The *TERT-CLPTM1L* locus for lung cancer predisposes to bronchial obstruction and emphysema

E. Wauters*,#,¶,ƒ, D. Smeets*,#,ƒ, J. Coolen+, J. Verschakelen+, P. De Leyn§, M. Decramer¶, J. Vansteenkiste¶, W. Janssens¶ and D. Lambrechts*,#

ABSTRACT: **Clinical studies suggest that bronchial obstruction and emphysema increase susceptibility to lung cancer. We assessed the possibility of a common genetic origin and investigated whether the lung cancer susceptibility locus on chromosome 5p15.33 increases the risk for bronchial obstruction and emphysema.**

Three variants in the 5p15.33 locus encompassing the *TERT* and *CLPTM1L* genes were genotyped in 777 heavy smokers and 212 lung cancer patients. Participants underwent pulmonary function tests and computed tomography of the chest, and completed questionnaires assessing smoking behaviour.

The rs31489 C-allele correlated with reduced forced expiratory volume in 1 s (p=0.006). Homozygous carriers of the rs31489 C-allele exhibited increased susceptibility to bronchial obstruction (OR 1.82, 95% CI 1.24–2.69; p=0.002). A similar association was observed for diffusing capacity of the lung for carbon monoxide (p=0.004). Consistent with this, CC-carriers had an increased risk of emphysema (OR 2.04, 95% CI 1.41–2.94; p=1.73 × 10^{-4}) and displayed greater alveolar destruction. Finally, CC-carriers also had an increased risk for lung cancer (OR 1.90, 95% CI 1.21–2.99; p=0.005), and were more susceptible to developing both lung cancer and bronchial obstruction than lung cancer alone (OR 2.11, 95% CI 1.04–4.26; p=0.038).

The rs31489 variant on 5p15.33 is associated with bronchial obstruction, presence and severity of emphysema, and lung cancer.

Around 50–70% of patients diagnosed with lung cancer exhibit some evidence of chronic obstructive pulmonary disease (COPD). Although cigarette smoking is generally considered to be responsible for this correlation, substantial epidemiological data show that the prevalence of COPD in lung cancer patients is 2–6-fold higher compared with the prevalence of COPD in a cohort of heavy smokers, and this was independent of age or pack-yrs smoked [1]. Moreover, reduced forced expiratory volume in 1 s (FEV1) has been shown to increase the risk for incident lung cancer independently of smoking history [2]. COPD is also associated with lung cancer in never-smokers [3], thereby suggesting that additional risk factors might explain the link between COPD and lung cancer.

Emphysema, which is characterised by alveolar destruction, is an important clinical phenotype of COPD. Indeed, COPD patients with emphysema have a poorer health-related quality of life and exhibit an increased risk of non-cancer related respiratory death [4]. Epidemiological data suggest that emphysema and bronchial obstruction independently increase susceptibility to lung cancer. For instance, a large study revealed a >2-fold increase of lung cancer risk in patients with bronchial obstruction, after adjusting for smoking history, and a >3-fold increase in patients with emphysema, after adjusting for both smoking history and COPD severity [2]. Although these data were not confirmed in a second smaller study [5], a 1.85-fold increased risk for lung cancer was subsequently reported in females with a prior

This article has supplementary material available from www.erj.ersjournals.com

history of bronchial obstruction, but as much as a 6.36-fold increased risk when emphysema was diagnosed ⩽9 yrs prior to the lung cancer [6]. Overall, this suggests that emphysema and lung cancer are intertwined, and that common genetic or other unidentified susceptibility factors may predispose to both diseases.

Variants in the genes encoding the nicotinic acetylcholine receptor (nAChR) subunits were the first to be implicated in lung cancer, COPD and emphysema [7–9]. Intriguingly, the same variants were also discovered as genetic predictors of smoking addiction and were subsequently shown to be associated with a more intense smoking behaviour [10, 11]. Although studies assessing the risk effect of nAChR variants on lung cancer, COPD or emphysema were corrected for smoking history, it remains difficult to interpret these associations independently from the patient's smoking behaviour. In contrast, the *TERT-CLPTM1L* locus, which is the second locus that was identified using genome-wide association (GWA) in lung cancer [7], is thought to associate directly with lung cancer, entirely independent of smoking behaviour. Indeed, the association between this locus and lung cancer occurs independently of tobacco usage since it has been replicated in never-smokers [12]. Furthermore, the *TERT-CLPTM1L* locus has also been identified in nonsmoking related cancers, such as cervix and prostate cancer [13], but was not identified in GWA studies involving smoking addiction [11]. The locus maps to two potential cancer susceptibility genes: the telomerase reverse transcriptase (*TERT*) gene, which encodes the telomerase protein that sustains telomere length [14], and the cleft lip palate transmembrane 1-like (*CLPTM1L*) gene, which encodes a protein linked to cisplatin resistance [15].

In an effort to assess the intriguing hypothesis that lung cancer, COPD and emphysema may have a common genetic origin, we tested whether variants in the *TERT-CLPTM1L* locus are also associated with bronchial obstruction or emphysema.

MATERIALS AND METHODS
Study design
This was an observational and prospective study carried out by the Respiratory Division of the University Hospital of Leuven (Leuven, Belgium). The protocol was run on an outpatient basis. The protocol was approved by the Ethics Committee of the University Hospitals Leuven (Leuven), and all subjects consented prior to study participation.

Study subjects
We recruited two independent study populations at the University Hospital of Leuven. The first population consisted of 777 heavy smokers, who were prospectively included between January 2007 and August 2010. Of these, 394 were recruited at the respiratory outpatient clinic because of symptoms suggestive of COPD [9]. In addition, 383 population-based subjects, who were Belgian participants of the Dutch–Belgian randomised lung cancer screening trial (NELSON), were included [16]. Inclusion criteria were a smoking history of ⩾15 pack-yrs, a minimum age of 50 yrs and the availability of complete pulmonary function test data (including diffusing capacity of the lung for carbon monoxide ($D_{L,CO}$) measurements). Patients with suspected asthma or diagnosis of asthma were excluded, as well as patients with other respiratory

diseases affecting pulmonary function. All COPD patients had a stable clinical condition with no exacerbation within 6 weeks before inclusion. From all study subjects, an extensive list of demographic variables (including age, sex and body mass index in $kg \cdot m^{-2}$), a computed tomography (CT) scan of the chest within 1 yr of enrolment and questionnaires determining smoking history were also collected.

The second population consisted of 212 newly diagnosed lung cancer patients, who were prospectively recruited at the Respiratory Oncology Unit (University Hospital of Leuven) between March 2010 and January 2011. The same inclusion criteria as for the first study population were applied, *i.e.* age >50 yrs, a smoking history of ⩾15 pack-yrs and the availability of complete pulmonary function test data prior to medical or surgical treatment. Additionally, the presence of primary lung cancer had to be confirmed by a pathologist. The following cancer subtypes were diagnosed: small cell carcinoma (10.9%), adenocarcinoma (36.3%), squamous cell carcinoma (29.2%) and other nonsmall cell lung carcinoma (mainly large cell and bronchoalveolar subtypes; 23.6%). No data on the radiographic extent of emphysema were obtained for this subgroup. All recruited study participants self-declared Belgian–Flemish ethnicity for three generations.

Pulmonary function testing
All pulmonary function measurements were performed using standardised equipment (Sensormedics Whole Body Plethysmograph; Viasys Healthcare, Vilvoorde, Belgium) according to American Thoracic Society/European Respiratory Society guidelines [17]. Spirometric values were post-bronchodilator measurements. $D_{L,CO}$ was determined by the single-breath carbon monoxide gas transfer method and corrected for alveolar volume but not for haemoglobin concentration [18]. All variables are given as absolute values expressed as % predicted of reference values. To determine whether the *TERT-CLPTM1L* region is a susceptibility locus for COPD, we performed a case–control study in which COPD cases were defined as participants with a post-bronchodilator FEV_1/forced vital capacity (FVC) ratio of <0.70 and FEV_1 <80% pred (*i.e.* Global Initiative for Chronic Obstructive Lung Disease (GOLD) stage II or higher) [19]. Subjects with a FEV_1/FVC>0.70 and FEV_1>80% predicted were selected as healthy smokers. The same definitions were used in three recently published GWA studies [8, 20, 21] to limit misclassification of healthy smokers as GOLD stage I patients.

CT imaging protocol
All 777 participants of the first study population received a chest CT scan allowing the semiquantitative assessment of alveolar destruction. The complete protocol used for CT imaging and quantification of emphysema has been described previously [22]. Briefly, a blinded radiologist specialised in thoracic imaging scored each of the CT scans for the presence and extent of emphysema at three levels in each lung. Emphysema was defined as an area of hypovascular low attenuation, graded at each level with an incremental 5% scale and averaged in a tissue score reflecting the extent of emphysema over both lungs. If emphysema was visually scored in any of the pre-defined fields, the patient was categorised as having emphysema. Based on National Emphysema Treatment Trial (NETT) criteria [22], four categories were generated, yielding an alveolar destruction

score ranging 0–3 (0: no emphysema; 1: emphysema affecting <20% of the lung; 2: emphysema affecting 20–50% of the lung; 3: emphysema affecting >50% of the lung).

Genotyping

Peripheral blood was sampled in potassium EDTA plastic Vacutainer® (BD, Erembodegem, Belgium) tubes and, after centrifugation, germ-line DNA was extracted from the precipitated leukocyte cell fraction according to standard procedures. DNA was aliquoted into 384-well plates and genotyped at the Vesalius Research Center (Leuven). Genotyping for the lung cancer susceptibility variants rs31489, rs4635969 and rs2736100 (identified in [23]) was performed in a blinded manner using iPLEX technology on a MassARRAY Compact Analyser (Sequenom Inc., San Diego, CA, USA), as reported previously [9]. The rs31489, rs2736100 and rs4635969 (National Center for Biotechnology Information dbSNP identifiers) variants are located in noncoding regions of chromosome 5, at Human Genome Build 36 locations 1395714, 1339516 and 1361552, respectively. Automated genotyping calls were generated using the MassARRAY RTTM software and were validated by manual review of the raw mass spectra. Quality control was performed by genotyping 13 samples in duplicate, with a duplicate concordance of 100%. To our knowledge, this is the first study assessing these variants as susceptibility factors for COPD or emphysema.

Statistical analysis

Data are presented as n (%) for categorical variables. Continuous variables are presented as mean±SD or median (interquartile range). Differences in baseline characteristics between patients with bronchial obstruction, emphysema or lung cancer, and healthy participants were compared using the unpaired t-test, Mann–Whitney U-test or Chi-squared test. Hardy–Weinberg disequilibrium was tested using standard Chi-squared analysis (one degree of freedom) and all three single-nucleotide polymorphisms (SNPs) fulfilled the criteria ($p > 0.05$). Differences in baseline characteristics between rs31489 genotypes were compared by univariate analysis, or the Kruskal–Wallis or Chi-square test. The λ statistic was used to calculate the genetic model of inheritance [24]. According to this method, values equal to 0, 0.5 and 1 correspond to a recessive, additive and dominant genetic model, respectively. For the identification of disease-associated genotypes (COPD, emphysema and lung cancer *versus* healthy smokers), we found a λ coefficient of <0.001, 0.089 and 0.055, respectively, suggesting a recessive inheritance risk model. Chi-squared analysis (one degree of freedom) and logistic regression were used to assess the association of rs31489 with presence of bronchial obstruction, emphysema or lung cancer under a recessive model without adjustment or after adjusting for other covariates. The p-value threshold for significance was adjusted for six multiple comparisons using the Bonferroni correction method, resulting in a significance threshold of $p < 0.008$. In particular, we corrected for testing the association between three variants and two variables (COPD and emphysema). We did not correct for additional variables, such as FEV1 and $D_{L,CO}$, as these are closely linked to COPD or emphysema status. We also did not correct for testing the association with lung cancer, as all three variants were identified through a GWA study, in which they were already extensively

replicated. All statistical analyses were performed using SPSS software version 17.0 (SPSS Inc., Chicago, IL, USA).

RESULTS

Population characteristics

In total, 777 heavy smokers free of lung cancer were included in this study. Details on demographics, smoking behaviour and pulmonary function status are shown in table 1. Briefly, the mean age of study participants was 64.9 yrs and 77.7% were male smokers with an average smoking history of 48.8 pack-yrs. Genotyping for the rs31489, rs4635969 and rs2736100 variants succeeded in 739 (95.1%), 776 (99.6%) and 764 (98.7%) participants, respectively. These variants were identified in lung cancer GWA studies and each represents tagging of SNPs of different linkage disequilibrium (LD) blocks: their pairwise r^2-values were 0.415, 0 and 0.003 for rs4635969 and rs31489, rs4635969 and rs2736100, and rs31489 and rs2736100, respectively (according to HapMap for CEU (Utah residents with ancestry from northern and western Europe)-based European ancestry [25]).

The rs31489 variant correlates with reduced pulmonary function

First, we assessed whether any of the three SNPs were associated with demographic parameters, smoking exposure or pulmonary function. Significant differences between the rs31489 genotypes were noticed at the pulmonary function level (table 1). Homozygous carriers of the rs31489 C-allele had significantly lower FEV1 values compared with AA carriers ($p = 0.006$ and $p = 0.014$ for mean absolute values and % pred values, respectively; table 1). Interestingly, a similar association was seen for $D_{L,CO}$, with mean $D_{L,CO}$ values being the lowest for carriers of the CC-genotype ($p = 0.004$; table 1). Adjustments for age, sex, height, pack-yrs and years quit confirmed that the associations of rs31489 with FEV1 and $D_{L,CO}$ were independent from other known risk factors ($p = 0.008$ and $p = 0.007$, respectively). No significant associations were observed for rs4635969 and rs2736100 ($p = 0.650$ and $p = 0.724$ for FEV1 and $p = 0.827$ and $p = 0.374$ for $D_{L,CO}$ values, respectively).

Association of rs31489 with bronchial obstruction

Since rs31489 correlated with reduced FEV1, we also assessed the association between rs31489 and COPD. A case–control association analysis, in which cases were defined as COPD patients diagnosed with moderate-to-severe bronchial obstruction (GOLD stage II or higher), and controls as participants with a FEV1/FVC >0.70 and FEV1 >80% pred, was performed. Baseline characteristics for the 653 participants, of whom 245 were healthy smokers and 408 were GOLD stage II or higher COPD patients, are shown in table 2. In a crude analysis, the rs31489 C-allele was more common in patients with bronchial obstruction (60.8% in obstructive *versus* 54.8% in healthy smokers, $p = 0.037$; table 3). Using the λ statistic to calculate the genetic model of inheritance for rs31489, the C-allele was identified as a recessive susceptibility allele [24]. At the genotypic level, 39.1% of COPD patients were CC-carriers compared with 28.6% of the healthy smokers (OR 1.60, 95% CI 1.13–2.27; $p = 0.010$ according to a recessive risk model). Logistic regression confirmed that CC-carriers significantly increased the risk for bronchial obstruction independently of

TABLE 1	Baseline characteristics for all heavy smokers free of lung cancer according to rs31489 genotypes				
	All subjects	**rs31489**			**p-value**
		CC	**CA**	**AA**	
Subjects	777	261 (35.3)	341 (46.1)	137 (18.5)	
Demographics					
Age yrs	64.9 ± 7.9	64.2 ± 7.8	65.1 ± 8.2	65.1 ± 7.4	0.287
Males	604 (77.7)	194 (74.3)	274 (80.4)	111 (81.0)	0.144
Height cm	170.3 ± 8.8	169.4 ± 9.0	171.2 ± 8.9	170.8 ± 7.8	0.037
Smoking history					
Exposure pack-yrs	48.8 ± 25.0	48.7 ± 25.7	49.0 ± 24.2	47.9 ± 25.9	0.919
Current smokers	370 (47.7)	127 (48.7)	159 (46.6)	66 (48.2)	0.876
Exposure yrs	41.7 ± 9.3	41.2 ± 8.8	41.7 ± 9.4	41.0 ± 9.5	0.731
Time without smoking in former smokers yrs	1.0 (0.0–8.0)	1.0 (0.0–8.0)	2.0 (0.0–10.0)	2.0 (0.0–10.0)	0.237
Pulmonary function tests					
Post-bronchodilator FEV_1 L	2.15 ± 1.04	2.03 ± 0.97	2.30 ± 1.10	2.21 ± 1.00	0.006
Post-bronchodilator FEV_1 % pred	72.9 ± 30.6	69.6 ± 29.1	77.0 ± 32.0	74.4 ± 29.5	0.014
Post-bronchodilator FVC L	3.57 ± 1.07	3.46 ± 1.00	3.71 ± 1.13	3.62 ± 1.04	0.019
Post-bronchodilator FVC % pred	96.8 ± 22.0	95.1 ± 21.3	98.7 ± 22.7	97.4 ± 21.2	0.146
FEV_1/FVC	0.58 ± 0.17	0.56 ± 0.17	0.60 ± 0.17	0.59 ± 0.17	0.058
$D_{L,CO}$ mmol·min^{-1}·kPa^{-1}	5.73 ± 2.37	5.42 ± 2.25	6.08 ± 2.42	5.87 ± 2.39	0.004
$D_{L,CO}$ % pred	65.3 ± 23.5	62.3 ± 22.6	68.7 ± 24.1	66.4 ± 22.8	0.004

Data are presented as n, n (%), mean±SD or median (interquartile range), unless otherwise stated. Percentages represent row percentages. Of the 777 heavy smokers (free of lung cancer) genotyping for rs31489 succeeded in 739 participants (95.1%). Due to technical limitations, six values (<1%) are missing for diffusing capacity of the lung for carbon monoxide ($D_{L,CO}$). For all other variables, there were no missing data. FEV_1: forced expiratory volume in 1 s; % pred: % predicted; FVC: forced vital capacity.

age, sex, height, pack-yrs and years quit (OR 1.82, 95% CI 1.24–2.69; p=0.002; table 3). The rs31489 variant did not, however, correlate with severity of COPD since genotype distributions between the different GOLD stages were similar (p=0.268; online supplementary material table S1). We did not find a significant association between rs4635969 or rs2736100 and bronchial obstruction (p=0.636 and p=0.598 for genotypes, respectively).

Association of rs31489 with emphysema and severity of emphysema

Since rs31489 correlated with $D_{L,CO}$, we also investigated the association of rs31489 with emphysema. Of the 408 COPD patients, 318 (78%) patients exhibited emphysema, as assessed by CT (see table 2 for baseline characteristics). A case–control analysis between these 318 emphysema patients and the 245 healthy smokers subsequently revealed that the C-allele was more common in emphysema patients than in healthy smokers (62.5 *versus* 54.8%, respectively; p=0.010; table 3). At the genotypic level, 42.1% of the emphysema patients were CC-carriers compared with 28.6% of the healthy smokers (OR 1.80, 95% CI 1.29–2.56; p=0.001 according to a recessive risk model). Logistic regression confirmed that the association between rs31489 and emphysema was independent of age, sex, height, pack-yrs and years quit (OR 2.04, 95% CI 1.41–2.94; p=1.73×10^{-4} in the recessive model; table 3). The association remained significant after including FEV_1 as a covariate in the analysis (OR 2.07, 95% CI 1.23–3.46; p=0.006). We did not find a significant association between rs4635969 or rs2736100 and emphysema (p=0.620 and p=0.687 for genotypes, respectively).

To assess severity of emphysema, emphysema was scored at three levels in each lung and the tissue score, which is the average of the scores in the individual six fields and is expressed in percentage of the lung affected, was calculated. CC-carriers exhibited a significantly higher score in each of the six fields compared with the other rs31489 genotypes (p<0.05 for every field; data not shown). Additionally, CC-carriers also exhibited a higher tissue score (p=0.003; online supplementary material table S2). Next, based on the NETT criteria [22], the alveolar destruction score was generated by categorising the tissue score into four classes of increasing emphysema severity, yielding an alveolar destruction score ranging 0–3 (0: no emphysema; 1: emphysema affecting <20% of the lung; 2: emphysema affecting 20–50% of the lung; 3: emphysema affecting >50% of the lung). CC-carriers more frequently exhibited emphysema in >50% of the lung than AA-carriers (18.8 *versus* 7.6%; p=0.010; online supplementary material table S2). Overall, these data suggest that rs31489 correlates with the severity of emphysema in COPD patients.

Association of rs31489 with COPD in lung cancer patients

Next, we assessed the rs31489 variant in an independent cohort of 212 newly diagnosed lung cancer patients and explored whether rs31489 contributes to a shared genetic predisposition of bronchial obstruction and lung cancer. Details on demographics, smoking behaviour and pulmonary function status of the 212 patients are shown in table 4. Briefly, all recruited lung cancer patients were heavy smokers with an average pack-yrs history of 43.6 yrs, and 72.6% were male with an average age of 66.2 yrs. Genotyping for rs31489, rs4635969 and rs2736100 was

TABLE 2	Baseline characteristics for healthy smokers, and chronic obstructive pulmonary disease (COPD) and emphysema patients		
	Healthy smokers	**COPD**	**Emphysema**
Subjects	245	408	318
Demographics			
Age yrs	61.5 ± 5.9	67.5 ± 8.6	67.7 ± 8.5
Males	183 (74.7)	324 (79.4)	253 (79.6)
Height cm	171.5 ± 9.3	169.1 ± 8.4	169.3 ± 8.4
Smoking history			
Exposure pack-yrs	42.5 ± 20.3	52.5 ± 26.3	53.2 ± 27.3
Current smokers	116 (47.9)	176 (43.8)	128 (40.4)
Exposure yrs	38.7 ± 7.9	43.5 ± 10.1	43.5 ± 10.1
Time without smoking in former smokers yrs	1.0 (0.0–9.0)	1.5 (0.0–8.0)	2.0 (0.0–8.2)
Pulmonary function tests			
Post-bronchodilator FEV_1 L	3.12 ± 0.71	1.36 ± 0.56	1.29 ± 0.54
Post-bronchodilator FEV_1 % pred	103.6 ± 14.8	47.7 ± 16.7	45.4 ± 16.4
Post-bronchodilator FVC L	4.10 ± 0.93	3.01 ± 0.86	3.02 ± 0.85
Post-bronchodilator FVC % pred	108.8 ± 15.1	83.6 ± 18.6	84.0 ± 19.0
FEV_1/FVC	0.76 ± 0.04	0.45 ± 0.12	0.42 ± 0.11
$D_{L,CO}$ mmol·min^{-1}·kPa^{-1}	7.50 ± 1.87	4.30 ± 1.72	3.92 ± 1.47
$D_{L,CO}$ % pred	83.6 ± 15.5	50.3 ± 17.9	45.9 ± 15.6

Data are presented as n, mean$\pm$SD, n (%) or median (interquartile range). Of the 408 heavy smokers with COPD (Global Initiative for Chronic Obstructive Lung Disease (GOLD) stage II or higher), 52 (12.7%) subjects were derived from the Dutch–Belgian randomised lung cancer screening trial (NELSON) study and 356 (87.3%) were recruited at the respiratory outpatient clinic. In the COPD group, 318 patients were diagnosed with emphysema. As expected, baseline characteristics, smoking history and pulmonary function tests differed significantly between COPD patients and healthy smokers (p<0.001). Only the number of current smokers and the time without smoking did not differ (p=0.305 and p=0.621, respectively). The same observations were made when comparing baseline characteristics between emphysema patients and healthy smokers. FEV_1: forced expiratory volume in 1 s; % pred: % predicted; FVC: forced vital capacity; $D_{L,CO}$: diffusing capacity of the lung for carbon monoxide.

successful in 203 (95.7%), 208 (98.1%) and 210 (99.1%) subjects, respectively.

In the first analysis, we compared rs31489 allele frequencies between 212 lung cancer patients and 245 healthy smokers (table 4). As expected, the at-risk C-allele was more common in patients with lung cancer (61.3 versus 54.8%; p=0.049; table 5). At the genotypic level, 39.4% of the lung cancer patients were CC-carriers compared with 28.6% of the healthy smokers (OR 1.65, 95% CI 1.11–2.45; p=0.018 according to a recessive risk model; table 5). Regression analysis revealed that the association between rs31489 and lung cancer was independent of age,

sex, height, pack-yrs smoked and years quit (OR 1.90, 95% CI 1.21–2.99; p=0.005). Since rs31489 was directly correlated with FEV_1 in the cohort of heavy smokers (table 1) and was significantly associated with COPD (table 3), we also assessed whether the at-risk rs31489 C-allele was associated with lung cancer independently of FEV_1. A similar regression analysis was performed, while correcting for FEV_1 and allowing for the interaction of FEV_1 with rs31489. CC-carriers of rs31489 displayed a significantly higher risk of developing lung cancer (p=0.006), suggesting a direct effect of rs31489 on lung cancer risk independent of FEV_1. Remarkably, the interaction between rs31489 and FEV_1 was also significant, indicating that smokers

TABLE 3	Association of rs31489 with chronic obstructive pulmonary disease (COPD) and emphysema						
Genotype	Healthy smokers	COPD	OR (95% CI)	p-value	Emphysema	OR (95% CI)	p-value
C	264 (54.8)	457 (60.8)	0.78 (0.62–0.98)	0.037	365 (62.5)	1.38 (1.08–1.76)	0.010
A	218 (45.2)	295 (39.2)			219 (37.5)		
CC	69 (28.6)	147 (39.1)	1.82 (1.24–2.69)	0.002[#]	123 (42.1)	2.04 (1.41–2.94)	1.73×10^{-4}[#]
CA	126 (52.3)	163 (43.4)			119 (40.8)		
AA	46 (19.1)	66 (17.6)			50 (17.1)		

Data are presented as n (%), unless otherwise stated. Percentages represent row percentages. Allele and genotype frequencies of rs31489 in COPD (Global Initiative for Chronic Obstructive Lung Disease (GOLD) stage II or higher; n=408) and emphysema patients (n=318) were compared with frequencies in healthy smokers (forced expiratory volume in 1 s (FEV_1)/forced vital capacity ratio of >0.70 and FEV_1 >80% predicted; n=245). [#]: under the assumption of a recessive genetic model as assessed by logistic regression and after correction for covariates.

TABLE 4	Baseline characteristics of heavy smokers diagnosed with lung cancer	
	Healthy smokers	**Lung cancer patients**
Subjects	245	212
Demographics		
Age yrs	61.5 ± 5.9	66.2 ± 9.8
Males	183 (74.7)	154 (72.6)
Height cm	171.5 ± 9.3	170.0 ± 7.5
Smoking history		
Exposure pack-yrs	42.5 ± 20.3	43.6 ± 31.5
Current smokers	116 (47.9)	114 (54.0)
Time without smoking in former smokers yrs	1.0 (0.0–9.0)	0.0 (0.0–9.0)
Pulmonary function tests		
Post-bronchodilator FEV_1 L	3.12 ± 0.71	2.27 ± 0.70
Post-bronchodilator FEV_1 % pred	103.6 ± 14.8	79.0 ± 21.0
Post-bronchodilator FVC L	4.10 ± 0.93	3.46 ± 0.92
Post-bronchodilator FVC % pred	108.8 ± 15.1	96.1 ± 19.0
FEV_1/FVC	0.76 ± 0.04	66.3 ± 12.6
$D_{L,CO}$ mmol·min^{-1}·kPa^{-1}	7.50 ± 1.87	5.78 ± 1.76
$D_{L,CO}$ % pred	83.6 ± 15.5	67.2 ± 17.3

Data are presented as n, mean $\pm$ SD, n (%) or median (interquartile range). Of the 245 healthy smokers, genotyping for rs31489 succeeded in 241 (98%) participants. Of the 212 lung cancer patients, genotyping for rs31489 succeeded in 203 (95.7%) participants. Due to technical limitations, six values for diffusing capacity of the lung for carbon monoxide ($D_{L,CO}$), including one in the group of cancer patients, are missing. For all other variables, complete data are presented. Lung cancer patients and healthy smokers were matched with respect to their sex, height and smoking history (p=0.619 and p=0.056 for sex and height, respectively; p=0.653, p=0.196 and p=0.621 for pack-yrs history, current smoking status and years quit in former smokers, respectively). All pulmonary function tests differed significantly (p<0.001). FEV_1: forced expiratory volume in 1 s; % pred: % predicted; FVC: forced vital capacity.

carrying the CC-genotype are more susceptible to having a reduced FEV_1 and to developing lung cancer (p=0.002 for interaction term). Consequently, when stratifying lung cancer patients into two categories, based on the presence or absence of COPD, we observed that rs31489 CC-carriers were almost twice as frequent in lung cancer patients with COPD as in lung cancer patients without COPD (65.0 *versus* 35.0%, respectively). Binary logistic regression analysis, while correcting for age,

sex, height, pack-yrs smoked and years quit, confirmed that CC-carriers were more susceptible to developing both lung cancer and bronchial obstruction than lung cancer alone (OR 2.11, 95% CI 1.04–4.26; p=0.038).

DISCUSSION

The current study identified rs31489 on chromosome 5p15.33 as a susceptibility variant for bronchial obstruction and emphysema. Indeed, rs31489 correlated significantly with reduced lung function. In particular, CC-carriers exhibited significantly lower FEV_1 values, as well as a reduced $D_{L,CO}$. In addition, rs31489 conferred an increased risk of developing bronchial obstruction, whereas CT scans obtained from the same patients provided evidence for an association between rs31489 and the presence and severity of emphysema. Each of these associations remained significant after adjustment for smoking behaviour and other clinical variables, including age and height. In addition, our data confirm that rs31489 also acts as a shared susceptibility factor for lung cancer and bronchial obstruction. Indeed, carriers of the rs31489 CC-genotype were significantly more common in lung cancer patients than in cancer-free healthy smokers. This association remained significant after incorporating FEV_1 as a covariate in the analysis. Intriguingly, rs31489 interacted significantly with FEV_1, indicating that rs31489 also correlates with FEV_1 in lung cancer patients. Stratification of lung cancer patients into those with and without bronchial obstruction confirmed that CC-carriers were two times more susceptible to developing both bronchial obstruction and lung cancer than lung cancer alone.

TABLE 5	Association of rs31489 with lung cancer *versus* healthy smokers			
Genotype	Healthy smokers	Lung cancer patients	OR (95% CI)	p-value[#]
C	264 (54.8)	249 (61.3)	1.30 (1.00–1.71)	0.049
A	218 (45.2)	157 (38.7)		
CC	69 (28.6)	80 (39.4)	1.65 (1.11–2.45)	0.018
CA	126 (52.3)	89 (43.8)		
AA	46 (19.1)	34 (16.7)		

Data are presented as n (%), unless otherwise stated. Percentages represent row percentages. Allele and genotype frequencies of rs31489 were compared between lung cancer cases (n=212) and healthy smokers (forced expiratory volume in 1 s (FEV_1)/forced vital capacity ratio of >0.70 and FEV_1 >80% predicted; n=245). [#]: under the assumption of a recessive model (CC *versus* CA/AA).

These findings are in agreement with previous observations by YOUNG and co-workers [26, 27] reporting that variants in the Hedgehog-interacting protein (*HHIP*) and nAChR subunit genes are more strongly associated with lung cancer and COPD than lung cancer alone. On a more general level, the more pronounced association of these loci with both COPD and lung cancer suggests that the pathogenetic pathways for smoking-related pulmonary disorders are shared.

The *TERT-CLPTM1L* locus contains numerous genetic variants that are strongly linked to each other and that are located in two genes, which could each be functionally affected by these variants [7, 12, 13, 23]. The first gene is *CLPTM1L*, in which the rs31489 variant is also located. Overexpression of this gene has been shown to enhance apoptosis in response to cisplatin in an ovarian tumour cell line [15]. The cytotoxic effects of cisplatin are attributed to its covalent binding to nuclear DNA, referred to as DNA adduct formation. Interestingly, there is recent evidence that a variant in the *CLPTM1L* gene (rs402710; pairwise with rs31489 $r^2=0.67$) favours DNA damage through DNA adduct formation [28]. Hence, variants in the *CLPTM1L* gene may be functionally involved in the development of COPD or lung cancer by enhancing DNA damage in response to noxious smoke-related particles. In addition, the region of LD wherein rs31489 resides also contains the 5′ end of the *TERT* gene. The function of this gene in human diseases, especially in carcinogenesis, is much better understood than that of *CLPTM1L* [14]. Telomerase, the enzyme product of the *TERT* gene, ensures that the ends of human chromosomes are covered by functional telomeres. An appropriate telomere length is essential for the cell's integrity since it prevents chromosomal end-to-end fusions and subsequent chromosomal aberrations [29]. Telomerase activity is typically higher in cells where a higher proliferative potential is needed, for example in germ-line stem cells. In differentiated somatic cells, however, telomerase levels are low or undetectable and telomeres get shorter each time a cell divides, until the point where telomere length becomes insufficient to prevent genomic instability and proliferative senescence is induced [30]. Intriguingly, rare germ-line mutations in *TERT* have been found in familial disorders, such as dyskeratosis congenita, aplastic anaemia and familial idiopathic pulmonary fibrosis [14], which are characterised by a limited cell and tissue renewal capacity and a higher rate of carcinogenesis. Lymphocytes from affected individuals are characterised by decreased telomerase activity and significantly shorter telomeres.

Based on these observations, it has been proposed that common variants in *TERT-CLPTM1L*, which were identified as susceptibility factors for cancer, are linked to reduced telomerase activity, short telomeres, limited tissue renewal capacity and chromosomal instability [31]. Indeed, a correlation between reduced telomere length and rs401681, which is strongly linked to rs31489 ($r^2=0.87$), has already been observed [13], whereas MIRABELLO *et al.* [32] also reported a direct correlation between rs31489 and shorter telomeres under a recessive genetic model. Shorter telomeres have also been observed in peripheral leukocytes from COPD patients [33] and a decreased telomere length was observed in the tissue of emphysematous lungs [34]. Intriguingly, a recent GWA study also identified genetic variants in *BICD1*, which are known to correlate with short telomeres, as susceptibility factors for

emphysema [35]. In addition to genetic predisposition, which we observed in this study, telomere shortening in COPD and emphysema may also result from an increased number of cell divisions due to ongoing inflammation [36]. Short telomeres could possibly also predispose to more severe emphysema, as they limit tissue renewal capacity and enhance alveolar destruction in response to environmental factors, such as smoking [30]. As already mentioned, critically short telomeres also lead to genomic instability and promote carcinogenesis [29], which could explain the clinically observed link between emphysema and lung cancer.

Strengths of the current study are the extensive characterisation of pulmonary function, including $D_{L,CO}$ measurements of all participants, and the availability of CT scans for all heavy smokers. In addition, we included a second study population to replicate the association of rs31489 with COPD. Since this population consisted of lung cancer patients with known lung function status, we were not only able to confirm the association of rs31489 with bronchial obstruction, but also to establish an independent correlation of rs31489 with lung cancer. However, since the latter study population was small, follow-up studies will be needed to confirm the independent association of the *TERT-CLPTM1L* locus with COPD, emphysema and lung cancer. Another limitation is that we used subjective visual scoring to assess the severity of emphysema rather than an objective and automated quantification method. Moreover, although several studies have shown that the *TERT-CLPTM1L* locus has functional effects on telomere length and DNA adduct formation, the causal variants within this locus are still not known. As a consequence, we were limited to assessing the association of three tagging SNPs, previously identified as susceptibility variants for lung cancer in this locus, with COPD and emphysema. Finally, although our analysis may be subject to population heterogeneity, we were not able to apply a genomic control method, since genome-wide data from unrelated variants in unrelated genes were not available for this study.

Overall, the conclusion of the current study is that the rs31489 variant on chromosome 5p15.33 is strongly and independently associated with bronchial obstruction, the presence and severity of emphysema, and lung cancer. On a more general level, these data support the hypothesis that lung cancer, bronchial obstruction and emphysema share common genetic bases. Additionally, this study proposes an intriguing hypothesis, suggesting that genetic variability in telomere shortening could at least partially explain the link between bronchial obstruction, emphysema and lung cancer.

REFERENCES

1 Young RP, Hopkins RJ, Christmas T, *et al.* COPD prevalence is increased in lung cancer, independent of age, sex and smoking history. *Eur Respir J* 2009; 34: 380–386.

2 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: 738–744.

3 Turner MC, Chen Y, Krewski D, *et al.* Chronic obstructive pulmonary disease is associated with lung cancer mortality in a prospective study of never smokers. *Am J Respir Crit Care Med* 2007; 176: 285–290.

4 Haruna A, Muro S, Nakano Y, *et al.* CT scan findings of emphysema predict mortality in COPD. *Chest* 2010; 138: 635–640.

5 Maldonado F, Bartholmai BJ, Swensen SJ, *et al.* Are airflow obstruction and radiographic evidence of emphysema risk factors for lung cancer? A nested case–control study using quantitative emphysema analysis. *Chest* 2010; 138: 1295–1302.

6 Schwartz AG, Cote ML, Wenzlaff AS, *et al.* Chronic obstructive lung diseases and risk of non-small cell lung cancer in women. *J Thorac Oncol* 2009; 4: 291–299.

7 Brennan P, Hainaut P, Boffetta P. Genetics of lung-cancer susceptibility. *Lancet Oncol* 2011; 12: 399–408.

8 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: e1000421.

9 Lambrechts D, Buysschaert 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: 486–493.

10 Le Marchand L, Derby KS, Murphy SE, *et al.* Smokers with the CHRNA lung cancer-associated variants are exposed to higher levels of nicotine equivalents and a carcinogenic tobacco-specific nitrosamine. *Cancer Res* 2008; 68: 9137–9140.

11 Thorgeirsson TE, Gudbjartsson DF, Surakka I, *et al.* Sequence variants at *CHRNB3-CHRNA6* and *CYP2A6* affect smoking behavior. *Nat Genet* 2010; 42: 448–453.

12 Wang Y, Broderick P, Matakidou A, *et al.* Role of 5p15.33 (*TERT-CLPTM1L*), 6p21.33 and 15q25.1 (*CHRNA5-CHRNA3*) variation and lung cancer risk in never-smokers. *Carcinogenesis* 2010; 31: 234–238.

13 Rafnar T, Sulem P, Stacey SN, *et al.* Sequence variants at the *TERT-CLPTM1L* locus associate with many cancer types. *Nat Genet* 2009; 41: 221–227.

14 Calado RT, Young NS. Telomere diseases. *N Engl J Med* 2009; 361: 2353–2365.

15 Yamamoto K, Okamoto A, Isonishi S, *et al.* A novel gene, *CRR9*, which was up-regulated in CDDP-resistant ovarian tumor cell line, was associated with apoptosis. *Biochem Biophys Res Commun* 2001; 280: 1148–1154.

16 van Iersel CA, de Koning HJ, Draisma G, *et al.* Risk-based selection from the general population in a screening trial: selection criteria, recruitment and power for the Dutch–Belgian randomised lung cancer multi-slice CT screening trial (NELSON). *Int J Cancer* 2007; 120: 868–874.

17 Miller MR, Hankinson J, Brusasco V, *et al.* Standardisation of spirometry. *Eur Respir J* 2005; 26: 319–338.

18 Rosenberg E. The 1995 update of recommendations for a standard technique for measuring the single-breath carbon monoxide diffusing capacity (transfer factor). *Am J Respir Crit Care Med* 1996; 154: 265–266.

19 Rabe KF, Hurd S, Anzueto A, *et al.* Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease: GOLD executive summary. *Am J Respir Crit Care Med* 2007; 176: 532–555.

20 Cho MH, Boutaoui N, Klanderman BJ, *et al.* Variants in *FAM13A* are associated with chronic obstructive pulmonary disease. *Nat Genet* 2010; 42: 200–202.

21 Wan ES, Cho MH, Boutaoui N, *et al.* Genome-wide association analysis of body mass in chronic obstructive pulmonary disease. *Am J Respir Cell Mol Biol* 2010; [Epub ahead of print DOI: 10.1165/rcmb.2010-0294OC].

22 National Emphysema Treatment Trial Research Group. Patients at high risk of death after lung-volume-reduction surgery. *N Engl J Med* 2001; 345: 1075–1083.

23 Landi MT, Chatterjee N, Yu K, *et al.* A genome-wide association study of lung cancer identifies a region of chromosome 5p15 associated with risk for adenocarcinoma. *Am J Hum Genet* 2009; 85: 679–691.

24 Minelli C, Thompson JR, Abrams KR, *et al.* The choice of a genetic model in the meta-analysis of molecular association studies. *Int J Epidemiol* 2005; 34: 1319–1328.

25 Altshuler DM, Gibbs RA, Peltonen L, *et al.* Integrating common and rare genetic variation in diverse human populations. *Nature* 2010; 467: 52–58.

26 Young RP, Hopkins RJ, Hay BA, *et al.* Lung cancer gene associated with COPD: triple whammy or possible confounding effect? *Eur Respir J* 2008; 32: 1158–1164.

27 Young RP, Whittington CF, Hopkins RJ, *et al.* Chromosome 4q31 locus in COPD is also associated with lung cancer. *Eur Respir J* 2010; 36: 1375–1382.

28 Zienolddiny S, Skaug V, Landvik NE, *et al.* The *TERT-CLPTM1L* lung cancer susceptibility variant associates with higher DNA adduct formation in the lung. *Carcinogenesis* 2009; 30: 1368–1371.

29 Hackett JA, Feldser DM, Greider CW. Telomere dysfunction increases mutation rate and genomic instability. *Cell* 2001; 106: 275–286.

30 Hao LY, Armanios M, Strong MA, *et al.* Short telomeres, even in the presence of telomerase, limit tissue renewal capacity. *Cell* 2005; 123: 1121–1131.

31 Baird DM. Variation at the *TERT* locus and predisposition for cancer. *Expert Rev Mol Med* 2010; 12: e16.

32 Mirabello L, Yu K, Kraft P, *et al.* The association of telomere length and genetic variation in telomere biology genes. *Hum Mutat* 2010; 31: 1050–1058.

33 Savale L, Chaouat A, Bastuji-Garin S, *et al.* Shortened telomeres in circulating leukocytes of patients with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2009; 179: 566–571.

34 Tsuji T, Aoshiba K, Nagai A. Alveolar cell senescence in patients with pulmonary emphysema. *Am J Respir Crit Care Med* 2006; 174: 886–893.

35 Kong X, Cho MH, Anderson W, *et al.* Genome-wide association study identifies *BICD1* as a susceptibility gene for emphysema. *Am J Respir Crit Care Med* 2011; 183: 43–49.

36 Hogg JC, Chu F, Utokaparch S, *et al.* The nature of small-airway obstruction in chronic obstructive pulmonary disease. *N Engl J Med* 2004; 350: 2645–2653.