ERJ Express. Published on April 2, 2008 as doi: 10.1183/09031936.00076207

Polymorphisms in the *Type IV Collagen Alpha3* Gene and the Risk of COPD

Kyung Mee Kim[*], Sun Ha Park[*], Jong Sik Kim[*], Won Kee Lee[#], Sung Ick Cha[¶], Chang Ho Kim[¶], Young Mo Kang[¶], Tae Hoon Jung[¶], In San Kim[*] and Jae Yong Park[*,¶]

[*]Department of Biochemistry, Kyungpook National University School of Medicine, Daegu, Korea

[#]Department of Preventive Medicine, Kyungpook National University School of Medicine, Daegu, Korea

[¶]Department of Internal Medicine, Kyungpook National University Hospital, Daegu, Korea

Correspondence: Jae Yong Park, Department of Internal Medicine, Kyungpook National University Hospital, Samduk 2a 50, Daegu, 700-412, Korea; Fax: 82-53-426-2046; E-mail: jaeyong@knu.ac.kr

Short title: *COL4A3* polymorphisms in COPD

ABSTRACT: A number of genome-wide linkage analyses have identified the 2q33.3-2q37.2 region as most likely to contain the genes that contribute to the susceptibility to chronic obstructive pulmonary disease (COPD). It was hypothesized that the type IV collagen alpha3 (*COL4A3*) gene, which is one of the genes located at the 2q33.3-2q37.2 region, may act as a low-penetrance susceptibility gene for COPD.

To test this hypothesis, the association of *COL4A3* -1162T>C, IVS2+12C>A, P141L, G162E, H451R, P574L, and *315C>A polymorphisms with the risk of COPD was investigated in a case-control study of 311 COPD patients and 386 controls.

The presence of at least one 451R allele was associated with a significantly higher risk of COPD compared with the 451 H/H genotype [adjusted odds ratio (OR) = 1.48, 95% confidence interval (CI) = 1.03-2.14, $P = 0.03$]. When the subjects were stratified according to age and COPD severity, the 451R allele was associated with a significantly higher risk of COPD only in younger individuals with severe COPD (adjusted OR = 3.02, 95% CI = 1.37-6.67, $P = 0.006$).

In conclusion, these findings suggest that the *COL4A3* gene contributes to the genetic susceptibility to COPD.

Key words: Chronic Obstructive Pulmonary Disease, Polymorphisms, Susceptibility, Type IV Collagen Alpha 3

Introduction

Chronic obstructive pulmonary disease (COPD) is a disease characterized by progressive irreversible airflow obstruction due to the narrowing of small airways by inflammation and loss of lung elastic recoil resulting from parenchymal destruction [1]. Cigarette smoking is the single most important environmental risk factor for the development of COPD, but only a fraction of chronic smokers develops symptomatic COPD [2-4]. These observations, together with the familial clustering of COPD [5], suggest that genetic factors contribute to the development of COPD [6-8]. The genes related to the protease-antiprotease hypothesis, oxidant-antioxidant hypothesis, inflammatory response, apoptosis, and xenobiotic metabolism have been investigated based on the known or presumed pathophysiology of COPD, but the results were often inconsistent across studies [9].

A number of genome-wide linkage analyses have identified chromosomal regions that are likely to contain genes that contribute to the susceptibility to COPD [10-13]. The 2q33.3-2q37.2 region showed LOD scores greater than 2.0 for the ratio of forced expiratory volume at one second (FEV_1) to forced vital capacity (FVC) in the Boston Early-Onset COPD Study families [10]. A similar region on chromosome 2q was linked to the FEV_1/FVC ratio in the families studied in Utah Genetic Reference Project [12]. In addition, it has been suggested that the 2q33.3-2q37.2 region of linkage contains genes that contribute to COPD susceptibility through gene-by-smoking interactions [13]. This region harbors a number of candidate genes for COPD, including interleukin-8 receptor-alpha and -beta (*IL8RA* and *IL8RB*), *CYP27A1*, type IV collagen alpha 3 (*COL4A3*), and *solute carrier family 11*.

Of the positional candidate genes for COPD, we investigated the potential role of the *COL4A3* gene in the susceptibility to COPD. COL4A3, which is one of six genetically distinct homologous alpha-chains (COL4 A1-A6), is abundantly expressed in normal alveolar basement membrane [14, 15]. It assembles into triple-helical collagen molecules

with the COL4A5 and COL4A5, and it plays a pivotal role in the regulation of cellular proliferation, adhesion, migration, and differentiation [14, 15]. Moreover, COL4A3 can inhibit the activation of neutrophils. COL4A3 increases cytoplasmic cAMP in neutrophils, resulting in the inhibition of their superoxide production and proteinase secretion [16, 17]. In addition, several studies have demonstrated that COL4A3 can inhibit endothelial cell proliferation and induce cell apoptosis [18, 19]. Taken together, these findings suggest that COL4A3 may be involved in the development of COPD by modulating the inflammatory response or alveolar wall apoptosis [20-22].

Several polymorphisms in the *COL4A3* gene have been deposited into public databases [23]. Although the functional effects of these polymorphisms have not yet been fully elucidated, we hypothesized that some of these variants may have an effect on COL4A3 expression or activity, and therefore may play a role in modulating the susceptibility to COPD. To test this hypothesis, we performed a case-control study to investigate the association between *COL4A3* genotypes/haplotypes and the risk of COPD.

Materials and Methods

Identification and selection of polymorphisms

In order to screen all the potentially functional variants in the *COL4A3* gene, the public database [23] was used to search for candidate variants in the promoter region, all exons including intron-exon boundaries (20 bp on either side of the introns), and the 3'-UTR of the gene because variants in these regions are most likely to affect the gene function. Seven polymorphisms [-1162T>C (the transcription start site was counted as +1, rs12613226), IVS2+12C>A (rs1882435), P141L (in exon 7, rs10178458), G162E (in exon 7, rs6436669), H451R (in exon 22, rs11677877), P574L (in exon 25, rs28381984), and *315C>A (the nucleotide 3' of the translation termination codon was denoted by *1 [24], rs2070735)] were captured and their frequencies were determined by direct sequencing in a preliminary study that included 27 healthy Korean individuals. The samples were collected after obtaining informed consent from each individual, which included 54 chromosomes, providing at least a 95% confidence level to detect alleles with frequencies > 5%. The primer sets used for sequencing were designed based on the GenBank reference sequence (accession no. NT_005403). The sequence variants were confirmed by two independent authors. Information regarding all SNPs, SNP IDs and allele frequencies was obtained from the NCBI homepage. The seven SNPs examined had minor allele frequencies > 10% in the 27 subjects. The P141L and G162E polymorphisms were in complete LD (|D'| = 1.00 and r^2 = 1.00). Therefore, six SNPs (-1162T>C, IVS2+12C>A, G162E, H451R, P574L, and *315C>A) were chosen for the association study (Fig. 1).

Study population

The patient group (n = 311) consisted of male COPD patients who visited at the respiratory center of the Kyungpook National University Hospital between July 2006 and December 2006. COPD was diagnosed according to the criteria established by the

NHLBI/WHO Global Initiative for COPD (GOLD, Ref. 1). The criteria for COPD were as follows: chronic respiratory symptoms and signs such as cough and dyspnea; post-bronchodilator FEV_1 < 80% of the predicted value, FEV_1/FVC < 70% and FEV_1 reversibility after inhaling 200 μg salbutamol < 12% of the pre-bronchodilator FEV_1. The severity of COPD was classified by the percentage predicted FEV_1, according to the guidelines established by the GOLD: mild (> 80%), moderate (50-80%), severe (30-50%) or very severe (< 30%). The control subjects (n = 386) were recruited from 765 males who visited the health check-up center at Kyungpook National University Hospital during the same period. The enrollment criteria for the controls were as follows: age > 45 years, no known disease and no history of any disease, and no airflow limitation. A total of 386 out of 718 males who met these criteria agreed to participate in this study (participation rate, 53.8%). Compared with the subjects that refused to participate, the enrolled subjects showed a similar age distribution and smoking exposure level. All the cases and the controls were ethnic Koreans and they resided in Daegu City or the surrounding regions.

LightCycler Genotyping

The genotypes of the six polymorphisms were determined by PCR and melting-curve analysis using fluorescence-labeled hybridization probes (LightCycler, Roche Diagnostic, Mannheim, Germany). A genotype success rate of more than 98% was achieved using the FRET method. Samples that could not be scored by the FRET method were re-genotyped by direct sequencing using an ABI PRISM 3700 genetic analyzer (Applied Biosystems). All genotyping analyses were performed “blind” with respect to the case/control status in order to ensure quality control. Approximately 10% of the samples were randomly selected to be genotyped again by a different investigator, and the results were 100% concordant.

Statistical Analysis

The cases and controls were compared using the Student's *t*-test for continuous variables and a χ^2 test for categorical variables. Hardy-Weinberg equilibrium was tested using a goodness-of-fit χ^2 test with one degree of freedom, as implemented through SAS Genetics. The strength of LD between pairs of polymorphisms was measured as *D'* and r^2 by HaploView [25]. LD blocks were inferred from the definition proposed by Gabriel et al. [26]. Unconditional logistic regression analysis was used to calculate odds ratios (ORs) and 95% confidence intervals (CIs), with adjustment for possible confounders (age and pack-years of smoking as continuous variables). In addition to the overall association analysis, we performed a stratified analysis according to age and COPD severity in order to further explore the association between *COL4A3* genotypes and the risk of COPD in each stratum. For the gene-smoking interaction analyses, we used three approaches to evaluate the consistency of the results: i) stratified analyses in specific categories of cumulative smoking exposure, ii) genotype/haplotype-smoking joint effects, and iii) logistic regression model including the interaction term between genotype and smoking. In the regression model, smoking was considered as both a discrete and continuous variable as follows: i) pack-years of smoking, ii) square root of pack-years, and iii) smoking exposure level; never-smoker/lighter-smoker/heavier-smoker. Because the interaction term was not statistically significant in any of these models, we only presented the result when the continuous cumulative smoking dose (pack-years of smoking) was used in the analysis. For these analyses, ever-smokers in both groups were categorized into two subgroups according to the median pack-year value: ever-smokers $\leq$ 33 pack-years (lighter smokers) and ever-smokers > 33 pack-years (heavier smokers). All of the analyses were performed using Statistical Analysis Software for Windows, version 8.12 (SAS institute, Gary, NC, USA).

Results

The baseline characteristics of the cases and controls enrolled in this study are shown in Table 1. The FEV_1 and FEV_1/FVC ratio were significantly lower in the COPD group than in the control group.

The genotype and polymorphic allele frequencies of the six *COL4A3* polymorphisms in the cases and controls are shown in Table 2. The genotype distributions of the six polymorphisms in the controls were in Hardy-Weinberg equilibrium. The frequency of the polymorphic allele at H451R was significantly higher in the cases than in the controls (0.17 vs 0.13, $P = 0.02$). Individuals with at least one 451R allele were at a significantly higher risk for COPD compared to carriers of the 451 H/H genotype (adjusted OR = 1.48, 95% CI = 1.03-2.14, $P = 0.03$). The genotype distributions of the other five polymorphisms studied were not significantly different between the cases and controls.

The association of the H451R genotypes with the risk of COPD was further examined after stratifying the subjects according to age, smoking status and the severity of COPD (Table 3). When stratified by median age, the H/R or R/R genotype was associated with a significantly increased risk of COPD in younger individuals (adjusted OR = 1.93, 95% CI = 1.10-3.39, $P = 0.02$); however, there was no significant association between these genotypes and risk of COPD in older individuals. When stratified according to the smoking status, the effect of the H/R or R/R genotype on the risk of COPD was significant in the smokers (adjusted OR = 1.57, 95% CI = 1.08-2.27, P = 0.02) but not in the never-smokers. When the ever-smokers were dichotomized by median pack-years of smoking, the effect of the H/R or R/R genotype on the risk of COPD was significant in the heavier smokers (adjusted OR = 1.84, 95% CI = 1.10-3.08, $P < 0.02$), whereas there was no significant association between these genotypes and risk of COPD in the lighter smokers. When the COPD cases were categorized by COPD severity, the presence of one or two

451R alleles was associated with a significantly increased risk for severe COPD (GOLD III-IV, adjusted OR = 1.62, 95% CI = 1.00-2.61, $P = 0.04$), whereas there was no significant association between the H451R genotypes and the risk of mild-to-moderate COPD (GOLD I-II). When age and COPD severity were considered together, the effect of the H/R or R/R genotype was only significant in the younger individuals with severe COPD (GOLD III-IV, adjusted OR = 3.02, 95% CI = 1.37-6.67, $P = 0.006$; $P = 0.001$, test for homogeneity).

In addition to the stratification analyses, the joint effects of the H451R genotypes and smoking status on the risk of COPD were also investigated (Table 4). When the group of never-smokers with the H/H genotype was used as the reference group, the group of heavier smokers with the H/R or R/R genotype was found to have the highest risk of COPD (adjusted OR = 9.51, 95% CI = 3.93-23.01, $P < 0.001$). Nevertheless, we did not observe statistically significant evidence for interactions between the H451R genotypes and the continuous cumulative smoking dose in the multivariate logistic regression analysis ($P = 0.32$ for the interaction term).

Discussion

In the present study, we investigated the potential association between polymorphisms in the *COL4A3* gene and the risk of COPD in a Korean population. The H451R polymorphism was associated with a significantly increased risk for COPD. The association between the H451R genotypes and risk of COPD was more evident in younger individuals and in patients with severe COPD. These findings suggest that the *COL4A3* gene may be involved in the development of COPD and that the H451R polymorphism in the *COL4A3* gene may be a useful marker for genetic susceptibility to COPD.

In the present study, the *COL4A3* H451R polymorphism interacted with tobacco smoking in determining the risk of COPD. The H/R or R/R genotype was significantly associated with the risk of COPD in the smokers but not in the never-smokers, which reflects a gene-environment interaction. Such an interaction is biologically plausible because smoking is known to be a major cause of COPD. Another interesting finding of this study is that the H451R polymorphism had a more pronounced association with severe COPD in the younger patients, which is consistent with the notion that genetic factors can play a major role in the early onset of disease [10, 11]. Individuals with aberrant *COL4A3* activity may be more prone to developing COPD at a younger age, thus the association would be more clearly observed in younger patient with severe COPD.

It remains to be elucidated whether the *COL4A3* H451R polymorphism itself affect protein folding, interaction sites, solubility or stability of the protein or whether it is in LD with either another *COL4A3* variant or an adjacent true susceptibility gene. Several computational algorithms have been introduced to predict the effect of nonsynonymous single nucleotide polymorphisms (nsSNPs) on protein structures or functions [27]. The PolyPhen algorithm was developed to identify functionally important nsSNPs by predicting whether an amino acid change is likely to be deleterious to the protein on the

basis of the three-dimensional structure and multiple alignments of homologous sequences [28]. Zhu et al. [29] applied PolyPhen to 166 molecular epidemiologic studies to examine the correlation between position-specific independent counts (PSIC) score and the OR associated with a particular nsSNP. They found a significant inverse correlation between the ORs and PSIC score difference, and between tolerance index and PSIC score difference. These results indicate that PolyPhen is helpful in predicting the functional significance of the nsSNP. We also used PolyPhen to infer the functional relevance of the three *COL4A3* nsSNPs studied (G162E, H451R and P574L) and found that the H451R polymorphism was the only polymorphism that may possibly be damaging, which is consistent with the finding of the current study that the 451R allele was associated with a significantly increased risk for COPD. This finding is, therefore, biologically plausible.

The interpretation of our data is limited by the lack of published evidence showing that COL4A3 contributes to the development of COPD. The pathologic hallmarks of COPD are small airway inflammation and emphysema. One of the most prevailing hypotheses for the pathogenesis of emphysema is an inflammatory cell theory, which suggests that cigarette smoke stimulates inflammatory cells, such as neutrophils and macrophages, to release proteinases, which leads to the disruption of the extracellular matrix (ECM) and basement membrane [1, 20, 30]. An alternative hypothesis is an apoptosis theory, which suggests that cigarette smoke decreases the expression of vascular endothelial growth factor and its receptor-2, resulting in epithelial and endothelial apoptosis with subsequent loss of ECM components and alveolar units [23, 31]. COL4A3 has been shown to inhibit the activation of neutrophils and thereby decreases their secretion of superoxide and proteinases, including elastase and type IV collagenase [16, 17]. In addition, several studies have demonstrated that the non-collagenous domain of COL4A3 inhibits angiogenesis by interacting with integrins, leading to cell apoptosis [18, 19]. Therefore, it is possible that COL4A3 may be involved in the degradation of the basement membrane and ECM by modulating either the inflammatory response or

apoptosis, and thus contribute to the development of COPD. However, this hypothesis must be verified in future studies.

COL4A3 polymorphisms may have an influence on disease progression. However, in this study, however, there was no significant difference in the genotype distributions of *COL4A3* polymorphisms according to the severity of COPD. In addition, the *COL4A3* genotypes were not significantly associated with the quantitative lung function measures (data not shown).

In conclusion, the *COL4A3* H451R polymorphism was significantly associated with the risk of COPD. The association between the H451R polymorphism and the risk of COPD appeared to be influenced by tobacco smoking. The effect of the polymorphism on the risk of COPD was more pronounced in the younger patients with severe COPD. These results suggest that the *COL4A3* gene may contribute to a genetic predisposition to COPD by influencing the response to cigarette smoke exposure through a gene-smoking interaction. However, because this is the first case-control study investigating the associations between *COL4A3* polymorphisms and the risk of COPD, additional studies are required to confirm our finding. Moreover, further study is needed to define a functional role for *COL4A3* in COPD.

Acknowledgment

This study was supported by a grant of the Korea Health 21 R&D Project, Ministry of Health & Welfare, Republic of Korea (A060165), and in part by the Brain Korea 21 Project in 2006.

References

1. Pauwels RA, Buist AS, Calverley PMA, Jenkins CR, Hurd SS. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2001, 163: 1256-76.

2. Fletcher C, Peto R. The natural history of chronic airflow obstruction. *Br Med J* 1977, 1: 1645-8.

3. Burrows B, Knudson RJ, Cline MG, Lebowitz MD. Quantitative relationships between cigarette smoking and ventilatory function. *Am Rev Respir Dis* 1977; 115: 195-205.

4. Lokke A, Lange P, Scharling H, Fabricius P, Vestbo J. Developing COPD: a 25 year follow up study of the general population. *Thorax* 2006;61:935-9.

5. Kueppers F, Miller RD, Gordon H, Hepper NG, Offord K. Familial prevalence of chronic obstructive pulmonary disease in a matched pair study. *Am J Med* 1977; 63: 336-42.

6. Bascom R. Differential susceptibility to tobacco smoke: possible mechanisms. *Pharmacogenetics* 1991, 1: 102-6.

7. Silverman EK, Chapman HA, Drazen JM, Weiss ST, Rosner B, Campbell EJ, O'Donnell WJ, Reilly JJ, Ginns L, Mentzer S, Wain J, Speizer FE. Genetic epidemiology of severe, early-onset chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 1998, 157: 1770-8.

8. Foss L, Pare PD, Sandford AF. Genetic risk factors for chronic obstructive pulmonary disease. *Swiss Med Wkly* 2002, 132: 27-37.

9. Hersh CP, DeMeo DL, Lange C, Litonjua AA, Reilly JJ, Kwiatkowski D, Laird N, Sylvia JS, Sparrow D, Speizer FE, Weiss ST, Silverman EK. Attempted replication of reported chronic obstructive pulmonary disease candidate gene associations. *Am J Respir Cell Mol Biol* 2005, 33: 71-8.

10. Silverman EK, Palmer LJ, Mosley JD, Barth M, Senter JM, Brown A, Drazen JM, Kwiatkowski DJ, Chapman HA, Campbell EJ, Province MA, Rao DC, Reilly JJ, Ginns LC, Speizer FE, Weiss ST. Genomewide linkage analysis of quantitative spirometric phenotypes in severe early-onset chronic obstructive pulmonary disease. *Am J Human Genet* 2002, 70: 1229-39.

11. Palmer LJ, Celedon JC, Chapman HA, Speizer FE, Weiss ST, Silverman EK. Genome-wide linkage analysis of bronchodilator responsiveness and post-bronchodilator spirometric phenotypes in chronic obstructive pulmonary disease. *Hum Mol Genet* 2003, 12: 1199-1210.

12. Malhotra A, Peiffer AP, Ryujin DT, Elsner T, Kanner RE, Leppert MF, Hasstedt SJ. Further evidence for the role of genes on chromosome 2 and chromosome 5 in the inheritance of pulmonary function. *Am J Respir Crit Care Med* 2003, 168: 556-61.

13. DeMeo DL, Celedon JC, Lange C, Reilly JJ, Chapman HA, Sylvia JS, Speizer FE, Weiss ST, Silverman EK. Genome-wide linkage of forced mid-expiratory flow in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2004, 170: 1294-1301.

14. Boutaud A, Borza DB, Bondar O, Gunwar S, Netzer KO, Singh N, Ninomiya Y, Sado Y, Noelken ME, Hudson BG. Type IV collagen of the glomerular basement membrane. Evidence that the chain specificity of network assembly is encoded by the noncollaenous NC1 domain. *J Biol Chem* 2000; 275: 30716-24.

15. Saito K, Naito I, Seki T, Oohashi T, Kimura E, Momota R, Kishiro Y, Sado Y, Yoshioka H, Ninomiya Y. Differential expression of mouse α5(IV) and α6(IV) collagen genes in epithelial basement membranes. *J Biochem* 2000; 128: 427-34.

16. Monboisse JC, Garnotel R, Bellon G, Ohno N, Perreau C, Borel JP, Kefalides NA. The α3 chain of type IV collagen prevents activation of human polymorphonuclear leukocytes. *J Biol Chem* 1994; 269: 25475-82.

17. Shahan TA, Ziaie Z, Pasco S, Fawzi A, Bellon G, Monboisse JC, Kefalides NA. Identification of CD47/integrin-associated protein and αvβ3 as two receptors for the α3 chain of type IV collagen on tumor cells. *Cancer Res* 1999; 59: 4584-90.

18. Maeshima Y, Yerramalla UL. Dhanabal M, Holthaus KA, Barbashov S, Kharbanda S, Reimer C, Manfredi M, Dickerson WM, Kalluri R. Extracellular matrix-derived peptide binds to αvβ3 integrin and inhibits angiogenesis. *J Biol Chem* 2001; 276: 31959-68.

19. Marneros AG, Olsen BR. The role of collage-derived proteolytic fragments in angiogenesis. *Matrix Biol* 2001; 20: 337-45.

20. Groneberg DA, Chung KF. Models of chronic obstructive pulmonary disease. *Respir Res* 2004: 5; 18.

21. Kasahara Y, Tuder RM, Taraseviciene-Stewart L, Le Cras TD, Abman S, Hirth PK, Waltenberger J, Veolkel NF. Inhibition of VEGF receptors causes lung cell apoptosis and emphysema. *J Clin Invest* 2000; 106: 1311-9.

22. Aoshiba K, Yokohori N, Nagai A. Alveolar wall apoptosis causes lung destruction and emphysematous changes. *Am J Respir Cell Mol Med* 2003; 28: 555-62.

23. http://www.ncbi.nlm.nih.gov/SNP.

24. den Dunnen JT, Antonarakis E. Nomenclature for the description of human sequence variations. *Hum Genet* 2001, 109: 121-4.

25. http://broad.mit.edu/mpg/haploview.

26. Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, Higgins J, DeFelice M, Lochner A, Faggart M, Liu-Cordero SN, Rotimi C, Adeyemo A, Cooper R, Ward R, Lander ES, Daly MJ, Altshuler D. The structure of haplotype blocks in the human genome. *Science* 2002, 296: 2225-9.

27. Rebbeck TR, Spitz M, Wu X. Assessing the function of genetic variants in candidate gene association studies. *Nat Rev Genet* 2004; 5: 589-97.

28. Sunyaev S, Ramensky V, Koch I, Lathe W 3rd, Kondrashov AS, Bork P. Prediction of deleterious human alleles. *Hum Mol Genet* 2001; 10: 591-7.

29. Zhu Y, Spitz MR, Amos CI, Lin J, Schabath MB, Wu X. An evolutionary perspective on single-nucleotide polymorphisms screening in molecular cancer epidemiology. *Cancer Res* 2004; 64: 2251-7.

30. Snider GL. Chronic obstructive pulmonary disease: risk factors, pathophysiology, and pathogenesis. *Ann Rev Med* 1989; 40: 411-29.

31. Shapiro SD. Vascular atrophy and VEGFR-2 signaling: old theories of pulmonary emphysema meet new data. *J Clin Invest* 2000; 106: 1309-10.

Fig. 1. Gene map, polymorphisms, and linkage disequilibrium (LD) coefficients. (A) Gene map and polymorphisms in *COL4A3* gene on chromosome 2q36-q37. Coding exons are marked by black blocks and 3'-UTR by white block. The first base of transcription start site is denoted as +1 (Ref. genomic sequence NT_005403). (B) LD block between *COL4A3* polymorphisms in 27 healthy Koreans. Black box indicates complete LD (|D'| = 1.0 and r^2 = 1.0). White boxes indicate strong recombination (upper CI ≤ 0.9), and the gray boxes indicate uninformative findings. Triangle indicates haplotype block. Numbers in squares are |D'| (x 100) values.

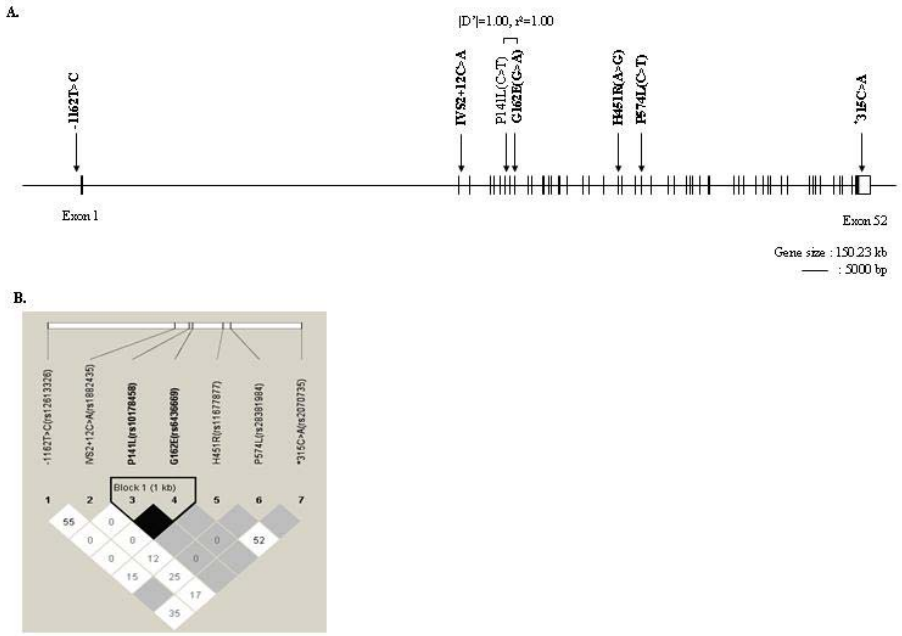

Table 1. Characteristics of the study population

Variable	COPD (n = 311)	Control (n = 386)	*P*
Age (years)	65.5 ± 8.1	60.4 ± 8.0	<0.001[*]
Smoking status			<0.001[#]
Current	171 (55.0)[¶]	205 (53.1)	
Former	129 (41.5)	128 (33.2)	
Never	11 (3.5)	53 (13.7)	
Pack-years[†]	42.6 ± 20.0	30.8 ± 16.9	<0.001[#]
FEV_1 (% Predicted)	63.4 ± 26.3	104.9 ± 16.8	<0.001[#]
FEV_1/FVC (%)	49.7 ± 13.1	80.7 ± 7.4	<0.001[#]

Note. All the cases and controls are male and ethnic Koreans.

FEV_1, forced expiratory volume in 1 second; FVC, forced vital capacity.

[*] *t*-test

[#] χ^2 test.

[¶] Numbers in parenthesis, column percentage.

[†] In current and former smokers.

Table 2. *COL4A3* genotypes of COPD cases and controls, and their association with the risk of COPD

Polymorphi (rs no.)	Genoty	Cases n (%)	Controls n (%)	$P^{\#}$	Minor allele frequency Cases	Contro	$P^{\#}$	Adjusted OR¶ (95% CI)	P¶
-1162T>C	TT	154 (49.5)	191 (49.5)	0.9	0.30	0.30	0.85	1.00	
(rs12613226	TC	130 (41.8)	158 (40.9)					0.98 (0.69-1.3	0.9
	CC	27 (8.7)	37 (9.6)					0.99 (0.55-1.8	0.9
IVS2+12C>	CC	101 (32.5)	135 (35.0)	0.7	0.42	0.41	0.54	1.00	
(rs1882435)	CA	157 (50.5)	188 (48.7)					1.20 (0.83-1.7	0.3
	AA	53 (17.0)	63 (16.3)					1.18 (0.72-1.9	0.5
G162E (G>	GG	241 (77.5)	307 (79.5)	0.3	0.12	0.11	0.74	1.00	
(rs6436669)	GE	68 (21.9)	73 (18.9)					1.11 (0.74-1.6	0.6

	EE	2 (0.6)	6 (1.6)					0.34 (0.06-1.9	0.2
H451R (A>	HH	213 (68.5)	294 (76.2)	0.0	0.17	0.13	0.02	1.00	
(rs11677877	HR	88 (28.3)	84 (21.8)					1.44 (1.00-2.1	0.0
	RR	10 (3.2)	8 (2.1)					1.99 (0.69-5.7	0.2
	HH	213 (68.5)	294 (76.2)	0.0				1.00	
	HR+RR	98 (31.5)	92 (23.8)					1.48 (1.03-2.1	0.0
P574L (C>	PP	86 (27.7)	123 (31.9)	0.3	0.46	0.44	0.60	1.00	
(rs28381984	PL	165 (53.1)	183 (47.4)					1.39 (0.95-2.0	0.0
	LL	60 (19.3)	80 (20.7)					1.25 (0.78-2.0	0.9
*315C>A†	CC	187 (60.1)	220 (57.0)	0.6	0.22	0.24	0.35	1.00	
(rs2070735)	CA	112 (36.0)	147 (38.1)					0.88 (0.62-1.2	0.4
	AA	12 (3.9)	19 (4.9)					0.55 (0.24-1.2	0.1

Two-sided $\chi 2$ test for either genotype distributions or allele frequencies between the cases and controls

¶ORs (95% CIs) and *P*-values were calculated by unconditional logistic analysis, adjusted for age and pack-years of smoking.

†The nucleotide 3' of the translation termination codon was denoted by *1.

Table 3. Association between *COL4A3* H451R (rs11677877A>G) genotypes and the risk of COPD

Variables	Cases		Controls		Adjusted OR (95 % CI) for HR + RR vs HH	*P*	P_H*
	HH	HR + RR	HH	HR + RR			
Age (years)							
≤ 60	51 (61.4)	32 (38.6)	155 (76.4)	48 (23.6)	1.93 (1.10-3.39) #	0.02	0.10
> 60	162 (71.1)	66 (28.9)	139 (76.0)	44 (24.0)	1.25 (0.78-2.01) #	0.35	
Smoking status							

Never	8 (72.7)	3 (27.3)	40 (75.5)	13 (24.5)	1.16 (0.27-5.01) ¶	0.85	
Ever†	205 (68.3)	95 (31.7)	254 (76.3)	79 (23.7)	1.57 (1.08-2.27) ¶	0.02	0.29
Pack-years of smoking†							
≤ 33	69 (74.2)	24 (25.8)	151 (76.3)	47 (23.7)	1.15 (0.65-2.06) ¶	0.63	
> 33	136 (65.7)	71 (34.3)	103 (76.3)	32 (23.7)	1.84 (1.10-3.08) ¶	0.02	0.09
Severity of COPD							
GOLD I + II	138 (69.7)	60 (30.3)	294 (76.2)	92 (23.8)	1.33 (0.88-2.01) ‡	0.18	
GOLD III + IV	75 (66.4)	38 (33.6)	294 (76.2)	92 (23.8)	1.62 (1.00-2.61) ‡	0.04	0.38
Age & Severity of COPD							
Age ≤ 60, GOLD I + II	35 (67.3)	17 (32.7)	155 (76.4)	48 (23.6)	1.44 (0.73-2.84) #	0.30	
GOLD III + IV	16 (51.6)	15 (48.4)	155 (76.4)	48 (23.6)	3.02 (1.37-6.67) #	0.006	0.001§
Age > 60, GOLD I + II	103 (70.5)	43 (29.5)	139 (76.0)	44 (24.0)	1.25 (0.75-2.10) #	0.39	
GOLD III + IV	59 (72.0)	23 (28.0)	139 (76.0)	44 (24.0)	1.31 (0.69-2.49) #	0.40	

* Test for homogeneity.

#Adjusted for pack-years of smoking; ¶Adjusted for age; †Current and former smokers; ‡Adjusted for age and pack-years of smoking.

§Compared with all the other groups combined.

Table 4. Interaction of *COL4A3* H451R genotypes and tobacco smoking on the risk of COPD

Smoking status	Genotype			
	H/H	Adjusted OR (95% CI)*	H/R + R/R	Adjusted OR (95% CI)*
Never smoker	8/40#	1.00 (reference)	3/13	1.23 (0.27-5.47)
Smoker				
≤ 33 pack-years	69/151	2.40 (1.05-5.49)¶	24/47	2.79 (1.11-7.05)†
> 33 pack-years	136/103	5.36 (2.37-12.13)‡	71/32	9.51 (3.93-23.01)‡

*Data were calculated by logistic regression, with H/H genotype in never-smokers as reference group and adjusted for age.

Cases no./Controls no.

¶$P = 0.04$.; †$P = 0.03$; ‡$P < 0.001$.

$P = 0.32$ for the interaction term between the genotype and continuous cumulative smoking dose (pack-years of smoking) in the multivariate model.