Genome-wide interaction study of gene-by-occupational exposure and effects on FEV_1 levels

Kim de Jong, PhD,[a,b] Judith M. Vonk, PhD,[a,b] Wim Timens, PhD,[b,c] Yohan Bossé, PhD,[d] Don D. Sin, PhD,[e] Ke Hao, PhD,[f] Hans Kromhout, PhD,[g] Roel Vermeulen, PhD,[g] Dirkje S. Postma, PhD,[b,h] and H. Marike Boezen, PhD[a,b] *Groningen and Utrecht, The Netherlands, Quebec City, Quebec, Canada, Vancouver, British Columbia, Canada, and New York, NY*

Background: Chronic obstructive pulmonary disease (COPD) is a complex disease characterized by impaired lung function and airway obstruction resulting from interactions between multiple genes and multiple environmental exposures. Thus far, genome-wide association studies have largely disregarded environmental factors that might trigger the development of lung function impairment and COPD, such as occupational exposures, which are thought to contribute to 15% to 20% of the COPD prevalence.

Objectives: We performed a genome-wide interaction study to identify novel susceptibility loci for occupational exposure to biological dust, mineral dust, and gases and fumes in relation to FEV_1 levels.

Methods: We performed an identification analysis in 12,400 subjects from the LifeLines cohort study and verified our findings in 1436 subjects from a second independent cohort, the Vlagtwedde-Vlaardingen cohort. Additionally, we assessed whether replicated single nucleotide polymorphisms (SNPs) were cis-acting expression (mRNA) quantitative trait loci in lung tissue.

Results: Of the 7 replicated SNPs that interacted with one of the occupational exposures, several identified loci were plausible candidates that might be involved in biological pathways leading to lung function impairment, such as *PCDH9 and GALNT13*. Two of the 7 replicated SNPs were cis-acting expression quantitative trait loci associated with gene expression of *PDE4D* and *TMEM176A* in lung tissue.

Conclusion: This genome-wide interaction study on occupational exposures in relation to the level of lung function identified several novel genes. Further research should determine whether the identified genes are true susceptibility loci for occupational exposures and whether these SNP-by-exposure interactions consequently contribute to the development of COPD. (J Allergy Clin Immunol 2015;■■■:■■■-■■■.)

Key words: *Genetics, occupational exposures, interactions, gene expression, lung function*

Chronic obstructive pulmonary disease (COPD) is a common disease with a high morbidity and mortality worldwide.[1] Despite the recognition of COPD as a major and increasing public health problem, there is still limited understanding about the cellular and molecular pathways driving the development of this disease.[2] COPD is a complex disease characterized by impaired lung function and airway obstruction resulting from interactions between multiple genes and multiple environmental exposures.[3]

Although tobacco smoking is still considered the main risk factor for reduced lung function, accelerated lung function decrease, and development of COPD, about 15% to 20% of all COPD cases have been attributed to occupational exposures.[4] Occupational exposure to vapors, gases, dusts, and fumes is common and has been associated with impaired lung function.[5-7]

Candidate gene approaches have shown that genetic variants can directly affect the level of lung function, decrease in lung function, and risk of COPD.[8] Moreover, genetic variation might indirectly affect respiratory outcomes by increasing the susceptibility to environmental exposures with known detrimental effects. For example, subjects carrying the "slow" genotype of the gene coding for the enzyme microsomal epoxide hydrolase had a significantly accelerated annual decrease in lung function when occupationally exposed to endotoxin.[9]

From [a]the Department of Epidemiology, [b]the Groningen Research Institute for Asthma and COPD (GRIAC), [c]the Department of Pathology and Medical Biology, and [h]the Department of Pulmonary Diseases, University of Groningen, University Medical Center Groningen; [d]Institut universitaire de cardiologie et de pneumologie de Québec, Department of Molecular Medicine, Laval University, Quebec City; [e]the Department of Medicine and Center for Heart Lung Innovation, University of British Columbia, Vancouver; [f]the Department of Genetics and Genomic Sciences, Icahn Institute of Genomics and Multiscale Biology, Icahn School of Medicine at Mount Sinai, New York; and [g]the Division of Environmental Epidemiology, Institute for Risk Assessment Sciences (IRAS), University of Utrecht.

Information on members of the LifeLines Cohort Study is provided in the acknowledgments section.

K.d.J. was funded by Research School GUIDE, University of Groningen, University Medical Center Groningen. The LifeLines cohort study was supported by the Dutch Ministry of Health, Welfare and Sport; the Ministry of Economic Affairs, Agriculture and Innovation; the province of Groningen; the European Union (regional development fund); the Northern Netherlands Provinces (SNN); the Netherlands Organisation for Scientific Research (NWO); University Medical Center Groningen (UMCG); the University of Groningen; de Nierstichting (the Dutch Kidney Foundation); and the Diabetes Fonds (the Diabetic Foundation). The Vlagtwedde-Vlaardingen cohort was supported by the Ministry of Health and Environmental Hygiene of The Netherlands and Netherlands Asthma Fund grant no. 187 and Netherlands Asthma Fund grant no. 3.2.02.51, the Stichting Astma Bestrijding, BBMRI-NL, and the European Respiratory Society COPD research award 2011 (to H.M.B.). Y.B. is the recipient of a Junior 2 Research Scholar award from the Fonds de recherche Québec–Santé (FRQS). The lung tissue database was funded by Merck Research Laboratories. The sponsors of this study had no role in the design of the current study, data collection, analysis, and interpretation, or in writing and submitting of the manuscript.

Disclosure of potential conflict of interest: W. Timens has received consulting fees from Pfizer; has received lecture fees from GlaxoSmithKline, Chiesi, and Roche Diagnostics; and has received research support from the Dutch Asthma Fund. D. D. Sin has received payment for serving on the International COPD Advisory Board for Merck Frosst and the International COPD Advisory Board for Almirall; has received research support from AstraZeneca; has received payment for lectures from GlaxoSmithKline, Nycomed, and AstraZeneca; and has received travel expenses from GlaxoSmithKline. D. S. Postma has received consulting fees from AstraZeneca, Chiesi, Boehringer Ingelheim, GlaxoSmithKline, Takeda, and TEVA and has received research support from Chiesi. The rest of the authors declare that they have no relevant conflicts.

Received for publication April 14, 2014; revised March 16, 2015; accepted for publication March 31, 2015.

Corresponding author: H. Marike Boezen, PhD, University of Groningen, University Medical Center Groningen, Department of Epidemiology, Hanzeplein 1, 9700 RB Groningen, The Netherlands. E-mail: h.m.boezen@umcg.nl.

0091-6749/$36.00

© 2015 American Academy of Allergy, Asthma & Immunology

http://dx.doi.org/10.1016/j.jaci.2015.03.042

Abbreviations used
COPD: Chronic obstructive pulmonary disease
DC: Dendritic cell
eQTL: Expression quantitative trait locus
GWAS: Genome-wide association study
GWIS: Genome-wide interaction study
JEM: Job Exposure Matrix
SNP: Single nucleotide polymorphism

Whereas candidate gene studies have mostly been driven by hypotheses relying on known biological pathways, hypothesis-free genome-wide association studies (GWASs) have sought to identify novel genetic loci associated with lung function levels and COPD. The first GWAS in this field, published in 2007 by Wilk et al,[10] identified a single nucleotide polymorphism (SNP) in *GSTO2* associated with both FEV_1 and forced vital capacity levels. Two years later, Pillai et al[11] published the first GWAS on COPD, identifying SNPs in the *HHIP* and *CHRNA3*/5 loci. Later GWASs showed associations with lung function levels and COPD for SNPs in *HHIP*, *AGER*, and *FAM13A*, among others.[12-15] Recent studies have identified even more loci by increasing sample size[16,17]; focusing on other phenotypes, such as lung function decrease[18] or percentage of emphysema[19]; or both.

Thus far, GWASs have largely disregarded environmental factors that might underlie the development of impaired lung function and eventually complex diseases, such as COPD.[3] Therefore the next step is to perform genome-wide interaction studies (GWISs) to identify genetic loci that affect susceptibility for the effects of known harmful exposures, such as occupational exposures. Recently, a genome-wide study was published that investigated genetic susceptibility to dust exposure in relation to FEV_1 level.[20] This study identified 1 genome-wide significant SNP in *SLC38A8*, yet findings were not verified in a second cohort and therefore remain largely inconclusive.

The current GWIS aimed to identify novel susceptibility loci for several types of occupational exposures (ie, biological dust, mineral dust, and gases and fumes) in relation to the FEV_1 level in a general population cohort. We used a second independent cohort to verify our initial findings. Furthermore, the functional meaning of newly identified SNPs interacting with occupational exposure on FEV_1 was extended to gene expression in lung tissue. Consequently, such findings might contribute to the understanding of the biological pathways driving lung function impairment. Moreover, they might shed light on the identification of susceptible subgroups within the general population and eventually might help to set exposure limits based on the most susceptible subgroups.

METHODS

Identification and replication samples

The identification sample included subjects from the Dutch general population–based LifeLines cohort study. The LifeLines cohort study is a 3-generation study for which first a random sample of persons aged between 25 and 50 years are invited to participate. These index cases subsequently invite their family members to also take part (parents, partners, parents in law, and children).[21] During the baseline visit, a blood sample was taken, and subjects filled in a standardized questionnaire and underwent a medical examination, including prebronchodilator spirometry according to American Thoracic Society guidelines. A subset of genetically unrelated LifeLines subjects was selected for genome-wide genotyping. DNA was isolated from whole blood. Genotypes and complete data on all covariates were available for 12,400 subjects (all with baseline visits between December 2006 and February 2011).

We included 1436 subjects having full data on genotypes and covariates from the Vlagtwedde-Vlaardingen cohort to verify our initial findings. The Vlagtwedde-Vlaardingen cohort is a prospective general population–based cohort. The cohort started with baseline measurements in 1965, 1967, and 1969 for which exclusively white subjects were randomly selected. After baseline, there were follow-up surveys every 3 years until the last survey in 1989-1990. In the current study we used prebronchodilator spirometry according to European Respiratory Society guidelines and phenotype data that were obtained from the last survey in 1989-1990. DNA was isolated from neutrophil depots in peripheral blood samples drawn in 1989-1990.

Both the LifeLines cohort study and the Vlagtwedde-Vlaardingen studies were approved by the Medical Ethics Committee of the University Medical Center Groningen, Groningen, The Netherlands. All subjects provided written informed consent.

Occupational exposure

Self-reported job titles and descriptions from both samples were coded according to the International Standard Classification of Occupations, version 1988.[22] These 4-digit codes were used to objectively estimate job-specific exposure to biological dust, mineral dust, and gases and fumes by using the ALOHA+ Job Exposure Matrix (JEM).[7,23] The ALOHA+ JEM classifies subjects based on the 4-digit job codes into no-, low-, and high-exposure categories (0, 1, and 2, respectively).

Genotyping and quality control

In both cohorts genotyping was performed with Illumina CytoSNP-12 arrays (Illumina, San Diego, Calif). SNPs that fulfilled the quality control criteria were included: genotype call rate, 95% or greater; minor allele frequency, 1% or greater; and Hardy-Weinberg equilibrium cutoff, P value of 10^{-4} or greater. Samples with call rates of less than 95% were excluded. In the LifeLines study first-degree relatives and nonwhite samples, based on self-report, outlier (Identity By State), and principal component analysis, were excluded.

After quality control, a total of 221,685 genotyped SNPs were available in both samples. We did not impute additional SNPs. Because of low numbers in the groups, we were not able to estimate effects for some SNPs. The exact number of SNPs with missing effect estimates depends on the minor allele frequency and prevalence of exposure; that is, the number of missing SNP-by-exposure estimates were 92 for biological dust, 22 for mineral dust, and 47 for gases and fumes. Consequently, the final number of SNPs in the (meta-) analysis was 221,593 for biological dust, 221,663 for mineral dust, and 221,638 for gases and fumes (for a flow diagram, see Fig E1 in this article's Online Repository at www.jacionline.org).

Statistical analysis

SNPs were tested in an additive genetic model. Exposure and SNP-by-exposure effects were included as dummy variables in the models. Associations of SNP-by-exposure interactions with FEV_1 levels were tested by using a linear regression model adjusted for sex, age, height, and ever smoking (no/yes) in the software package PLINK (PLINK version 1.07).[24] Finally, the model was specified as follows:

FEV_1 = SNP (additive) + Low exposure (no/yes) + High exposure (no/yes) + SNP*Low exposure + SNP*High exposure + Covariates.

To have a clear exposure contrast, we focused on SNP-by-high exposure interactions only.

We selected SNPs that interacted with high exposure to biological dust, mineral dust, or gases/fumes in the identification analysis ($P < 1 \times 10^{-5}$, Fig 1). Subsequently, we assessed whether the selected SNPs interacted with high occupational exposure in the replication cohort ($P < .05$). Finally,

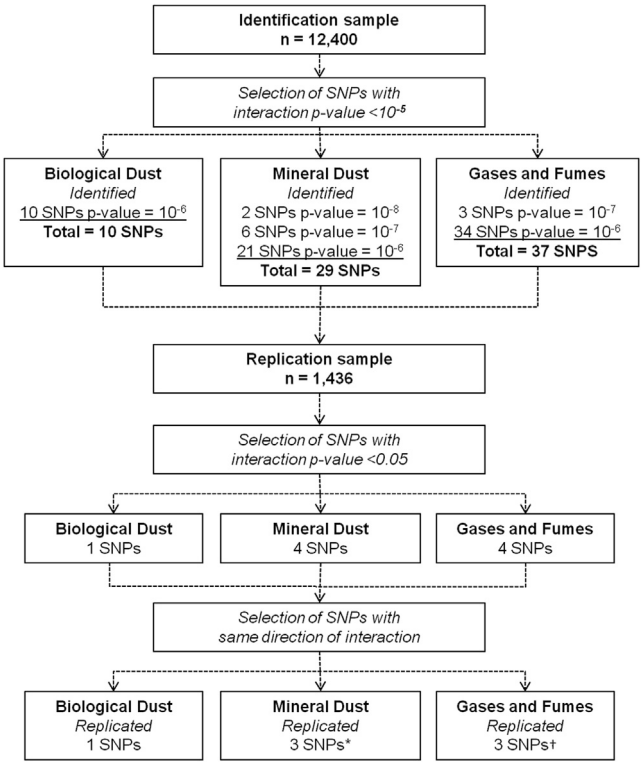


FIG 1. Flow chart showing selection of SNPs. *One SNP reached genome-wide significance after meta-analysis. †Two SNPs reached genome-wide significance after meta-analysis.

identified ($P < 1 \times 10^{-5}$) and replicated ($P < .05$) SNPs that had an interaction effect in the same direction in both samples were meta-analyzed by using effect estimates weighted by the inverse of the corresponding SEs and further investigated in gene expression analysis. Meta-analyzed SNP-by-high exposure interactions with a P value of less than the Bonferroni-corrected significance threshold P value of 2.26×10^{-7} (ie, 0.05/221,593 SNPs for biological dust, 221,663 SNPs for mineral dust, and 221,638 SNPs for gases and fumes) were considered genome-wide significant.

To assess the robustness of the SNP-by-high exposure interactions on FEV_1 levels in different subgroups, we additionally assessed interactions in models stratified by smoking status (never/ever smoker) and age (<40 and ≥40 years). We did not officially test for (3-way) interactions with smoking status and age.

Gene expression analysis

In gene expression analysis we assessed whether the replicated SNPs were cis-acting expression (mRNA) quantitative trait loci (cis-eQTLs) in lung tissue. Lung tissue was collected from patients who underwent lung resection surgery at 3 participating sites: the University of Groningen, Laval University, and the University of British Columbia.[25] Characteristics for the 1111 subjects included in the initial data set are presented in Table E1 in this article's Online Repository at www.jacionline.org. DNA samples were genotyped with Illumina Human1M-Duo BeadChip arrays, and gene expression profiles were obtained by using a custom Affymetrix array (Affymetrix, Santa Clara, Calif). Imputed SNP data were available for 1095 of the 1111 subjects included in the initial data set.[25] Gene expression levels were log-transformed (2-Log) and adjusted for the first 25 principal components.

Linear regression analysis was used to test for association between the SNP genotypes and gene expression levels. We defined a cis-eQTL as an SNP that was significantly associated with expression levels of a gene within a 2-Mb distance of that SNP with a P value of less than the Bonferroni-corrected threshold ($P = .05$/number of probe sets within the 4-Mb window).

TABLE I. Characteristics of the subjects included in the identification (LifeLines) and replication (Vlagtwedde-Vlaardingen) cohorts

	LifeLines	Vlagtwedde-Vlaardingen
No. with nonmissing data	12,400	1,436
Male subjects, no. (%)	5,123 (41)	772 (54)
Age (y), median (minimum-maximum)	47 (18-89)	53 (35-79)
Smoking status		
Never, no. (%)	5,070 (41)	431 (30)
Ever, no. (%)	7,330 (59)	1,005 (70)
Among ever smokers		
Current smokers, no. (%)	2,781 (38)	525 (52)
Pack years, median (minimum-maximum)*	10 (0-100)	19 (0-262)
Lung function, mean (SD)		
FEV_1 (% predicted)†	102 (14)	93 (16)
FEV_1/IVC ratio‡	76 (7)	74 (9)
COPD (FEV_1/FVC ratio <70%), no. (%)	1,895 (15)	399 (28)
High level of exposure, no. (%)		
Biological dust	505 (4)	126 (9)
Mineral dust	590 (5)	307 (21)
Gases and fumes	739 (6)	140 (10)

FVC, Forced vital capacity; *IVC*, inspiratory vital capacity.
*There were missing data on pack years smoked for 603 subjects in the LifeLines cohort and 34 in the Vlagtwedde-Vlaardingen cohort.
†FEV_1 as percent predicted according to the reference equations of Quanjer et al.[26]
‡FEV_1/IVC ratio for the Vlagtwedde-Vlaardingen cohort.

Patients provided written informed consent, and the study was approved by the ethics committees of the Institut universitaire de cardiologie et de pneumologie de Québec and the UBC-Providence Health Care Research Institute Ethics Board for Laval and UBC, respectively. The study protocol was consistent with the Research Code of the University Medical Center Groningen and Dutch national ethical and professional guidelines.

RESULTS

Characteristics of the study populations and the prevalence of exposures can be found in Table I.[26] Genomic inflation factors for the identification and replication samples suggest little population stratification (identification sample $\lambda = 1.05$, replication sample $\lambda = 1.02$, see Fig E2 in this article's Online Repository at www.jacionline.org). The number of subjects in each exposure category (no/low/high), genotype frequencies (see Tables E2 and E3 in this article's Online Repository at www.jacionline.org), more detailed information about each model (ie, the SNP and exposure main effects; see Tables E4-E6 in this article's Online Repository at www.jacionline.org), and a retrospective power calculation (see the Methods section in this article's Online Repository at www.jacionline.org) can be found in this article's Online Repository. Regional association plots for SNPs with identified and replicated associations can also be found in Fig E3 in this article's Online Repository at www.jacionline.org.

SNP-by-exposure interactions

Biological dust. Ten SNPs interacted with high exposure to biological dust on FEV_1 levels in the identification sample (P value for interaction $< 1 \times 10^{-5}$, Fig 1). The interaction with 1 SNP was replicated significantly in the second sample ($P < .05$), and this interaction was in the same direction in both cohorts (Table II).[27] SNP rs17490056 is located at an intergenic region

TABLE II. Interactions between SNPs (additive effect for minor allele A1) and high exposure to biological dust, mineral dust, and gases and fumes on FEV_1 levels (in milliliters)

				Annotation	Identification					Replication					Meta-analysis*			
SNP	CHR	A1	A2	Nearest gene	No.	B_{int}	SE_{int}	P_{int}	MAF	No.	B_{int}	SE_{int}	P_{int}	MAF	I	Q	B_f	P_f
Biological dust																		
rs17490056	13	T	C	*PCDH9* (150 Kb 3′)	12,400	−136.9	30.7	8.26E-06	0.50	1,433	−121.2	61.5	4.89E-02	0.48	0	0.82	−133.8	1.12E-06
Mineral dust																		
rs13278529	8	G	T	*ZMAT4* (55 Kb 3′)	12,400	178.4	38.4	3.49E-06	0.15	1,436	164.8	59.7	5.87E-03	0.19	0	0.85	174.4	**6.77E-08**
rs473892	6	T	C	*OLIG3* (136 Kb 5′)	12,400	126.3	27.3	3.89E-06	0.46	1,436	93.9	43.4	3.06E-02	0.49	0	0.53	117.1	4.15E-07
rs6751439	2	A	G	*GALNT13* (intronic)	12,400	−184.8	40.5	5.16E-06	0.13	1,415	−134.8	65.1	3.85E-02	0.12	0	0.51	−170.8	6.86E-07
Gases and fumes																		
rs159497	5	C	T	*PDE4D* (57 Kb 3′)	12,400	131.1	26.3	6.44E-07	0.46	1,435	136.3	62.4	2.90E-02	0.46	0	0.94	131.9	**5.42E-08**
rs516732	5	C	T	*ODZ2* (1.6 Mb 5′)	12,400	−115.4	25.0	3.76E-06	0.47	1,436	−158.5	61.1	9.61E-03	0.47	0	0.51	−121.6	**1.42E-07**
rs2888674	7	A	G	*TMEM176A* (8.7 Kb 3′)	12,400	122.2	26.6	4.36E-06	0.46	1,436	130.5	62.1	3.58E-02	0.45	0	0.90	123.5	4.33E-07

Linear regression models were adjusted for sex, age, height, and smoking status (never/ever). Minor allele frequencies are given for A1.

A1, Effect allele; *A2*, reference allele; B_f, β value from the fixed-effects model; B_{int}, β value for the interaction effect; *CHR*, chromosome number; I^2, percentage of variation across studies; *MAF*, minor allele frequency given for A1; *No.*, final number of subjects included in the analysis; P_f, *P* values from the fixed-effects model; P_{int}, *P* value for the interaction effect; *Q*, *P* value for Cochrane Q[27]; SE_{int}, SE for the interaction effect.

*The meta-analyses was performed on the interaction effect (B_{int}) of each sample weighted by the inverse of the corresponding standard errors (SE_{int}). The genome-wide significance *P* value was 2.26E-07 (*P* = .05/number of SNPs in the meta-analysis: biological dust, 221,593; mineral dust, 221,663; and gases and fumes, 221,638). Genome-wide significant *P* values are indicated in boldface.

at chromosome 13 near *PCDH9*. The minor allele of this SNP was associated with a lower FEV_1 level in subjects with high exposure to biological dust (Fig 2). The interaction effect did not reach genome-wide significance after meta-analysis of the effect estimates from both cohorts (Table II).

Mineral dust. We identified 29 SNPs that interacted with high exposure to mineral dust on FEV_1 levels (Fig 1). Interactions with 4 SNPs were significantly replicated ($P < .05$; rs5994031, rs13278529, rs473892, and rs6751439), of which 3 interactions were in the same direction in both cohorts (rs13278529, rs473892, and rs6751439; Table II).

Of the 3 SNPs, there was 1 intronic SNP located in *GALNT13* (Table II). The minor allele of SNP rs6751439 in *GALNT13* was associated with a lower FEV_1 level in the subjects with high exposure to mineral dust (Fig 3, *C*). The other 2 SNPs that interacted with high exposure to mineral dust on FEV_1 levels were both intergenic variants located near the genes *ZMAT4* and *OLIG3*. The minor alleles of rs13278529 (nearby *ZMAT4*) and rs473892 (nearby *OLIG3*) were associated with a higher FEV_1 level in subjects with high exposure to mineral dust (Fig 3, *A* and *B*). SNP rs13278529 near *ZMAT4* reached genome-wide significance after meta-analysis of the effects from both cohorts, and meta-analyzed interaction effects of the other 2 SNPs were borderline significant (Table II).

Gases and fumes. We identified 37 SNPs that interacted with high exposure to gases and fumes on FEV_1 levels (Fig 1). Interactions with 4 SNPs were significantly replicated ($P < .05$; rs6515493, rs159497, rs516732, and rs2888674), of which 3 SNPs had interactions with gases and fumes in the same direction in both cohorts (rs159497, rs516732, and rs2888674; Table II). These 3 SNPs were located in intergenic regions near the genes *PDE4D*, *ODZ2*, and *TMEM176A*. Effects of the SNPs in the groups with no, low, and high exposure are shown in Fig 4. The minor alleles of rs159497 (nearby *PDE4D*) and rs2888674 (nearby *TMEM176A*) were associated with a higher FEV_1 level in subjects with high exposure to gases and fumes (Fig 4, *A* and *C*), and the minor allele of rs516732 (nearby *ODZ2*) was associated with a lower FEV_1 level in subjects with high exposure to gases and fumes (Fig 4, *B*). The interaction effects of SNP rs159497 near *PDE4D* and rs516732 near *ODZ2* reached genome-wide significance after meta-analysis, and the meta-analyzed interaction effect of rs2888674 near *TMEM176A* was borderline significant (Table II).

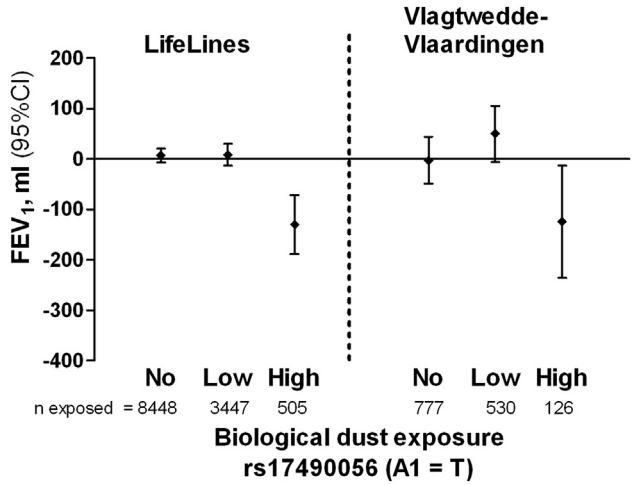


FIG 2. Additive associations (for minor allele A1) between the SNP and the FEV_1 level in subjects with no, low, and high exposure to biological dust.

eQTLs

We investigated whether the 7 identified and replicated SNPs were associated with gene expression levels in lung tissue from 1095 patients. Two SNPs showed cis-eQTL associations with *P* values of less than the Bonferroni-corrected threshold (see Table E7 in this article's Online Repository at www.jacionline.org). rs159497, the SNP that significantly interacted with high exposure to gases and fumes, was associated with expression of *PDE4D* ($P = 1.81 \times 10^{-4}$; Fig 5, *A*). SNP rs2888674, which also interacted with high exposure to gases and fumes, was significantly associated with expression levels of both its left- and right-neighboring genes *TMEM176A* and *ABP1* ($P = 1.73 \times 10^{-16}$ and $P = 1.54 \times 10^{-5}$, respectively; Fig 5, *B* and *C*).

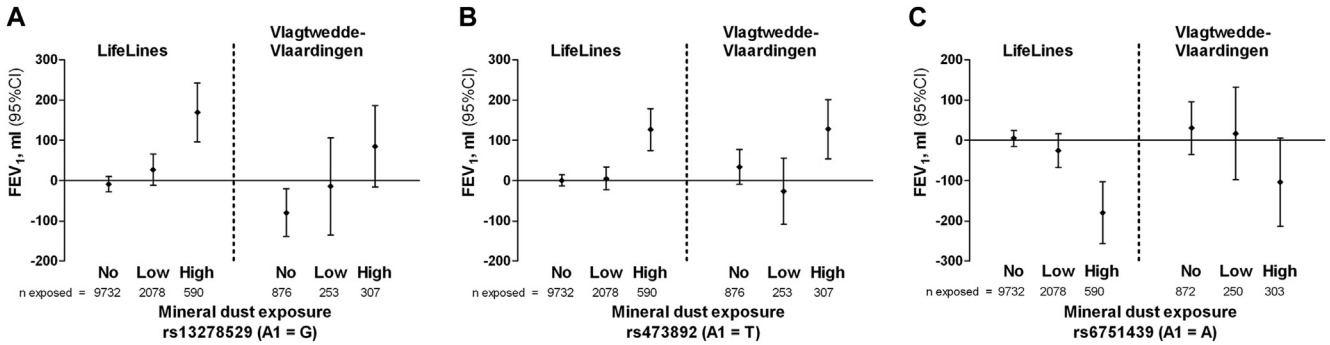


FIG 3. Additive associations (for minor allele A1) between SNPs and FEV_1 levels in subjects with no, low, and high exposure to mineral dust. SNPs presented are rs13278529 **(A)**, rs473892 **(B)**, and rs6751439 **(C)**.

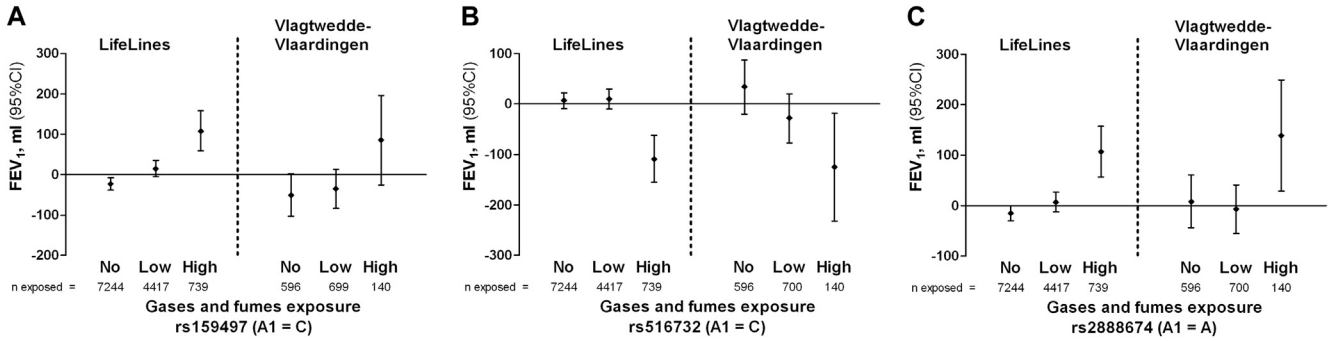


FIG 4. Additive associations (for minor allele A1) between SNPs and FEV_1 levels in subjects with no, low, and high exposure to gases and fumes. SNPs presented are rs159497 **(A)**, rs516732 **(B)**, and rs2888674 **(C)**.

Robustness of associations in stratified analysis

To assess the robustness of our findings, we assessed SNP-by-high exposure interactions on FEV_1 levels separately in never and ever smokers and in subjects aged less than 40 and 40 or more years. When stratified by smoking status, similar associations with FEV_1 in never and ever smokers from the identification sample were seen for 3 SNP-by-high exposure interactions (SNPs rs17490056, rs13278529, and rs6751439). For the remaining 4 SNP-by-exposure interactions, 2 associations were more pronounced in never smokers (rs473892 and rs516732) and 2 were more pronounced in ever smokers (rs159497 and rs2888674, Table III). When the models were stratified by age (<40 and ≥40 years), associations remained of similar magnitude or were more pronounced in the subjects aged 40 years and older (identification sample, Table III). Similar results for the stratified analyses were seen in the replication sample (results not shown).

DISCUSSION

This is the first genome-wide gene-by-occupational exposure interaction study on the level of lung function with replication in an independent cohort. We identified 7 SNPs that were associated with FEV_1 levels in subjects with high exposure to one of the following occupational exposures: biological dust, mineral dust, or gases and fumes. All SNPs had relevant main effects in the highly exposed subjects, and generally, there were no effects in the subjects without exposure. This supports the hypothesis that there are individual differences in genetic susceptibility to occupational exposure to biological dust, mineral dust, and gases and fumes.

Most identified and replicated loci are novel and have not been studied in relation to lung function impairment or COPD to date (Table IV).[18,28-43] Several identified SNPs were located in or near genes that might be involved in biological pathway leading to lung function impairment (ie, *GALNT13* and *PCDH9*, respectively). GALNT13 belongs to the GalNAcT family of enzymes, which initiate O-glycosylation of mucins. Terminal sugars might affect the physical and/or biological properties of mucins, and altered glycosylation has been found in patients with airways disease, such as cystic fibrosis.[44,45] Polymorphisms in *GALNT13* have been associated with survival time among mice with acute lung injury induced by acrolein.[28] Moreover, a recently published GWIS identified 2 intronic SNPs in *GALNT13* that interacted with *in utero* tobacco smoke exposure on childhood asthma; however, these interactions could not be replicated.[29] *PCDH9* is a protocadherin belonging to the cadherin superfamily of adhesion molecules. A family member, *PCDH1*, has been identified as a susceptibility gene for bronchial hyperresponsiveness.[30] In a recent GWAS 2 SNPs near *PCDH9* (rs17077331 and rs17077335) were associated with FEV_1/forced vital capacity ratio decrease in nonasthmatic patients, although this association was only driven by a single study.[18] The identified SNPs were not cis-eQTLs for *GALNT13* and *PCDH9* in lung tissue in our study but might be associated with lung function impairment through other biological mechanisms. For example, SNPs might change protein structure and consequently alter GALNT13 enzyme activity or the functionality of PCDH9.

In the cis-eQTL analysis we identified 2 intergenic SNPs that were associated with expression of nearby genes in lung tissue. The strongest cis-eQTL association was found for rs2888674

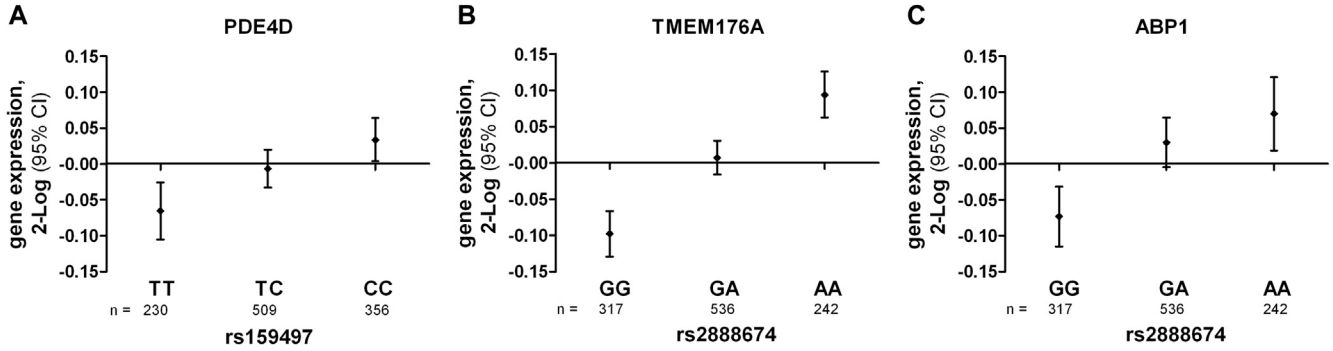


FIG 5. Gene expression levels stratified by genotype for cis-eQTL SNPs rs159497 **(A)** and rs2888674 (**B** and **C**). SNP rs2888674 was associated with gene expression levels of both its left- and right-neighboring genes *TMEM176A* (Fig 5, *B*) and *ABP1* (Fig 5, *C*), respectively.

TABLE III. Interactions between SNPs (additive effect for minor allele A1) and high exposure to biological dust, mineral dust, and gases and fumes on FEV_1 levels (in milliliters) in the LifeLines cohort (identification sample) stratified for smoking status (never/ever smoker) and age (<40 and ≥40 years old)

				Annotation	Never smokers,* n = 5070 (41%)			Ever smokers,* n = 7330 (59%)			Age <40 y,† n = 2848 (23%)			Age ≥40 y,† n = 9552 (77%)		
SNP	**CHR**	**A1**	**A2**	**Nearest gene**	B_{int}	SE_{int}	P_{int}	B_{int}	SE_{int}	P_{int}	B_{int}	SE_{int}	P_{int}	B_{int}	SE_{int}	P_{int}
Biological dust																
rs17490056	13	T	C	*PCDH9* (150 Kb 3′)	−137.9	43.4	1.49E-03	−128.2	42.7	2.69E-03	−137.4	62.3	2.76E-02	−134.9	35.2	1.28E-04
Mineral dust																
rs13278529	8	G	T	*ZMAT4* (55 Kb 3′)	183.0	58.7	1.83E-03	176.6	50.6	4.82E-04	143.3	75.7	5.85E-02	185.5	44.4	2.93E-05
rs473892	6	T	C	*OLIG3* (136 Kb 5′)	199.1	43.3	4.32E-06	89.3	35.3	1.14E-02	133.3	55.4	1.61E-02	124.9	31.5	7.43E-05
rs6751439	2	A	G	*GALNT13* (intronic)	−163.9	59.0	5.51E-03	−198.3	55.0	3.14E-04	−137.5	78.3	7.89E-02	−200.7	47.1	2.05E-05
Gases and fumes																
rs159497	5	C	T	*PDE4D* (57 Kb 3′)	6.8	43.0	8.74E-01	189.9	33.5	1.45E-08	108.9	52.6	3.84E-02	140.7	30.3	3.38E-06
rs516732	5	C	T	*ODZ2* (1.6 Mb 5′)	−185.7	41.3	7.14E-06	−79.9	31.5	1.13E-02	−92.5	50.4	6.68E-02	−118.2	28.6	3.56E-05
rs2888674	7	A	G	*TMEM176A* (8.7 kb 3′)	43.8	43.8	3.18E-01	169.3	33.7	4.98E-07	90.8	48.3	6.01E-02	135.9	31.6	1.73E-05

A1, Effect allele; *A2*, reference allele; B_{int}, β value for the interaction effect; *CHR*, chromosome number; P_{int}, *P* value for the interaction effect; SE_{int}, SE for the interaction effect.
*Model adjusted for sex, height, and age (continuous).
†Model adjusted for sex, height, age (continuous), and smoking status (never/ever).

associated with expression of *TMEM176A*. Inhibition of the expression of *TMEM176A* has been shown to increase expression of the costimulatory molecules CD86 and CD40 in immature mouse bone marrow–derived dendritic cells (DCs), thereby causing DC maturation and consequently immune response stimulation, such as T-cell differentiation and survival. This suggests that *TMEM176A* plays a role in the maintenance of the immature state of the DCs.[31] In human subjects Freeman et al found that more severe COPD (according to FEV_1 as percentage predicted) is associated with increased expression of CD40 and CD86 costimulatory molecules on specific types of pulmonary DCs.[46] Nevertheless, the role of DC maturation in patients with (smoking-related) COPD remains controversial, with conflicting data from several *in vivo* and *in vitro* studies.[47]

In the current study the minor allele for rs2888674 was associated with higher expression of *TMEM176A* and higher FEV_1 levels in subjects exposed to gases and fumes. We hypothesize that higher expression of *TMEM176A* might lead to decreased expression of costimulatory molecules and subsequent immune and inflammatory response to environmental stimuli, such as gases and fumes. Mechanistic studies focusing on downstream cellular processes after *TMEM176A* overexpression in combination with gas and fume exposure might help to unravel how this relates to DC maturation and subsequent inflammatory responses in the airways. Studies assessing colocalization of *TMEM176A* and maturation markers of DCs in various stages of COPD (according to level of lung function) versus healthy control subjects might help to further explain our findings.

The second cis-eQTL is rs159497, which is associated with higher expression of the nearby gene *PDE4D*. *PDE4D* is a regulator of airway smooth muscle contractility, and there is evidence from the literature supporting a role for *PDE4D* in lung pathology; *PDE4D* has been shown to be a susceptibility gene for asthma[32] and to be associated with a lower FEV_1 level in ever smokers.[33] PDE4 enzymes metabolize cyclic AMP to 5′ AMP, and inhibition of these PDE4 enzymes decreases the activity of inflammatory cells in association with bronchodilation. This seems contradictory to our finding that the minor allele for rs2221132 was associated with increased expression of *PDE4D* in lung tissue, as well as a higher FEV_1 levels in subjects with high exposure to gases and fumes. We have studied expression in whole lung tissue, and this expression might be different in specific cell types, such as epithelial or alveolar cells, on which gases and fumes might act. Moreover, the SNP could be associated with a specific alternative splice variant of *PDE4D*. Alternative splice variants have been shown to be differentially expressed and regulated in lung tissue.[48,49]

TABLE IV. Putative function of the identified and replicated genes

SNP	CHR	Nearest gene	Putative function of nearest gene
rs17490056	13	*PCDH9* (150 Kb 3′)	Protocadherin 9 *(PCDH9)*: PCDH9 expression was shown in epithelial regions of the nasal cavity.[34] PCDH9 is a member of the family of δ1-protocadherins, which are thought to be involved in intracellular signaling. A family member, *PCDH1*, has been identified as a susceptibility gene for bronchial hyperresponsiveness.[30] In a recent GWAS 2 SNPs near *PCDH9* (rs17077331 and rs17077335) were associated with FEV_1/FVC ratio decrease in nonasthmatic patients, although this association was only driven by a single study (Imboden et al[18]).
rs13278529	8	*ZMAT4* (55 Kb 3′)	Zinc finger, matrin-type 4 *(ZMAT4)*: SNPs in *ZMAT4* have been associated with fasting blood glucose levels[35] and refractive error, a leading cause of visual impairment.[36]
rs473892	6	*OLIG3* (136 Kb 5′)	Oligodendrocyte transcription factor 3 *(OLIG3)*: *OLIG3* is involved in specification of class A and B neurons. The *OLIG3/TNFAIP3* locus is associated with autoimmune diseases, such as rheumatoid arthritis[37] and celiac disease.[38]
rs6751439	2	*GALNT13* (intronic)	UDP-N-acetyl-α-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 13 *(GALNT13, GalNAcT13)*: *GALNT13* belongs to the GalNAcT family of enzymes, which initiate O-glycosylation of mucins. O-glycosylation might affect the physical and/or biological properties of mucins. Sites for O-glycosylation vary between tissues, during inflammation, and in the setting of cancer.[39] Polymorphisms in *GALNT13* have been associated with survival time among mice with acute lung injury induced by acrolein.[28] A recently published GWIS identified 2 intronic SNPs in *GALNT13* that interacted with *in utero* tobacco smoke exposure on childhood asthma; however, these interactions could not be replicated.[29]
rs159497	5	*PDE4D* (57 Kb 3′)	Phosphodiesterase-4D *(PDE4D)*: *PDE4D* is an important cyclic AMP–metabolizing enzyme expressed in immune and inflammatory cells, airway smooth muscle and pulmonary nerves. The PDE4 enzyme plays a significant role in modulating the activity of cyclic AMP, an important second messenger that mediates the relaxation of airway smooth muscle.[40] Evidence from the literature supports a role for *PDE4D* in lung pathogeneses; *PDE4D* has been shown as susceptibility gene for asthma[32] and was associated with FEV_1 levels in ever smokers.[33]
rs516732	5	*ODZ2* (1.6 Mb 5′)	Teneurin-2 *(ODZ2)*: Teneurin-2 is a transmembrane glycoprotein. Teneurins have been mainly studied in relation to neural development, where they seem to regulate the establishment of proper connectivity within the nervous system by promoting neurite outgrowth and cell adhesion. Recently, Teneurin-2 was shown to be expressed in the majority of malignant mesothelioma cell lines but not in lung adenocarcinoma cell lines investigated.[41,42] Mesothelioma is a rare form of cancer developing in the mesothelium, the protective body that covers many of the internal organs of the body, such as the pleura (the outer lining of the lungs and internal chest wall). Mesothelioma is most commonly caused by (occupational) exposure to asbestos.
rs2888674	7	*TMEM176A* (8.7 Kb 3′)	Transmembrane protein 176A *(TMEM176A)*: *TMEM176A* was shown to be highly expressed in the lungs and more specifically to be highly expressed in bone marrow–derived DCs (BMDCs) and in classical DCs (cDC). Moreover, *TMEM176A* was downregulated in mouse BMDCs after inflammatory stimulation by LPS or polyinosinic:polycytidylic acid. Additionally, inhibition of *TMEM176A* expression in immature BMDCs increased expression of the costimulatory molecules CD86 and CD40. These findings suggested that *TMEM176A*, together with *TMEM176B*, plays a role in maintaining the immature state of DCs (which seems to no longer be important in more mature DCs).[31] Another study showed that *TMEM16A* expression levels in the lungs were low, yet proteins levels were increased in patients with lung carcinoma.[43]

CHR, Chromosome; *FVC*, forced vital capacity.

Compared with the recently published GWIS that studied SNP interactions with general dust exposure in one cohort,[20] we replicated findings in a second independent cohort. Both the identification and replication cohorts included only white subjects, and occupational exposure assessment was performed by using the same method in both samples. Finally, we extended our findings to gene expression analysis in lung tissue.

We investigated several types of occupational exposures: biological dust, mineral dust, and gases and fumes. Job exposure was objectively estimated by using the ALOHA+ JEM, which is specifically designed for population-based studies. Especially in large population-based studies, such as GWASs in which large numbers are required to be more specific (individual), exposure assessment is often not feasible. Compared with self-reports, JEM-based exposure estimates are less likely to be affected by recall bias and differential misclassification.[50] JEM-based exposures are generally affected by nondifferential misclassification caused by heterogeneity of exposure in a given job. However, this type of misclassification generally results in an underestimation of effects. A limitation of JEM-based exposure assessments is that chemical specificity is lacking. Like other exposures, such as tobacco smoke or ambient air pollution, the 3 exposures under study (biological dust, mineral dust, and gases and fumes) are complex mixtures including different substances all with their own chemical and biological properties.

The major differences between the 2 samples included in the current study were the higher prevalence of occupational exposures and current smokers and the higher cumulative exposure to personal cigarette smoking (pack years smoked) in the replication sample (Vlagtwedde-Vlaardingen) compared with the identification sample (LifeLines). This might be explained by the historical timing of both studies; the identification sample included data from measurements performed between 2006 and 2011, whereas the replication sample included data from measurements performed in 1989-1990. The results did only marginally change when we additionally adjusted the models for pack years smoked (not shown). In the stratified analysis we found that some SNP-by-exposure interactions were of similar

magnitude in never and ever smokers, indicating that these interactions were not modified by smoking. Although we did not formally test for interaction with smoking (exposure*SNP*-ever smoking), some interactions were more pronounced in ever smokers, suggesting that tobacco smoking might modify the associations between SNP-by-occupational exposure interactions and FEV_1 levels.

Interactions were of similar magnitude or slightly more pronounced in subjects 40 years and older compared with subjects less than 40 years old. More pronounced effects in older subjects might reflect larger cumulative exposure and accelerated decline, although this hypothesis could not be studied in the current cross-sectional sample. Further studies with longitudinal data are warranted to elucidate whether the identified and replicated interactions are associated with the decrease in lung function and the development of COPD. Ideally, these further studies would also include more specific exposure assessment (information on specific chemicals), as well as information on exposure history and duration.

Because our GWIS is explorative, we wanted to keep the risk of not detecting a true association low and therefore used a more liberal P value threshold for identification of SNPs in the first sample ($P < 1 \times 10^{-5}$). When we assessed these interactions in the second independent sample, we found more significant interactions than expected based on chance only. Three SNP-by-high exposure interactions had a P value of less than the Bonferroni-corrected threshold P value of 2.26×10^{-7} after meta-analysis. Among these, the P value of one SNP approximated the even stricter GWAS threshold of 5.0×10^{-8}. Moreover, the difference in additive SNP effects between the unexposed and highly exposed subjects were of clinical relevance, with effects between 100 and 200 mL of FEV_1 per minor allele for most replicated SNPs. Finally, the cis-eQTL analysis showed that 2 SNPs identified in the interaction analysis were significantly associated with expression levels of neighboring genes. These are unique data and provide additional support for a possible role of these genes in lung function impairment.

In conclusion, this is the first genome-wide study with replication in a second independent cohort that investigated interactions of SNPs with several types of occupational exposures on the level of lung function. We identified several plausible candidates that might be involved in biological pathways leading to lung function impairment (ie, *PCDH9*, *GALNT13*, *PDE4D*, and *TMEM176A*). Further research should determine whether the identified (novel) genes are true susceptibility loci for lung function impairment caused by occupational exposure to biological dust, mineral dust, and gases and fumes and whether these SNP-by-exposure interactions consequently increase the risk to develop COPD. This information might eventually contribute to the understanding of cellular and molecular pathways driving the development of COPD. Moreover, because occupational exposures are common but modifiable, this knowledge can be used to set exposure limits considering susceptible subgroups.

We thank Rob Bieringa, Joost Keers, René Oostergo, Rosalie Visser, and Judith Vonk for their work related to data collection and validation in the LifeLines Cohort Study. We also thank the study participants, the staff from the LifeLines Cohort Study and Medical Biobank Northern Netherlands, and the participating general practitioners and pharmacists. We thank Jan Schouten for his work related to data collection and validation in the Vlagtwedde-Vlaardingen cohort. Specimens for the lung eQTL study collected at Laval University were from the "Institut universitaire de cardiologie et de pneumologie de Québec" (IUCPQ) site of the Respiratory Health Network Tissue Bank of the "Fonds de recherche du Québec–Santé" (www.tissuebank.ca).

Group author: the LifeLines Cohort study

The LifeLines Cohort Study: B. Z. Alizadeh,[a] R. A. de Boer,[b] H. M. Boezen,[a] M. Bruinenberg,[c] L. Franke,[d] P. van der Harst,[b] H. L. Hillege,[a,b] M. M. van der Klauw,[e] G. Navis,[f] J. Ormel,[g] D. S. Postma,[h] J. G. M. Rosmalen,[g] J. P. Slaets,[i] H. Snieder,[a] R. P. Stolk,[a] B. H. R. Wolffenbuttel,[e] and C. Wijmenga[d]

From [a]the Department of Epidemiology; [b]the Department of Cardiology; [c]the LifeLines Cohort Study; [d]the Department of Genetics; [e]the Department of Endocrinology; [f]the Department of Internal Medicine, Division of Nephrology; [g]the Department of Psychiatry, Interdisciplinary Center of Psychopathology of Emotion Regulation (ICPE); [h]the Department of Pulmonology; and [i]the University Center for Geriatric Medicine, University of Groningen, University Medical Center Groningen.

Clinical implications: Using a hypothesis-free approach, we identified several plausible candidate genes that might underlie susceptibility to several types of occupational exposures in relation to lung function impairment (ie, *PCDH9*, *GALNT13*, *PDE4D*, and *TMEM176A*).

REFERENCES

1. World Health Organization. Chronic obstructive pulmonary disease (COPD)—fact sheet Available at: http://www.who.int/mediacentre/factsheets/fs315/en/index.html Accessed April 16, 2012.
2. Barnes P, Kleinert S. COPD—a neglected disease. Lancet 2004;364:564-5.
3. Boezen HM. Genome-wide association studies: what do they teach us about asthma and chronic obstructive pulmonary disease? Proc Am Thorac Soc 2009;6:701-3.
4. Balmes J, Becklake M, Blanc P, Henneberger P, Kreiss K, Mapp C, et al. American Thoracic Society statement: occupational contribution to the burden of airway disease. Am J Respir Crit Care Med 2003;167:787-97.
5. Zock JP, Sunyer J, Kogevinas M, Kromhout H, Burney P, Ant JM. Occupation, chronic bronchitis, and lung function in young adults. An international study. Am J Respir Crit Care Med 2001;163:1572-7.
6. Sunyer J, Kogevinas M, Kromhout H, Antó J, Roca J, Tobias A, et al. Pulmonary Ventilatory Defects and Occupational Exposures in a Population-based Study in Spain. Am J Respir Crit Care Med 1998;157:512-7.
7. de Jong K, Boezen HM, Kromhout H, Vermeulen R, Postma DS, Vonk JM, et al. Pesticides and other occupational exposures are associated with airway obstruction: the LifeLines cohort study. Occup Environ Med 2014;71:88-96.
8. Bosse Y. Updates on the COPD gene list. Int J Chron Obstruct Pulmon Dis 2012;7:607-31.
9. Hang J, Zhou W, Wang X, Zhang H, Sun B, Dai H, et al. Microsomal epoxide hydrolase, endotoxin, and lung function decline in cotton textile workers. Am J Respir Crit Care Med 2005;171:165-70.
10. Wilk JB, Walter RE, Laramie JM, Gottlieb DJ, O'Connor GT. Framingham Heart Study genome-wide association: results for pulmonary function measures. BMC Med Genet 2007;8(suppl 1):S8.
11. Pillai S, Ge D, Zhu G, Kong X, Shianna K, Need A, et al. A genome-wide association study in chronic obstructive pulmonary disease (COPD): identification of two major susceptibility loci. PLoS Genet 2009;5:e1000421.
12. Wilk JB, Chen TH, Gottlieb DJ, Walter RE, Nagle MW, Brandler BJ, et al. A genome-wide association study of pulmonary function measures in the Framingham Heart Study. PLoS Genet 2009;5:e1000429.
13. Repapi E, Sayers I, Wain LV, Burton PR, Johnson T, Obeidat M, et al. Genome-wide association study identifies five loci associated with lung function. Nat Genet 2010;42:36-44.
14. Hancock D, Eijgelsheim M, Wilk J, Gharib S, Loehr L, Marciante K, et al. Meta-analyses of genome-wide association studies identify multiple loci associated with pulmonary function. Nat Genet 2010;42:45-52.
15. Cho MH, Boutaoui N, Klanderman BJ, Sylvia JS, Ziniti JP, Hersh CP, et al. Variants in FAM13A are associated with chronic obstructive pulmonary disease. Nat Genet 2010;42:200-2.
16. Soler Artigas M, Loth DW, Wain LV, Gharib SA, Obeidat M, Tang W, et al. Genome-wide association and large-scale follow up identifies 16 new loci influencing lung function. Nat Genet 2011;43:1082-90.

17. Cho MH, McDonald ML, Zhou X, Mattheisen M, Castaldi PJ, Hersh CP, et al. Risk loci for chronic obstructive pulmonary disease: a genome-wide association study and meta-analysis. Lancet Respir Med 2014;2:214-25.
18. Imboden M, Bouzigon E, Curjuric I, Ramasamy A, Kumar A, Hancock DB, et al. Genome-wide association study of lung function decline in adults with and without asthma. J Allergy Clin Immunol 2012;129:1218-28.
19. Manichaikul A, Hoffman EA, Smolonska J, Gao W, Cho MH, Baumhauer H, et al. Genome-wide study of percent emphysema on computed tomography in the general population. The Multi-Ethnic Study of Atherosclerosis Lung/SNP Health Association Resource Study. Am J Respir Crit Care Med 2014;189:408-18.
20. Liao SY, Lin X, Christiani DC. Gene-environment interaction effects on lung function—a genome-wide association study within the Framingham heart study. Environ Health 2013;12:101.
21. Stolk R, Rosmalen JG, Postma D, de Boer R, Navis G, Slaets JP, et al. Universal risk factors for multifactorial diseases: LifeLines: a three-generation population-based study. Eur J Epidemiol 2008;23:67-74.
22. International Labour Organization. The revised international standard classification of occupations (ISCO-88). Geneva: International Labour Organization; 1990.
23. Matheson MC, Benke G, Raven J, Sim MR, Kromhout H, Vermeulen R, et al. Biological dust exposure in the workplace is a risk factor for chronic obstructive pulmonary disease. Thorax 2005;60:645-51.
24. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007;81:559-75.
25. Hao K, Bosse Y, Nickle DC, Pare PD, Postma DS, Laviolette M, et al. Lung eQTLs to help reveal the molecular underpinnings of asthma. PLoS Genet 2012;8:e1003029.
26. Quanjer PH, Tammeling GJ, Cotes JE, Pedersen OF, Peslin R, Yernault JC. Lung volumes and forced ventilatory flows. Report Working Party Standardization of Lung Function Tests, European Community for Steel and Coal. Official Statement of the European Respiratory Society. Eur Respir J Suppl 1993;16:5-40.
27. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ 2003;327:557-60.
28. Leikauf GD, Concel VJ, Liu P, Bein K, Berndt A, Ganguly K, et al. Haplotype association mapping of acute lung injury in mice implicates activin a receptor, type 1. Am J Respir Crit Care Med 2011;183:1499-509.
29. Scholtens S, Postma DS, Moffatt MF, Panasevich S, Granell R, Henderson AJ, et al. Novel childhood asthma genes interact with in utero and early-life tobacco smoke exposure. J Allergy Clin Immunol 2014;133:885-8.
30. Koppelman GH, Meyers DA, Howard TD, Zheng SL, Hawkins GA, Ampleford EJ, et al. Identification of PCDH1 as a novel susceptibility gene for bronchial hyperresponsiveness. Am J Respir Crit Care Med 2009;180:929-35.
31. Condamine T, Le Texier L, Howie D, Lavault A, Hill M, Halary F, et al. Tmem176B and Tmem176A are associated with the immature state of dendritic cells. J Leukoc Biol 2010;88:507-15.
32. Himes BE, Hunninghake GM, Baurley JW, Rafaels NM, Sleiman P, Strachan DP, et al. Genome-wide association analysis identifies PDE4D as an asthma-susceptibility gene. Am J Hum Genet 2009;84:581-93.
33. Obeidat M, Wain LV, Shrine N, Kalsheker N, Soler Artigas M, Repapi E, et al. A comprehensive evaluation of potential lung function associated genes in the SpiroMeta general population sample. PLoS One 2011;6:e19382.
34. Redies C, Vanhalst K, Roy F. delta-Protocadherins: unique structures and functions. Cell Mol Life Sci 2005;62:2840-52.
35. Meigs JB, Manning AK, Fox CS, Florez JC, Liu C, Cupples LA, et al. Genome-wide association with diabetes-related traits in the Framingham Heart Study. BMC Med Genet 2007;8(suppl 1):S16.
36. Verhoeven VJ, Hysi PG, Wojciechowski R, Fan Q, Guggenheim JA, Hohn R, et al. Genome-wide meta-analyses of multiancestry cohorts identify multiple new susceptibility loci for refractive error and myopia. Nat Genet 2013;45:314-8.
37. Plant D, Farragher T, Flynn E, Martin P, Eyre S, Bunn D, et al. A genetic marker at the OLIG3/TNFAIP3 locus associates with methotrexate continuation in early inflammatory polyarthritis: results from the Norfolk Arthritis Register. Pharmacogenomics J 2012;12:128-33.
38. Trynka G, Zhernakova A, Romanos J, Franke L, Hunt KA, Turner G, et al. Coeliac disease-associated risk variants in TNFAIP3 and REL implicate altered NF-kappaB signalling. Gut 2009;58:1078-83.
39. Gill DJ, Chia J, Senewiratne J, Bard F. Regulation of O-glycosylation through Golgi-to-ER relocation of initiation enzymes. J Cell Biol 2010;189:843-58.
40. Hansen G, Jin S, Umetsu DT, Conti M. Absence of muscarinic cholinergic airway responses in mice deficient in the cyclic nucleotide phosphodiesterase PDE4D. Proc Natl Acad Sci U S A 2000;97:6751-6.
41. Ziegler A, Cerciello F, Bigosch C, Bausch-Fluck D, Felley-Bosco E, Ossola R, et al. Proteomic surfaceome analysis of mesothelioma. Lung Cancer 2012;75:189-96.
42. Ziegler A, Corvalan A, Roa I, Branes JA, Wollscheid B. Teneurin protein family: an emerging role in human tumorigenesis and drug resistance. Cancer Lett 2012;326:1-7.
43. Cuajungco MP, Podevin W, Valluri VK, Bui Q, Nguyen VH, Taylor K. Abnormal accumulation of human transmembrane (TMEM)-176A and 176B proteins is associated with cancer pathology. Acta Histochem 2012;114:705-12.
44. Rose MC, Voynow JA. Respiratory tract mucin genes and mucin glycoproteins in health and disease. Physiol Rev 2006;86:245-78.
45. Xia B, Royall JA, Damera G, Sachdev GP, Cummings RD. Altered O-glycosylation and sulfation of airway mucins associated with cystic fibrosis. Glycobiology 2005;15:747-75.
46. Freeman CM, Martinez FJ, Han MK, Ames TM, Chensue SW, Todt JC, et al. Lung dendritic cell expression of maturation molecules increases with worsening chronic obstructive pulmonary disease. Am J Respir Crit Care Med 2009;180:1179-88.
47. Van Pottelberge GR, Bracke KR, Joos GF, Brusselle GG. The role of dendritic cells in the pathogenesis of COPD: liaison officers in the front line. COPD 2009;6:284-90.
48. Richter W, Jin SL, Conti M. Splice variants of the cyclic nucleotide phosphodiesterase PDE4D are differentially expressed and regulated in rat tissue. Biochem J 2005;388:803-11.
49. Conti M, Richter W, Mehats C, Livera G, Park JY, Jin C. Cyclic AMP-specific PDE4 phosphodiesterases as critical components of cyclic AMP signaling. J Biol Chem 2003;278:5493-6.
50. Mannetje A', Kromhout H. The use of occupation and industry classifications in general population studies. Int J Epidemiol 2003;32:419-28.

METHODS

Retrospective power calculation

The current study is well powered to detect gene-environment interactions with an interaction effect of 100 mL of FEV_1. The power to detect an interaction effect of 100 mL of FEV_1 is 99% for minor allele frequencies, ranging from 5% to 45% with an additive SNP effect given the sample size of 12,400 subjects with a marginal SNP effect of 5 mL of FEV_1 (from analysis on SNP rs675149), a marginal exposure effect of 67 mL of FEV_1 (from analysis on SNP rs675149), and a population mean FEV_1 of 3379 mL with an SE of 84 mL. These calculations were made with the software package Quanto (USC, version 1.2.4).

REFERENCE

E1. Hao K, Bosse Y, Nickle DC, Pare PD, Postma DS, Laviolette M, et al. Lung eQTLs to help reveal the molecular underpinnings of asthma. PLoS Genet 2012;8:e1003029.

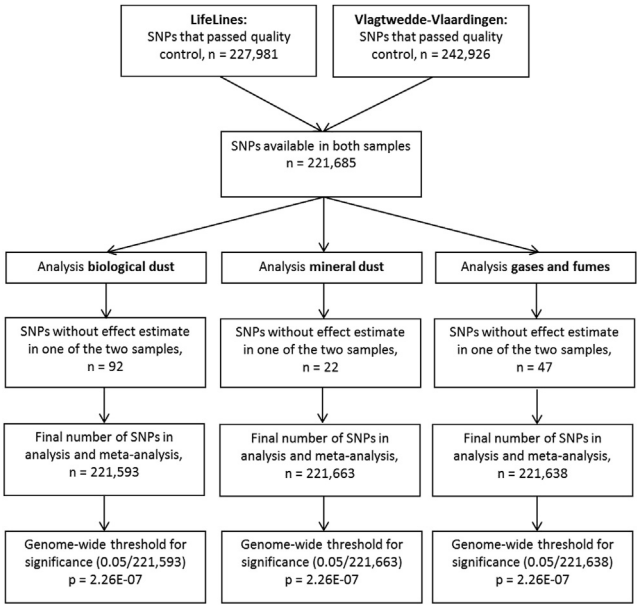


FIG E1. Flow diagram showing the number of SNPs in each analysis.

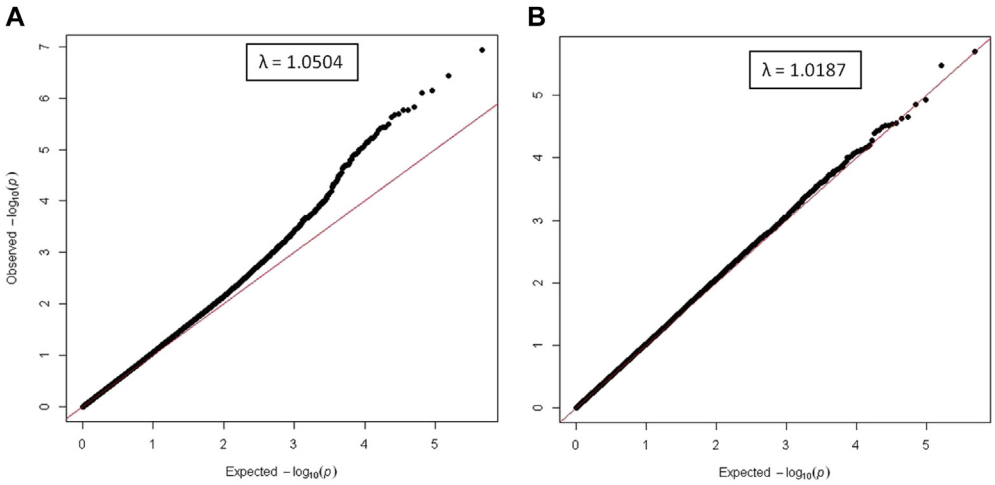


FIG E2. QQ plots for the marginal association between SNPs and FEV_1 levels in identification **(A)** and replication **(B)** samples. The analysis was adjusted for sex, age, height, and ever smoking (no/yes).

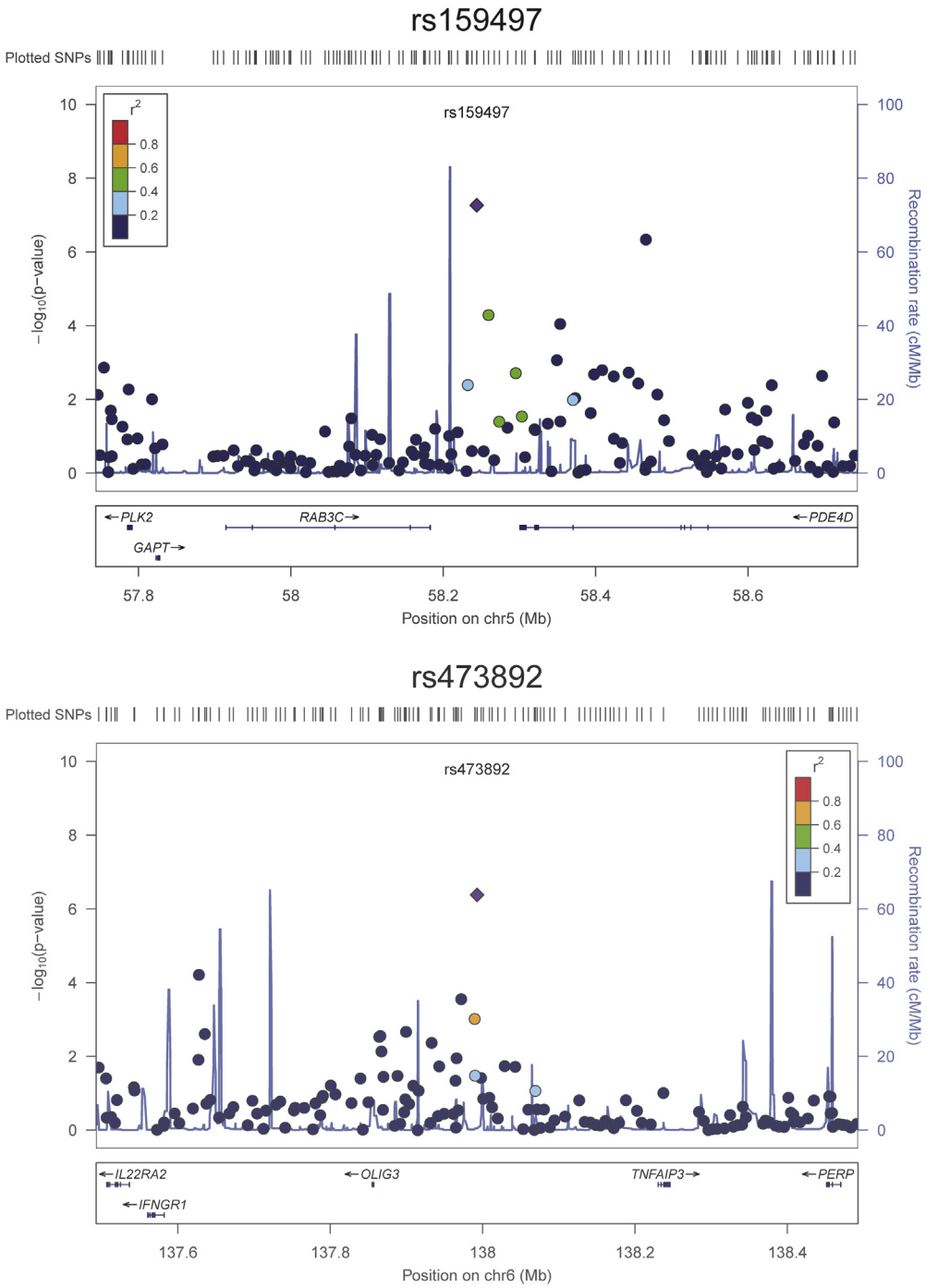


FIG E3. Regional association plots showing the 7 identified and replicated loci associated with FEV_1 levels in interaction with biological dust (rs17490056), mineral dust (rs13278529, rs473892, and rs6751439), and gases and fumes (rs159497, rs516732, and rs2888674) by using the online available software LocusZoom. Each SNP (sentinel SNP; *purple diamond*) ± 500 kb was plotted, with the *y-axis* showing the meta-analyzed (fixed-effects) −log *P* value, and the *x-axis* showing the chromosomal position (genome build hg18). Linkage disequilibrium with the sentinel SNP was plotted by using the 1000 Genomes Project population June 2010 CEU.

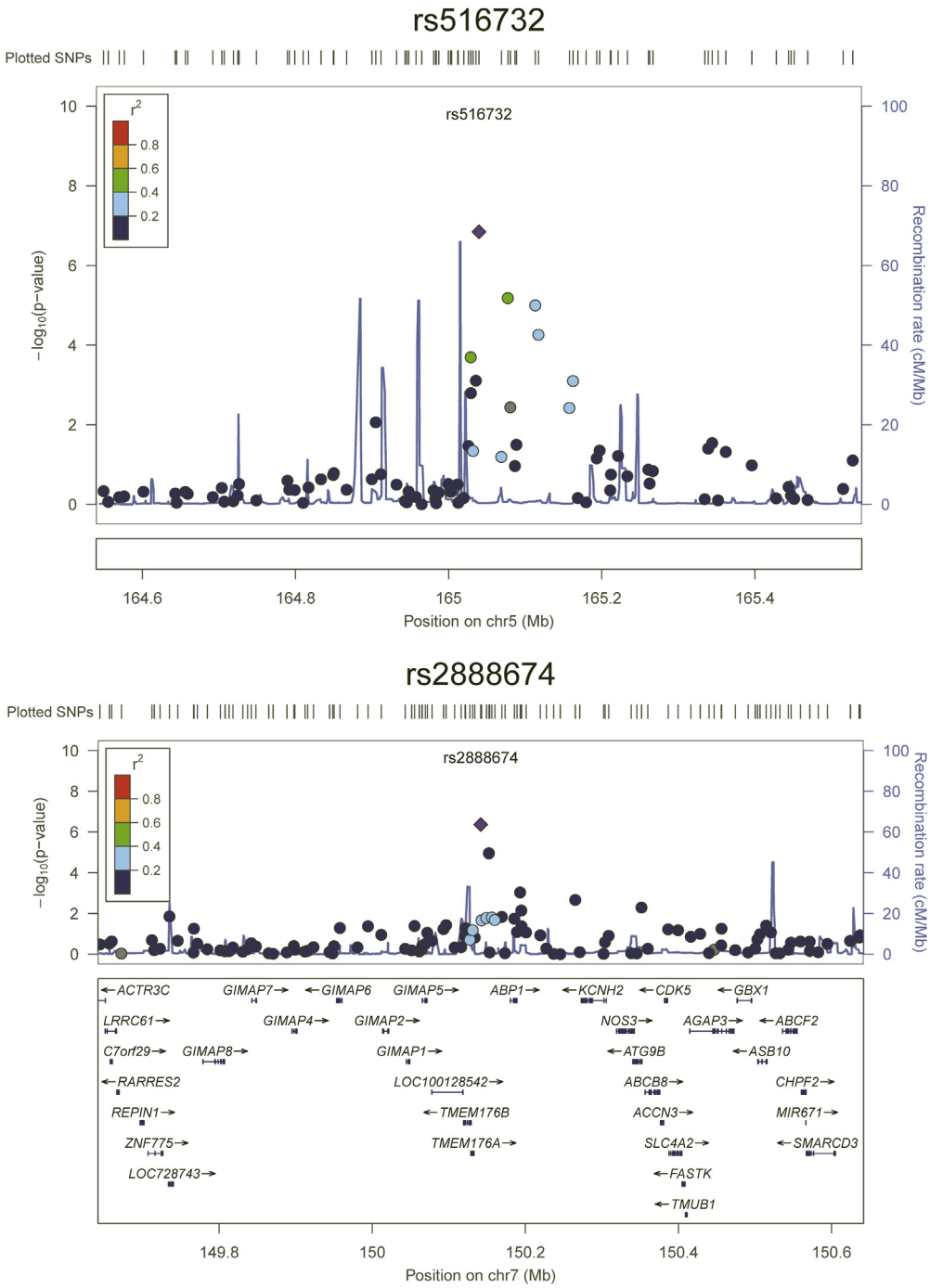
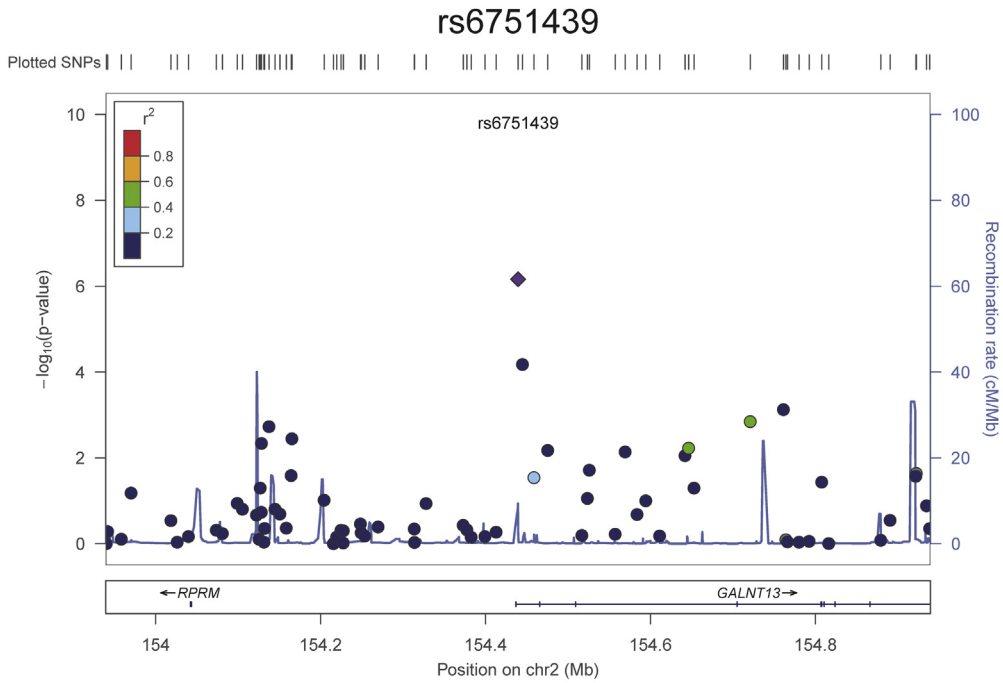


FIG E3. *(Continued).*



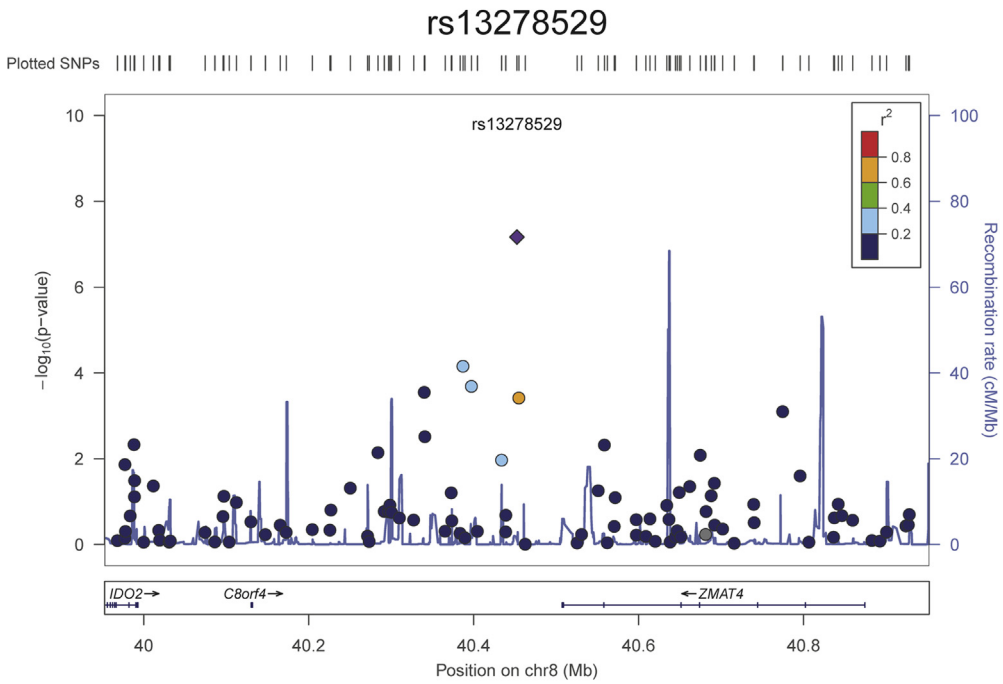


FIG E3. *(Continued).*

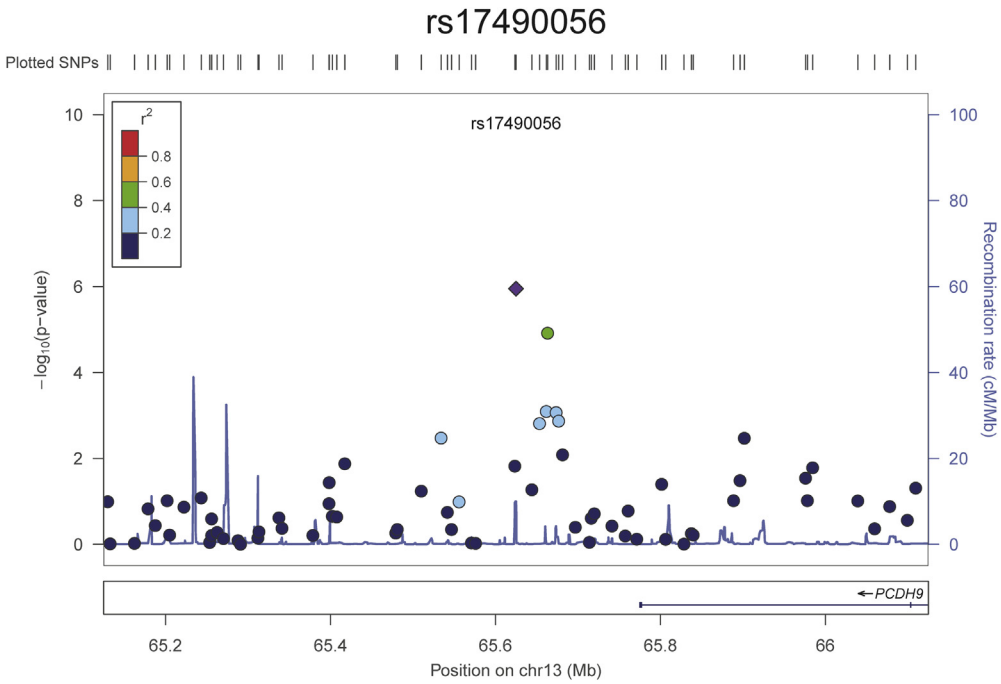


FIG E3. *(Continued).*

TABLE E1. Demographic and clinical characteristics of the 1111 subjects included in the initial lung tissue database[E1]

	Laval	Groningen	University of British Columbia
Quantitative traits, mean ± SD	n = 409	n = 363	n = 339
Age (y)	63.3 ± 10.0	51.8 ± 15.2	60.0 ± 14.3
Pre-BD FEV_1 (% predicted)	80.5 ± 18.9	60.5 ± 30.0	80.4 ± 22.4
Pre-BD FVC (% predicted)	89.8 ± 16.1	75.0 ± 26.5	88.9 ± 19.7
Categorical traits (%)			
Male subjects	56.0	53.6	53.4
Smoking status			
Nonsmoker	8.7	28.6	8.8
Exsmoker	69.0	54.9	56.8
Current smoker	22.0	16.5	34.4

FVC, Forced vital capacity.

TABLE E2. Genotype frequencies in subjects with no, low, and high exposure included in the identification analyses

Identification		No exposure				Low exposure				High exposure			
SNP	**A1**	**Total**	**WT**	**HZ**	**HM**	**Total**	**WT**	**HZ**	**HM**	**Total**	**WT**	**HZ**	**HM**
Biological dust													
rs17490056	T	8448	2129 (25)	4207 (50)	2112 (25)	3447	905 (26)	1646 (48)	896 (26)	505	129 (26)	257 (51)	119 (24)
Mineral dust													
rs13278529	G	9732	7124 (73)	2423 (25)	185 (2)	2078	1498 (72)	524 (25)	56 (3)	590	421 (71)	153 (26)	16 (3)
rs473892	T	9732	2888 (30)	4838 (50)	2006 (21)	2078	607 (29)	1025 (49)	446 (22)	590	191 (32)	270 (46)	129 (22)
rs6751439	A	9732	7424 (76)	2149 (22)	159 (2)	2078	1545 (74)	495 (24)	38 (2)	590	437 (74)	141 (24)	12 (2)
Gases and fumes													
rs159497	C	7244	2053 (28)	3652 (50)	1539 (21)	4417	1238 (28)	2273 (51)	906 (21)	739	215 (29)	383 (52)	141 (19)
rs516732	C	7244	2032 (28)	3586 (50)	1626 (22)	4417	1235 (28)	2230 (50)	952 (22)	739	212 (29)	346 (47)	181 (24)
rs2888674	A	7244	2098 (29)	3585 (49)	1561 (22)	4417	1320 (30)	2149 (49)	948 (22)	739	200 (27)	394 (53)	145 (20)

Values are presented as numbers (percentages). A1 is the minor (mutant/effect) allele.
HM, Homozygote mutant for allele A1; *HZ*, heterozygote; *WT*, wild-type.

TABLE E3. Genotype frequencies in subjects with no, low, and high exposure included in the replication analyses

Replication		No exposure				Low exposure				High exposure			
SNP	**A1**	**Total**	**WT**	**HZ**	**HM**	**Total**	**WT**	**HZ**	**HM**	**Total**	**WT**	**HZ**	**HM**
Biological dust													
rs17490056	T	777	179 (23)	403 (52)	195 (25)	530	131 (25)	269 (51)	130 (25)	126	32 (25)	60 (48)	34 (27)
Mineral dust													
rs13278529	G	876	646 (74)	206 (24)	24 (3)	253	197 (78)	51 (20)	5 (2)	307	224 (73)	75 (24)	8 (3)
rs473892	T	876	281 (32)	428 (49)	167 (19)	253	82 (32)	128 (51)	43 (17)	307	104 (34)	148 (48)	55 (18)
rs6751439	A	862	655 (76)	194 (23)	13 (2)	250	188 (75)	56 (22)	6 (2)	303	239 (79)	57 (19)	7 (2)
Gases and fumes													
rs159497	C	596	165 (28)	297 (50)	134 (22)	699	207 (30)	344 (49)	148 (21)	140	37 (26)	73 (52)	30 (21)
rs516732	C	596	163 (27)	307 (51)	126 (21)	700	200 (29)	348 (50)	152 (22)	140	42 (30)	68 (49)	30 (21)
rs2888674	A	596	186 (31)	291 (49)	119 (20)	700	207 (30)	342 (49)	151 (22)	140	37 (26)	72 (51)	31 (22)

Values are presented as numbers (percentages). A1 is the minor (mutant/effect) allele.
HM, Homozygote mutant for allele A1; *HZ*, heterozygote; *WT*, wild-type.

TABLE E4. Estimates for the effects of SNP, biological dust exposure, and their interaction on FEV_1 levels (in milliliters) by using linear regression models adjusted for sex, age, height, and smoking status (never/ever)

			Identification				Replication			
				95% CI				95% CI		
SNP	**A1**	**Predictor**	**B**	**Lower**	**Upper**	***P* value**	**B**	**Lower**	**Upper**	***P* value**
rs17490056	T	SNP (additive)	7.3	−6.8	21.4	3.12E-01	−2.7	−49.5	44.0	9.09E-01
		Low exposure	−20.5	−52.6	11.6	2.11E-01	−138.2	−227.3	−49.0	2.43E-03
		High exposure	56.4	−16.4	129.2	1.29E-01	33.4	−116.7	183.6	6.63E-01
		SNP*Low exposure	1.0	−24.9	26.9	9.41E-01	52.4	−20.4	125.1	1.58E-01
		SNP*High exposure	−136.9	−197.1	−76.8	8.26E-06	−121.2	−241.7	−0.7	4.89E-02

TABLE E5. Estimates for the effects of SNP, mineral dust exposure, and their interaction on FEV_1 levels (in milliliters) by using linear regression models adjusted for sex, age, height, and smoking status (never/ever)

			Identification				Replication			
				95% CI				95% CI		
SNP	**A1**	**Predictor**	**B**	**Lower**	**Upper**	***P* value**	**B**	**Lower**	**Upper**	***P* value**
rs13278529	G	SNP (additive)	−9.4	−28.3	9.5	3.30E-01	−79.7	−139.4	−20.0	8.94E-03
		Low exposure	−65.7	−91.8	−39.6	8.07E-07	−16.4	−89.8	57.0	6.62E-01
		High exposure	−173.8	−220.2	−127.3	2.42E-13	−119.5	−191.3	−47.6	1.14E-03
		SNP*Low exposure	36.2	−7.3	79.6	1.03E-01	65.4	−68.9	199.7	3.40E-01
		SNP*High exposure	178.4	103.0	253.7	3.49E-06	164.8	47.7	281.8	5.87E-03
rs473892	T	SNP (additive)	0.4	−12.8	13.7	9.48E-01	33.7	−9.5	76.9	1.26E-01
		Low exposure	−59.3	−95.9	−22.6	1.53E-03	55.6	−46.7	157.9	2.87E-01
		High exposure	−230.8	−293.3	−168.3	4.75E-13	−147.7	−242.9	−52.6	2.39E-03
		SNP*Low exposure	5.0	−26.5	36.4	7.57E-01	−60.2	−153.3	32.9	2.06E-01
		SNP*High exposure	126.3	72.7	179.9	3.89E-06	93.9	8.9	178.9	3.06E-02
rs6751439	A	SNP (additive)	4.5	−15.3	24.3	6.55E-01	30.7	−35.1	96.4	3.61E-01
		Low exposure	−46.9	−72.6	−21.2	3.51E-04	7.8	−67.1	82.7	8.39E-01
		High exposure	−66.6	−112.3	−20.9	4.27E-03	−38.5	−109.5	32.6	2.89E-01
		SNP*Low exposure	−29.5	−75.5	16.5	2.08E-01	−13.7	−146.4	119.1	8.40E-01
		SNP*High exposure	−184.8	−264.2	−105.4	5.16E-06	−134.8	−262.4	−7.2	3.85E-02

TABLE E6. Estimates for the effects of SNP, gas and fume exposure, and their interaction on of FEV_1 levels (in milliliters) by using linear regression models adjusted for sex, age, height, and smoking status (never/ever)

			Identification				Replication			
				95% CI				95% CI		
SNP	A1	Predictor	B	Lower	Upper	*P* value	B	Lower	Upper	*P* value
rs159497	C	SNP (additive)	−22.7	−38.1	−7.3	3.94E-03	−50.8	−103.1	1.5	5.73E-02
		Low exposure	−71.5	−100.8	−42.2	1.71E-06	−67.0	−150.7	16.8	1.17E-01
		High exposure	−222.8	−282.1	−163.5	1.91E-13	−167.3	−313.6	−20.9	2.53E-02
		SNP*Low exposure	37.6	12.3	62.8	3.54E-03	15.6	−55.5	86.7	6.68E-01
		SNP*High exposure	131.1	79.5	182.7	6.44E-07	136.3	14.1	258.6	2.90E-02
rs516732	C	SNP (additive)	6.6	−8.7	21.8	3.99E-01	33.6	−19.8	87.0	2.17E-01
		Low exposure	−39.8	−69.1	−10.5	7.71E-03	7.6	−76.8	92.1	8.59E-01
		High exposure	6.3	−53.1	65.6	8.36E-01	107.3	−34.2	248.8	1.38E-01
		SNP*Low exposure	3.4	−21.5	28.4	7.87E-01	−62.1	−134.1	10.0	9.15E-02
		SNP*High exposure	−115.4	−164.4	−66.5	3.76E-06	−158.5	−278.2	−38.7	9.61E-03
rs2888674	A	SNP (additive)	−14.9	−30.2	0.4	5.58E-02	8.4	−43.9	60.8	7.52E-01
		Low exposure	−57.1	−85.8	−28.3	1.02E-04	−36.4	−118.3	45.6	3.85E-01
		High exposure	−217.0	−277.6	−156.4	2.31E-12	−163.8	−309.4	−18.1	2.77E-02
		SNP*Low exposure	22.1	−2.6	46.9	7.95E-02	−15.3	−86.3	55.8	6.73E-01
		SNP*High exposure	122.2	70.1	174.3	4.36E-06	130.5	8.8	252.1	3.58E-02

TABLE E7. SNPs acting as cis-eQTL on genes within a 2-Mb distance

SNP	Interaction with:	Cis-gene	*P* value*	Threshold†
rs159497	Gases and fumes	*PDE4D*	1.81E-04	1.35E-03
rs2888674	Gases and fumes	*TMEM176A*	1.73E-16	3.50E-04
		ABP1	1.54E-05	3.50E-04

*P value from the linear regression model.
†Bonferroni-corrected threshold calculated as $P = .05$/number of probe sets within the 4-Mb window.