found to be associated with an increased risk of both lung cancer and COPD in many studies in the past.^{22–28} Furthermore, studies have shown the *GSTM1* null genotype is associated with increased DNA adducts in smokers with lung cancer.²⁹ However, none of the lung cancer association studies subphenotyped their cases for the presence of COPD, so it remains possible that an interactive modifier or confounding effect from COPD may have gone undetected. This is because the frequency of COPD is estimated to be between 50% and 70% in lung cancer cases compared with 10% and 30% in smoking controls, depending on where they were recruited from (Figure 1).^{3–7} In this study, a recently described pseudo-SNP, known to function as a *GSTM1* deletion marker (Figure 2),³⁰ was used and its frequency was compared in smoking groups with COPD, lung cancer, and normal lung function.

Materials and methods

Study subjects

All subjects recruited were of Caucasian ancestry based on their grandparents' descent (all four grandparents of Caucasian descent). Subjects recruited into the study were aged 40–80 years with a minimum smoking history of 15 pack years and COPD confirmed by a respiratory specialist based on prebronchodilator spirometric criteria. All subjects were recruited between 2001 and 2007. Control subjects were recruited based on the following criteria: aged 45–80 years and with a minimum smoking history of 15 pack years. Control subjects were volunteers who were identified through either a community postal advertisement or while attending community-based retired military/servicemen's

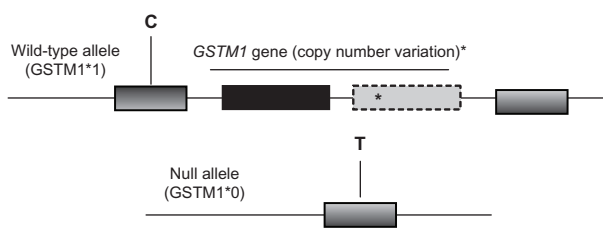


Figure 2 *GSTM1* (glutathione S-transferase mu 1) gene structure, the gene deletion in the null genotype and rs366631 pseudo-single nucleotide polymorphism.³⁰

Note: *Multiple copies of the *GSTM1* gene have been reported in non-Caucasian ethnicities.³⁰

clubs located in the same patient catchment as those serving the lung cancer and COPD hospital clinics. Lung cancer cases were aged >40 years and in 95% of cases their diagnosis was confirmed through histological or cytological specimens. Nonsmokers (those who had smoked < 100 cigarettes over their lifetime) with lung cancer were excluded from the study and only primary lung cancer cases with the following pathological diagnosis were included: adenocarcinoma, squamous cell cancer, small-cell cancer, and nonsmall-cell cancer (generally large cell or bronchoalveolar subtypes). Lung function measurement (prebronchodilator) was performed within 3 months of lung cancer diagnosis, prior to surgery, and in the absence of pleural effusions or lung collapse on plain chest radiographs. In a minority of lung cancer cases, computed tomography evidence of advanced disease and/or poor respiratory reserve meant histological confirmation was not possible in 5% of cases and spirometry was not performed in 7% of cases.

Spirometry was done in 100% of those with COPD and the smoking controls, and in 93% of those with lung cancer. Lung function conformed to American Thoracic Society standards for acceptability and reproducibility, with the highest value of the best three blows used for classification of COPD status. Spirometry was performed using a portable spirometer (Easy-One™ Spirometer; ndd Medizintechnik AG, Zurich, Switzerland). COPD was defined according to the Global Initiative for Chronic Obstructive Lung Diseases (GOLD) spirometric criteria of stage II or above (FEV_1 /forced vital capacity [FVC] <70% and FEV_1 % predicted \leq 80%) although using prebronchodilator measurements.³¹ For lung cancer cases that had already undergone surgery, preoperative lung function performed by the hospital lung function laboratory was sourced from medical records. Controls with COPD based on spirometry (GOLD stage I or above) constituted 50% of the community volunteers and were excluded from further analysis. At the time of recruitment, all participants gave written informed consent and underwent blood sampling for DNA extraction, spirometry, and an investigator-administered questionnaire. The study was approved by the Multi Center Ethics Committee (New Zealand).

Study design

The present retrospective case-control study compared smokers of the same ethnicity and comparable demographic variables (specifically age, sex, and smoking history). The healthy (resistant) controls in the current study were carefully chosen to be representative of the majority of smokers (60%–80%)

who have maintained normal or near-normal lung function despite decades of smoking.^{32,33} Using healthy smokers as controls allows the association of *GSTM1* with COPD and lung cancer to be independently assessed. Moreover, COPD-related genetic associations among those with lung cancer (sub-phenotyping for COPD) can also be identified. This is because the resistant smokers best reflect those smokers least likely to develop lung cancer or COPD. Similarly, the cases in this study are defined according to generally accepted criteria for genetic studies of COPD³⁴ and exclude those with GOLD stage I criteria, where aging effects or other respiratory “insults” (eg, asthma) rather than smoking may be relevant. This approach resembles a pharmacogenetic approach or experiment where smoking exposure is comparable but phenotype misclassification of cases and controls is minimized, thereby improving the power to detect differences between affected (COPD) and unaffected (resistant) smokers.³⁵

Genotyping

Genomic DNA was extracted from whole blood samples using standard salt-based methods and purified genomic DNA was aliquoted (10 ng μL^{-1} concentration) into 96-well plates. Samples were genotyped using Taqman® SNP Genotyping Assays (Applied Biosystems, Foster City, CA) utilizing minor groove-binder probes as previously described.^{17–19} The present study investigated the frequencies of the *GSTM1* null genotype (deletion) in the 1p13.3 region by genotyping the SNP rs366631 (assay ID: Custom TaqMan SNP Genotyping Assays) according to the methods of Huang et al who showed this pseudo SNP can be used as a *GSTM1* deletion marker (Figure 2).³⁰

Analysis

Patient characteristics in the cases and controls were compared by analysis of variance for continuous variables and Chi-square test for discrete variables (Cochran–Mantel–Haenszel, odds ratio [OR]). Genotype frequencies for the cases and controls were examined separately and compared with the scientific literature. Population admixture across groups was performed using structure analysis on genotyping data from 40 randomly selected SNPs.³⁶ Distortions in the genotype frequencies were identified between cases and controls using two-by-two contingency tables. Given that the functional biology and previously published literature supports a difference between those homozygote null genotype versus the rest, this model was tested between cases and control smokers. Based on the literature, with expected *GSTM1* null genotype frequencies in controls of 0.55 and in cases of 0.66 (based on

OR 1.2 from the meta-analyses), a sample size in excess of 420 in each group is sufficiently powered (at the 90% level) to achieve significance at the 0.05 level (two-tailed).

Results

Demographic variables

Characteristics of the lung cancer cases, COPD cases, and healthy control smokers are summarized in Table 1 and described in detail elsewhere.^{3,18,19} The demographic variables and histological subtypes of the lung cancer cases (identified in 94% of cases) are comparable to previously published data.³⁷ The COPD cases have higher pack year exposure than the lung cancer cases and healthy control smokers ($P < 0.05$). This reflects outliers with high smoking histories in the COPD group (skewed distribution) and no difference exists after log transformation of pack years (data not shown). All groups are comparable with respect to age started smoking, years smoked, years since quitting, and cigarettes/day (Table 1). Overall, the authors believe the three groups have comparable smoking exposure that may be critical in a gene–environment interaction where a smoking

dose threshold and/or variable penetrance may alter a genetic effect and its related association (OR and P value). The lower frequency of current smokers in the lung cancer and COPD cases, compared with the healthy smoker group (35%, 40%, and 48%, respectively), may reflect the presence of symptoms from pulmonary complications of smoking (eg, exertional breathlessness from COPD).

The lung cancer cases, COPD cases, and smoking controls were also comparable with respect to other aeropollutant exposures (Table 1). The lung cancer cases reported higher rates of a family history of lung cancer compared with the COPD cases and healthy smokers (19%, 11%, and 9%, respectively). Mean height was slightly lower in those with lung cancer compared with those in the control group ($P < 0.05$), but height was not different when sex adjustment was made. As expected, lung function was worse in the lung cancer and COPD cases compared with the healthy smoker controls. Testing of lung function (as described above) was achieved in 93% of lung cancer cases and allowed stratification of results to test for an interactive modifier or confounding effect from coexisting COPD in those with lung cancer.

Table 1 Summary of the characteristics for the lung cancer, COPD, and control smokers

Parameter Mean (1 SD)	Lung cancer (n = 454)	COPD (n = 669)	Control smokers (n = 488)
% male	53%	59%	60%
Age (years)*	69 (10)	66 (9)	65 (10)
Height (m)	1.67 (0.08)	1.68 (0.09)	1.69 (0.09)
Smoking history			
Current smoking (%)	35%	40%	48%
Age started (years)	18 (4)	17 (3)	17 (3)
Years smoked	41 (12)	42 (11)	35 (11)
Pack years*	41 (25)	47(20) ^a	40 (19)
Cigarettes/day	20 (10)	23 (9)	24 (11)
Years since quitting	11.4 (6.7)	9.8 (7.4)	13.9 (8.1)
History of other exposures			
Work dust exposure*	63%	59%	47%
Work fume exposure	41%	40%	38%
Asbestos exposure*	23%	22%	16%
Family history			
COPD	33%	37%	28%
Lung cancer*	19%	11%	9%
Lung function			
FEV ₁ (L)*	1.86 (0.48)	1.25 (0.48)	2.86 (0.68)
FEV ₁ % predicted*	73%	46%	99%
FEV ₁ /FVC*	64% (13)	46% (8)	78% (7)
Spirometric COPD ^{b*}	51%	100%	0%

Notes: ^aNo significant difference after log transformation of pack years due to skewed distribution; ^bAccording to Global Initiative for Chronic Obstructive Lung Diseases stage II or above criteria. * $P < 0.05$.

Abbreviations: SD, standard deviation; COPD, chronic obstructive pulmonary disease; FEV₁, forced expiratory volume in 1 second; FVC, forced vital capacity.

Genotype results

Genotype frequencies for the *GSTM1* null genotype (TT genotype of the rs366631 C/T pseudo-SNP) are shown in Table 2. The genotype frequencies were consistent with those reported in the literature^{21–28} and from the International Hapmap Project.³⁸ Moreover, this SNP assay was validated using an alternative restricted fragment length polymorphism-based genotyping method (n = 476 samples) with 100% concordance, suggesting significant genotyping error is unlikely. No evidence for population stratification was found between the case and control groups using 40 unlinked SNPs from unrelated genes (mean $\chi^2 = 3.3$, $P = 0.58$).³⁶

The *GSTM1* homozygote null genotype (TT genotype of the rs366631) was more frequent in smokers with COPD alone (OR 1.30, 95% confidence interval [CI] 1.02–1.66, $P = 0.031$) and lung cancer (OR 1.26, 95% CI 0.96–1.65, $P = 0.083$) than in matched smokers with normal lung function (62%, 61%, and 56% respectively; Table 2). However, when smokers with lung cancer were subgrouped according to the presence of COPD, then the association with all COPD subjects (n = 882, OR 1.34, 95% CI 1.07–1.70, $P = 0.010$) and those with COPD and lung cancer (n = 213, OR 1.50, 95% CI 1.06–2.12, $P = 0.02$) continued to be significant while that with lung cancer only became much weaker (OR 1.11, 95% CI 0.78–1.56, $P = 0.55$). *GSTM1* homozygote null (TT) genotype frequencies were 63%, 65%, and 58% respectively versus

Table 2 *GSTMI* (glutathione *S*-transferase mu 1) null genotype frequency for the rs366631 single nucleotide polymorphism in healthy smokers, COPD, and lung cancer (total and subgrouped by COPD) cohorts compared with smoking controls

Primary cohorts (call rate %)	GSTM*0 (TT)	GSTM*1 (CT, CC) ^b	Odds ratio ^a (95% CI) ^a P value ^a
Controls (n = 476) (98%)	265 (56%)	211 (44%)	–
COPD* (n = 669) (99%)	415 (62%)	254 (38%)	1.30 (1.02–1.66) 0.03
Lung cancer (n = 447) (99%)	274 (61%)	173 (39%)	1.26 (0.96–1.65) 0.083
Lung cancer subgroup analyses^c			
Lung cancer with COPD* (n = 213)	139 (65%)	74 (35%)	1.50 (1.06–2.12) 0.018
Lung cancer only (n = 203)	118 (58%)	85 (42%)	1.11 (0.78–1.56) 0.55
COPD subgroup analysis			
COPD* and lung cancer with COPD* (n = 882)	554 (63%)	328 (37%)	1.34 (1.07–1.70) 0.010

Notes: ^aGSTM*0 versus GSTM*1 compared to matched smoking controls (Cochran–Mantel–Haenszel); ^bCC is only found in those with African descent; ^cspirometry available in 422/454 (93%), COPD* = Global Initiative for Chronic Obstructive Lung Diseases stage II or above criteria.

Abbreviations: LC, lung cancer; COPD, chronic obstructive pulmonary disease; CI, confidence interval.

56% in smoking controls (Table 2). These associations were independent of age, sex, height, lung function, and smoking history in a logistic regression analysis. These results were not altered when those with no histological confirmation of lung cancer were excluded from the analysis (data not shown).

When the *GSTMI* null genotype association was examined in the lung cancer cases subdivided according to histological subtype (Table 3), no effect for any histological type was found, although an underpowered sample size may be relevant here. Also, no association was found between this *GSTMI* null genotype and lung function or smoking exposure in the lung cancer cases (Table 3). When the COPD cases were subgrouped according to the *GSTMI* null genotype (Table 4), no associations were found with the demographic variables, smoking history, GOLD level, or lung function. The latter suggests the *GSTMI* genotype does not correlate with severity of COPD.

Discussion

Numerous genetic epidemiological studies have reported a positive association between the *GSTMI* null genotype

Table 3 *GSTMI* (glutathione *S*-transferase mu 1) genotype in the lung cancer cases and relationship to demographic variables

Cohort lung cancer	GSTM*0 (n = 270)	GSTM*1 (n = 168)	P value
Demographics			
Male	141 (52%)	83 (49%)	–
Female	129 (48%)	85 (51%)	0.62
Mean age (years)	67 (41–91)	68 (44–91)	0.58
Mean height (cm)	167 (143–192)	167 (149–187)	0.78
Smoking history			
Mean pack years	41	41	0.41
Mean cigarettes/day	20 (3–73)	20 (2–60)	0.30
Mean total years smoked	41 (2–70)	41 (5–66)	0.72
Currently smoking	106 (39%)	50 (30%)	0.051
Ex smoker	146 (54%)	107 (64%)	0.059
Non smoker	18 (7%)	11 (6%)	0.99
Self reported COPD			
Yes	76 (28%)	51 (30%)	0.67
No/not known	194 (72%)	117 (70%)	0.83
Spirometry (prebronchodilator)			
Mean FEV ₁ absolute (L)	1.67 (0.44–4.94)	1.71 (0.45–4.06)	0.59
Mean FEV ₁ (% predicted)	66.2% (13%–140%)	69.1% (17%–147%)	0.37
Mean FVC absolute (L)	2.78 (0.83–6.87)	2.83 (0.95–5.91)	0.42
Mean FEV ₁ /FVC	60.3% (27%–90%)	60.6 (18%–90%)	0.71
GOLD criteria			
GOLD no	70 (26%)	51 (30%)	0.34
GOLD stage I	36 (13%)	28 (17%)	0.33
GOLD stage II	109 (41%)	53 (32%)	0.83
GOLD stages III and IV	53 (20%)	36 (21%)	0.71
Histology			
Small cell	47 (17%)	31 (18%)	0.80
Nonsmall cell	23 (9%)	20 (12%)	0.25
Adenocarcinoma	118 (44%)	70 (42%)	0.69
Squamous cell	66 (24%)	41 (24%)	0.99
No histology (but CT)	16 (6%)	7 (4%)	0.46

Note: Wilcoxon independent groups or Fisher's exact test.

Abbreviations: COPD, chronic obstructive pulmonary disease; FEV₁, forced expiratory volume in 1 second; FVC, forced vital capacity; GOLD, Global Initiative for Chronic Obstructive Lung Diseases; CT, computed tomography.

and both COPD and lung cancer.^{22–28} Studies have also shown that such associations may be explained by biological (functional) studies showing that the null genotype confers reduced *GSTMI* expression and increased DNA adduct levels in smokers.^{21,29} In this study, the *GSTMI* null genotype was found to be associated with an increased risk of both COPD and lung cancer (OR 1.30 and 1.26, respectively). The *GSTMI* null association with COPD reported here was replicated in the lung cancer cases with COPD (OR 1.50) and is consistent with many other genetic COPD association studies, including large meta-analytical

Table 4 *GSTM1* (glutathione *S*-transferase mu 1) genotype in the COPD cases and relationship to demographic variables

Cohort COPD	<i>GSTM1</i> *0 (n = 335)	<i>GSTM1</i> *1 (n = 213)	P value
Demographics			
Male	185 (55%)	130 (61%)	0.19
Female	150 (45%)	83 (39%)	0.19
Mean age (years)	65 (39–84)	65 (42–85)	0.77
Mean height (cm)	168 (141–190)	169 (147–188)	0.35
Smoking history			
Mean pack years	48 (9–225)	46 (11–153)	0.34
Mean cigarettes/day	23 (5–100)	23 (8–70)	0.67
Mean total years smoked	42 (13–71)	42 (17–64)	0.50
Mean age started smoking	17 (6–48)	17 (10–37)	0.28
Currently smoking	122 (36%)	99 (47%)	0.02
Mean age stopped smoking	55 (30–74)	53 (34–75)	0.34
Other exposures			
Asbestos	63 (19%)	41 (19%)	0.91
Dusts	190 (57%)	112 (53%)	0.38
Fumes	129 (39%)	93 (44%)	0.25
Self-reported COPD			
Yes	233 (70%)	152 (71%)	0.70
No/not known	102 (30%)	61 (29%)	0.63
Spirometry (prebronchodilator)			
Mean FEV ₁ absolute (L)	1.25 (0.2–2.72)	1.23 (0.28–3.56)	0.76
Mean FEV ₁ (% predicted)	44% (8%–79%)	45% (14%–98%)	0.93
Mean FVC absolute (L)	2.6 (0.51–5.58)	2.7 (0.84–5.85)	0.18
Mean FEV ₁ /FVC	46% (18%–69%)	45% (18%–69%)	0.35
GOLD criteria			
GOLD stage I	0	1 (0.5%)	0.39
GOLD stage II	136 (41%)	81 (38%)	0.59
GOLD stage III	109 (33%)	73 (34%)	0.71
GOLD stage IV	90 (27%)	58 (27%)	0.92
GOLD stages III and IV	199 (59%)	131 (62%)	0.65

Abbreviations: COPD, chronic obstructive pulmonary disease; FEV₁, forced expiratory volume in 1 second; FVC, forced vital capacity; GOLD, Global Initiative for Chronic Obstructive Lung Diseases.

studies.^{22–25} However, after accounting for the presence of COPD, the association between the *GSTM1* null genotype and lung cancer alone was considerably weaker (OR 1.11). This observation needs to be verified but raises the important possibility that the *GSTM1* association with lung cancer^{26–28} results from a modifier or confounding effect with COPD. This effect from unrecognized COPD in the previous lung cancer studies is possible because the frequency of COPD is estimated to be two- to threefold greater in lung cancer cases compared with “unaffected” smoking controls, depending on how they were recruited.^{3–6,17} It is

concluded that if coexisting COPD continues to be ignored in lung cancer association studies,^{26–28} then this apparent modifier or confounding effect from COPD will likely go unrecognized.

The possibility that COPD has a modifier or confounding effect on genetic associations with lung cancer is not new.^{19,39} The authors suggested this possibility in 2008, showing, in the same smoking cohorts as the present study, that the nicotinic receptor subunit 3/5 alpha gene (*CHRNA 3/5*, chromosome 15q25) link with lung cancer^{11–13} may be more strongly associated with COPD than lung cancer.¹⁹ The *CHRNA 3/5* association with COPD has been subsequently confirmed in several genome-wide association studies¹⁶ and a validation case-control study.⁴⁰ The implications of this modifier or confounding effect are significant, as any previous association reported between lung cancer cases and smoking controls could be actually mediated through a COPD modifier effect. Such an effect would not be limited to genetic association studies but any biomarker studies including expression assays where noncancer specimens (eg, bronchial biopsies) have been compared between lung cancer cases and unaffected controls.⁴¹ One of the difficulties with identifying coexisting COPD in those diagnosed with lung cancer is that COPD results from the insidious loss of lung function, with subtle exertional breathlessness that goes undiagnosed in 50%–80% of those affected.⁴² This means spirometry to detect airflow limitation, or computed tomography to detect emphysema, is required to recognize chronic smokers with COPD. Historically, few, if any, studies of lung cancer patients have measured pulmonary function tests to subphenotype for COPD. This has relevance when trying to develop risk models for lung cancer⁴³ or identify pathogenic pathways to target chemopreventive treatments.⁴⁴ By subphenotyping lung cancer cases for COPD, the authors have shown that several of the recently identified COPD-related SNPs (eg, *HHIP*, *FAM13A*, and *ADAM19*) are also relevant in lung cancer susceptibility.^{17,18,45} Indeed, the authors suggest that this overlapping effect provides evidence that susceptibility among smokers to COPD and lung cancer are related by much more than just smoking exposure and also by overlapping pathogenetic pathways conferring susceptibility or resistance to smoking.⁴⁴ There is an abundance of molecular biological evidence describing these overlapping biological pathways.^{2,44} These studies suggest that the pathogenic processes initiated by smoking that lead to COPD may also contribute to the pathogenic processes leading to lung cancer.^{10,17–19} Specifically, the events leading to COPD, which involve matrix remodeling with excess growth factors and matrix metalloproteinases,

are also important in promoting epithelial mesenchymal transition, a premalignant change found in both COPD and lung cancer.^{44–48} If this is true, and COPD is part of the causal pathway for the development of lung cancer, then the COPD effect described here is a modifier effect (rather than a strictly confounding effect). This means the *GSTMI* null genotype primarily increases a smoker's risk of COPD, a proportion of which might also progress to lung cancer. Based on the overlap in pathogenic pathways reported in the molecular studies⁴⁴ and the overlapping associations of some the genetic variants in genome-wide association studies of COPD and lung cancer,⁴⁵ the authors favor a causal relationship. If, alternatively, there is no causal relationship linking COPD and lung cancer, then COPD might be correctly considered a true confounder. If this is true, then any study reporting an association between *GSTMI* and lung cancer could be attributed to a confounding effect from unrecognized COPD.^{26–29,49} In this situation, where COPD does not share pathogenic (causal) pathways with lung cancer, any association derived from comparing lung cancer with smoking controls might suffer a COPD confounding effect.

Conclusion

COPD results from smoking/aeropollutant exposure and a genetic predisposition that represents the net effect of a variable number of genetic variants (SNPs, copy number variations, methylation, and others) that are conferring either a protective or susceptibility effect. There is good functional data to support a role for the *GSTMI* null genotype in smoking-related diseases where *GSTMI* activity is markedly reduced and damage to DNA is more common.^{21,29} Although many studies have reported a weak association (susceptibility effect) for both COPD and lung cancer, the authors' data suggests the *GSTMI* null effect may be primarily through conferring susceptibility to COPD. It is proposed that there is good evidence to suggest that previously published lung cancer genetic association studies may be vulnerable to a COPD modifier or confounder effect and that spirometry (and computed tomography, if funding allows) be routinely measured in these genetic studies to better distinguish genes underlying COPD, lung cancer, both, or neither smoking-related complication.

Acknowledgments

This paper is dedicated to the memory of Dr Tom Petty and to the smokers with lung cancer who agreed to participate in this study. We also acknowledge the support of Dr Peter Black who died suddenly in January 2010. This study was supported by grants from the University of Auckland,

Auckland Medical Research Foundation, and Health Research Council of New Zealand.

Disclosure

The authors report no conflicts of interest in this work.

References

- Center for Disease Control and Prevention. Annual smoking-attributable mortality, years of potential life lost, and productivity losses – United States, 1997–2001. *MMWR Morb Mortal Wkly Rep.* 2005;54(25): 625–628.
- Punturieri A, Szabo E, Croxton TL, Shapiro SD, Dubinett SM. Lung cancer and chronic obstructive pulmonary disease: needs and opportunities for integrated research. *J Natl Cancer Inst.* 2009;101(8): 554–559.
- Young RP, Hopkins RJ, Christmas T, Black PN, Metcalf P, Gamble GD. COPD prevalence is increased in lung cancer independent of age, sex and smoking history. *Eur Respir J.* 2009;34(2):380–386.
- de Torres J, Bastarrika G, Wisnivesky JP, et al. Assessing the relationship between lung cancer risk and emphysema detected on low dose CT of the chest. *Chest.* 2007;132(6):1932–1938.
- Wilson DO, Weissfeld JL, Balkan A, et al. Association of radiographic emphysema and airflow obstruction with lung cancer. *Am J Respir Crit Care Med.* 2008;178(7):738–744.
- Mannino DM, Aguayo SM, Petty TL, Redd SC. Low lung function and incident lung cancer in the United States: data from the first NHANES follow-up. *Arch Intern Med.* 2003;163(12):1475–1480.
- Tockman MS, Anthonisen NR, Wright EC, Donithan MG. Airways obstruction and the risk for lung cancer. *Ann Intern Med.* 1987;106(4): 512–518.
- Hubert HB, Fabsitz RR, Feinleib M, Gwinn C. Genetic and environmental influences on pulmonary function in adult twins. *Am Rev Respir Dis.* 1982;125(4):409–415.
- Lichtenstein P, Holm NV, Verkasalo PK, et al. Environmental and heritable factors in the causation of cancer – analyses of cohorts of twins from Sweden, Denmark and Finland. *N Eng J Med.* 2000;343(2): 78–85.
- Young RP, Hopkins RJ. Link between COPD and lung cancer. *Respir Med.* 2010;104(5):758–759.
- Amos CI, Wu X, Broderick P, et al. Genome-wide association scan of tag SNPs identifies a susceptibility locus for lung cancer at 15q25.1. *Nat Genet.* 2008;40(5):616–622.
- Hung RJ, McKay JD, Gaborieau V, et al. A susceptibility locus for lung cancer maps to nicotinic acetylcholine receptor subunit genes on 15q25. *Nature.* 2008;452(7187):633–637.
- Thorgeirsson TE, Geller F, Sulem P, et al. A variant associated with nicotine dependence, lung cancer and peripheral arterial disease. *Nature.* 2008;452(7187):638–642.
- Wilk JB, Chen TH, Gottlieb DJ, et al. A genome-wide association study of pulmonary function measures in the Framingham Heart Study. *PLoS Genet.* 2009;5(3):e1000429.
- Hancock DB, Eijgelsheim M, Wilk JB, et al. Meta-analyses of genome-wide association studies identify multiple loci associated with pulmonary function. *Nat Genet.* 2010;42(1):45–52.
- Pillai SG, Ge D, Zhu G, et al. A genome-wide association study in chronic obstructive pulmonary disease (COPD): identification of two major susceptibility loci. *PLoS Genet.* 2009;5(3):e1000421.
- Young RP, Whittington CF, Hopkins RJ, et al. Chromosome 4q31 locus in COPD also associated with lung cancer. *Eur Respir J.* 2010;36(6): 1375–1382.
- Young RP, Hopkins RJ, Hay BA, Whittington CF, Epton MJ, Gamble GD. *FAM13A* locus in COPD independently associated with lung cancer – evidence of a molecular genetic link between COPD and lung cancer. *Appl Clin Genet.* 2011;4(1):1–10.

19. Young RP, Hopkins RJ, Hay BA, Epton MJ, Black PN, Gamble GD. Lung cancer gene associated with COPD: triple whammy or possible confounding effect? *Eur Respir J*. 2008;32(5):1158–1164.
20. Rebbeck TR. Molecular epidemiology of the human glutathione S-transferase genotype GSTM1 and GSTT1 in cancer susceptibility. *Cancer Epidemiol Biomarkers Prev*. 1997;6(9):733–743.
21. Butler MW, Hackett NR, Salit J, et al. Glutathione S-transferase copy number variation alters lung gene expression. *Eur Respir J*. 2011; Feb 24. [Epub ahead of print]. doi:10.1183/09031936.00029210.
22. Solomonska J, Wijmenga C, Postma DS, Boezen HM. Meta-analyses on suspected chronic obstructive pulmonary disease genes: a summary of 20 years' research. *Am J Respir Crit Care Med*. 2009;180(7):618–631.
23. Hu G, Yao W, Zhou Y, et al. Meta- and pooled analyses of the effect of glutathione S-transferase M1 and T1 deficiency on chronic obstructive pulmonary disease. *Int J Tuberc Lung Dis*. 2008;12(12):1474–1481.
24. Imboden M, Downs SH, Senn O, et al. Glutathione S-transferase genotypes modify lung function decline in the general population: SAPALDIA cohort study. *Respir Res*. 2007;8:2.
25. Cheng SL, Yu CJ, Chen CJ, Yang PC. Genetic polymorphism of epoxide hydrolase and glutathione S-transferase in COPD. *Eur Respir J*. 2004;23(6):818–824.
26. Ye Z, Song H, Higgins JPT, Pharoah P, Danesh J. Five glutathione S-transferase gene variants in 23,452 cases of lung cancer and 30,397 controls: meta-analysis of 130 studies. *PLoS Med*. 2006;3(4):e91.
27. Vineis P, Anttila S, Benhamou S, et al. Evidence of gene-gene interactions in lung carcinogenesis in a large pooled analysis. *Carcinogenesis*. 2007;28(9):1902–1905.
28. Carlsten C, Sagoo GS, Frodsham AJ, Burke W, Higgins JP. Glutathione S-transferase M1 (GSTM1) polymorphism and lung cancer: a literature-based systematic HuGE review and meta-analysis. *Am J Epidemiol*. 2008;167(7):759–774.
29. Lee MS, Su L, Christiani DC. Synergistic effects of NAT2 slow and GSTM1 null genotypes on carcinogen DNA damage in the lung. *Cancer Epidemiol Biomarkers Prev*. 2010;19(6):1492–1497.
30. Huang RS, Chen P, Wisel S, et al. Population-specific GSTM1 copy number variation. *Hum Mol Genet*. 2009;18(2):366–372.
31. Global Initiative for Chronic Obstructive Lung Diseases. Spirometry for healthcare providers. 2010. Available from: http://www.goldcopd.org/uploads/users/files/GOLD_Spirometry_2010.pdf. Accessed August 23, 2011.
32. Løkke A, Lange P, Scharling H, Fabricius P, Vestbo J. Developing COPD: a 25 year follow up study of the general population. *Thorax*. 2006;61(11):935–939.
33. Kohansal R, Martinez-Cambor P, Agusti A, Buist AS, Mannino DM, Soriano JB. The natural history of chronic airflow obstruction revisited: an analysis of the Framingham offspring cohort. *Am J Respir Crit Care Med*. 2009;180(1):3–10.
34. Crapo C, Silverman E. COPDGen[®] Study. 2011. Available from: <http://www.copdgene.org/>. Accessed August 23, 2011.
35. Moskvina V, Holmans P, Schmidt KM, Craddock N. Design of case-controls studies with unscreened controls. *Ann Hum Genet*. 2005; 69(Pt 5):566–576.
36. Pritchard J, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. *Genetics*. 2000;155(2):945–959.
37. Yang P, Allen MS, Aubry MC, et al. Clinical features of 5,628 primary lung cancer patients; experience at Mayo Clinic from 1997 to 2003. *Chest*. 2005;128(1):452–462.
38. International HapMap Project. October 2002. Available from: <http://hapmap.ncbi.nlm.nih.gov/>. Accessed August 23, 2011.
39. Yang P, Li Y, Jiang R, et al. A rigorous and comprehensive validation: common genetic variations and lung cancer. *Cancer Epidemiol Biomarkers Prev*. 2010;19(1):240–244.
40. Lambrechts D, Buyschaert I, Zanen P, et al. The 15q24/25 susceptibility variant for lung cancer and chronic obstructive pulmonary disease is associated with emphysema. *Am J Respir Crit Care Med*. 2010;181(5):486–493.
41. Gustafson AM, Soldi R, Anderlind C, et al. Airway P13K pathway activation is an early and reversible event in lung cancer development. *Sci Transl Med*. 2010;2(26):26ra25.
42. Stav D, Raz M. Prevalence of chronic obstructive pulmonary disease among smokers aged 45 and up in Israel. *Isr Med Assoc J*. 2007;9(11):800–802.
43. Young RP, Hopkins RJ, Hay BA, et al. A gene based risk score for lung cancer susceptibility in smokers and ex-smokers. *Postgrad Med J*. 2009;85(1008):515–524.
44. Young RP, Hopkins R, Eaton TE. Pharmacological actions of statins: potential utility in COPD. *Eur Respir Rev*. 2009;18(114):222–232.
45. Young RP, Hopkins RJ, Whittington CF, Hay BA, Epton MJ, Gamble GD. Individual and cumulative effects of GWAS susceptibility loci in lung cancer: associations after sub-phenotyping for COPD. *Plos One*. 2011;6(2):e16476.
46. Sohal SS, Reid D, Soltani A, et al. Reticular basement membrane fragmentation and potential epithelial mesenchymal transition is exaggerated in the airways of smokers with chronic obstructive pulmonary disease. *Respirology*. 2010;15(6):930–938.
47. Dasari V, Gallup M, Lemjabbar H, Maltseva I, McNamara N. Epithelial-mesenchymal transition in lung cancer: Is tobacco the “smoking gun”? *Am J Respir Cell Mol Biol*. 2006;35(1):3–9.
48. Lee G, Walser TC, Dubinett SM. Chronic inflammation, chronic obstructive pulmonary disease, and lung cancer. *Curr Opin Pulm Med*. 2009;15(4):303–307.
49. Brennan P, Hainaut P, Boffetta P. Genetics of lung-cancer susceptibility. *Lancet Oncol*. 2011;12(4):399–408.

The Application of Clinical Genetics

Publish your work in this journal

The Application of Clinical Genetics is an international, peer-reviewed open access journal that welcomes laboratory and clinical findings in the field of human genetics. Specific topics include: Population genetics; Functional genetics; Natural history of genetic disease; Management of genetic disease; Mechanisms of genetic disease; Counselling and

Submit your manuscript here: <http://www.dovepress.com/the-application-of-clinical-genetics-journal>

ethical issues; Animal models; Pharmacogenetics; Prenatal diagnosis; Dysmorphology. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Dovepress