

Single-nucleotide polymorphisms in SMAD3 are associated with chronic obstructive pulmonary disease

Ting Yang^{1,2}, Binwu Ying³, Xingbo Song³, Sizhong Zhang⁴, Hong Fan^{1,2,3}, Dan Xu^{1,2}, Tao Wang^{1,2}, Daishun Liu^{1,2} and Fuqiang Wen^{1,2}

¹Division of Pulmonary Disease, State Key Laboratory of Biotherapy; ²Department of Respiratory Medicine; ³Department of Laboratory Medicine; ⁴Department of Medical Genetics, West China Hospital, Sichuan University, Chengdu, Sichuan 610041, People's Republic of China

Corresponding author: Fuqiang Wen. Email: wenfuqiang@126.com

Abstract

The aim of this study was to determine the frequencies of single-nucleotide polymorphisms (SNPs) in the Smad3 gene in the Chinese population and their possible association with chronic obstructive pulmonary disease (COPD). The frequency distribution of nine SNPs in the introns of the Smad3 gene was determined in both the COPD patients and control groups by the TaqMan polymerase chain reaction method using a minor groove binder probe. The genotype distribution of the rs28683050 polymorphism in control subjects was as follows: CC, 67.6%; CT, 27.0% and TT, 5.4%; and in COPD patients was as follows: CC, 47.0%; CT, 38.8% and TT, 14.2%. Therefore, a significant difference in allelic and genotypic frequencies between the COPD patients and control subjects is observed. Their genotypic distribution was not significantly different from that predicted by Hardy-Weinberg's equilibrium. The frequency of the TT genotype in the COPD patients was significantly higher than that in controls (14.2% versus 5.4%, odds ratio [OR] = 3.762, 95% confidence interval: 1.650–8.581, $P = 0.002$), and the frequency of the T allele in the COPD patients was significantly higher than that in controls (81.1% versus 66.4%, OR = 1.774, 95% confidence interval: 1.354–2.324, $P = 0.001$). These data show that the rs28683050 polymorphism of the Smad3 gene may be associated with COPD.

Keywords: allele, chronic obstructive pulmonary disease, single-nucleotide polymorphisms, Smad3, TGF- β , genetics

Experimental Biology and Medicine 2010; 235: 599–605. DOI: 10.1258/ebm.2010.009268

Introduction

Cigarette smoking is the major risk factor for chronic obstructive pulmonary disease (COPD); however, only approximately 15% of smokers develop clinically relevant airflow obstruction.¹ Multiple studies in diverse populations have shown a large genetic contribution to the variability in pulmonary function and for the familial aggregation of COPD patients.^{2,3} As expected, segregation analysis suggests that multiple genes may be involved. At present, however, only a single gene, α 1-antitrypsin, a potent inhibitor of inflammatory cell protease in the lung, has been unequivocally implicated in the development of COPD.⁴ Recently, the associations between COPD and polymorphisms in several other genes have been studied, which include α 1-antichymotrypsin,⁵ microsomal epoxide hydrolase,⁶ vitamin D-binding protein,⁷ tumor necrosis factor- α ,⁸ SERPINE2⁹ and a-nicotinic acetylcholine receptor (CHRNA 3/5).^{10,11}

Transforming growth factor- β (TGF- β) plays an important role in the pathogenesis of COPD. Signals from the

activated TGF- β receptor complex are transduced to the nucleus of airway cells by SMAD proteins, which represent a family of transcription factors that have recently been implicated to play a major role as intracellular mediators of inflammation. The Smad family consists of the receptor-regulated Smads, a common pathway Smad, and inhibitory Smads. Receptor-regulated Smads are phosphorylated by the TGF- β type I receptor. They include Smad2 and Smad3, which are recognized by TGF- β and activin receptors, and Smads 1, 5, 8 and 9, which are recognized by bone morphogenetic protein receptors. Smad4 is a common pathway of Smad, which is also defined as co-operating Smad and is not phosphorylated by the TGF- β type I receptor. Inhibitory Smads (anti-Smads) include Smad6 and Smad7, which downregulate TGF- β signaling. To date, the Smads are the only TGF- β receptor substrates with a demonstrated ability to propagate signals, and with regard to the growing number of investigations of Smad-mediated effects in the airways, Smads may prove

to be an important target for future development of new therapeutic strategies for asthma and COPD.

The human Smad3 gene has been assigned to chromosome 15q21–22 and spans over 5.7 kb with nine exons and eight introns.^{12–14} So far, data on the association between DNA polymorphisms in the Smad3 gene and COPD have not been reported. The current study was to screen for DNA sequence variants in the Smad3 gene in Chinese Hans – which account for 95% of the Chinese population – and to determine the association between the polymorphisms and COPD (Table 1).

To examine the interaction between a Smad3 gene polymorphism and COPD, we carried out a case-control study.

Materials and methods

Subjects

Two hundred and nineteen patients with COPD were recruited from West China Hospital of Sichuan University. The definition of COPD was consistent with that in the American Thoracic Society (ATS) consensus statement. The patients had a history of chronic or recurrent productive cough for >2 y and decreased maximum expiratory flow that had been slowly progressive and irreversible. The presence of other lung or cardiac diseases as the cause of patient symptoms was excluded by clinical and radiographic examinations. The criteria of enrollment were as follows: (1) individuals with a forced expiratory volume in 1 s (FEV1) <70% of predicted, an FEV1/forced vital capacity (FVC) ratio of 70% and an increase in FEV1 of <12% 15 min after the inhalation of 400 μ g Fenoterol HBr MDI (Berotec; Boehringer Ingelheim, Ridgefield, CT, USA); and (2) patient consent to participate in the study. Most of the patients were receiving oral methylxanthine, inhaled anticholinergic agents and inhaled β 2-agonists as needed. The patient's name, age, sex, family history, smoking habits, the number of cigarettes smoked, the duration of diseases and chest radiographic findings were recorded. Pulmonary function testing (CHESTAC-33-8800, Tokyo, Japan) was performed according to the ATS performance requirements. One hundred and forty-eight unrelated, age-matched healthy subjects, who had no known medical illness or family disorders and were taking no medications, acted as control subjects. This study was approved by the ethics committee of West China

Hospital, Sichuan University, and signed informed consent forms were obtained from all subjects.

Candidate single-nucleotide polymorphism selection and polymorphism genotyping

The single-nucleotide polymorphisms (SNPs) of Smad3 used for the test are shown in Figure 1. Their chromosomal location are illustrated in Figure 1. SNPs of Smad3 were extracted from the database of SNP (<http://www.ncbi.nlm.nih.gov/SNP/index.html>) (Figure 1).

Genotyping method

Genomic DNA was extracted from the buffy coat using a QIAamp DNA Blood Kit (Qiagen Inc, Stanford, CA, USA). Analysis of genetics polymorphisms SNP genotyping was performed by TaqMan allelic discrimination assay.¹⁵ Reagents were purchased from Applied Biosystems (Foster City, CA, USA). TaqMan probes were designed and synthesized by Applied Biosystems, and distinguished the SNPs at the end of a polymerase chain reaction (PCR). One allelic probe was labeled with the fluorescent FAM dye and the other with the fluorescent VIC dye. PCR was performed by TaqMan Universal Master Mix without UNG (Applied Biosystems) with PCR primers at concentrations of 200 nmol/L. Reactions were performed in 384-well formats in a total reaction volume of 3 μ L using 3.0 ng of genomic DNA.

The PCR amplification protocol for the TaqMan assays included denaturation at 95°C for 10 min, followed by 40 cycles at 92°C for 15 s, 60°C for 1 min and 72°C for 45 s, followed by elongation at 72°C for 5 min. The TaqMan assays were then read on a 7900HT Fast Real-Time PCR System and alleles were called using the SDS software (Applied Biosystems).

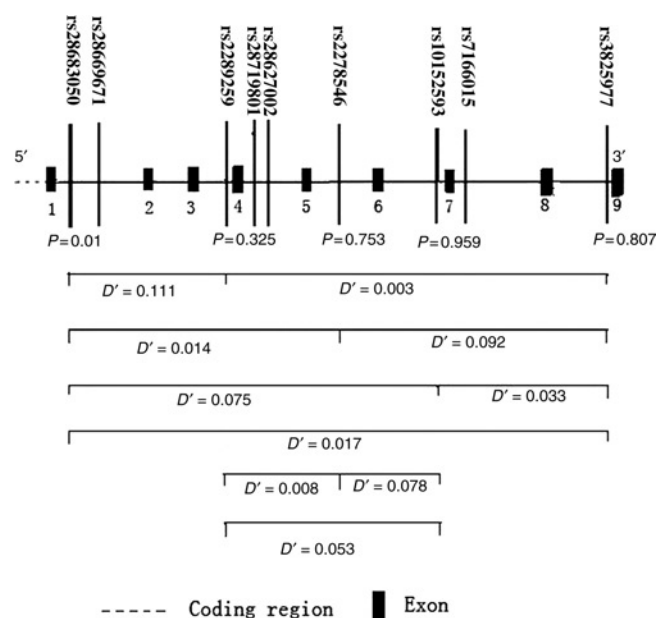


Figure 1 A schematic overview of the nine investigated single-nucleotide polymorphisms in the Smad3 gene

Table 1 Nine SNPs in the Smad3 gene in the study

Reference SNP ID	Public location position	Region in gene	A allele	B allele
rs28683050	20,142	Intron 1	C	T
rs28669671	20,438	Intron 1	C	T
rs2289259	99,656	Intron 3	T	C
rs28719801	101,163	Intron 4	A	G
rs28627002	102,928	Intron 4	C	T
rs2278546	108,408	Intron 5	A	G
rs10152593	117,241	Intron 6	C	T
rs7166015	119,931	Intron 7	C	T
rs3825977	123,054	Intron 8	C	T

SNP, single-nucleotide polymorphism

Statistical analysis

Hardy–Weinberg analyses were performed using the Hardy–Weinberg equilibrium (HWE) program of the LINKUTIL software package.¹⁶ Data analyses were performed with the Statistical Package for the Social Science 13.0 (SPSS Inc., Chicago, IL, USA), and the significance level for statistical tests was taken to be 0.05. The two-sided Student's *t*-test was used for checking significant differences in clinical data between the COPD patients and the control subjects. Differences between the patients with COPD and the controls with respect to the allele frequencies and genotype distributions were analyzed by the χ^2 test or the Fisher's exact test when necessary. Haplotype frequencies for pairs of alleles, as well as χ^2 values for allele associations, were estimated by SHEsis software.¹⁷ Linkage disequilibrium (LD) coefficients $D' = D/D_{\max}$ were calculated by SHEsis software,¹⁷ and $D' > 0.70$ was taken to represent a significant difference.

Results

General characteristics

The study population consisted of 219 patients with COPD and 148 control subjects; its characteristics are described in Table 2. All the individuals included in the study were from southwestern China. The COPD cases and control subjects did not significantly differ in sex, age and smoking history characteristics. The parameters used for FEV1 and FEV1/FVC were significantly decreased in the COPD case subjects compared with the controls ($P < 0.01$).

Distribution of the nine SNPs of the Smad3 gene in COPD patients and controls

To determine the prevalence of the nine SNPs of the Smad3 gene, we screened them in all 219 COPD patients and in the 148 controls. Genotype and allele frequencies by case-control status are given in Table 3. There was no deviation from HWE in controls and COPD patients (see Table 3).

In our study, four known SNPs in the Smad3 gene, including rs28719801, rs28627002, rs28669671 and rs7166015, were not found in the COPD patients and control subjects, indicating that they are rare in the Chinese Han population.

Allelic or genotypic frequency distributions of rs3825977, rs2278546, rs2289259 and rs10152593 polymorphism in Smad3 gene between the COPD patients and controls are shown in Table 3. No significant difference of allelic or

genotypic frequencies of rs3825977, rs2278546, rs2289259 and rs10152593 polymorphism in the Smad3 gene between the COPD patients and control subjects were observed.

Rs28683050 has a significant difference in allelic or genotypic frequencies between the COPD patients and control subjects.

The genotype frequency of CC, CT and TT was 67.6%, 27.0% and 5.4% in controls, respectively, and 7.0%, 38.8% and 14.2% in COPD patients, respectively ($\chi^2 = 57.38$, $P = 0.01$). The allele frequency of T and C was 81.1% and 18.9% in controls and 66.4% and 33.6% in COPD patients. ($\chi^2 = 18.93$, $P < 0.05$).

The frequency of the TT genotype in the COPD patients was significantly higher than that in controls (14.2% versus 5.4%, odds ratio [OR] = 3.762, 95% confidence interval [CI]: 1.650–8.581, $P = 0.002$). The frequency of the T allele in the COPD patients was significantly higher than that in controls (81.1% versus 66.4%, OR = 1.774, 95% CI: 1.354–2.324, $P = 0.001$).

Relationship between cigarette smoking and the distribution of the Smad3 gene polymorphisms

The allele and genotype distributions were compared in smokers (COPD patients and controls) to determine whether the prevalence of different alleles or genotypes was associated with smoking. Results showed that no significant difference of allelic or genotypic frequencies of the nine SNPs polymorphism in the Smad3 gene between the COPD patients and control subjects were observed (Table 4).

Dominant and recessive genetic models

We applied both dominant and recessive genetic models in these statistical analyses. Our analyses of Fisher's exact test showed that all SNPs in this study in the Smad3 gene were not significantly associated with COPD in the recessive model and the dominant model (Table 5).

LD in nine SNPs of the Smad3 gene

Haplotype analysis, testing associations using several polymorphisms, sometimes demonstrates genetic influences that are not detected by the analysis of single polymorphisms. The relation in the nine SNPs of the Smad3 gene and their effects on COPD were analyzed. The extent of *D* in pairwise combinations of alleles in different variation sites was estimated by means of maximum likelihood from the frequency of diploid genotypes in the COPD and control groups. Haplotype frequencies and the coefficient of LD (*D'*) are given in Table 6 and Figure 1. There was no strong disequilibrium in these SNPs of the Smad3 gene.

We used estimated haplotype frequencies higher than 0.01 to construct phased multilocus genotypes for the Smad3 gene (Table 7). In Table 7 we found significant difference in some haplotypes of the Smad3 gene, e.g. ccaccaac ($P < 0.05$, OR = 0.249 [95% CI = 0.132–0.469]), ccaccagt ($P < 0.05$, OR = 2.768 [95% CI = 1.234–6.209]), ccactcaat ($P < 0.05$, OR = 0.331 [95% CI = 0.131–0.838]), ccattcagc ($P < 0.05$, OR = 0.419 [95% CI = 0.213–0.826]) and ccattcaat ($P < 0.05$, OR = 0.417 [95% CI = 0.229–0.759]).

Table 2 Description of study population

Item	Controls (n = 148)	Cases (n = 219)	P value
Age in years, median (range)	69 (50–80)	70 (50–90)	>0.1
Male, n (%)	109 (73.6)	158 (72.1)	>0.1
Smoking history			
0–20 pack years	40	56	>0.05
≥ 20 pack years	108	163	>0.05
FEV1/predicted (%)	93.7 ± 3.4	49.0 ± 0.4	<0.01
FEV1/FVC	78.0 ± 4.6	59.2 ± 8.3	<0.01

FEV1, forced expiratory volume in 1 s; FVC, forced vital capacity

Table 3 Frequency distributions of the Smad3 gene in patients with COPD versus controls

Polymorphism	Genotype	Control, n (%)	COPD, n (%)	χ^2	Allele	Control, n (%)	COPD, n (%)	χ^2	HWE
rs10152593	CC	54 (36.5)	78 (35.6)	0.084	C	174 (58.8)	257 (58.1)	0.001	$\chi^2_1 = 0.94$
	TT	28 (18.9)	40 (18.2)						$P_1 = 0.33$
	TC	66 (44.6)	101 (46.1)	$P = 0.959$	T	122 (41.2)	181 (41.9)	$P > 0.05$	$\chi^2_2 = 0.53$
rs2289259	TT	20 (13.5)	27 (12.3)	2.249	T	106 (35.8)	137 (31.3)	1.638	$P_2 = 0.47$
	TC	66 (44.6)	83 (37.9)						$\chi^2_1 = 0.13$
	CC	62 (41.9)	109 (49.8)	$P = 0.325$	C	190 (64.2)	301 (68.7)	$P > 0.05$	$P_1 = 0.72$
rs2278546	AA	75 (50.6)	103 (47.0)	0.568	A	205 (69.3)	291 (65.8)	0.679	$\chi^2_2 = 3.10$
	AG	55 (37.2)	85 (38.8)						$P_2 = 0.08$
	GG	18 (12.2)	31 (14.2)	$P = 0.753$	G	91 (31.7)	147 (34.2)	$P > 0.05$	$\chi^2_1 = 2.40$
rs3825977	CC	20 (13.5)	35 (16.0)	0.43	C	122 (41.2)	187 (42.7)	0.158	$P_1 = 0.12$
	CT	82 (55.4)	117 (53.4)						$\chi^2_2 = 3.70$
	TT	46 (31.1)	67 (30.6)	$P = 0.807$	T	174 (58.8)	251 (57.3)	$P > 0.05$	$P_2 = 0.06$
rs28719801	AA	148 (100)	219 (100)		A	296 (100)	438 (100)		$\chi^2_1 = 3.04$
	AG	0 (0)	0 (0)	ND				ND	$P_1 = 0.08$
	GG	0 (0)	0 (0)		G	0 (0)	0 (0)	ND	$\chi^2_2 = 1.85$
rs28627002	TT	148 (100)	219 (100)		T	296 (100)	438 (100)		$P_2 = 0.17$
	CT	0 (0)	0 (0)	ND				ND	
	CC	0 (0)	0 (0)		C	0 (0)	0 (0)	ND	
rs28669671	TT	148 (100)	219 (100)		T	296 (100)	438 (100)		
	CT	0 (0)	0 (0)	ND				ND	
	CC	0 (0)	0 (0)		C	0 (0)	0 (0)	ND	
rs7166015	TT	148 (100)	219 (100)		C	296 (100)	438 (100)		
	CT	0 (0)	0 (0)	ND				ND	
	CC	0 (0)	0 (0)		T	0 (0)	0 (0)	ND	
rs28683050	CC	100 (67.6)	103 (47.0)	57.38	T	240 (81.1)	291 (66.4)	18.93	$\chi^2_1 = 2.10$
	CT	40 (27.0)	85 (38.8)						$P_1 = 0.15$
	TT	8 (5.4)	31 (14.2)	$P = 0.01$	C	56 (18.9)	147 (33.6)	$P < 0.05$	$\chi^2_2 = 3.68$
									$P_2 = 0.06$

COPD, chronic obstructive pulmonary disease; HWE, Hardy–Weinberg equilibrium; ND, not determined
 χ^2_1 is the HWE χ^2 of the controls, χ^2_2 is the HWE χ^2 of the COPD patients

Table 4 Frequency distributions of the Smad3 gene in smoked COPD versus smoked controls

Polymorphism	Genotype	Control, n (%)	COPD, n (%)	χ^2	Allele	Control, n (%)	COPD, n (%)
rs10152593	CC	19 (17.6)	30 (18.4)	2.551	C	84 (38.9)	143 (43.9)
	CT	46 (42.6)	83 (50.9)				
	TT	43 (39.8)	50 (30.7)	$P > 0.05$	T	132 (61.1)	183 (56.1)
rs2289259	TT	15 (13.9)	23 (14.1)	2.857	T	79 (36.6)	104 (31.9)
	TC	49 (45.4)	58 (35.6)				
	CC	44 (40.7)	82 (50.3)	$P > 0.05$	C	137 (63.4)	222 (68.1)
rs2278546	AA	53 (49.0)	81 (49.7)	0.568	A	147 (68.1)	221 (67.8)
	AG	41 (38.0)	59 (36.2)				
	GG	14 (13.0)	23 (14.1)	$P > 0.05$	G	69 (31.9)	105 (32.2)
rs3825977	CC	14 (13.0)	27 (16.6)	0.43	T	80 (37.0)	139 (42.6)
	CT	52 (48.1)	85 (52.1)				
	TT	42 (38.9)	51 (31.3)	$P > 0.05$	C	136 (63.0)	187 (57.4)
rs28719801	GG	108 (100)	163 (100)		A	216 (100)	326 (100)
	AG	0 (0)	0 (0)	ND			
	AA	0 (0)	0 (0)		G	0 (0)	0 (0)
rs28627002	TT	108 (100)	163 (100)		T	216 (100)	326 (100)
	CT	0 (0)	0 (0)	ND			
	CC	0 (0)	0 (0)		C	0 (0)	0 (0)
rs28669671	TT	108 (100)	163 (100)		T	216 (100)	326 (100)
	CT	0 (0)	0 (0)	ND			
	CC	0 (0)	0 (0)		C	0 (0)	0 (0)
rs7166015	CC	108 (100)	163 (100)		C	216 (100)	326 (100)
	CT	0 (0)	0 (0)	ND			
	TT	0 (0)	0 (0)		T	0 (0)	0 (0)
rs28683050	CC	72 (66.7)	101 (62.0)	1.691	C	173 (80.09)	246 (75.46)
	CT	29 (26.9)	44 (27.0)				
	TT	7 (6.5)	18 (11.0)	$P > 0.05$	T	43 (19.91)	80 (24.54)

COPD, chronic obstructive pulmonary disease; ND, not determined

Table 5 Fisher's exact test in each SNP examined in this study

SNP ID	Symbol	Genetic model	Genotype	Phenotype		Fisher's exact test	
				COPD	Control	OR (95% CI)	P value
rs10152593		Dominant	CC + CT	141	94	1.038 (0.673–1.603)	0.865
			TT	78	54		
			CC	40	28		
rs2289259		Recessive	CT + TT	179	120	0.958 (0.561–1.636)	0.874
			CC + CT	192	128		
			TT	27	20		
rs2278546		Dominant	CC	109	62	0.728 (0.478–1.108)	0.138
			CT + TT	110	86		
			AG + GG	116	73		
rs28719801		Recessive	AA	103	75	0.864 (0.569–1.312)	0.493
			GG	31	18		
			AG + AA	188	130		
rs28627002		Dominant	GG	219	148	0.840 (0.451–1.565)	0.582
			AG + AA	0	0		
			AA	0	0		
rs28669671		Recessive	AG + GG	219	148	ND	ND
			CT + CC	0	0		
			TT	219	148		
rs7166015		Dominant	CC	0	0	ND	ND
			CT + TT	219	148		
			CT + CC	0	0		
rs3825977		Recessive	TT	219	148	ND	ND
			CT + TT	219	148		
			AG + GG	0	0		
rs28683050		Dominant	AA	219	148	ND	ND
			GG	0	0		
			AG + AA	219	148		
rs2278546		Recessive	CT + TT	184	128	1.217 (0.672–2.205)	0.516
			CC	35	20		
			TT	67	46		
rs2289259		Dominant	CT + CC	152	102	1.023 (0.651–1.607)	0.921
			CT + CC	145	101		
			TT	18	7		
rs2278546		Recessive	CC	101	72	0.558 (0.225–1.386)	0.204
			CT + TT	62	36		
			CT + TT	62	36		

SNP, single-nucleotide polymorphism; COPD, chronic obstructive pulmonary disease; OR, odds ratio; CI, confidence interval; ND, not determined

These results suggest that these haplotypes may be associated with COPD.

Discussion

In COPD, a marked fibrosis of the airways is observed, which is partly mediated by TGF- β ¹⁸ that has been proposed to be a potent fibrogenic factor in numerous diseases.¹⁹ The effects of the extensively studied TGF- β 1 include fibroblast proliferation and an increased production and deposition of extracellular matrix proteins. Enzymes involved in extracellular

matrix degradation and destruction can also be stimulated by TGF- β 1.^{19–21} While much interest has focused on the TGF- β 1-induced effects in COPD, little attention has been paid to the intracellular signal transduction pathways. The pathways of the TGF- β family include the family of transcription factors termed SMADs that are directly activated by the receptors. The most predominant member of the TGF superfamily in COPD, TGF- β 1, mediates its effects by the regulatory SMAD2 and SMAD3 proteins, which are phosphorylated by the type I receptor after binding of TGF- β 1. The phosphorylated SMADs bind the common SMAD4 and translocate into the nucleus. In contrast, the inhibitory SMAD6

Table 6 Disequilibrium statistics in nine SNPs of the Smad3 gene

D'	rs28669671	rs7166015	rs3825977	rs2289259	rs28627002	rs28719801	rs2278546	rs10152593
rs28683050	0	0	0.017	0.111	0	0	0.014	0.075
rs28669671	–	0	0	0	0	0	0	0
rs7166015	–	–	0	0	0	0	0	0
rs3825977	–	–	–	0.003	0	0	0.092	0.033
rs2289259	–	–	–	–	0	0	0.008	0.053
rs28627002	–	–	–	–	–	0	0	0
rs28719801	–	–	–	–	–	–	0	0
rs2278546	–	–	–	–	–	–	–	0.078

SNP, single-nucleotide polymorphism

Table 7 Estimate of pairwise haplotype frequencies in nine SNPs of the Smad3 gene

Haplotype	Case (freq.)	Control (freq.)	χ^2	Fisher's <i>P</i>	Pearson's <i>P</i>	Odds ratio (95% CI)
c c a c c c a a c	14.38 (0.033)	36.25 (0.122)	20.964	4.79E-06*	4.75E-06*	0.249 (0.132–0.469)
c c a c c c a a t	34.35 (0.078)	18.77 (0.063)	0.835	0.360864	0.360813	1.314 (0.731–2.362)
c c a c c c a g c	7.35 (0.017)	3.03 (0.010)	–	–	–	–
c c a c c c a g t	29.11 (0.066)	7.77 (0.026)	6.572	0.010389*	0.010379*	2.768 (1.234–6.209)
c c a c t c a a c	17.88 (0.041)	5.87 (0.020)	2.782	0.095395	0.095333	2.196 (0.852–5.660)
c c a c t c a a t	6.83 (0.016)	13.96 (0.047)	5.965	0.01463*	0.014617*	0.331 (0.131–0.838)
c c a c t c a g c	5.91 (0.013)	5.07 (0.017)	–	–	–	–
c c a c t c a g t	9.42 (0.022)	9.00 (0.030)	0.456	0.499614	0.499599	0.727 (0.288–1.839)
c c a t c c a a c	38.64 (0.088)	16.43 (0.055)	3.239	0.071981	0.071929	1.727 (0.947–3.149)
c c a t c c a a t	41.06 (0.094)	36.50 (0.123)	1.242	0.265058	0.264985	0.763 (0.473–1.229)
c c a t c c a g c	14.48 (0.033)	22.99 (0.078)	6.659	0.009896*	0.009887*	0.419 (0.213–0.826)
c c a t c c a g t	17.89 (0.041)	11.17 (0.038)	0.097	0.755201	0.755194	1.129 (0.526–2.427)
c c a t t c a a c	13.31 (0.030)	17.47 (0.059)	3.219	0.072844	0.072791	0.517 (0.249–1.075)
c c a t t c a a t	19.06 (0.044)	29.87 (0.101)	8.603	0.003371*	0.003367*	0.417 (0.229–0.759)
c c a t t c a g c	5.66 (0.013)	0.00 (0.000)	–	–	–	–
c c a t t c a g t	15.67 (0.036)	5.86 (0.020)	1.805	0.179119	0.17904	1.914 (0.731–5.012)
t c a c c c a a c	9.94 (0.023)	0.00 (0.000)	–	–	–	–
t c a c c c a a t	18.06 (0.041)	5.09 (0.017)	3.677	0.055239	0.055196	2.568 (0.947–6.958)
t c a c c c a g c	15.16 (0.035)	0.00 (0.000)	10.9	0.000968	0.000967	–
t c a c c c a g t	2.58 (0.006)	6.51 (0.022)	–	–	–	–
t c a c t c a a c	0.00 (0.000)	5.03 (0.017)	–	–	–	–
t c a c t c a a t	8.30 (0.019)	0.00 (0.000)	–	–	–	–
t c a c t c a g c	5.19 (0.012)	0.00 (0.000)	–	–	–	–
t c a c t c a g t	2.53 (0.006)	5.65 (0.019)	–	–	–	–
t c a t c c a a c	19.53 (0.045)	6.57 (0.022)	2.904	0.088426	0.088367	2.148 (0.874–5.279)
t c a t c c a a t	22.48 (0.051)	13.22 (0.045)	0.277	0.59856	0.59856	1.206 (0.600–2.427)
t c a t c c a g c	5.06 (0.012)	0.00 (0.000)	–	–	–	–
t c a t c c a g t	10.92 (0.025)	5.72 (0.019)	–	–	–	–
t c a t t c a a c	8.51 (0.019)	0.00 (0.000)	–	–	–	–
t c a t t c a a t	18.67 (0.043)	0.00 (0.000)	13.506	0.00024	0.00024	–
t c a t t c a g c	0.00 (0.000)	3.31 (0.011)	–	–	–	–
t c a t t c a g t	0.06 (0.000)	4.91 (0.017)	–	–	–	–

SNP, single-nucleotide polymorphism; CI, confidence interval

The SNP order of constructing haplotype was as follows: rs28683050, rs28669671 rs7166015, rs3825977, rs2289259, rs28627002, rs28719801, rs2278546, rs10152593

and SMAD7^{22,23} function as a negative feedback loop of the TGF- β -signal by either competing for SMAD4 or terminating the signal and therefore may represent important regulators of TGF- β 1-induced effects.

The data above confirmed the central role of TGF- β and the Smad pathway for proceeding to progressive fibrosis. In these animals, the administration of interleukin-1 β caused extensive inflammation and tissue damage in both the Smad3 null as well as wild-type mice. If anything, the Smad3 null mouse showed an even greater degree of inflammation. However, after 20 d, only the wild-type animal had progressed to fibrosis, whereas the Smad3 null mouse had no indication of fibrogenesis and was not markedly different from animals receiving only control vector.²⁴ The Smad3 null mouse has been used to demonstrate that, in the absence of Smad3, the ALK5 receptor cannot transmit messages through to the nucleus and thus cannot upregulate matrix gene expression in either bleomycin- or TGF- β gene-mediated fibrotic stimulation.^{25,26} Taken together, these data indicate that only when there is an intact TGF- β RI signaling mechanism and intact Smad3 signaling can the progressive nature of fibrosis proceed.

In large genome-wide association studies on COPD, Dr Silverman and Spitz identified two SNPs at the CHRNA 3/5 locus, identified earlier as a risk factor for both lung cancer and nicotine dependence, to be associated with COPD. Smad3 plays an important role in the pathogenesis of COPD, and the CHRNA 3/5 gene and the Smad3 gene

are both located in chromosome 15q.^{10,11} Taking these backgrounds into account, we hypothesize that SNPs in the Smad3 may render susceptibility to COPD, and thereby to contribute in predicting its occurrence during the course of the disease. Although introns were originally believed to be non-functional because they do not code for proteins, it has been suggested that some of these sequences do indeed have relevance. Introns have also been implicated in regulating gene expression and DNA–protein interactions; mutations in intron sequences may affect these functions.²⁷ Otsuka *et al.*²⁸ performed the tag SNP analysis using 111 atopic dermatitis (AD) families (384 members) in Japan. Thirty SNPs to cover the Smad3 gene region were chosen for genotyping using an Illumina GoldenGate assay. One intron SNP (rs4147358) was associated with AD in the pedigree disequilibrium test (PDT) and transmission disequilibrium test (TDT) analysis (significance level of $P < 0.05$). The rs4147358 was further genotyped in independent case-control samples and it was found that rs4147358 was associated with AD; the association was strengthened when the AD samples were limited to those of patients with IgE ≥ 250 IU/mL. In another case-control study, three intron SNPs of the Smad3 gene (rs2289790, rs2289791 and rs3825977) were found to be in association with diffuse cutaneous scleroderma in a UK Caucasian population.²⁹ Just like COPD, the fibrosis induced by transforming growth factor ss-Smads signaling also plays a major

role for pathogenesis of AD and diffuse cutaneous scleroderma. So in this study, we have investigated nine intron SNPs of the *Samd3* gene in COPD patients and controls. And we only found rs28683050, as a known SNP in intron 1 of the *Smad3* gene, has a significant difference in allelic or genotypic frequencies between the COPD patients and control subjects. Its frequencies of the TT genotype and T allele in the COPD patients were significantly higher than that in controls. Rs2289259 was not found to be in association with COPD, which was also not associated with AD in the PDT and TDT analysis.²⁸ Contrary to the study by Pushpakom *et al.*,²⁹ rs3825977 was not associated with COPD in our study, and the different ethnic backgrounds may be one of the reasons.

In summary, we have performed a comprehensive investigation of *Smad3* gene polymorphisms. The SNP showing significant differences may be useful for predicting COPD susceptibility or possible preventive intervention. No previous research reported the SNPs in the current study have a function, and the SNPs are not in spliced junction regions. The SMAD3 SNP (rs28683050) found to be associated with COPD in this study has no known function, and it is likely that the SNPs are in LD with other functional SNPs that determine disease susceptibility. In the current study, no exon SNP of SMAD3 was included. Considering they may also render susceptibility to COPD, further studies need to be conducted to clarify the association between exon SNPs of SMAD3 and COPD and the functional characterization of these SNPs.

Author contributions: TY, BY and FW participated in research design; TY, BY, XS and DL contributed to the performance of laboratory measurements; TY, BY, DX and TW were involved in data collection and analysis; and TY, BY, HF, SZ and FW wrote the manuscript. TY and BY contributed equally to this work.

ACKNOWLEDGEMENTS

This study was supported by Grant 30971327 from the National Natural Science Foundation of China; and Grants 00-722 and 06-834 from the China Medical Board of New York to Dr FQW.

REFERENCES

- Davis RM, Novotny TE. The epidemiology of cigarette smoking and its impact on chronic obstructive pulmonary disease. *Am Rev Respir Dis* 1989;**140**(Part 2):S82-4
- McCloskey SC, Patel BD, Hinchliffe SJ, Reid ED, Wareham NJ, Lomas DA. Siblings of patients with severe chronic obstructive pulmonary disease have a significant risk of airflow obstruction. *Am J Respir Crit Care Med* 2001;**164**(Part 1):1419-24
- Molfino NA. Genetics of COPD. *Chest* 2004;**125**:1929-40
- Poller W, Meisen C, Olek K. DNA polymorphisms of the alpha 1-antitrypsin gene region in patients with chronic obstructive pulmonary disease. *Eur J Clin Invest* 1990;**20**:1-7
- Samilchuk EI, Chuchalin AG. Mis-sense mutation of alpha 1-antichymotrypsin gene and chronic lung disease. *Lancet* 1993;**342**:624
- Smith CA, Harrison DJ. Association between polymorphism in gene for microsomal epoxide hydrolase and susceptibility to emphysema. *Lancet* 1997;**350**:630-3
- Schellenberg D, Pare PD, Weir TD, Spinelli JJ, Walker BA, Sandford AJ. Vitamin D binding protein variants and the risk of COPD. *Am J Respir Crit Care Med* 1998;**157**(Part 1):957-61
- Huang SL, Su CH, Chang SC. Tumor necrosis factor-alpha gene polymorphism in chronic bronchitis. *Am J Respir Crit Care Med* 1997;**156**:1436-9
- Zhu G, Warren L, Aponte J, Gulsvik A, Bakke P, Silverman EK, Anderson WH, Lomas DA, Pillai SG; ICGN Investigators. The SERPINE2 gene is associated with chronic obstructive pulmonary disease in two large populations. *Am J Respir Crit Care Med* 2007;**176**:167-73
- Spitz MR, Amos CI, Dong Q, Lin J, Wu X. The CHRNA5-A3 region on chromosome 15q24-25.1 is a risk factor both for nicotine dependence and for lung cancer. *J Natl Cancer Inst* 2008;**100**:1552-6
- Pillai SG, Ge D, Zhu G, Kong X, Shianna KV, Need AC, Feng S, Hersh CP, Bakke P, Gulsvik A, Ruppert K, Lødrup Carlsen KC, Roses A, Anderson W, Rennard SI, Lomas DA, Silverman EK, Goldstein DB; ICGN investigators. A genome-wide association study in chronic obstructive pulmonary disease (COPD): identification of two major susceptibility loci. *PLoS Genet* 2009;**5**:e1000421
- Dennler S, Huet S, Gauthier JM. A short amino-acid sequence in MH1 domain is responsible for functional differences between Smad2 and Smad3. *Oncogene* 1999;**18**:1643-8
- Massague J. TGF-beta signal transduction. *Annu Rev Biochem* 1998;**67**:753-91
- Nagarajan RP, Liu J, Chen Y. Smad3 inhibits transforming growth factor-beta and activin signaling by competing with Smad4 for FAST-2 binding. *J Biol Chem* 1999;**274**:31229-35
- Livak KJ. Allelic discrimination using fluorogenic probes and the 5' nuclease assay. *Genet Anal* 1999;**14**:143-9
- HWE program. See [ftp://linkage.rockefeller.edu/software/utilities](http://linkage.rockefeller.edu/software/utilities)
- Shi YY, He L. SHEsis, a powerful software platform for analyses of linkage disequilibrium, haplotype construction, and genetic association at polymorphism loci. *Cell Res* 2005;**15**:97-8
- Takizawa H, Tanaka M, Takami K, Ohtoshi T, Ito K, Satoh M, Okada Y, Yamasawa F, Nakahara K, Umeda A. Increased expression of transforming growth factor-beta1 in small airway epithelium from tobacco smokers and patients with chronic obstructive pulmonary disease (COPD). *Am J Respir Crit Care Med* 2001;**163**:1476-83
- Massague J. The transforming growth factor-beta family. *Annu Rev Cell Biol* 1990;**6**:597-641
- Overall CM, Wrana JL, Sodek J. Independent regulation of collagenase, 72-kDa progelatinase, and metalloendoproteinase inhibitor expression in human fibroblasts by transforming growth factor-beta. *J Biol Chem* 1989;**264**:1860-9
- Murphy G, Docherty AJ. The matrix metalloproteinases and their inhibitors. *Am J Respir Cell Mol Biol* 1992;**7**:120-5
- Imamura T, Takase M, Nishihara A, Oeda E, Hanai J, Kawabata M, Miyazono K. Smad6 inhibits signalling by the TGF-beta superfamily. *Nature* 1997;**389**:622-6
- Nakao A, Afrakhte M, Moren A, Nakayama T, Christian JL, Heuchel R, Itoh S, Kawabata M, Heldin NE, Heldin CH. Identification of Smad7, a TGFbeta-inducible antagonist of TGF-beta signalling. *Nature* 1997;**389**:631-5
- Bonniaud P, Margetts PJ, Ask K, Flanders K, Gaudie J, Kolb M. TGF-beta and Smad3 signaling link inflammation to chronic fibrogenesis. *J Immunol* 2005;**175**:5390-5
- Zhao J, Shi W, Wang YL, Chen H, Bringas P Jr, Datto MB, Frederick JP, Wang XF, Warburton D. Smad3 deficiency attenuates bleomycin-induced pulmonary fibrosis in mice. *Am J Physiol Lung Cell Mol Physiol* 2002;**282**:L585-93
- Bonniaud P, Kolb M, Galt T, Robertson J, Robbins C, Stampfli M, Lavery C, Margetts PJ, Roberts AB, Gaudie J. Smad3 null mice develop airspace enlargement and are resistant to TGF-beta-mediated pulmonary fibrosis. *J Immunol* 2004;**173**:2099-108
- Tabor HK, Risch NJ, Myers RM. Candidate-gene approaches for studying complex genetic traits: practical considerations. *Nat Rev Genet* 2002;**3**:391-7
- Otsukaa K, Takeshita S, Enomoto H, Takahashi T, Hirota T, Tamari M, Ebe K, Otsuka F, Shibasaki M, Arinami T, Noguchi E. SMAD3 as an atopic dermatitis susceptibility gene in the Japanese population. *J Dermatol Sci* 2009;**55**:193-204
- Pushpakom SP, Herrick AL, Kumar S, Worthington JE. Diffuse systemic sclerosis: role of SMAD3 polymorphisms. *Rheumatology* 2005;**44**(Suppl. 1):I16-I16

(Received September 2, 2009, Accepted February 18, 2010)