ERJ Express. Published on September 13, 2006 as doi: 10.1183/09031936.00045206

Tumor necrosis factor and lymphotoxin A polymorphisms and lung function in smokers

Authors: Goh Tanaka[1], Andrew J Sandford[1], Kelly Burkett[1], John E Connett[2], Nicholas R Anthonisen[3], Peter D Paré[1], Jian-Qing He[1]

Institutions: 1. The James Hogg iCAPTURE Centre for Cardiovascular and Pulmonary Research, St. Paul's Hospital, University of British Columbia, Vancouver, British Columbia, Canada, 2. Division of Biostatistics, School of Public Health, University of Minnesota, Minneapolis, Minnesota, USA, 3. Faculty of Medicine, University of Manitoba, Winnipeg, Manitoba, Canada

Correspondence: Dr. Andrew J Sandford; The James Hogg iCAPTURE Centre for Cardiovascular and Pulmonary Research, St. Paul's Hospital, 1081 Burrard Street, Vancouver, B.C., Canada. V6Z 1Y6.

Email: asandford@mrl.ubc.ca, TEL: (604) 806-9008, FAX: (604) 806-8351

Funding support: This work was supported by grants from the Canadian Institutes of Health Research and National Institutes of Heath Grant 5R01HL064068-04. The Lung Health Study was supported by contract N01-HR-46002 from the Division of Lung Diseases of the National Heart, Lung, and Blood Institute.

A running head: *TNF* polymorphisms, smoking and FEV_1

Word Count: 3,195 words

Abstract

Genetic variants in the tumor necrosis factor (*TNF*) gene have been investigated in chronic obstructive pulmonary disease (COPD). However there are many instances of non-replication of these associations and this might be due to insufficient power or other factors. In this study, we have examined a large number of subjects to elucidate whether genetic variations of *TNF* and/or lymphotoxin A (*LTA*), which is clustered with *TNF,* are associated with lung function in smokers. We designed two nested case-control studies in the National Heart, Lung, and Blood Institute Lung Health Study (LHS) which enrolled 5,887 smokers. The first design included continuous smokers who had the fastest (n=279) and the slowest (n=304) decline of lung function during the 5 year follow up period, and the second included the subjects who had the lowest (n=533) and the highest (n=532) post bronchodilator FEV_1% predicted at the start of the LHS. We selected and genotyped 10 tagging SNPs within the *TNF* and *LTA* region.

Unlike the previous associations between *TNF*-308 and COPD in Asians, we found no association between either of the two phenotypes and the *LTA* and *TNF* polymorphisms.

These results support the findings of previous studies in late-onset COPD in Caucasian populations.

Word counts: 199 words

Key words: FEV_1, lymphotoxin A, polymorphism, smoking, tumor necrosis factor

Introduction

The pathologic characteristics of chronic obstructive pulmonary disease (COPD) include chronic inflammation of the airways, parenchyma, and pulmonary vasculature. Polymorphonuclear leukocytes, macrophages, and T lymphocytes found at the site of disease are believed to release mediators that promote and maintain inflammation which leads to tissue damage and remodeling. Tumor necrosis factor (TNF) which is released primarily from macrophages is thought to play a critical part in the progression of COPD by increasing the expression of various pro-inflammatory mediators such as IL8 (1). Huang et al found that -308A allele in the *TNF* promoter was associated with an increased risk for the development of bronchitis in the Taiwanese population (2). Subsequently, many investigators have sought to implicate polymorphisms in *TNF* alpha in the pathogenesis of COPD. Some of these reports have found an association between polymorphisms of *TNF* and subgroups of COPD (3-5). However there are many instances of non-replication of these associations (6-9). This inconsistency might result from false positive and false negative studies due to small sample size, insufficiently defined phenotypes, lack of adjustment for important covariates and/or genetic heterogeneity of populations. A similar inconsistency of results has been observed in functional studies of the -308 polymorphism and other *TNF* alleles (10).

Another closely related candidate gene for COPD is lymphotoxin A (*LTA*; previously designated TNFB). *LTA* is clustered with *TNF* within a 6.1 kb region on chromosome 6p21.3, and polymorphisms of *LTA* and *TNF* have been reported to be in strong linkage disequilibrium (LD) in many populations (11). LTA is involved in the normal development of lymphoid tissue and also acts as an inducer of the inflammatory response (12). A polymorphism of *LTA* at position +252, which is thought to be involved in the modulation of gene expression (13, 14), has been reported as a susceptibility variant in asthma and other diseases such as myocardial infarction (14, 15). In a

recent detailed histological analysis of the small airways in patients with COPD Hogg et al (16) found that the percentage of airways containing lymphoid follicles was strongly associated with the late stages of airway obstruction.

In this study, we have examined a large number of subjects in two nested case-control studies designed to elucidate whether genetic variations of *TNF* and/or *LTA* are associated with lung function decline or lung function level in smokers who had mild to moderate airway obstruction.

Material and methods

Study subjects

All subjects were of European descent selected from the participants in the NHLBI Lung Health Study (LHS). The LHS enrolled 5,887 smokers who had spirometric evidence of mild to moderate lung function impairment from 10 North American medical centers (17).

Two nested case control studies were conducted as previously described (18). In the first design individuals who had the most and least rapid rate of decline of lung function were selected from among those who continued to smoke for the duration of the 5 year follow up. Those whose FEV_1% predicted decreased by ≥ 3.0%/year during the 5 year period (fast decline group, n=279) were compared with those whose FEV_1% predicted increased ≥ 0.4%/year (non decline group, n=304). In the second design 532 subjects who had the highest post bronchodilator FEV_1% predicted (≥ 88.9%; high function group) and 533 subjects who had the lowest post bronchodilator FEV_1% predicted (≤ 67.0%; low function group) at the start of the LHS were compared. One hundred and forty subjects included in fast or non decline groups had baseline lung function within the criteria described above, and therefore they were included in the baseline lung function study.

TagSNP selection and genotyping

For the selection of the *TNF* and *LTA* single nucleotide polymorphisms (SNPs), we used information from SeattleSNPs (19). TagSNPs were chosen using the lDSelect program (version 1.0) (20). A LD threshold of r^2>0.8 and a minor allele frequency of >10 % were set in the program. Seven SNPs in the *LTA* gene sequence (GenBank accession# AY070490), and 3 SNPs in the *TNF* gene sequence (GenBank accession# AY066019) were chosen (fig. 1). Because we could not establish an assay to genotype the SNP located at 559 in the *LTA* gene sequence and there was no alternative SNP within the same bin, we excluded this SNP from the study. The SNP located at 352 (*TNF*-238) in the *TNF* gene sequence whose allele frequency was less than 10 % was included in the study, because this SNP had been reported to be a susceptibility locus in many diseases (10). Ultimately, we analyzed 10 SNPs in the region using the TaqMan 5' exonuclease assay (21).

Statistical Analysis

Hardy-Weinberg equilibrium tests and LD estimation were done using the genetics package for R (22). For descriptive purposes haplotype frequencies were estimated with an EM algorithm using the R haplo.stats package.

For the 2-by-2 contingency tables of dominant and recessive models, the Fisher Exact test was performed. For the 3-by-2 codominant model, the chi-square statistic and asymptotic p-values were calculated. Since some cell counts were low, significance was also assessed by permutation testing. P-values from asymptotics and permutation tests were similar and therefore the asymptotic p-values are given. For the additive model, the Armitage trend test was performed. All single-locus association tests were performed using R.

For both the rate of decline and baseline samples, multivariate logistic regression was used to adjust for potential risk and confounding factors. Age, sex, pack-years of smoking, and research center were examined as potential factors. Generalized Additive Models were first used to examine the relationships between the log odds of having poor lung function (*i.e.* low baseline lung function or a fast rate of decline) and the continuous covariates of age and pack-years. In the case of pack-years, a linear relationship was not appropriate. A quadratic term for pack-years was added to the model to account for the non-linearity. All modeling for the single-SNP analyses was performed using Splus version 6.2 (Insightful Corp. Seattle, WA).

Haplotype association was tested using hapassoc (23), a contributed R package available at www.r-project.org. Haplotypes of the 6 loci in *LTA* and the 4 loci in *TNF* were considered. Since the *LTA* and *TNF* SNPs are in LD with each other, haplotypes of 3 SNPs at a time were also tested for all sets of 3 consecutive SNPs across the two genes (haplotype windows). All covariates previously considered in the single SNP association regression models were included in the logistic regression models for haplotype association.

Power Calculations

Unadjusted Analyses

Power of the two study designs was estimated using the two independent proportions and many proportions functions in PASS 2005 (Number Cruncher Statistical Systems. Kaysville, Utah). The sample sizes used were those of the baseline (high 532, low 533) and rate of decline (fast 279, non 304) groups.

To estimate power for the codominant models, the odds ratio (OR) for two copies of the variant was set to be the square of the OR for one copy. Power was calculated for the 2 degree of

freedom chi-square test. These ORs and the observed proportions in the control groups were used to determine likely values of *W*, the measure of effect size (24), for calculating power.

Adjusted Analyses

PASS was used to calculate sample size for multiple logistic regression. The method assumes that the effect of one dichotomous covariate is of interest, while controlling for the effect of other covariates.

Results

Characteristics of subjects

The characteristics of the study subjects are shown in tables 1 and 2. Some potentially confounding factors were significantly different between the groups as has been reported in previous LHS studies (18).

Single SNP analysis

We analyzed each of the selected SNPs and the rate of lung function decline (table 3). A multivariate logistic regression model was used to adjust for potential risk and confounding factors including age, gender, pack-years of smoking, and research center in the analyses. There were no differences between the fast decline group and the non decline group in the distribution of genotype frequencies for these SNPs. Further analysis using dominant, recessive, and additive models, confirmed the lack of any association between the rate of lung function decline and *LTA* or *TNF* polymorphisms (data not shown).

Similarly for the baseline level of lung function we found no significant difference in genotype frequencies between the high and low baseline groups (table 4). Additionally, we analyzed the baseline level of lung function excluding the subjects who were involved in fast or

non decline groups. We confirmed the lack of association in this analysis (data not shown). All the SNPs were in Hardy-Weinberg equilibrium except for rs3093543 in the high function group of the baseline lung function study (p=0.018).

Haplotype analysis

The distribution of haplotypes formed by the 6 *LTA* polymorphisms and in the 4 *TNF* polymorphisms was analyzed in the two study designs (tables 5 and 6). As in the single SNP analyses, we adjusted for potential risk and confounding factors. There was no significant association between the *LTA* or *TNF* haplotypes and the two phenotypes. Furthermore we performed a 3 SNP haplotype analysis using a sliding window, however we did not find any significant associations (data not shown).

Power Calculations

Unadjusted Analyses

Figures 2 a and b give the power curves for the rate of decline and baseline datasets respectively. For the baseline group sample sizes, there would be greater than 80% power to detect an OR of at least 1.75 assuming a minor allele frequency of 0.10 or an OR of 1.5 assuming a minor allele frequency of 0.20. Since the proportions of individuals in the dominant genotype category range from 0.1 to 0.6, there is reasonable power to detect an OR of 1.5 for most SNPs genotyped. However, for a recessive model, the OR would have to be greater than 2 to have greater than 80% power for the 0.03 recessive genotype category proportion. Since the rate of decline groups are smaller than those of the baseline groups, the ORs would have to be higher still to have reasonable power. For a proportion of 0.10 non decliners in a genotype category, the OR would have to be 2 to have a greater than 80% power to detect the difference in proportions.

For a codominant model, at an effect size of at least W=0.10 there is 80% power or greater to detect the difference in genotypic proportions in the baseline dataset of 1096 individuals. Under most distributions of *TNF-LTA* genotypes, this value corresponds to an OR of 1.4 per allele. For the rate of decline dataset of 595 individuals, the necessary effect size is W=0.13 or greater. For example, this effect size could correspond to a minor allele frequency (MAF) of 0.3 and an OR of 1.4 per allele, or a MAF of 0.25 and an OR of 1.6 per allele. For all genotype distributions observed for *TNF-LTA*, there would be adequate power to detect an OR of 2 per allele in the rate of decline dataset.

Adjusted Analyses

The power for the rate of decline models was calculated using R^2 values of 0 and 0.03, with an OR of 2 and a variety of percentages of the sample with the risk genotype that corresponded to the values in the unadjusted power analyses. The power computed for an R^2 of 0.03 was very close to that of an R^2 of 0 and that calculated using the difference in proportion functions given in the previous section and in figures 2 a and b.

Discussion

In this paper, we selected 10 polymorphisms within the *TNF* and *LTA* region and genotyped these variants in two nested case-control studies: one using the rate of decline of lung function as the phenotype the other using the level of lung function as the phenotype. We analyzed the relationship between these two COPD phenotypes and polymorphisms in the *LTA* and *TNF* genes and found no evidence for association.

Polymorphisms of *TNF*, especially at the *TNF*-308 locus, have been reported as susceptibility variants in many infectious and autoimmune diseases (10). An association between COPD and *TNF* polymorphisms was first reported by Huang et al (2). These investigators genotyped 42 adult males with chronic bronchitis and 42 sex, age and smoking matched control subjects as well as 99 randomly sampled schoolchildren. They found that the frequency of the *TNF*-308A allele was significantly higher in the cases than the controls ($p<0.001$, odds ratio = 11.1, 95% CI = 2.9-42.6). Subsequently, Sakao et al reported a significant association of the 308A allele in a Japanese population of smokers with a FEV_1 less than 80% predicted (4). However no association of *TNF* SNPs and COPD phenotypes was found by another Japanese group (6) or in a study of a Thai population (7).

On the other hand, in a Caucasian population, Keatings et al showed that COPD patients who were AA homozygotes for *TNF*-308 demonstrated less reversibility of air flow obstruction following a β2-agonist ($p<0.05$) (3). However, in the same study (3), no difference was found in the distribution of the *TNF*-308 genotypes between the COPD patients and controls. This lack of association was consistent with the results from studies of other Caucasian populations (8, 9).

An important challenge in case control studies is the possibility of false positive or negative associations owing to small sample size, population stratification, multiple testing, and differences

in phenotypic definition (25). The inconsistent results of previous studies of *TNF* polymorphisms in COPD might be due to some or all of these factors. Most previous studies have had small sample sizes and therefore the negative results may be due to low power. Since power is such an important consideration in the interpretation of studies which report negative results we performed a detailed analysis of the power to detect associations in this study. The results of this analysis show that for most of the allele frequencies and genetic models we studied we should have > 80% power to detect a genotype/haplotype relative risk of ~ 1.7 or greater. Therefore, there was enough power to detect the difference of genotypic proportions shown in the previous studies in which positive association between phenotypes of COPD and *TNF*-308 had been reported (2,4). However, we cannot exclude the possibility that the true odds ratio for the polymorphism is <1.7. Meta-analysis may be an effective tool to detect such variants, although true sources of variability that may exist between populations and publication bias make this problematic.

Another explanation of inconsistent results in association studies is differences in phenotypic definition. Most previous studies used different phenotypes such as diagnosis of chronic bronchitis, COPD or emphysema. In our study, we used two well defined phenotypes based on lung function which are closely linked to the diagnosis and evaluation of COPD, but we did not investigate other phenotypes such as emphysema. TNF was shown to be crucial to the development of emphysema in a study of TNF receptor knock out mice exposed to smoke (26). Therefore there is a possibility that the specific sub-phenotype of emphysema is associated with *TNF* polymorphisms.

Furthermore, the inconsistent results in association studies may come from the possibility that the reported variants did not contribute to the diseases but are in LD with causal variants. In addition, compared with a SNP level approach, the gene-based approach for replication has the advantage of being less susceptible to erroneous findings due to genetic differences between

populations (27). This is the major reason that we have expanded our previous study of the *TNF*-308 and *LTA*+252 SNPs (28) to 10 SNPs in *TNF* and *LTA*. Although *TNF* polymorphisms especially *TNF*-308 have been most studied in COPD, whether these polymorphisms are functional or not is a controversial issue as a result of both ex vivo and in vitro studies (10). Recently Knight et al developed a method called haploChIP employing chromatin immunoprecipitation and mass spectrometry to detect the amount of phosphorylated RNA polymerase II bound to two different alleles of a gene, the quantity of which is related to the transcriptional activity (29). Using this technique, they showed that functionally specific haplotypes of the *TNF*/*LTA* locus did not correlate with allele-specific transcription of *TNF* but did correlate with evidence of transcription of *LTA*. They speculated that polymorphisms of *TNF* were in LD with causal variants of *LTA*. To avoid overlooking potential causal variants in the two candidate genes, we used a stringent LD threshold to select a set of highly informative markers which covered from 2,000 bp upstream of the 5' end of *LTA* to 1,500 bp downstream of 3' end of *TNF* including the whole of the *TNF* promoter site.

Most recently, an association was found between the *TNF*-308 polymorphism and both quantitative and qualitative phenotypes related to COPD in the Boston Early-Onset COPD study that included 949 individuals from 127 families (9). However, the same authors were unable to replicate this association in a case-control study of usual later onset COPD which included 304 patients and 441 controls (9). The authors suggested that genetic factors related to severe early-onset COPD might be different from usual later onset COPD as a possible explanation of their inconsistent results. Furthermore they emphasized the necessity of replication in an independent cohort because of the possibility that positive associations might be due to multiple testing. Our results, however, are not a direct replication of their results from the case-control study. The study designs we employed are not true case-controls but rather comparison of

phenotypic extremes within a cohort of smokers selected for evidence of mild/moderate airflow obstruction. Therefore, we are investigating disease severity genes. On the other hand, the selection of phenotypic extremes is an accepted strategy for identifying risk alleles for complex genetic disease (30). Although it is difficult to estimate how this design would influence our ability to identify genetic risk, it might be expected to increase rather than decrease our power since it is likely that individuals at phenotypic extremes harbor susceptibility alleles.

We designed the study to be relatively robust to type II error, but there are still limitations to our study, which relate to phenotypic and ethnic differences compared with previous studies. On one hand, we did not use emphysema as a phenotype. Additionally, our study was limited to Caucasian individuals and many of the previous associations between *TNF*-308 and COPD were reported in Asian populations.

In conclusion, we selected a set of highly informative markers based on the r^2 LD statistic within the *TNF* and *LTA* region. To reduce the type II error in the association study, we analyzed a relatively large number of samples for two well defined phenotypes, the rate of decline and baseline level of the lung function, and we adjusted for factors which had an effect on the phenotypes. Unlike the previous associations between *TNF*-308 and COPD in Asians, we found no association between either of the two phenotypes and polymorphisms of *LTA* and *TNF*. These results support the findings of other previous studies in late-onset COPD in Caucasian populations.

Acknowledgments

The authors gratefully acknowledge the NHLBI for the recruitment and characterization of this study. The authors wish to thank Paula Lindgren for her insightful review of the manuscript. AJS is the recipient of a Canada Research Chair in genetics and a Michael Smith Foundation for Health Research Scholar Award.

References

1. Keatings VM, Collins PD, Scott DM, Barnes PJ. Differences in interleukin-8 and tumor necrosis factor-alpha in induced sputum from patients with chronic obstructive pulmonary disease or asthma. *Am J Respir Crit Care Med* 1996;153:530-534.
2. Huang SL, Su CH, Chang SC. Tumor necrosis factor-alpha gene polymorphism in chronic bronchitis. *Am J Respir Crit Care Med* 1997;156:1436-1439.
3. Keatings VM, Cave SJ, Henry MJ, Morgan K, O'Connor CM, FitzGerald MX, Kalsheker N. A polymorphism in the tumor necrosis factor-alpha gene promoter region may predispose to a poor prognosis in COPD. *Chest* 2000;118:971-975.
4. Sakao S, Tatsumi K, Igari H, Shino Y, Shirasawa H, Kuriyama T. Association of tumor necrosis factor alpha gene promoter polymorphism with the presence of chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2001;163:420-422.
5. Kucukaycan M, Van Krugten M, Pennings HJ, Huizinga TW, Buurman WA, Dentener MA, Wouters EF. Tumor Necrosis Factor-alpha +489G/A gene polymorphism is associated with chronic obstructive pulmonary disease. *Respir Res* 2002;3:29.
6. Ishii T, Matsuse T, Teramoto S, Matsui H, Miyao M, Hosoi T, Takahashi H, Fukuchi Y, Ouchi Y. Neither IL-1beta, IL-1 receptor antagonist, nor TNF-alpha polymorphisms are associated with susceptibility to COPD. *Respir Med* 2000;94:847-851.
7. Chierakul N, Wongwisutikul P, Vejbaesya S, Chotvilaiwan K. Tumor necrosis factor-alpha gene promoter polymorphism is not associated with smoking-related COPD in Thailand. *Respirology* 2005;10:36-39.
8. Seifart C, Dempfle A, Plagens A, Seifart U, Clostermann U, Muller B, Vogelmeier C, von Wichert P. TNF-alpha-, TNF-beta-, IL-6-, and IL-10-promoter polymorphisms in patients with chronic obstructive pulmonary disease. *Tissue Antigens* 2005;65:93-100.

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: 71-78.
10. Bayley JP, Ottenhoff TH, Verweij CL. Is there a future for TNF promoter polymorphisms? *Genes Immun* 2004;5:315-329.
11. Belfer I, Buzas B, Hipp H, Dean M, Evans C, Lorincz I, Max MB, Goldman D. Haplotype structure of inflammatory cytokines genes (IL1B, IL6 and TNF/LTA) in US Caucasians and African Americans. *Genes Immun* 2004;5:505-512.
12. Pfeffer K. Biological functions of tumor necrosis factor cytokines and their receptors. *Cytokine Growth Factor Rev* 2003;14:185-191.
13. Messer G, Spengler U, Jung MC, Honold G, Blomer K, Pape GR, Riethmuller G, Weiss EH. Polymorphic structure of the tumor necrosis factor (TNF) locus: an NcoI polymorphism in the first intron of the human TNF-beta gene correlates with a variant amino acid in position 26 and a reduced level of TNF-beta production. *J Exp Med* 1991;173:209-219.
14. Ozaki K, Ohnishi Y, Iida A, Sekine A, Yamada R, Tsunoda T, Sato H, Hori M, Nakamura Y, Tanaka T. Functional SNPs in the lymphotoxin-alpha gene that are associated with susceptibility to myocardial infarction. *Nat Genet* 2002;32:650-654.
15. Moffatt MF, Cookson WO. Tumour necrosis factor haplotypes and asthma. *Hum Mol Genet* 1997;6:551-554.
16. Hogg JC, Chu F, Utokaparch S, Woods R, Elliott WM, Buzatu L, Cherniack RM, Rogers RM, Sciurba FC, Coxson HO, Pare PD. The nature of small-airway obstruction in chronic obstructive pulmonary disease. *N Engl J Med* 2004;350:2645-2653.

17. Anthonisen NR, Connett JE, Kiley JP, Altose MD, Bailey WC, Buist AS, Conway WAJ, Enright PL, Kanner RE, O'Hara P, et a. Effects of smoking intervention and the use of an inhaled anticholinergic bronchodilator on the rate of decline of FEV1. The Lung Health Study. *JAMA* 1994;272:1497-1505.

18. He JQ, Connett JE, Anthonisen NR, Pare PD, Sandford AJ. Glutathione S-transferase variants and their interaction with smoking on lung function. *Am J Respir Crit Care Med* 2004;170:388-394.

19. SeattleSNPs. NHLBI Program for Genomic Applications, SeattleSNPs, Seattle, WA http://pga.gs.washington.edu. Date last updated: December 6 2005. Date last accessed: March 20 2006.

20. Carlson CS, Eberle MA, Rieder MJ, Yi Q, Kruglyak L, Nickerson DA. Selecting a maximally informative set of single-nucleotide polymorphisms for association analyses using linkage disequilibrium. *Am J Hum Genet* 2004;74:106-120.

21. Ranade K, Chang MS, Ting CT, Pei D, Hsiao CF, Olivier M, Pesich R, Hebert J, Chen YD, Dzau VJ, Curb D, Olshen R, Risch N, Cox DR, Botstein D. High-throughput genotyping with single nucleotide polymorphisms. *Genome Res* 2001;11:1262-1268.

22. R Development Core Team. R: A language and environment for statistical computing. Vienna, Austria, 2005. (http://www.R-project.org)

23. Burkett K, McNeney B, Graham J. A note on inference of trait associations with SNP haplotypes and other attributes in generalized linear models. *Hum Hered* 2004;57:200-206.

24. Cohen J. Statistical Power Analysis for the Behavioral Sciences. Lawrence Erlbaum Associates, New Jersey, 1988.

25. Weiss ST, Silverman EK, Palmer LJ. Case-control association studies in pharmacogenetics. *Pharmacogenomics J* 2001;1:157-158.

26. Churg A, Wang RD, Tai H, Wang X, Xie C, Wright JL. Tumor necrosis factor-alpha drives 70% of cigarette smoke-induced emphysema in the mouse. *Am J Respir Crit Care Med* 2004;170:492-498.

27. Neale BM, Sham PC. The future of association studies: gene-based analysis and replication. *Am J Hum Genet* 2004;75:353-362.

28. Sandford AJ, Chagani T, Weir TD, Connett JE, Anthonisen NR, Pare PD. Susceptibility genes for rapid decline of lung function in the lung health study. *Am J Respir Crit Care Med* 2001;163:469-473.

29. Knight JC, Keating BJ, Rockett KA, Kwiatkowski DP. In vivo characterization of regulatory polymorphisms by allele-specific quantification of RNA polymerase loading. *Nat Genet* 2003;33:469-475.

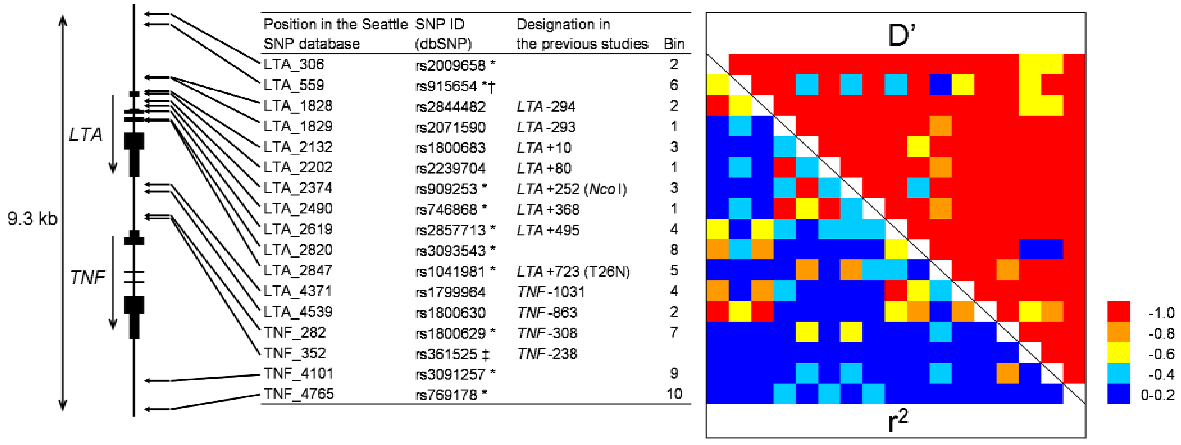
30. Lander ES, Schork NJ. Genetic dissection of complex traits. *Science* 1994;265:2037-2048.

Figure Legends

Figure 1

SNPs whose allele frequency was more than 10% in European American Coriell samples are shown.

- Tag SNPs, † rs915654 was excluded from the study, ‡ rs361525 (*TNF*-238) whose allele frequency was less than 10 % was included in the study (see methods section).



Position in the Seattle SNP database	SNP ID (dbSNP)	Designation in the previous studies	Bin
LTA_306	rs2009658 *		2
LTA_559	rs915654 *†		6
LTA_1828	rs2844482	*LTA* -294	2
LTA_1829	rs2071590	*LTA* -293	1
LTA_2132	rs1800683	*LTA* +10	3
LTA_2202	rs2239704	*LTA* +80	1
LTA_2374	rs909253 *	*LTA* +252 (*Nco* I)	3
LTA_2490	rs746868 *	*LTA* +368	1
LTA_2619	rs2857713 *	*LTA* +495	4
LTA_2820	rs3093543 *		8
LTA_2847	rs1041981 *	*LTA* +723 (T26N)	5
LTA_4371	rs1799964	*TNF* -1031	4
LTA_4539	rs1800630	*TNF* -863	2
TNF_282	rs1800629 *	*TNF* -308	7
TNF_352	rs361525 ‡	*TNF* -238	
TNF_4101	rs3091257 *		9
TNF_4765	rs769178 *		10

Figure 2

a) Power curves for the rate of decline dataset sample sizes of 279 fast decliners and 304 non decliners over a range of odds ratios (OR) and a range of proportion of controls having

susceptibility genotype. b) Power curves for the baseline dataset sample sizes of 532 high baseline and 533 low baseline over a range of odds ratios (OR) and a range of proportion of controls having susceptibility genotype.

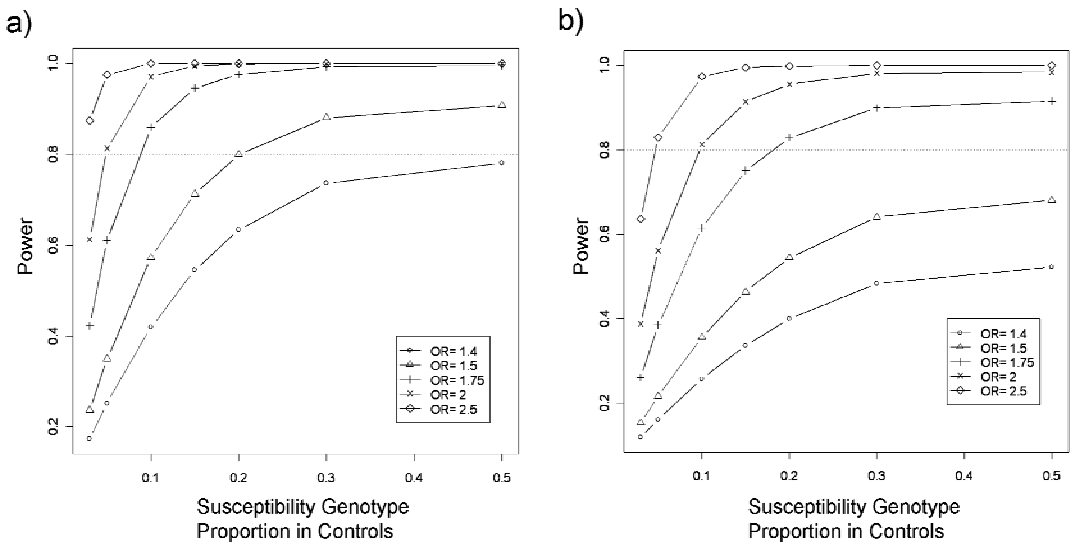


Table 1: Characteristics of subjects in the rate of lung function decline study

	Fast decline (n=279)	Non decline (n=304)	p value
Male/Female	163/117	203/101	0.04
Age (years)	49.5 ± 0.4	47.6 ± 0.4	0.0005
Smoking history (pack-yrs)	43.0 ± 1.1	38.3 ± 1.0	0.003
Smoking amount (cigarettes/d)	25.6 ± 0.6	22.4 ± 0.6	0.0003
ΔFEV1, % predicted/yr*	-4.13 ± 0.06	1.08 ± 0.04	<0.0001
ΔFEV1, ml/yr	-153.5 ± 2.6	15.1 ± 1.5	<0.0001
Baseline FEV1, % predicted†	72.5 ± 0.5	75.7 ± 0.5	<0.0001

Data are presented as means ± SE

* Change in lung function over 5 year period per year, † Lung function at the start of the Lung Health Study

Table 2: Characteristics of subjects in the baseline lung function study

	Low function (n = 533)	High function (n = 532)	p value
Male/Female	329/204	353/179	0.116
Age (years)	50.7 ± 0.3	46.2 ± 0.3	<0.0001
Smoking history (pack-yrs)	45.3 ± 0.8	35.3 ± 0.8	<0.0001
ΔFEV1, % predicted/yr*	-1.28 ± 0.08	-0.54 ± 0.07	<0.0001
Baseline FEV1, % predicted†	62.6 ± 0.1	91.8 ± 0.1	<0.0001

* Change in lung function over 5 year period per year, † Lung function at the start of the Lung Health Study as measured FEV1% predicted post bronchodilator

Table 3: Genotypic distribution of the single SNPs of the *LTA* and *TNF* genes in the rate of lung function decline study

				Unadjusted		Adjusted *	
SNP	Genotype	Fast decline (%)	Non decline (%)	OR (95% CI)	P-value	OR (95% CI)	P-value
rs2009658	C/C	204 (73.4)	210 (69.3)	1	0.51	1	0.84
	C/G	68 (24.5)	87 (28.7)	0.80 (0.55-1.19)		0.90 (0.61-1.33)	
	G/G	6 (2.2)	6 (2.0)	1.03 (0.27-3.92)		0.83 (0.25-2.77)	
rs909253	A/A	113 (41.7)	128 (43.1)	1	0.35	1	0.5
(*LTA*+252)	A/G	133 (49.1)	132 (44.4)	1.14 (0.79-1.64)		1.07 (0.74-1.55)	
	G/G	25 (9.2)	37 (12.5)	0.77 (0.41-1.40)		0.75 (0.41-1.37)	
rs746868	G/G	98 (35.1)	125 (41.4)	1	0.17	1	0.24
(*LTA*+368)	G/C	139 (49.8)	127 (42.1)	1.40 (0.96-2.03)		1.37 (0.94-2.00)	
	C/C	42 (15.1)	50 (16.6)	1.07 (0.64-1.80)		1.08 (0.64-1.80)	
rs2857713	T/T	153 (55.0)	159 (52.5)	1	0.83	1	0.93
(*LTA*+495)	T/C	105 (37.8)	121 (39.9)	0.90 (0.63-1.29)		0.98 (0.68-1.41)	
	C/C	20 (7.2)	23 (7.6)	0.90 (0.45-1.80)		0.88 (0.45-1.74)	
rs3093543	A/A	241 (86.4)	269 (89.4)	1	0.33	1	0.17
	A/C	38 (13.6)	32 (10.6)	1.32 (0.78-2.27)		1.45 (0.85-2.48)	
	C/C	0 (0.0)	0 (0.0)	-		-	
rs1041981	C/C	116 (41.9)	129 (42.6)	1	0.57	1	0.73
(*LTA*+723)	C/A	134 (48.4)	137 (45.2)	1.09 (0.76-1.56)		1.02 (0.71-1.47)	
	A/A	27 (9.7)	37 (12.2)	0.81 (0.45-1.47)		0.81 (0.45-1.45)	
rs1800629	G/G	185 (67.0)	201 (66.6)	1	0.97	1	0.97
(*TNF*-308)	G/A	82 (29.7)	92 (30.5)	0.97 (0.67-1.41)		0.96 (0.65-1.40)	
	A/A	9 (3.3)	9 (3.0)	1.09 (0.37-3.16)		0.96 (0.35-2.65)	
rs361525	G/G	239 (86.0)	268 (88.2)	1	0.51	1	0.72
(*TNF*-238)	G/A	39 (14.0)	36 (11.8)	1.21 (0.73-2.04)		1.10 (0.66-1.83)	
	A/A	0 (0.0)	0 (0.0)	-		-	
rs3091257	G/G	212 (76.8)	228 (75.0)	1	0.73	1	0.63
	G/A	59 (21.4)	72 (23.7)	0.88 (0.58-1.33)		0.89 (0.59-1.36)	
	A/A	5 (1.8)	4 (1.3)	1.34 (0.29-6.86)		1.71 (0.43-6.79)	
rs769178	G/G	228 (83.2)	254 (84.9)	1	0.34	1	0.37
	G/T	42 (15.3)	44 (14.7)	1.06 (0.65-1.73)		1.18 (0.73-1.92)	
	T/T	4 (1.5)	1 (0.3)	4.44 (0.44-220.06)		3.70 (0.40-34.41)	

* Multivariate logistic regression was used to adjust for age, sex, pack-years of smoking, a quadratic term for pack-years and research center as potential risk and confounding factors.

Table 4: Genotypic distribution of the single SNPs of the *LTA* and *TNF* genes in the baseline lung function study

SNP	Genotype	Low function (%)	High function (%)	Unadjusted		Adjusted *	
				OR (95% CI)	P-value	OR (95% CI)	P-value
rs2009658	C/C	346 (69.8)	356 (69.3)	1	0.62	1	0.71
	C/G	137 (27.6)	139 (27.0)	1.01 (0.76-1.35)		1.04 (0.76-1.43)	
	G/G	13 (2.6)	19 (3.7)	0.70 (0.31-1.53)		0.74 (0.33-1.64)	
rs909253	A/A	241 (47.0)	241 (46.3)	1	0.97	1	0.91
(*LTA*+252)	A/G	214 (41.7)	221 (42.4)	0.97 (0.74-1.27)		0.95 (0.71-1.26)	
	G/G	58 (11.3)	59 (11.3)	0.98 (0.64-1.50)		1.01 (0.65-1.58)	
rs746868	G/G	176 (35.0)	188 (36.6)	1	0.86	1	0.82
(*LTA*+368)	G/C	248 (49.3)	246 (47.9)	1.08 (0.81-1.43)		1.10 (0.81-1.49)	
	C/C	79 (15.7)	80 (15.6)	1.05 (0.71-1.56)		1.08 (0.71-1.63)	
rs2857713	T/T	272 (53.8)	270 (52.2)	1	0.21	1	0.34
(*LTA*+495)	T/C	185 (36.6)	210 (40.6)	0.87 (0.67-1.14)		0.86 (0.64-1.15)	
	C/C	49 (9.7)	37 (7.2)	1.31 (0.81-2.14)		1.24 (0.75-2.05)	
rs3093543	A/A	456 (87.4)	462 (88.0)	1	0.88	1	0.83
	A/C	61 (11.7)	57 (10.9)	1.08 (0.73-1.62)		1.13 (0.74-1.74)	
	C/C	5 (1.0)	6 (1.1)	0.84 (0.20-3.35)		1.19 (0.31-4.51)	
rs1041981	C/C	245 (47.4)	241 (46.3)	1	0.94	1	0.87
(*LTA*+723)	C/A	215 (41.6)	221 (42.5)	0.96 (0.73-1.25)		0.93 (0.70-1.25)	
	A/A	57 (11.0)	58 (11.2)	0.97 (0.63-1.48)		1.03 (0.66-1.61)	
rs1800629	G/G	353 (68.5)	370 (70.6)	1	0.49	1	0.59
(*TNF*-308)	G/A	151 (29.3)	139 (26.5)	1.14 (0.86-1.51)		1.13 (0.84-1.53)	
	A/A	11 (2.1)	15 (2.9)	0.77 (0.31-1.82)		0.77 (0.33-1.84)	
rs361525	G/G	460 (87.6)	471 (89.7)	1	0.52	1	0.55
(*TNF*-238)	G/A	63 (12.0)	53 (10.1)	1.22 (0.81-1.83)		1.07 (0.69-1.65)	
	A/A	2 (0.4)	1 (0.2)	2.05 (0.11-120.98)		4.36 (0.28-69.07)	
rs3091257	G/G	394 (77.1)	425 (81.4)	1	0.11	1	0.28
	G/A	112 (21.9)	89 (17.0)	1.36 (0.98-1.88)		1.31 (0.93-1.84)	
	A/A	5 (1.0)	8 (1.5)	0.67 (0.17-2.36)		0.80 (0.24-2.65)	
rs769178	G/G	412 (83.6)	416 (82.1)	1	0.66	1	0.99
	G/T	76 (15.4)	83 (16.4)	0.92 (0.65-1.32)		0.97 (0.66-1.43)	
	T/T	5 (1.0)	8 (1.6)	0.63 (0.16-2.21)		0.94 (0.29-3.06)	

* Multivariate logistic regression was used to adjust for age, sex, pack-years of smoking, a quadratic term for pack-years and research center as potential risk and confounding factors.

Table 5: Haplotype analysis of the *LTA* and *TNF* genes in the rate of lung function decline study

	Haplotype frequency		Unadjusted			Adjusted		
Haplotype	Fast decline	Non decline	Wald test chi-square statistic	df	P-value	Wald test chi-square statistic	df	P-value
LTA								
CAGCAC	0.118	0.112	5.29	4	0.26	5.05	4	0.28
CACTAC	0.400	0.377						
CGGTAA	0.339	0.349						
GAGCAC	0.076	0.109						
GAGCCC	0.068	0.054						
TNF								
GGGG	0.659	0.683	1.78	4	0.78	1.53	4	0.82
GGGT	0.091	0.077						
GAGG	0.069	0.058						
AGGG	0.055	0.049						
AGAG	0.125	0.132						

Haplotype frequencies were estimated with an EM algorithm implemented in haplo.stats.

Table 6: Haplotype analysis of the *LTA* and *TNF* genes in the baseline lung function study

	Haplotype frequency		Unadjusted			Adjusted		
Haplotype	Low function	High function	Wald test chi-square statistic	df	P-value	Wald test chi-square statistic	df	P-value
LTA								
CAGCAC	0.115	0.104	1.14	4	0.89	0.96	4	0.92
CACTAC	0.404	0.397						
CGGTAA	0.318	0.324						
GAGCAC	0.096	0.106						
GAGCCC	0.068	0.066						
TNF								
GGGG	0.678	0.691	3.98	4	0.41	2.19	4	0.70
GGGT	0.088	0.094						
GAGG	0.064	0.050						
AGGG	0.053	0.062						
AGAG	0.118	0.099						

Haplotype frequencies were estimated with an EM algorithm implemented in haplo.stats.