ERJ Express. Published on September 24, 2009 as doi: 10.1183/09031936.00054408

Novel strategy to identify genetic risk factors for COPD severity: a genetic isolate

CC van Diemen[1], DS Postma[2], YS Aulchenko[3], PJLM Snijders[3], BA Oostra[3], CM van Duijn[3], HM Boezen[1]

Departments of [1] Epidemiology, [2] Pulmonology, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands; [3] Epidemiology & Biostatistics, Erasmus Medical Center Rotterdam, The Netherlands.

Corresponding author:

Prof. HM Boezen, PhD, Department of Epidemiology, E3.29

University Medical Center Groningen

PO Box 30.001, 9700 RD Groningen, The Netherlands

Phone: +3150 361 0899, Fax: +3150 361 4493

Email: h.m.boezen@epi.umcg.nl

Word Count: 2587

Key words: chronic obstructive pulmonary disease, genetically isolated population, lung function, single nucleotide polymorphism

Running title: lung function genetics in a genetic isolate

Copyright 2009 by the European Respiratory Society.

Abstract

Background Studies using genetic isolates with limited genetic variation may be useful in COPD genetics, but are thus far lacking. We studied associations between Single Nucleotide Polymorphisms (SNPs) in candidate genes and lung function in COPD in a genetic isolate.

Methods In 91 subjects with COPD GOLD stage ≥ I, members of an extended pedigree including 6,175 people from the Genetic Research in Isolated Population study, we analyzed 32 SNPs in 13 candidate genes: *ADAM33*, *TGFβ1*, *MMP1*, *MMP2*, *MMP9*, *MMP12*, *TIMP1*, *SFTPA1*, *SFTPA2*, *SFTPB*, *SFTPD*, *GSTP1* and *HMOX1*, and studied their relation to FEV_1, IVC, and FEV_1/IVC levels using restricted maximum likelihood linear mixed modeling, accounting for pedigree structure. We replicated significant associations in the general Vlagtwedde/Vlaardingen study.

Results Six SNPs in *TGFB1*, *SFTPA1*, *SFTPA2* and *SFTPD* were significantly associated with FEV_1/IVC in subjects with COPD GOLD stage ≥ I. Two SNPs in *TGFB1 (*C-509T and Leu10Pro), Leu50Val in *SFTPA1*, and Ala160Thr in *SFTPD* showed suggestive evidence of association with FEV_1/IVC in subjects with GOLD stage ≥ II. The *TGFB1* associations were replicated in GOLD stage ≥ II patients from the Vlagtwedde/Vlaardingen population, with similar effect sizes.

Discussion We show that a genetic isolate can be used to determine genetics of lung function, which can be replicated in COPD patients from an independent population.

Introduction

Chronic Obstructive Pulmonary Disease (COPD) is the third cause of death worldwide and is expected to increase in prevalence in the forthcoming decades [1,2]. The disease has a large personal, societal, and economic impact. COPD is characterized by chronic airway inflammation, airway remodeling, and airflow limitation that is not fully reversible. Since not all smokers develop COPD, genetic susceptibility has to play a role in development of this disease, in addition to environmental factors. The genetic determinants for COPD are difficult to study, since COPD is a disease that becomes clinically manifest only at later ages, when parents of COPD patients have already died, and their children are likely to young to manifest airway obstruction. This limits the option to perform family based genetic research. Moreover, published studies frequently use various definitions of disease status, which makes it difficult to compare their results. Therefore, it makes sense to choose a robust phenotype to define COPD like the level of lung function, which can be more easily compared between studies. Moreover, a low level of lung function is a predictor for mortality from COPD [3-5].

Another complicating factor in studies on genetics of COPD is that COPD is considered a complex genetic trait, i.e. multiple, possibly interacting, genetic and environmental factors are involved. Therefore it has advantages to try and identify risk genes in populations that are relatively genetically and environmentally homogeneous, such as genetically isolated populations. Due to the small number of founders and drift in genetically isolated populations, the genetic variation is reduced [6]. However, these processes raise the question whether findings can be extrapolated to the general population. Previous simulation studies suggest that this is the case for common variants with a frequency of >1% [6], but no empirical evidence is available.

We conducted a candidate gene study for level of airflow limitation in patients with COPD who were ascertained as part of the Genetic Research in Isolated Populations (GRIP) study that is conducted in a young genetically isolated population from the southwestern

part of the Netherlands. All patients were genotyped using 32 Single Nucleotide Polymorphisms (SNPs) in 13 candidate genes for COPD, chosen based on their previously published association with either COPD, level of lung function, or lung function decline as reported in the general population. Extensive genealogy information was collected resulting in an extremely large and complex pedigree of 6,175 members. Finally, we studied 1390 Caucasians from the general Dutch population, including 351 patients with COPD, to establish whether our findings could be replicated in the general population. In both studies, we investigated whether the severity of the disease, as reflected by lung function reduction, is genetically influenced in established COPD.

Methods

Study populations

Our study is part of the GRIP program [7,8]. GRIP is based in a recent genetically isolated population from the southwestern of the Netherlands, which was founded in the middle of the 18th century by approximately 150 individuals and was genetically isolated until the last few decades. The population now includes approximately 20,000 inhabitants in 8 adjacent communities. GRIP participants are generally related via multiple lines of descent and are inbred via multiple consanguineous loops [9,10].

We invited subjects with general practitioner's diagnosed COPD to the research center to undergo spirometry and complete a questionnaire [11]. Spirometry was performed by trained pulmonary research technicians using a Pneumotagograph (Viasys, formerly Jaeger Spirometer system). Predicted values for FEV_1 were calculated using adjusted Quanjer-equations for Caucasian subjects [12]. We isolated DNA from blood using Puregene® DNA Purification Kits (Gentra, Inc, Minneapolis, USA). All participants gave written informed consent.

To verify the findings from GRIP in the general population, we used cross-sectional data from the general population-based Vlagtwedde/Vlaardingen cohort. Questionnaires, spirometry and DNA were collected [13,14]. For this study, we selected 351 subjects according to the Global Initiative for Chronic Obstructive Lung Disease (GOLD) criteria with GOLD stage ≥ I COPD at the last 1989/1990 survey, of whom 167 had GOLD stage ≥ II [15].

Genotyping

We have genotyped SNPs in candidate genes for lung function and COPD, based on their previously published significant associations (table 1). The selected SNPs were either the most significant SNPs in previous studies, tagging SNPs for the gene, or SNPs with a known functional effect on gene expression or function. Genotyping was performed using

Applied Biosystems TaqMan® SNP Genotyping Assays (Nieuwekerk aan de IJssel, The Netherlands). Sequences of primers and probes are available on request.

Statistical analysis

To analyze pedigree data, we used Measured Genotype (MG) approach [16], which models quantitative traits as $y_i = \mu + kg_i + \sum_j \beta_j c_{ji} + G_i + e_i$

where y_i: the phenotype of the *i*-th individual, *g:* the vector of genotypes at the marker under study, *k*: is the marker genotype effect, c_{ij}: the value of the *j*-th covariate or fixed effect for the individual *i*, β_j: an estimate of the *j*-th fixed effect or covariate, and G_i and e_i are random additive polygenic and residual effects, respectively. The random effects are assumed to follow multivariate normal distribution with mean zero. The variance for the polygenic effects is defined as $\Phi\sigma_G^2$, where Φ is the relationship matrix and σ_G^2 the additive genetic variance due to polygenes. For the residual random effects, the variance is defined as $I\sigma_e^2$, where I is the identity matrix and σ_e^2 the residual variance.
Because the pedigree under analyses was very large, we used fast GRAMMAR approximation to the full MG approach [17]. The GRAMMAR consist of a fast though conservative test at screening stage, followed up with full MG analysis of polymorphisms which pass the relaxed (P<0.1) screening significance threshold. All analysis involving pedigree were performed using ASReml

v2.0 [18]– a package for linear mixed model analysis using restricted maximum likelihood.

In the Vlagtwedde/Vlaardingen population, we tested significant associations using linear regression analyses. All analyses were adjusted for age, height and sex.

Results

GRIP study population

We ascertained 157 individuals who were diagnosed with COPD by their general practitioners. Spirometry measures confirmed COPD in 91 subjects, i.e. subjects with COPD GOLD stage ≥ I (defined as FEV_1/IVC<70%) [15]. The rest of the subjects could not be defined as having COPD according to their spirometry and were therefore excluded from the analyses. We determined the familial relationship of these 91 subjects in the larger GRIP study database. This resulted in a large extended pedigree structure of 6,175 members. The characteristics of the GRIP COPD population and the Vlagtwedde/Vlaardingen replication cohort are shown in Table 2.

Association of genes with lung function parameters in GRIP and replication in Vlagtweddde/Vlaardingen

We first analyzed the effects of SNPs in the studied genes on FEV_1 % predicted, IVC, and FEV_1/IVC in the 91 subjects with COPD GOLD stage ≥ I. None of the SNPs was associated with FEV_1 % predicted or IVC. Six SNPs in *TGFB1*, *SFTPA1*, *SFTPA2* and *SFTPD* were significantly associated with FEV_1/IVC (table 3). None of these associations were replicated in subjects from the Vlagtwedde/Vlaardingen cohort with COPD GOLD stage ≥ I (data not shown).

We additionally analyzed the effects of SNPs in the studied genes using a more stringent definition of COPD, namely GOLD stage ≥ II (defined as FEV_1/IVC<70% and FEV_1%pred <80). This resulted for the GRIP population in 67 cases. In these subjects, two SNPs in *TGFB1 (*C-509T and Leu10Pro), Leu50Val in *SFTPA1*, and Ala160Thr in *SFTPD* showed suggestive evidence of association with FEV_1/IVC ($p<0.10$, table 3). The *TGFB1* C-509T and Leu10Pro associations were replicated in GOLD ≥ II subjects from the Vlagtwedde/Vlaardingen population (n=167), with similar effect sizes (see table 3).

Discussion

Our study is the first to use a genetically isolated population to analyze genetic effects on level of lung function in COPD. Interestingly, we found significant effects of SNPs in COPD candidate genes on the severity of COPD, assessed by lung function in subjects with COPD even though our study population is small. Our results show that levels of FEV_1/IVC, measures of airway obstruction, are genetically influenced in established COPD. This means that even within patients with phenotypical COPD, we can identify genotypes that are associated with severity of the disease. This is of clinical importance since low lung function levels have been shown to predict mortality of COPD not only in the general population, but also within COPD patients [3-5].

The *TGFβ1* SNPs that were associated with FEV_1/IVC in our populations have previously been associated with development of COPD or with lower levels of FEV_1 and FEV_1/VC in several [19-21], but not all previous studies [14,22,23]. Our results (in both the genetically isolated and general population) thus confirm the former studies that implicate a role of *TGFβ1* in the severity of airflow limitation. The *SFTPA1* and *SFTPD* SNPs have been associated with COPD previously [24,25]. We now for the first time show that these SNPs may also play a role in severity of COPD. This is plausible, since surfactant proteins decrease surface tension at the air–liquid interface and, therefore, reduce the tendency of alveoli to collapse during expiration. The latter contributes to the severity of airway obstruction, as measured by FEV_1/IVC.

We found no significant associations of *ADAM33, MMP1*, *MMP2*, *MMP9*, *MMP12*, *TIMP1*, *SFTPB*, *GSTP1* and *HMOX1* with level of lung function in COPD patients. This does not imply that these genes do not play a role in COPD whatsoever. So far, no studies have analyzed genetic effects on the severity of airway obstruction *within* patients with established COPD. Our study shows that SNPs in *TGFβ1*, *SFTPA*1, and *SFTPD* may be important in progression of COPD, whereas the SNPs in other genes, i.e. *ADAM33,*

MMP1, *MMP2*, *MMP9*, *MMP12*, *TIMP1*, *GSTP1* and *HMOX1,* may simply constitute genes that are important in development of COPD.

One important advantage of testing genes in a genetically isolated population is that it provides the opportunity to find genes associated with disease in a relatively small sample size due to increased homogeneity of the population, as recently demonstrated for multiple sclerosis [26]. Thus for a lower cost and effort, one can test many genes as to their significance in contributing to disease severity, which subsequently can then be replicated in a larger sample of the general population. The most important requirement for such studies is that the genetic isolate is representative for the general population or disease-specific study populations. This is indeed the case since we showed that in selected subjects with COPD from the general population, we can replicate the associations found in the young genetic isolate for a substantial part. Thus, we are able to translate findings in a genetic isolate to the general population, but correct and comparable phenotyping of the study populations is still crucial to replicate association between populations.

We were unable to replicate results of any of the SNPs in subjects with GOLD stage ≥ I from the Vlagtwedde/Vlaardingen population. When looking more closely, it appeared that the GRIP COPD patients with GOLD stage ≥ I had more severe COPD, i.e. lower lung function and more symptoms, than COPD patients with similar stage of disease in the Vlagtwedde/Vlaardingen population. A more strict definition of COPD (GOLD stage ≥ II) in Vlagtwedde/Vlaardingen and GRIP gives a phenotypically better comparison. Indeed, when analyzing subjects with subjects with GOLD ≥ II from Vlagtwedde/Vlaardingen population, SNPs *TGFβ1* C-509T and *TGFβ1* Leu10Pro were significantly associated with FEV_1/IVC, as they were in the GRIP GOLD ≥ II COPD patients.

Since the percentage of subjects with amongst other chronic cough was different in both cohorts, we re-ran our analyses with straight forward linear regression models with chronic cough in the model to check for stability of the effect estimates. Analyses on FEV_1/IVC in

the GRIP GOLD stage ≥ II population, taking for example chronic cough into account, resulted in similar regression estimates for the SNPs in TGFβ1 and SFTPA1, but smaller p-values and slightly higher explained variances, while the suggestive associations of the other SNPs disappeared. Additional adjustment for chronic cough in the Vlagtwedde/Vlaardingen GOLD stage ≥ II population resulted in similar significant regression estimates for the SNPs in TGFβ1 with FEV_1/IVC. Therefore, our effect estimates appear to be stable within both GOLD stage ≥ II groups, irrespective of differences in characteristics between the GRIP and Vlagtwedde/Vlaardingen GOLD stage ≥ II population.

Several explanations may exist for the lack of replication for *SFTPA1* and *SFTPD* (Met11Thr) SNPs with FEV_1/IVC in the Vlagtwedde/Vlaardingen COPD GOLD stage ≥ II. First, the original GRIP findings on these genes could be false-positive. Indeed, multiple (though correlated) outcomes and SNPs were studied in GRIP. Another, more biological, explanation for the lack of replication may be that the prevalence of certain alleles in genetically isolated populations differs from a general population as a result of genetic drift and founder effects. Indeed, the genotype frequencies for the *SFTPA1* Leu50Val SNP were significantly different between the two populations, but not for the other SNPs (see table E1). A third explanation may be that differences exist in characteristics between the study populations. The GRIP population had more severe COPD and was slightly older than the Vlagtwedde/ Vlaardingen COPD population.

In addition, differences in environment may affect the lack of replication of the *SFTP* genes. The genetically isolated population shares the same environment, similar socio-economic status, and the same general practitioners. We cannot rule out that the COPD patients in GRIP have a higher prevalence of chronic bronchitis and airway disease whereas the airway obstruction in the Vlagtwedde/Vlaardingen population may have been caused by emphysema [27-29]. Further research is needed to separately assess these

phenomena, since CT scans are necessary, which we unfortunately do not have of these patients.

In conclusion, this study provides two important messages. Firstly, we found significant effects of SNPs on the severity of COPD, i.e. level of lung function in patients with established COPD in a relatively small genetically isolated population with a large pedigree structure. Secondly, we replicated two of these associations in COPD patients selected from the general population on the condition that they were phenotypically similar. These findings are important since more severe airway obstruction is associated with progression and mortality of COPD. Future studies using this genetic isolate should focus on progression of COPD, since this population seems to be highly suitable to determine genetic risk factors for severity of airway obstruction in established COPD that can be translated to the general population.

Acknowledgements

We thank L. Testers for his help in logistics of the fieldwork and DNA collection, and M. Farenhorst, Th. van Hoogdalem J. Post, A. Verbokkem, and K. Vink-Klooster for collecting the lung function data. We would like to thank P. Veraart, H. Kornman and E. Boeren for their contribution to genealogical research. The Dutch Asthma Foundation funded the collection of lung function data (NAF 3.4.04.041). C.C. van Diemen is assigned through the Dutch Asthma Foundation (NAF3.2.02.51). GRIP is supported by grants from the Netherlands Organization for Scientific Research (NWO: Pionier grant C.M. van Duijn), the Center for Medical Systems Biology (CMSB). We would like to thank all participants of the GRIP study for their cooperation as well as the general practitioners that made this work possible. All research assistants of the GRIP study are acknowledged for their help in data collection.

Reference List

(1) WHO. World Health Report 2002. Available from http://www.who.int/whr/2002

(2) Murtagh E, Heaney L, Gingles J, Shepherd R, Kee F, Patterson C, et al. The prevalence of obstructive lung disease in a general population sample: the NICECOPD study. *Eur J Epidemiol* 2005;20(5):443-53.

(3) Ekberg-Aronsson M, Pehrsson K, Nilsson JA, Nilsson PM, Lofdahl CG. Mortality in GOLD stages of COPD and its dependence on symptoms of chronic bronchitis. *Respir Res* 2005 ;6:98.

(4) Hospers JJ, Postma DS, Rijcken B, Weiss ST, Schouten JP. Histamine airway hyper-responsiveness and mortality from chronic obstructive pulmonary disease: a cohort study. *Lancet* 2000;356(9238):1313-7.

(5) Sin DD, Wu L, Man SF. The relationship between reduced lung function and cardiovascular mortality: a population-based study and a systematic review of the literature. *Chest* 2005 Jun;127(6):1952-9.

(6) Pardo LM, Mackay I, Oostra B, van Duijn CM, Aulchenko YS. The effect of genetic drift in a young genetically isolated population. *Ann Hum Genet* 2005;69(Pt 3):288-95.

(7) Aulchenko YS, Heutink P, Mackay I, Bertoli-Avella AM, Pullen J, Vaessen N, et al. Linkage disequilibrium in young genetically isolated Dutch population. *Eur J Hum Genet* 2004;12(7):527-34.

(8) Njajou OT, Alizadeh BZ, Aulchenko Y, Zillikens MC, Pols HA, Oostra BA, et al. Heritability of serum iron, ferritin and transferrin saturation in a genetically isolated

population, the Erasmus Rucphen Family (ERF) Study. *Hum Hered* 2006;61(4):222-8.

(9) Liu F, Elefante S, van Duijn CM, Aulchenko YS. Ignoring distant genealogic loops leads to false-positives in homozygosity mapping. *Ann Hum Genet* 2006;70(Pt 6):965-70.

(10) Liu F, rias-Vasquez A, Sleegers K, Aulchenko YS, Kayser M, Sanchez-Juan P, et al. A genomewide screen for late-onset Alzheimer disease in a genetically isolated Dutch population. *Am J Hum Genet* 2007;81(1):17-31.

(11) Boezen HM, Vonk JM, van Aalderen WM, Brand PL, Gerritsen J, Schouten JP, et al. Perinatal predictors of respiratory symptoms and lung function at a young adult age. *Eur Respir J* 2002;20(2):383-90.

(12) 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.

(13) van Diemen CC, Postma DS, Vonk JM, Bruinenberg M, Schouten JP, Boezen HM. A disintegrin and metalloprotease 33 polymorphisms and lung function decline in the general population. *Am J Respir Crit Care Med* 2005;172(3):329-33.

(14) van Diemen CC, Postma DS, Vonk JM, Bruinenberg M, Nolte IM, Boezen HM. Decorin and TGF-beta1 polymorphisms and development of COPD in a general population. *Respir Res* 2006;7:89.

(15) From the Global Strategy for the Diagnosis, Management and Prevention of COPD, Global Initiative for Chronic Obstructive Lung Disease (GOLD) 2006. Available from: http://www.goldcopd.org. 2007.

(16) Boerwinkle E, Chakraborty R, Sing CF. The use of measured genotype information in the analysis of quantitative phenotypes in man. I. Models and analytical methods. *Ann Hum Genet* 1986 May;50(Pt 2):181-94.

(17) Aulchenko YS, de Koning DJ, Haley C. Genomewide rapid association using mixed model and regression: a fast and simple method for genomewide pedigree-based quantitative trait Loci association analysis. *Genetics* 2007;177(1):577-85.

(18) Gilmour ABGBCSWaRT. ASReml user guide release 1.0. NSW Agriculture, Orange, Australia ed. 2002.

(19) Celedon JC, Lange C, Raby BA, Litonjua AA, Palmer LJ, DeMeo DL, et al. The transforming growth factor-beta1 (TGFB1) gene is associated with chronic obstructive pulmonary disease (COPD). *Hum Mol Genet* 2004;13(15):1649-56.

(20) Su ZG, Wen FQ, Feng YL, Xiao M, Wu XL. Transforming growth factor-beta1 gene polymorphisms associated with chronic obstructive pulmonary disease in Chinese population. *Acta Pharmacol Sin* 2005;26(6):714-20.

(21) Wu L, Chau J, Young RP, Pokorny V, Mills GD, Hopkins R, et al. Transforming growth factor-beta1 genotype and susceptibility to chronic obstructive pulmonary disease. *Thorax* 2004;59(2):126-9.

(22) Ogawa E, Ruan J, Connett JE, Anthonisen NR, Pare PD, Sandford AJ. Transforming growth factor-beta1 polymorphisms, airway responsiveness and lung function decline in smokers. *Respir Med* 2007;101(5):938-43.

(23) Yoon HI, Silverman EK, Lee HW, Yoo CG, Lee CT, Chung HS, et al. Lack of association between COPD and transforming growth factor-beta1 (TGFB1) genetic polymorphisms in Koreans. *Int J Tuberc Lung Dis* 2006;10(5):504-9.

(24) Guo X, Lin HM, Lin Z, Montano M, Sansores R, Wang G, et al. Surfactant protein gene A, B, and D marker alleles in chronic obstructive pulmonary disease of a Mexican population. *Eur Respir J* 2001;18(3):482-90.

(25) van Diemen CC, Postma DS, Vonk JM, Boezen HM. Polymorphisms in surfactant proteins and FEV1 decline and development of COPD in the general population. *Eur Respir J* 2006; 28[Suppl 50], 143s-144s.

(26) Aulchenko YS, Hoppenbrouwers IA, Ramagopalan SV, Broer L, Jafari N, Hillert J, et al. Genetic variation in the KIF1B locus influences susceptibility to multiple sclerosis. *Nat Genet* 2008;40(12):1402-3.

(27) DeMeo DL, Hersh CP, Hoffman EA, Litonjua AA, Lazarus R, Sparrow D, et al. Genetic Determinants of Emphysema Distribution in the National Emphysema Treatment Trial. *Am J Respir Crit Care Med* 2007; 176(1):42-8.

(28) Martinez FJ, Foster G, Curtis JL, Criner G, Weinmann G, Fishman A, et al. Predictors of mortality in patients with emphysema and severe airflow obstruction. *Am J Respir Crit Care Med* 2006;173(12):1326-34.

(29) Martinez FJ, Curtis JL, Sciurba F, Mumford J, Giardino ND, Weinmann G, et al. Gender Differences in Severe Pulmonary Emphysema. *Am J Respir Crit Care Med* 2007; 176(3):234-52.

(30) Gosman MM, Boezen HM, van Diemen CC, Snoeck-Stroband JB, Lapperre TS, Hiemstra PS, et al. A disintegrin and metalloprotease 33 and chronic obstructive pulmonary disease pathophysiology. *Thorax* 2007;62(3):242-7.

(31) Heidinger K, Konig IR, Bohnert A, Kleinsteiber A, Hilgendorff A, Gortner L, et al. Polymorphisms in the human surfactant protein-D (SFTPD) gene: strong evidence

that serum levels of surfactant protein-D (SP-D) are genetically influenced. *Immunogenetics* 2005;57(1-2):1-7.

(32) Hersh CP, DeMeo DL, Lazarus R, Celedon JC, Raby BA, Benditt JO, et al. Genetic Association Analysis of Functional Impairment in Chronic Obstructive Pulmonary Disease. *Am J Respir Crit Care Med* 2006; 173(9):977-84.

(33) Leth-Larsen R, Garred P, Jensenius H, Meschi J, Hartshorn K, Madsen J, et al. A common polymorphism in the SFTPD gene influences assembly, function, and concentration of surfactant protein D. *J Immunol* 2005;174(3):1532-8.

(34) Joos L, He JQ, Shepherdson MB, Connett JE, Anthonisen NR, Pare PD, et al. The role of matrix metalloproteinase polymorphisms in the rate of decline in lung function. *Hum Mol Genet* 2002;11(5):569-76.

(35) Rutter JL, Mitchell TI, Buttice G, Meyers J, Gusella JF, Ozelius LJ, et al. A single nucleotide polymorphism in the matrix metalloproteinase-1 promoter creates an Ets binding site and augments transcription. *Cancer Res* 1998;58(23):5321-5.

(36) Price SJ, Greaves DR, Watkins H. Identification of novel, functional genetic variants in the human matrix metalloproteinase-2 gene: role of Sp1 in allele-specific transcriptional regulation. *J Biol Chem* 2001;276(10):7549-58.

(37) Ito I, Nagai S, Handa T, Muro S, Hirai T, Tsukino M, et al. Matrix metalloproteinase-9 promoter polymorphism associated with upper lung dominant emphysema. *Am J Respir Crit Care Med* 2005;172(11):1378-82.

(38) Lose F, Thompson PJ, Duffy D, Stewart GA, Kedda MA. A novel tissue inhibitor of metalloproteinase-1 (TIMP-1) polymorphism associated with asthma in Australian women. *Thorax* 2005;60(8):623-8.

(39) Siedlinski M, van Diemen CC, Postma DS, Boezen HM. Heme oxygenase 1 variations and lung function decline in smokers: proof of replication. *J Med Genet* 2008;45(6):400.

(40) Ishii T, Matsuse T, Teramoto S, Matsui H, Miyao M, Hosoi T, et al. Glutathione S-transferase P1 (GSTP1) polymorphism in patients with chronic obstructive pulmonary disease. *Thorax* 1999;54(8):693-6.

(41) Sundberg K, Johansson AS, Stenberg G, Widersten M, Seidel A, Mannervik B, et al. Differences in the catalytic efficiencies of allelic variants of glutathione transferase P1-1 towards carcinogenic diol epoxides of polycyclic aromatic hydrocarbons. *Carcinogenesis* 1998;19(3):433-6.

(42) Vibhuti A, Arif E, Deepak D, Singh B, Qadar Pasha MA. Genetic polymorphisms of GSTP1 and mEPHX correlate with oxidative stress markers and lung function in COPD. *Biochem Biophys Res Commun* 2007;359(1):136-42.

Table 1: Candidate genes and single nucleotide polymorphisms genotyped in the study population

Gene	*Description of gene*	*SNPs genotyped rs numbers - alternative name*		*Functional SNP*	References
ADAM33	A Disintegrin And Metalloprotease 33: exact function unknown; identified by genome wide screen as susceptibility gene for asthma. Associated with decline in FEV_1 and development of COPD in the general population and severity of inflammation in COPD patients.	rs17548913 rs17548907 rs3918396 rs528557 rs597980 rs2280091 rs2280090 rs2787094	ADAM33F+1 ADAM33Q-1 ADAM33S1 ADAM33S2 ADAM33ST+5 ADAM33T1 ADAM33T2 ADAM33V4	- - - - - - - -	[13;30]
TGFβ1	Transforming Growth Factor-β1:a chemotactic cytokine for fibroblasts, inducing synthesis of matrix proteins and glycoproteins and inhibiting collagen degradation by induction of protease inhibitors and reduction of metalloproteases; TGF-β1 levels are increased in COPD; SNPs have been associated with COPD	rs1800469 rs1982073 rs6957	TGFB1C-509T TGFB1Leu10Pro TGFB13UTR	increased TGFB increased TGFB -	[13,14;19-21]
SFTPA1	Surfactant protein A1: Surfactant proteins (SP) are involved in the first response to microorganisms in the lung, regulation of inflammation and structure of alveoli. SP reduce surface tension at the air–liquid interface and therefore prevent alveolar collapse during expiration.	rs1059047 rs1136450 rs4253527	SPA1Val19Ala SPA1Leu50Val SPA1Arg219Trp	- - -	[24,31-33]
SFTPA2	Surfactant protein A2: idem SP-A1, homologous gene	rs1059046 rs17886395 rs1965707	SPA2Asn9Thr, SPA2Pro91Ala, SPA2Ser140Ser	- - -	
SFTPB	Surfactant protein B: hydrophobic component of pulmonary surfactant	rs1130866	SPBIle131Thr	altered affinity	
SFTPD	Surfactant protein D: a C-type lectin present in pulmonary surfactant and several other mucosal surfaces. It modulates innate immunity, allergic response, expression of matrix metalloproteases, alveolar wall remodeling, emphysema, fibrosis and lipid and macrophage homeostasis. Associated with COPD.	rs721917 rs2243639 rs1799750	SPDMet11Thr SPDThr160Ala MMP1G-1607GG	altered SP-D protein assembly, function and levels - additional ETS transcription factor binding site, increased expression	[34,35]
MMP1	Matrix Metalloprotease 1: an interstitial collagenase involved in tissue remodeling and repair associated with lung development and inflammation. Levels are increased in sputum of COPD patients compared to healthy controls. Associated with lung function decline.				

MMP2	Matrix Metalloprotease 2: a type IV collagenase specifically cleaving type IV collagen, the major structural component of basement membranes	rs243865	MMP2C-1306T	loss of SP-1 transcription factor binding site, less expression	[36]
MMP9	Matrix Metalloprotease 9: a gelatinase B involved in tissue remodeling; smokers with airway obstruction show higher MMP9 expression than smokers without COPD and non-smokers	rs3918278 rs6065912 rs8113877	mmp9_rs3918278 mmp9_rs6065912 mmp9_rs8113877	tagging tagging tagging	[34,37]
MMP12	Matrix Metalloprotease 12: a human macrophage elastase involved in degradation of extracellular matrix in lungs of patients with COPD. Associated with lung function decline.	rs2276109 rs652438	MMP12A-82G MMP12Asn357Ser	AP-1transcription factor binding site, increased MMP12 -	[34]
TIMP1	Tissue Inhibitor of Matrix Metalloprotease 1: inhibitor of several MMPs, including MMP1, MMP9 and MMP12. X-chromosomal. Associated with asthma.	rs11551797 rs4898	timp1Ile158 timp1Phe124	- -	[38]
HMOX1	Hemoxygenase 1: role in oxidant-antioxidant balance in the lung. Genetic variation associated with COPD.	rs2071747	HO1Asp7His	-	[39]
GSTP1	Glutathione S-transferase P1: role in oxidant-antioxidant balance in the lung. Associated with COPD.	rs1695 rs1138272	gstp1Ile105Val gstp1Ala114Val	increased enzyme activity -	(40-42)

Table 2: Characteristics of the GRIP and Vlagtwedde study populations

	GRIP ***FEV_1/IVC <70%*** ***N=91***	***Vla/Vla*** ***FEV_1/IVC <70%*** ***N=351***	***P value***	***GRIP*** ***FEV_1/IVC <70%,*** ***FEV_1pp <80%*** ***N=67***	***Vla/Vla*** ***FEV_1/IVC*** ***<70%,*** ***FEV_1pp <80%*** ***N=167***	P value *
Age	66 (41-84)	58.0 (35-76)	<0.001	66.0 (43-82)	59.0 (35-76)	<0.001
Sex m/f, n	47/44	244/107	0.001	36/31	122/45	0.004
Smoking,						
% Never	3.4	18.8	0.001	3.1	16.2	0.026
% Ex	38.6	35.9		38.5	33.5	
% Current	58.0	45.3		58.4	50.3	
Pack-years	34.8 (0-120)	21.4 (0-262)	0.001	39.0 (0-120)	26.0 (0-262)	0.015
FEV_1 % of predicted	69.4 (26.4-110.5)	80.7 (36.0-115.0)	<0.001	63.5 (26.4-79.0)	69.9 (36.0-79.8)	0.001
FEV_1/FVC	56.2 (27.7-68.4)	n.a.	n.a.	52.8 (27.7-67.9)	n.a.	n.a.
FEV1/IVC	54.5 (20.7-69.8)	64.9 (29.0-69.9)	<0.001	50.8 (20.7-67.7)	59.2 (29.4-69.8)	<0.001
Chronic cough, %	58.2	14.5	<0.001	60.6	22.2	<0.001
Chronic phlegm, %	50.5	10.5	<0.001	51.5	15.0	<0.001

*P value for difference between GRIP and Vla/Vla study populations with FEV_1/IVC <70%, FEV_1pp <80% derived from X^2-test for comparison of discrete variables; Mann-Whitney U test for test between continuous variables; data are presented as median (range). Abbreviations: GRIP Genetic Research in Isolated populations; Vla/Vla Vlagtwedde/Vlaardingen; GOLD Global initiative for Obstructive Lung Diseases; m male; f female; FEV_1 Forced Expiratory Volume in 1 second; FVC Forced Vital Capacity; FEV_1pp=FEV_1 as percentage of predicted

Table 3: Associations of SNPs with FEV_1/IVC in GOLD stage ≥ I in GRIP and GOLD stage ≥ II in GRIP and Vlagtwedde/Vlaardingen

SNP	*Comparison*	*N*	*Estimate GRIP GOLD ≥ I n=91*	*P value*	*N*	*Estimate GRIP GOLD ≥ II n=67*	*P value*	*N*	*Estimate Vla/Vla GOLD ≥ II n=167*	*P value*
ADAM33	wt	15	ref	-	12	ref	-	27	ref	-
ST+5	het vs. wt	46	3.8	0.220	31	2.8	0.424	86	2.9	**0.082**
	hom vs. wt	29	0.3	0.919	24	1.9	0.594	51	0.8	0.668
TGFβ1	wt	41	ref	-	29	ref	-	94	ref	-
C-509T	het vs. wt	38	-3.8	0.102	30	-3.7	0.146	60	-1.3	0.298
	hom vs. wt	11	-6.6	**0.063**	8	-9.4	**0.017**	8	-5.0	**0.070**
TGFβ1	wt	32	ref	-	22	ref	-	68	ref	-
Leu10Pro	het vs. wt	40	-4.6	**0.061**	32	-4.7	**0.081**	65	-0.8	0.952
	hom vs. wt	13	-5.8	**0.088**	8	-10.8	**0.007**	17	-4.5	**0.028**
SFTPA1	wt	60	ref	-	45	ref	-	123	ref	-
Leu50Val	het vs. wt	19	-2.7	0.329	15	-1.5	0.623	20	2.7	0.159
	hom vs. wt	4	13.6	**0.015**	1	18.9	**0.076**	11	1.8	0.474
SFTPA2	wt	58	ref	-	42	ref	-	117	ref	-
Pro91Ala	het vs. wt	29	0.5	0.833	22	-0.1	0.986	41	-1.1	0.423
	hom vs. wt	3	-10.2	**0.099**	3	-7.2	0.232	3	0.4	0.923
SFTPD	wt	33	ref	-	22	ref	-	44	ref	-
Met11Thr	het vs. wt	35	-4.4	**0.090**	29	-4.0	0.161	85	-0.9	0.512
	hom vs. wt	19	-3.2	0.291	13	-4.3	0.226	31	-0.3	0.888
SFTPD	wt	33	ref	-	26	ref	-	54	ref	-
Ala160Thr	het vs. wt	41	5.6	**0.025**	30	5.4	**0.055**	73	2.1	0.112
	hom vs. wt	12	2.00	0.582	8	1.15	0.778	29	-1.6	0.376

* Model of inheritance: general model: het vs. wt: effect of heterozygous genotype compared to wild type genotype; hom vs. wt: effect of homozygous mutant genotype compared to wild type genotype. Abbreviations: GRIP Genetic Research in Isolated populations; Vla/Vla Vlagtwedde/Vlaardingen; GOLD Global initiative for Obstructive Lung Diseases;ADAM33 A Disintegrin and Metalloprotease 33; TGFβ1 Transforming Growth Factor β1; SFTP Surfactant Protein; ref reference; se standard error

Table E1: Genotype frequencies of the significant SNPs in the GRIP GOLD ≥ II population compared to the genotype frequencies in the Vlagtwedde/Vlaardingen GOLD ≥ II population

SNP and genotype		***GRIP N=67 N (%)***	***Vla/Vla N=167 N (%)***	**P value**
ADAM33	AA	12 (17.9)	27 (16.5)	0. 690
ST+5	AG	31 (46.3)	86 (52.4)	
	GG	24 (35.8)	51 (31.1)	
TGFβ1	GG	29 (43.3)	94 (58.0)	0.051
C-509T	GA	30 (44.8)	60 (37.0)	
	AA	8 (11.9)	8 (4.9)	
TGFβ1	AA	22 (34.9)	68 (45.3)	0.368
Leu10Pro	AG	33 (52.4)	65 (43.3)	
	GG	8 (12.7)	17 (11.3)	
SFTPA1	GG	45 (73.8)	123 (79.9)	**0.045**
Leu50Val	GC	15 (24.6)	20 (13.0)	
	CC	1 (3.1)	11 (7.1)	
SFTPA2	GG	42 (62.7)	117 (72.7)	0.242
Pro91Ala	GC	22 (32.8)	41 (25.5)	
	CC	3 (4.5)	3 (1.9)	
SFTPD	TT	22 (34.4)	44 (27.5)	0.522
Met11Thr	TC	29 (45.3)	85 53.1)	
	CC	13 (20.3)	31 (19.4)	
SFTPD	AA	26 (40.6)	54 (34.6)	0.484
Ala160Thr	AG	30 (46.9)	73 (46.8)	
	GG	8 (12.5)	29 (18.6)	

Abbreviations: GRIP Genetic Research in Isolated populations; Vla/Vla Vlagtwedde/Vlaardingen; ADAM33 A Disintegrin and Metalloprotease 33; TGFβ1 Transforming Growth Factor β1; SFTP Surfactant Protein