Single-nucleotide polymorphisms in the TSPYL-4 and NT5DC1 genes are associated with susceptibility to chronic obstructive pulmonary disease

YI GUO[1*], YI GONG[2*], GUOCHAO SHI[1], KUN YANG[1], CHUNMING PAN[3], MIN LI[1], QINGYUN LI[1], QIJIAN CHENG[1], RANRAN DAI[1], LIANG FAN[1] and HUANYING WAN[1]

[1]Department of Respiratory Medicine, Ruijin Hospital, Shanghai Jiao Tong University (SJTU), School of Medicine; [2]Department of Respiratory Medicine, Huashan Hospital, Fudan University; [3]Ruijin Hospital, State Key Laboratory of Medical Genomics, Molecular Medicine Center, Shanghai Jiao Tong University (SJTU), School of Medicine, Shanghai 200025, P.R. China

Received February 12, 2012; Accepted June 18, 2012

DOI: 10.3892/mmr.2012.964

Abstract. The risk of developing chronic obstructive pulmonary disease (COPD) is partially determined by genetic and environmental factors. Many published candidate gene studies show conflicting results due to ethnic differences and sample sizes. The number of these studies carried out in Chinese populations is small. To investigate candidate genes and haplotypes for susceptibility to COPD in a southern Han Chinese population, we performed genotyping of DNA samples in 200 COPD patients and 250 control subjects by analyzing 54 single-nucleotide polymorphisms (SNPs) in 23 genes associated with the development of COPD and/or pulmonary function identified by genome-wide association studies (GWAS). We also performed linkage disequilibrium (LD) and haplotype analysis according to the results of genotyping. The frequencies of the SNP [rs3749893 of testis-specific protein Y-encoded-like 4 (TSPYL-4) gene] G allele and SNP [rs1052443 of 5'-nucleotidase domain containing 1 (NT5DC1) gene] A allele were significantly higher in the cases studied compared to the control subjects (P=0.032, P<0.05, OR=0.692, 95% CI 0.495-0.970; P=0.0205, P<0.05, OR=0.670, 95% CI 0.477-0.941, respectively). Results showed that two blocks of SNPs (rs1052443 and rs3749893; rs11155242 and rs6937121) had sufficient precision to allow construction of a haplotype block. We constructed the TSPYL-4 and NT5DC1 haplotypes of the cases and controls, but no significant difference between the two groups was found. rs3749893 A allele of TSPYL-4 and rs1052443 C allele of NT5DC1 were associated with a protective effect against the deterioration of pulmonary function. In conclusion, TSPYL-4 and NT5DC1 gene polymorphisms are associated with susceptibility to COPD and pulmonary function.

Correspondence to: Professor Huanying Wan, Department of Respiratory Medicine, Ruijin Hospital, Shanghai Jiao Tong University (SJTU), School of Medicine, Shanghai 200025, P.R. China
E-mail: wanhuanying@hotmail.com

*Contributed equally

Key words: chronic obstructive pulmonary disease, single-nucleotide polymorphism, genome-wide association studies, genotype, testis-specific protein Y-encoded-like 4, 5'-nucleotidase domain containing 1

Introduction

Chronic obstructive pulmonary disease (COPD) is expected to be the third leading cause of mortality and the fifth leading cause of morbidity by the year 2020 (1). The disease is mainly characterized by the presence of chronic airflow limitation that progresses slowly over a period of years and is largely irreversible (2,3). In China, it is becoming an increasingly common problem. A survey of 20,245 participants in seven regions of China conducted in 2007 indicated that the prevalence of COPD in adults aged over 40 years was 8.2% (4). However, the disease remains under-recognized and under-diagnosed, and we need to further understand the pathogenesis, particularly in the earlier mild and moderate stages of COPD. Although cigarette smoking is the major risk factor for COPD, only a minority (20%) of smokers develop the disease clinically (5). Hodge *et al* (6) revealed in their study that apoptosis of airway epithelial cells and inflammation of the airway mucosa persisted even after smoking cessation in patients with COPD.

Evidence suggests that the risk of developing COPD is partially determined by genetic and environmental factors (7). A number of candidate gene studies have therefore been carried out in recent years. Family studies and linkage analysis in early-onset COPD pedigrees have highlighted a genetic predisposition (8-11), and genome-wide association studies (GWAS) for COPD or pulmonary function identified some susceptibility loci (12-16), but with varying degrees of reproducibility. Conflicting results may be due to ethnic differences and sample sizes.

The number of these studies carried out in Chinese individuals is small. Past candidate gene studies may focus on a single gene or on a few genes in combination, with these genes identified based on prior knowledge or suspected mechanisms of disease pathogenesis. Nonetheless, elucidating the genetics of these disorders is severely hampered by genetic heterogeneity, the low penetrance of individual disease alleles and the potential for gene-gene and gene-environment interactions. It is probable that groups of genes rather than single genes are involved in disease development.

The aim of the present study was to investigate candidate genes and haplotypes in susceptibility to COPD in a south Han Chinese population.

Materials and methods

Subjects. A total of 200 male COPD patients visiting the Department of Respiratory Disease of the Shanghai Ruijin Hospital, China, between December 2008 and December 2009 were recruited. COPD was diagnosed according to the criteria established by the NHLBI/WHO Global Initiative for COPD (GOLD) (17). Criteria were as follows: age ≥40 years; chronic respiratory symptoms and signs, such as cough and dyspnea; airflow limitation as indicated by forced expiratory volume in 1 sec (FEV1)/forced vital capacity (FVC) <70% and FEV1 reversibility after inhalation of 400 μg salbutamol to <12% of the pre-bronchodilator FEV1. Patients were excluded if they had a diagnosis of asthma, lung cancer or radiographic abnormalities suggestive of other significant respiratory diseases, such as bronchiectasis or pulmonary tuberculosis.

A total of 250 control male subjects were enrolled at the General Health Check-up Center in Shanghai No. 10 Hospital during the same period. Their characteristics were mentioned in a previous study (18). The cases and control subjects were from an ethnic Chinese, southern Han population who resided in Shanghai City or the surrounding regions, and were matched for age, gender and smoking history. The study protocol was approved by the medical ethics committee of Shanghai Ruijin Hospital, School of Medicine, Shanghai Jiaotong University, and all the participants gave written informed consent.

DNA extraction and genotyping of study samples. We collected 4 ml of peripheral blood from each participant for DNA preparation. Genomic DNA was extracted using a Blood DNA Extraction kit (Tiangen Biotech, Co. Ltd., Beijing, China). Any sample with a DNA concentration <10 ng/μl was excluded.

In total, 54 single-nucleotide polymorphisms (SNPs) were found in 23 genes associated with the development of COPD and/or pulmonary function, as identified by publications of previous GWAS and by searching the dbSNP database of NCBI (Table I). Genotyping was performed using the Mass-Array™ Technology Platform of Sequenom Inc. (San Diego, CA, USA). As a result of a quality control measure, we excluded 30 SNPs: the minor allele frequency (MAF) of 10 SNPs was <0.03; these were rs8034191, rs17036052, rs17035960, rs11097901, rs11728716, rs10516526, rs11727189, rs17036090, rs17331332 and rs17036341. Eleven SNPs were not compatible in the same multiplex PCR system; these were rs7710510, rs1042718, rs1042717, rs3753661, rs3766934, rs1903003, rs7671167, rs1980057, rs11168048, rs7735184 and rs16865421. Nine SNPs showed deviation from the Hardy-Weinberg equilibrium (HWE): rs2070600, rs2395730, rs6830970, rs13147758, rs17019336, rs2035901, rs10498230, rs6712954 and rs6734100. Therefore, 24 SNPs were selected for the investigation. The sequence information of these 24 SNPs is shown in Table II.

Genotyping by multiplex PCR. Genotyping was performed by multiplex PCR, which was a variant of PCR enabling the simultaneous amplification of numerous targets of interest in one reaction using more than one pair of primers (19). We used Mass-array Assay Design 2.0 software to design multiplex primers: 1st-PCR primer, 2nd-PCR primer and UEP primer for each SNP; primers of the 24 SNPs are shown in Table II.

Statistical analysis. P-values for genotype and allele frequencies were obtained using the χ^2 test with SPSS 13.0 software ($P<0.05$). We excluded the SNPs in which MAF was <0.03. The relative risk associated was estimated as an odds ratio (OR) with a 95% confidence interval (CI), as analyzed by the Woolf method. Each SNP was tested for deviation from HWE (http://ihg2.helmholtz-muenchen.de/cgi-bin/hw/hwa1.pl). SNPs were excluded from the analysis if they were out of HWE ($P\leq0.05$). Haplotype frequencies and linkage disequilibrium (LD) analysis were evaluated using the Phase and Haploview software.

Results

Study population characteristics. The study population characteristics for those subjects with successful genotyping are shown in Table III. Due to a lack of certain data, the case group comprised 160 subjects and the control group 177 subjects. The two groups were matched for age, gender and percentage of smokers. FEV1 and FVC of the case group were significantly decreased compared to the control group ($P<0.05$).

Result of genotyping. As a result of a quality control measure, a total of 24 SNPs were finally compared between the case and control groups. The frequencies of the SNP [rs3749893 of testis-specific protein Y encoded-like 4 (TSPYL-4) gene] G allele and SNP [rs1052443 of 5'-nucleotidase domain containing 1 (NT5DC1) gene] A allele were significantly higher in the cases studies compared to the control subjects ($P=0.032$, $P<0.05$, OR=0.692, 95% CI 0.495-0.970; $P=0.0205$, $P<0.05$, OR=0.670, 95% CI 0.477-0.941, respectively). The details are shown in Table IV.

Linkage disequilibrium and haplotype analysis. Using Haploview, haplotype blocks were constructed separately according to the confidence interval method of Gabriel *et al* (20) for the cases and controls. This method uses both an estimate of d' and a measure of its precision (confidence bounds) to construct haplotype blocks (Fig. 1). Blocks with pairwise d'<1 have actual d' values in the squares. Although the estimated pairwise LD between a number of the SNPs was high, the precision of the estimates was not sufficiently high to fulfill the criteria for the construction of haplotype blocks utilising all the SNPs. Two blocks of SNPs (rs1052443 and rs3749893; rs11155242 and rs6937121) were identified that had sufficient precision to allow construction of a haplotype block.

Table I. Summary of positive single-nucleotide polymorphisms (SNPs) in the previous genome-wide association studies.

Gene	Gene description	Chromosome	SNP	MAF in HC[a]	Allele
ADAM19 (14)	Metallopeptidase domain 19	5	rs2277027	0.178	A/C
		5	rs1422795	0.178	A/G
ADCY2 (14)	Adenylate cyclase 2 (brain)	5	rs6555465	0.433	A/G
		5	rs7710510[c]	0.432	C/T
AGER (14)	Advanced glycosylation end-product-specific receptor	6	rs2070600	0.289	A/G
ARDB2 (36)	β2-adrenergic receptor	5	rs1042718[c]	0.344	A/C
		5	rs1042717[c]	0.322	A/G
CHRNA3 (13)	Cholinergic receptor nicotinic, α3	15	rs1051730	0.033	C/T
		15	rs8034191	0.011[b]	C/T
DAAM2 (14)	Dishevelled activator of morphogenesis 2	6	rs2395730	0.375	A/G
EPHX1 (36)	Epoxide hydrolase 1, microsomal (xenobiotic)	1	rs3753661[c]	0.244	G/T
		1	rs3766934[c]	0.244	G/T
FAM13A (14)	Family with sequence similarity 13, member A	4	rs2869967	0.478	C/T
		4	rs6830970	0.444	A/G
		4	rs1903003[c]	0.433	C/T
		4	rs7671167[c]	0.488	C/T
FLJ20184 (14)	Rho guanine nucleotide exchange factor (GEF) 38	4	rs17036052	0[b]	C/T
		4	rs17035960	0[b]	C/T
GPR126 (14)	G protein-coupled receptor 126	6	rs7776375	0.439	A/G
		6	rs6937121	0.389	G/T
		6	rs11155242	0.133	A/C
		6	rs3817928	0.133	A/G
GSTCD (14)	Glutathione S-transferase, C-terminal domain containing	4	rs11097901	0[b]	C/T
		4	rs11728716	0.011[b]	A/G
		4	rs10516526	0[b]	A/G
HHIP (12,14,16)	Hedgehog interacting protein	4	rs13147758	0.289	A/G
		4	rs17019336	0.300	A/T
		4	rs2353397	0.289	C/T
		4	rs2035901	0.300	A/G
		4	rs6537302	0.244	A/T
		4	rs12504628	0.298	C/T
		4	rs1032295	0.211	T/G
		4	rs1980057[c]	0.300	T/C
HTR4 (14,16)	5-hydroxytryptamine (serotonin) receptor 4	5	rs3995090	0.256	A/C
		5	rs6889822	0.333	A/G
		5	rs11168048[c]	0.211	C/T
		5	rs7735184[c]	0.267	G/T
NT5DC1 (14)	5'-nucleotidase domain containing 1	6	rs1052443	0.389	A/C
INTS12 (14)	Integrator complex subunit 12	4	rs11727189	0[b]	G/T
		4	rs17036090	0[b]	C/T
NPNT (14)	Nephronectin	4	rs17331332	0[b]	A/G
		4	rs17036341	0[b]	C/G
PID1 (14)	Phosphotyrosine interaction domain containing 1	2	rs1435867	0.107	C/T
		2	rs10498230	0.122	C/T
PPT2 (34)	Palmitoyl-protein thioesterase 2	6	rs10947233	0.239	G/T
PTCH1 (35)	Patched 1	9	rs16909898	0.078	A/G
		9	rs10512249	0.080	C/T
SERPINE2 (36,37)	Serpin peptidase inhibitor, clade E (nexin, plasminogen activator inhibitor type 1), member 2	2	rs7583463	0.292	A/C
		2	rs6712954	0.086	A/G
		2	rs6734100	0.189	C/G
		2	rs16865421[c]	0.218	A/G
THSD4 (16)	Thrombospondin, type I, domain containing 4	15	rs12899618	0.133	A/G
TNS1 (16)	Tensin 1	2	rs2571445	0.400	C/T
TSPYL4 (14)	Testis-specific protein Y encoded-like-4	6	rs3749893	0.389	A/G

[a]MAF in HC, minor allele frequency in Han Chinese. These SNPs were not included in genotyping, since their MAF was <0.03. [b]These SNPs were excluded, since they were not compatible in the same multiplex PCR system.

Table II. Sequence of 24 single-nucleotide polymorphisms (SNPs) and their primers in multiplex PCR.

Gene	SNP_ID	Sequence	1st-PCR Primer	2nd-PCR Primer	UEP_SEQ
ADAM19	rs2277027	GTGTCCTCAT[A/C]AAAAATTCAT	ACGTTGGATGTTGGCTTGGCCATATGTGTC	ACGTTGGATGGAAGAATTAAGGGCTGTGGC	ccccGATCAAAATGTGTCCTCAT
	rs1422795	GTCCTTTCTC[A/G]GTTGGAGGCGC	ACGTTGGATGGCGTTGTCATGGTACTTCTG	ACGTTGGATGCCATATTCTACCCTCTGGTC	gCTACCCTCTGGTCCTTTCTC
ADCY2	rs6555465	GGGATTCAGA[A/G]GCTGGATATG	ACGTTGGATGACCTTCACCTTAGAGACAGC	ACGTTGGATGCTGGGTGGCTTCCAATTTTC	TATCCATGTCATATCCAGC
CHRNA3	rs1051730	GCCCCAGGCTA[C/T]AAACACGACA	ACGTTGGATGCAGCAGTTGTACTTGATGTC	ACGTTGGATGCTCAAGGACTATTGGGAGAG	ATCATCAAAGCCCCAGGCTA
FAM13A	rs2869967	CTGGAAGGGT[C/T]AAGGGATGTATT	ACGTTGGATGAAGAGCAGAGCTCTTGGAAC	ACGTTGGATGTGAGCCCCCATTTTCTAACC	tacCCCAGGATCCTGAATACATCCCTT
GPR126	rs11155242	TTCTGATTCA[A/C]AATGTTTGTT	ACGTTGGATGGGCTACTTTCTATCCATTTC	ACGTTGGATGAGCTGTTCAAAGCTTTCTGC	TAATGCATTATTCAACAAACATT
	rs7776375	CCCCCAACAC[A/G]GGTACCAAAA	ACGTTGGATGACTTTGCACCGTGTGGTAAC	ACGTTGGATGGTCAAATTAATGATCCCCCC	attcATGATCCCCCCCAACAC
	rs6937121	AATTTGTCCA[G/T]TCCTTTTAAC	ACGTTGGATGCAGTAAAATTCTGATAGCC	ACGTTGGATGGAAAGATTGGCACAACTGTC	gCAAATCTATTATCTAATTTGTCCA
	rs3817928	GATGACCCACCA[A/G]TTCAGTCCCTC	ACGTTGGATGGGTGCCAAGGTAAAGAGATG	ACGTTGGATGTGCAGCGTGTAATGTCTCAG	ttATGTTTCACTTGTGAGGGACTGAA
HHIP	rs6537302	CCATCATCTA[A/T]CATACTAACA	ACGTTGGATGGTGAAGTGCTTTGTAAGCTG	ACGTTGGATGGCATATTTTCCTTCTTCCCC	cccCCCCTCCTTTTTATGTTAGTATG
	rs12504628	CATTACCCCA[C/T]TAAAGGTA	ACGTTGGATGCCCAGAAAATAAGAGTCTGC	ACGTTGGATGGGATTTGAGATTTAGAGTGCC	TTAGAGTGCCATTACCCCA
	rs1032295	TGGAGGAGAG[T/G]TGTATCAAGT	ACGTTGGATGTTTGCCCAGAGCTGCTAAAG	ACGTTGGATGTGGCAAAGAACATCTGGGAG	ttCAGTGTTTGGAGGAGAG
	rs2353397	CACCATTATT[C/T]GTATTCTTGT	ACGTTGGATGGCTACCATAGCTGTTATATAC	ACGTTGGATGGGTACAGCTGCAAATAGCTC	aatcGCTCATTTCCACCATTATT
HTR4	rs3995090	TCCTGTAGCT[A/C]ATCTCCATGA	ACGTTGGATGAAACTCAGCCTCTTCTTCCC	ACGTTGGATGGGAAGTGGCCATATATATTC	atgCAACATAGACATATCATGGAGAT
	rs6889822	TAGTGCACAA[A/G]GGGTAGCCAA	ACGTTGGATGTGGGAGATCAAGAGTGAAGG	ACGTTGGATGCCATTCACTCAAATAGCAGG	cTTGTTGAATTTTAGTGCACAA
NT5DC1	rs1052443	ACTGCTTGCC[A/C]ATCAGTTTGT	ACGTTGGATGGCCTATAGAGAAGTTGGGTC	ACGTTGGATGGTCTGTGAACAGGTACATGG	aaAGGTACATGGTACAACAAACTGAT
PID1	rs1435867	ATATTTCCTT[C/T]CTGATTTTAC	ACGTTGGATGACAGTGACTCATCAAAGCTC	ACGTTGGATGGAGTTGGGGAGAACAGATAG	ccccTGCTAAACTTGTATATTTCCTT
PPT2	rs10947233	GGCTGGATTT[G/T]GTCCTTTGGTC	ACGTTGGATGAGAGGTTGGCAAACTGTGAC	ACGTTGGATGAATGAATGTGCCTGTGTTCC	AGTAGCAGGCTGGATTT
PTCH1	rs16909898	GAAGACAGGA[A/G]GAGCCTTAAG	ACGTTGGATGGAAGCAATCTGATGAACTCC	ACGTTGGATGCCAAGGTAATCTGCCACAAC	TCTGCCACAACTTAAGGCTC
	rs10512249	TGGTGGTTGA[C/T]CTCTCACTGG	ACGTTGGATGTCAGCCCAAAGAGAATGTGC	ACGTTGGATGGGTGTCCTTTTCTCGTTATG	TTATGCTTTGGTGGTTGA
SERPINE2	rs7583463	GTTCTATTAA[A/C]CCTCACTGTA	ACGTTGGATGCCTTATGAAAGCACATGGAG	ACGTTGGATGACTGAAAACACACGTGTGTC	TATGGATGAAAGTACAGTGAGG
THSD4	rs12899618	GAGCCCTGAT[A/G]AAAGAAAAAT	ACGTTGGATGCCCAATACTCTGGCTGGAAT	ACGTTGGATGACAGTGTCCTGTGTGCTATG	ccTATTGCTTTTATGAGCCCTGAT
TNS1	rs2571445	TGGCTTCGGC[A/C/G/T]GGCGGGCCAT	ACGTTGGATGCAGCCATGCTGGGATTGATG	ACGTTGGATGAACAGTGGGCACCAACACTC	ccCTCCCCCTAGTCCTGGCTTCGGC
TSPYL-4	rs3749893	AGAAAACATC[a/G]ACTTACATTT	ACGTTGGATGCCTAAGTCCTTGAATTCACC	ACGTTGGATGGGTCACTCCTAGGATAATTG	cgAGGATAATTGAGAGAAAACATC

Primer letters in lower case indicate primer bases (added to balance molecular weight).

Table III. Demographics of chronic obstructive pulmonary disease patients and control subjects.

	Case	Control
No.	160	177
Age (years)	54.00±13.28	52.00±4.84
BMI (kg/m^2)	21.62±3.77	23.81±2.54
FEV1 (L)	1.03±0.56[a]	2.88±0.63
FVC (L)	2.10±0.92[a]	3.35±0.90
FEV1/FVC (%)	51.00±14.02[a]	80.00±10.03
Smoking percentage	75.50	74.71

Data were presented as the means ± SD. FEV1, forced expiratory volume in 1 sec; FVC, forced vital capacity; BMI, body mass index. [a]P<0.05 vs. control using the Mann-Whitney U test.

Table IV. Allele frequencies and genotypes of the candidate single-nucleotide polymorphisms (SNPs) in the case and control groups.

SNP	Gene	SNP_ID	Case group					Control group					Case (%)	Control (%)	Allele P-value	OR (95% CI)
			a	b	c	d	HWE	a	b	c	d	HWE				
1	ADAM19	rs2277027	115	40	5	160	Yes	132	40	5	177	Yes	15.63	14.12	0.580	1.13 (0.736-1.722)
2	ADAM19	rs1422795	112	41	5	158	Yes	129	38	5	172	Yes	16.14	13.95	0.430	1.19 (0.774-1.820)
3	ADCY2	rs6555465	45	69	41	155	Yes	50	94	33	177	Yes	48.71	45.20	0.370	1.15 (0.848-1.563)
4	CHRNA3	rs1051730	149	10	1	160	Yes	164	10	0	174	Yes	3.75	2.87	0.530	1.32 (0.561-3.091)
5	FAM13A	rs2869967	40	87	29	156	Yes	46	93	33	172	Yes	53.53	47.97	0.150	1.25 (0.919-1.698)
6	GPR126	rs7776375	70	69	21	160	Yes	69	87	20	176	Yes	34.69	36.08	0.710	0.94 (0.685-1.292)
7	GPR126	rs11155242	135	23	2	160	Yes	144	31	2	177	Yes	8.44	9.89	0.520	0.84 (0.496-1.422)
8	GPR126	rs6937121	64	71	21	156	Yes	64	85	22	171	Yes	36.22	37.72	0.690	0.94 (0.682-1.289)
9	GPR126	rs3817928	130	28	2	160	Yes	137	35	2	174	Yes	10.00	11.21	0.610	0.88 (0.537-1.443)
10	HHIP	rs12504628	79	63	15	157	Yes	90	63	19	172	Yes	29.62	29.36	0.940	1.01 (0.724-1.416)
11	HHIP	rs1032295	98	51	8	157	Yes	101	63	9	173	Yes	21.34	23.41	0.520	0.89 (0.615-1.281)
12	HHIP	rs2353397	82	66	12	160	Yes	97	69	10	176	Yes	28.10	25.30	0.410	1.16 (0.821-1.628)
13	HHIP	rs6537302	97	49	9	155	Yes	96	69	12	177	Yes	21.61	26.27	0.160	0.77 (0.540-1.108)
14	HTR4	rs3995090	78	70	12	160	Yes	86	80	11	177	Yes	29.38	28.81	0.870	1.03 (0.737-1.433)
15	HTR4	rs6889822	60	79	16	155	Yes	72	88	17	177	Yes	35.81	34.46	0.720	1.06 (0.771-1.460)
16	NT5DC1	rs1052443	85	64	7	156	Yes	78	75	20	173	Yes	25.00	33.24	0.021*	0.67 (0.477-0.941)
17	PID1	rs1435867	126	29	5	160	Yes	132	41	4	177	Yes	12.19	13.84	0.520	0.86 (0.550-1.356)
18	PPT2	rs10947233	101	52	7	160	Yes	98	64	9	171	Yes	20.63	24.43	0.240	0.80 (0.558-1.158)
19	PTCH1	rs16909898	138	21	1	160	Yes	150	25	2	177	Yes	7.19	8.19	0.630	0.87 (0.491-1.534)
20	PTCH1	rs10512249	133	21	1	155	Yes	149	26	2	177	Yes	7.42	8.47	0.620	0.87 (0.491-1.524)
21	serpine2	rs7583463	73	67	16	156	Yes	92	70	10	172	Yes	31.73	26.16	0.120	1.31 (0.935-1.840)
22	THSD4	rs12899618	121	32	2	155	Yes	135	41	1	177	Yes	11.61	12.15	0.830	0.95 (0.593-1.523)
23	TNS1	rs2571445	51	69	33	153	Yes	61	73	35	169	Yes	44.12	42.31	0.640	1.08 (0.788-1.471)
24	TSPYL4	rs3749893	86	67	7	160	Yes	79	77	19	175	Yes	25.31	32.86	0.032*	0.69 (0.495-0.970)

a, Homozygous dominant no.; b, heterozygous no.; c, homozygous recessive no.; d, total no. HWE, Hardy-Weinberg equilibrium.

As the two SNPs (rs3749893 and rs1052443) are situated on chromosome 6, we constructed the TSPYL-4 and NT5DC1 haplotypes of cases and controls. Haplotypes with frequencies >2% were selected for the analysis. No significant difference was observed between the two groups (Table V).

Genotype of TSPYL-4 and NT5DC1 SNPs, and pulmonary function. The SNPs of TSPYL-4 and NT5DC1 were associated with the development of COPD according to the results of our study, while COPD is characterized by an airflow limitation that is not fully reversible. We also investigated the relationship between the TSPYL-4 and NT5DC1 gene polymorphisms, and the pulmonary function (FEV1, FVC and FEV1/FVC) was then investigated using ANOVA.

In COPD patients, for rs3749893 of TSPYL-4, the mean FEV1/FVC levels were significantly higher in AA carriers

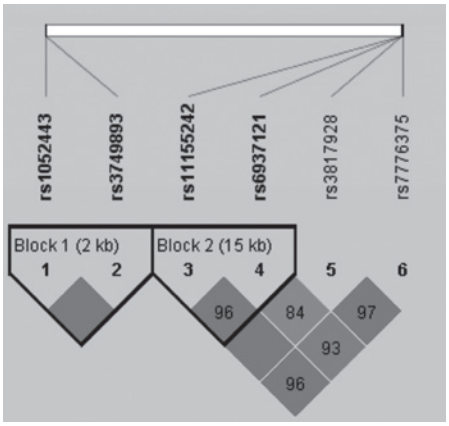

Figure 1. Linkage disequilibrium plot of SNP haplotypes. Haplotypes were constructed from genotyping data from Caucasians using the Gabriel block method. Significant d' values are shown. There were two blocks of SNPs (rs1052443 and rs3749893; rs11155242 and rs6937121).

than in AG carriers (57.00±2.16 vs. 51.99±12.27, P=0.043, P<0.05) and FEV1/FVC levels were also significantly higher in AG carriers than in GG carriers (51.99±12.27 vs. 47.09±12.8, P=0.016, P<0.05). For rs1052443 of NT5DC1, CC carriers were associated with significantly higher FEV1/FVC levels compared to CA carriers (57.00±2.16 vs. 52.09±12.36, P=0.037, P<0.05) and CA carriers had significantly higher FEV1/FVC levels than AA carriers (52.09±12.36 vs. 46.83±12.65, P=0.011, P<0.05). It appears that the rs3749893 A allele of TSPYL-4 and the rs1052443 C allele of NT5DC1 are associated with a protective effect against the deterioration of pulmonary function in our COPD patients. In the control group, similar trends were observed, but these did not reach the level of significance. The details are shown in Table VI.

Discussion

The present study identified that SNP rs3749893 of TSPYL-4 and rs1052443 of NT5DC1 genes was significantly associated with susceptibility to COPD in a south Han Chinese population. In addition, the two SNPs constitute a haploview block. Recently, hundreds of GWAS were published, involving a number of diseases, such as asthma, obesity, diabetes and mental illness (21-26). GWAS for COPD were also carried out and related SNPs were reported; however, most of the studies were performed in Caucasians, not in Asians. In our study, we succeeded in replicating these SNPs in a southern Han Chinese population and found that two SNPs are associated with susceptibility to COPD.

We also demonstrated that the two gene polymorphisms played a significant role in pulmonary function (FEV1/FVC). The rs3749893 A allele of TSPYL-4 and the rs1052443 C allele of NT5DC1 are associated with a protective effect

Table V. Results of the haplotype analysis in the case and control groups.

Chromosome	Haplotype		Percentage (%)		P-value	OR (95% CI)
			Case	Control		
6	0	2	53.750	44.820	0.10152	0.69891 (0.455-1.074)
	1	2	21.620	21.240	0.93233	0.97768 (0.581-1.646)
	1	0	33.548	25.007	0.08464	0.66051 (0.412-1.060)

Table VI. Univariate analysis of variance comparing mean FEV1/FVC levels according to the different genotypes of the TSPYL-4 and NT5DC1 single-nucleotide polymorphisms.

	rs no.	Genotype	Subject (n)	FEV1/FVC[a]	P-value
COPD	rs3749893	GG	84	47.09±12.8	Reference
		AG	67	51.99±12.27[c]	0.016
		AA	7	57.00±2.16[b]	0.043
	rs1052443	AA	85	46.83±12.65	Reference
		CA	64	52.09±12.36[e]	0.011
		CC	7	57.00±2.16[d]	0.037
Control	rs3749893	GG	79	87.77±10.54	Reference
		AG	77	88.89±9.71	0.272
		AA	19	91.17±9.67	0.513
	rs1052443	AA	78	87.91±10.54	Reference
		CA	75	88.79±9.83	0.200
		CC	20	91.17±9.42	0.588

[a]Data were presented as the means ± SD. [b]Significant difference vs. rs3749893 AG carriers (COPD). [c]Significant difference vs. rs3749893 GG carriers (COPD). [d]Significant difference vs. rs1052443 CA carriers (COPD). [e]Significant difference vs. rs1052443 AA carriers (COPD). COPD, chronic obstructive pulmonary disease.

against the deterioration of pulmonary function in our COPD patients. Our finding is similar to that of the study by Hancock *et al* (14). Those authors conducted a meta-analysis of GWAS, which revealed that the two gene polymorphisms were associated with pulmonary function. However, their study mainly referred to ethnicities other than Chinese.

The TSPYL gene is significantly homologous to TSPY, which is expressed in the normal germ cells of fetal and adult testes and ectopically in tumor germ cells; designated as TSPY-like (TSPYL). TSPYL was assigned as a new member of the TSPY-SET-NAP1L1 family (27), which includes TSPYL1, TSPYL2, TSPYL3, TSPYL4, TSPYL5 and TSPYL6 (28). Human TSPYL is mapped to chromosome 6, and murine TSPYL to chromosome 10 (27). The TSPYL gene lacks introns and contains a coding region of 1,314 bases. The mRNA is approximately 3,200 bases in length, and the mature TSPYL protein is 437 aa (29). Expression of TSPYL was observed in all tissues, as well as at early onset during development. Vogel *et al* (27) investigated its expression in different tissues by northern blot analysis and RT-PCR. TSPYL is transcribed in all probed murine tissues, including the ovary and liver from females, as well as testes, spleen, brain, kidney, prostate, lung, liver and heart from males.

The functions of the TSPYL gene in the pathogenesis of COPD are far from being sufficiently studied. TSPYL may play a role in development by altering the regulation of specific developmental genes and contributing to region-specific chromatin remodeling (29). TSPYL is the putative gene for gonadoblastoma. The expression of TSPYL4 in human lung tissue and its function remain unclear. Published studies on the TSPYL4 gene are limited. Other members, such as TSPYL1, were found to be associated with the 46,XY disorder of sex development, male infertility and sudden infant death with dysgenesis of the testes syndrome (SIDDT) (30). TSPYL5 is one of the frequent targets of epigenetic silencing in primary glial tumors (31). In their study, Jung *et al* (32) reported that this gene is important in the development of gastric cancer, indicating that it is one of the potent tumor suppressor genes associated with DNA methylation. Kim *et al* (33) were the first to show that the TSPYL5 gene is partly involved in cell growth and resistance to cytotoxic agents via regulation of the cell level of the P21WAF/Cip1 and PTEN/AKT pathways.

NT5DC1 shows significant structural homology to several established members of the haloacid dehydrogenase (HAD) super-family, particularly phosphoserine phosphatase (PSP). The catalytic mechanism of NT5DC1 is also closely correlated with that of PSPs (34,35). Its expression and role in human lung tissue remain to be elucidated.

Limitations of our study are missing data, which may make it difficult to classify patients according to lung function severity, smoking index and family history. This lack of data may have resulted in bias towards the null, although we used strict criteria for both the case patients and control subjects to minimize this possibility.

In conclusion, these are the first reported SNPs in TSPYL-4 and NT5DC1 associated with COPD in a southern Han Chinese population. The two gene polymorphisms are crucial in pulmonary function (FEV1/FVC). The rs3749893 A allele of TSPYL-4 and the rs1052443 C allele of NT5DC1 are associated with a protective effect against the deterioration of pulmonary function. SNP rs3749893 of TSPYL-4 and rs1052443 of NT5DC1 constitute a haplotype block, which transmits as a whole unit. However, the expression and function of TSPYL-4 and NT5DC1 genes in human pulmonary tissue remain to be elucidated by further experimentation.

Acknowledgements

This study was funded by the 11th Chinese National Five-year Development Plan (2008BAI52B00). The authors thank Shaojun Yin and Changhui Wang, Shanghai No. 10 Hospital, Xiaoyan Jin, Changning District Central Hospital, and Boying Zhang, Luwan District Central Hospital, and acknowledge the help of the recruited volunteers. They also thank Yanyan Song from the Department of Biostatistics, Shanghai Jiao Tong University, School of Medicine. They also gratefully acknowledge the numerous sample donors for making this study possible.

References

1. Murray CJ and Lopez AD: Alternative projections of mortality and disability by cause 1990-2020: global burden of disease study. Lancet 349: 1498-1504, 1997.
2. Anthonisen NR, Connett JE, Kiley JP, Altose MD, Bailey WC, Buist AS, Conway WA Jr, Enright PL, Kanner RE, O'Hara P, *et al*: Effects of smoking intervention and the use of an inhaled anticholinergic bronchodilator on the rate of decline of FEV1. The lung health study. JAMA 272: 1497-1505, 1994.
3. Siafakas NM, Vermeire P, Pride NB, Paoletti P, Gibson J, Howard P, Yernault JC, Decramer M, Higenbottam T, Postma DS, *et al*: Optimal assessment and management of chronic obstructive pulmonary disease (COPD). The European Respiratory Society Task Force. Eur Respir J 8: 1398-1420, 1995.
4. Wang X, Li L, Xiao J, Jin C, Huang K, Kang X, Wu X and Lv F: Association of ADAM33 gene polymorphisms with COPD in a northeastern Chinese population. BMC Med Genet 10: 132-138, 2009.
5. Cigarette smoking and health. American Thoracic Society. Am J Respir Crit Care Med 153: 861-865, 1996.
6. Hodge S, Hodge G, Holmes M and Reynolds PN: Increased airway epithelial and T-cell apoptosis in COPD remains despite smoking cessation. Eur Respir J 25: 447-454, 2005.
7. Silverman EK: Progress in chronic obstructive pulmonary disease genetics. Proc Am Thorac Soc 3: 405-408, 2006.
8. Silverman EK, Chapman HA, Drazen JM, Weiss ST, Rosner B, Campbell EJ, O'Donnell WJ, Reilly JJ, Ginns L, Mentzer S, Wain J and Speizer FE: Genetic epidemiology of severe, early-onset chronic obstructive pulmonary disease. Risk to relatives for airflow obstruction and chronic bronchitis. Am J Respir Crit Care Med 157: 1770-1778, 1998.
9. Silverman EK, Mosley JD, Palmer LJ, Barth M, Senter JM, Brown A, Drazen JM, Kwiatkowski DJ, Chapman HA, Campbell EJ, *et al*: Genome-wide linkage analysis of severe, early-onset chronic obstructive pulmonary disease: airflow obstruction and chronic bronchitis phenotypes. Hum Mol Genet 11: 623-632, 2002.
10. DeMeo DL, Celedón JC, Lange C, Reilly JJ, Chapman HA, Sylvia JS, Speizer FE, Weiss ST and Silverman EK: Genome-wide linkage of forced mid-expiratory flow in chronic obstructive pulmonary disease. Am J Respir Crit Care Med 170: 1294-1301, 2004.
11. Silverman EK, Palmer LJ, Mosley JD, Barth M, Senter JM, Brown A, Drazen JM, Kwiatkowski DJ, Chapman HA, Campbell EJ, *et al*: Genomewide linkage analysis of quantitative spirometric phenotypes in severe early-onset chronic obstructive pulmonary disease. Am J Hum Genet 70: 1229-1239, 2002.
12. Wilk JB, Chen TH, Gottlieb DJ, Walter RE, Nagle MW, Brandler BJ, Myers RH, Borecki IB, Silverman EK, Weiss ST and O'Connor GT: A genome-wide association study of pulmonary function measures in the Framingham Heart Study. PLoS Genet 5: e1000429, 2009.

13. Pillai SG, Ge D, Zhu G, Kong X, Shianna KV, Need AC, Feng S, Hersh CP, Bakke P, Gulsvik A, *et al*: A genome-wide association study in chronic obstructive pulmonary disease (COPD): identification of two major susceptibility loci. PLoS Genet 5: e1000421, 2009.
14. Hancock DB, Eijgelsheim M, Wilk JB, Gharib SA, Loehr LR, Marciante KD, Franceschini N, van Durme YM, Chen TH, Barr RG, *et al*: Meta-analyses of genome-wide association studies identify multiple loci associated with pulmonary function. Nat Genet 42: 45-52, 2009.
15. Cho MH, Boutaoui N, Klanderman BJ, Sylvia JS, Ziniti JP, Hersh CP, DeMeo DL, Hunninghake GM, Litonjua AA, Sparrow D, *et al*: Variants in FAM13A are associated with chronic obstructive pulmonary disease. Nat Genet 42: 200-202, 2010.
16. Repapi E, Sayers I, Wain LV, Burton PR, Johnson T, Obeidat M, Zhao JH, Ramasamy A, Zhai G, Vitart V, *et al*: Genome-wide association study identifies five loci associated with lung function. Nat Genet 42: 36-44, 2009.
17. Rabe KF, Hurd S, Anzueto A, Barnes PJ, Buist SA, Calverley P, Fukuchi Y, Jenkins C, Rodriguez-Roisin R, van Weel C, Zielinski J; Global Initiative for Chronic Obstructive Lung Disease: Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease: GOLD executive summary. Am J Respir Crit Care Med 176: 532-555, 2007.
18. Gong Y, Fan L, Wan H, Shi Y, Shi G, Feng Y, Liu J, Ni L, Pan C and Zhang R: Lack of association between the TGF-β(1) gene and development of COPD in Asians: a case-control study and meta-analysis. Lung 189: 213-223, 2011.
19. Edwards MC and Gibbs RA: Multiplex PCR: advantages, development, and applications. Genome Res 3: S65-S75, 1994.
20. Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, Higgins J, DeFelice M, Lochner A, Faggart M, *et al*: The structure of haplotype blocks in the human genome. Science 296: 2225-2229, 2002.
21. Klein RJ, Zeiss C, Chew EY, Tsai JY, Sackler RS, Haynes C, Henning AK, SanGiovanni JP, Mane SM, Mayne ST, *et al*: Complement factor H polymorphism in age-related macular degeneration. Science 308: 385-389, 2005.
22. Cantor RM, Lange K and Sinsheimer JS: Prioritizing GWAS results: a review of statistical methods and recommendations for their application. Am J Hum Genet 86: 6-22, 2010.
23. Melen E, Himes BE, Brehm JM, Weiss ST and Lasky-Su J: Limited evidence for shared genetic factors between asthma and obesity based on known candidate genes. Am J Respir Crit Care Med 181: A1305, 2010.
24. Frayling TM: Genome-wide association studies provide new insights into type 2 diabetes aetiology. Nat Rev Genet 8: 657-662, 2007.
25. Hinney A, Nguyen TT, Scherag A, Friedel S, Brönner G, Müller TD, Grallert H, Illig T, Wichmann HE, Rief W, Schäfer H and Hebebrand J: Genome wide association (GWA) study for early onset extreme obesity supports the role of fat mass and obesity associated gene (FTO) variants. PLoS One 2: e1361, 2007.
26. The Psychiatric GWAS Consortium Steering Committee: A framework for interpreting genome-wide association studies of psychiatric disorders. Mol Psychiatry 14: 10-17, 2009.
27. Vogel T, Dittrich O, Mehraein Y, Dechend F, Schnieders F and Schmidtke J: Murine and human TSPYL genes: novel members of the TSPY-SET-NAP1L1 family. Cytogenet Cell Genet 81: 265-270, 1998.
28. von Lindern M, van Baal V, Wiegant J, Raap A, Hagemeijer A and Grosveld G: Can, a putative oncogene associated with myeloid leukemogenesis, may be activated by fusion of its 3' half to different genes: characterization of the set gene. Mol Cell Biol 12: 3346-3355, 1992.
29. Puffenberger EG, Hu-Lince D, Parod JM, Craig DW, Dobrin SE, Conway AR, Donarum EA, Strauss KA, Dunckley T, Cardenas JF, *et al*: Mapping of sudden infant death with dysgenesis of the testes syndrome (SIDDT) by a SNP genome scan and indentification of TSPYL loss of function. Proc Natl Acad Sci USA 101: 11689-11694, 2004.
30. Vinci G, Brauner R, Tar A, *et al*: Mutations in the TSPYL1 gene associated with 46,XY disorder of sex development and male infertility. Fertil Steril 92: 1347-1350, 2009.
31. Kim TY, Zhong S, Fields CR, Fields RC, Kim JH and Robertson KD: Epigenomic profiling reveals novel and frequent targets of aberrant DNA methylation-mediated silencing in malignant glioma. Cancer Res 66: 7490-7501, 2006.
32. Jung Y, Park J, Bang YJ and Kim TY: Gene silencing of TSPYL5 mediated by aberrant promoter methylation in gastric cancers. Lab Invest 88: 153-160, 2008.
33. Kim EJ, Lee SY, Kim TR, Choi SI, Cho EW, Kim KC and Kim IJ: TSPYL5 is involved in cell growth and the resistance to radiation in A549 cells via the regulation of p21(WAF1/Cip1) and PTEN/AKT pathway. Biochem Biophys Res Commun 392: 448-453, 2010.
34. Bitto E, Bingman CA, Wesenberg GE, McCoy JG and Phillips GN Jr: Structure of pyrimidine 5'-nucleotidase type 1. Insight into mechanism of action and inhibition during lead poisoning. J Biol Chem 281: 20521-20529, 2006.
35. Kim WJ, Hersh CP, DeMeo DL, Reilly JJ and Silverman EK: Genetic association analysis of COPD candidate genes with bronchodilator responsiveness. Respir Med 103: 552-557, 2009.
36. Zhu G, Warren L, Aponte J, Gulsvik A, Bakke P, Anderson WH, Lomas DA, Silverman EK, Pillai SG; International COPD Genetics Network (ICGN) Investigators: The SERPINE2 gene is associated with chronic obstructive pulmonary disease in two large population. Am J Respir Crit Care Med 176: 167-173, 2007.
37. DeMeo D, Mariani T, Lange C, Lake S, Litonjua A, Celedón J, Reilly J, Chapman HA, Sparrow D, Spira A, *et al*: The SERPINE2 gene is associated with chronic obstructive pulmonary disease. Proc Am Thorac Soc 3: 502, 2006.