



ORIGINAL ARTICLE

Genetic variation of six specific SNPs of chronic obstructive pulmonary disease among Chinese population

J. Jing^{a,b,#}, D. Xu^{b,#}, Z. Li^b, J. Wang^c, J. Dai^a, F.S. Li^{b,c,*}

^a The Fourth Clinical Medical College of Xinjiang Medical University, China

^b The COPD Laboratory of Clinical Research Base, Chinese Medicine Hospital of Xinjiang Uygur Autonomous Region, China

^c The Clinical Research Base Laboratory, Chinese Medicine Hospital of Xinjiang Uygur Autonomous Region, China

Received 5 March 2021; accepted 13 February 2022

Available online 29 April 2022

KEYWORDS

Genetic variation;
SNP;
Chronic obstructive pulmonary disease;
Chinese population;
Case-control study;
MDR analysis;
Single nucleotide polymorphisms

Abstract

Background: Chronic obstructive pulmonary disease (COPD) is a chronic bronchitis (or) emphysema with a high disability and fatality rate. This study aimed to explore the correlation between the six selected single nucleotide polymorphisms (SNPs) and the risk of COPD in the Chinese population.

Methods: The Agena MassARRAY platform was used to select six SNPs from 629 subjects for genotyping. The correlation between SNPs and COPD risk was evaluated using calculated odds ratios (ORs) and 95% confidence intervals (CIs). Multi-factor dimensionality reduction (MDR) was performed to analyze the impact of SNP interactions on COPD risk. Multiple comparisons were performed using Bonferroni- correction.

Results: Our results indicated that rs4719841 and rs7934083 variants were associated with a reduced risk of COPD. The analysis results of age, gender and non-smokers showed that rs4719841 and rs7934083 were associated with reducing the risk of COPD. In addition, the results showed that the genetic models of rs4719841, rs7934083 and rs7780562 were related to the forced vital capacity, respiratory rate per second, and respiratory rate / forced vital capacity of COPD patients, respectively. The results of the MDR analysis showed that the three-locus model (rs4719841, rs7934083, and rs78750958) is the best for COPD risk assessment.

Conclusion: This study shows that rs4719841 and rs7934083 are associated with the risk of COPD in the Chinese population, which provides some insights for early screening, prevention, and diagnosis of COPD in high-risk populations.

© 2022 Sociedade Portuguesa de Pneumologia. Published by Elsevier España, S.L.U. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

* Corresponding author.

E-mail address: fengsen602@163.com (F.S. Li).

These authors contributed equally to this work.

Introduction

Chronic obstructive pulmonary disease (COPD) is a common, preventable, and treatable chronic lung disease that is one of the leading causes of death globally, killing more than 3 million people each year.¹ It affects 329 million people worldwide, accounting for approximately 5% of the global population,² imposes a huge economic burden on the health care system and has a significant social impact on patients and their families. Progress has been made in treating symptoms and preventing acute exacerbations; however little progress has been made in improving disease progression or affecting mortality. Therefore, a better understanding of the complex disease mechanisms that lead to COPD is

urgently needed.³ In addition to smoking being a common cause of COPD, genetic factors play a significant role.⁴

Long-chain noncoding RNA molecules are approximately 200 nucleotides in length. They are well conserved, and defined as noncoding RNA molecules.⁵ However, some studies have shown that the peptide/protein encoded by ncRNAs.⁶ Lncrna is involved in many human diseases, such as prostate cancer,⁷ cardiovascular diseases⁸ and atherosclerosis.⁹ A ncRNA, *Loc105375199* is located at 7p15.2, which is located on the short arm of human chromosome 7 and has band 2 in zone 15, with a total of nine exons. At present, there are few reports on *Loc105375199*; however, many studies have suggested that the gene on 8p22 may be involved in the occurrence and development of various lung

Table 1 The basic information of the participants of the present study.

Characteristics	Cases (N = 315)	Controls (N = 314)	p-value
Age			0.306 ^a
> 70years	188(60%)	177(56%)	
≤ 70years	127(40%)	137(44%)	
Mean ±SD	72.1 ± 8.5	69.9 ± 8.8	
Gender			0.908 ^b
Male	239(76%)	237(75%)	
female	76(24%)	77(25%)	
Smoking status			
yes	147(47%)	52(17%)	0.057 ^b
no	166(53%)	118(38%)	0.156 ^b
Complication status			0.930 ^b
yes	93(30%)	0	
no	174(55%)	0	
Wheezing status			0.113 ^b
yes	153(49%)	0	
no	123(39%)	0	
Gasp			< 0.001 ^b
yes	115(37%)	0	
no	166(53%)	0	
Chest tightness status			0.445 ^b
yes	102(32%)	0	
no	179(57%)	0	
Recent respiratory infection status			0.788 ^b
yes	164(52%)	0	
no	117(37%)	0	
BMI			0.015 ^a
>24	251(80%)	67(21%)	
≤24	29(9%)	78(25%)	
Respiratory rate (per minute)	22.31±2.462		
Pulse rate (per inute)	86.33±11.675		
Forced vital capacity (L)	2.26±3.717		
Quantity per second (L)	1.22±0.600		
Rate per second (s)	39.91±29.329		
GOLD spirometric grade, n (%)			
I	34(10.8%)		
II	147(46.7%)		
III	102(32.4%)		
IV	32(10.2%)		

SD: Standard deviation; COPD:chronic obstructive pulmonary disease; BMI: Body mass index.

$p < 0.05$ indicates statistical significanc..

^a p value was calculated from independent sample Student's t-test.

^b p value was calculated from two-sided Chi-squared test;.

Table 2 Basic information of the selected SNPs in this study.

SNP_ID	Gene	Chr	Position	Allele A/B	MAF		p-HWE	OR(95%CI)	p-value	HaploReg
					Case	Control				
rs10245353	Adjacent location of <i>LOC105375199</i>	7	25,858,614	C/A	0.360	0.412	1.000	0.8 (0.64–1.01)	0.0558	SiPhy cons, DNase, Motifs changed, Selected eQTL hits
rs2290263	Adjacent location of <i>LOC105375199</i>	7	25,887,278	G/A	0.059	0.032	1.000	1.9 (1.09–3.31)	0.0219	Promoter histone marks, Enhancer histone marks, DNase, Proteins bound, Motifs changed
rs4719841	<i>LOC105375199</i>	7	25,997,536	A/G	0.308	0.427	0.491	0.6 (0.47–0.75)	<0.0001	Motifs changed
rs7780562	<i>LOC105375199</i>	7	26,022,414	A/C	0.271	0.311	0.895	0.83 (0.65–1.05)	0.1241	Motifs changed
rs7934083	<i>MIR4300HG</i>	11	81,492,479	T/C	0.106	0.200	0.053	0.47 (0.34–0.65)	<0.0001	Motifs changed
rs78750958	<i>MIR4300HG</i>	11	81,669,557	G/A	0.298	0.228	1.000	1.44 (1.12–1.86)	0.0044	DNase, Motifs changed

SNP: Single nucleotide polymorphism; HWE: Hardy-Weinberg equilibrium; OR: Odds ratio; 95% CI: 95% confidence interval; COPD: chronic obstructive pulmonary disease. HWE p-value obtained from Chi-squared test.

p values were calculated from Chi-squared test regarding to the allele distribution frequencies among COPD and healthy controls. Bonferroni's multiple adjustment was applied, with $p < 0.0017(0.05/30)$. Bold indicates a statistically significant SNP ($p < 0.0017$).

diseases such as COPD.^{10–14} Although there are no detailed reports on COPD and *Loc105375199*, previous studies suggest that *Loc105375199* may be involved in COPD occurrence. Therefore, it is necessary to study the relationship between the *Loc105375199* polymorphism and the risk of COPD to provide a new perspective for preventing, screening, and treating COPD.

MIR4300HG is a non-coding RNA located on human chromosome 11. We found very few reports regarding *MIR4300HG* and most of them were related to spine diseases.¹⁵ In addition, many studies have shown that many genes on chromosome 11 are associated with COPD. T Ishii et al. revealed that polymorphism of the glutathione S-transferase P1 (*GSTP1*) gene located on chromosome 11 is related to COPD.¹⁶ Paul Stoll et al. found that patients with COPD had high levels of BDNF in their serum, which is encoded by the *BDNF* gene on chromosome 11.¹⁷ Multiple studies have shown that the cytokine *IL18* located on chromosome 11 plays an important role in the inflammation in COPD.^{18–20} Therefore, we speculate that the *MIR4300HG* gene may be associated with the susceptibility to COPD.

To verify these speculations, we conducted a case-control study on the association between *Loc105375199/MIR4300HG* genetic polymorphisms and COPD susceptibility.

Methods

Study participants and sample collection

The project was approved by the medical ethics committee of the Chinese Medicine Hospital of Xinjiang Uygur Autonomous Region., All participants signed an informed consent form before the experiment. Data on environment, occupation, smoking history, tumors and general situation, according to the unified epidemiological questionnaire, was obtained from all patients. In this study, according to the GOLD (Global Chronic Obstructive Pulmonary Disease Action) report,²¹ 315 clinical patients diagnosed with COPD were included in the case group, of which 147 were smokers (smoking more than two cigarettes per day, or had quit smoking before enrollment) and, 166 people were non-smokers (never smoked). The diagnosis of COPD by hospital pulmonologists is based on clinical examination and spirometry. The severity of COPD is classified according to fixed forced expiratory volume in 1 s (FEV1) / forced vital capacity (FVC <0.7) and the predicted value of FEV1. The healthy controls were 314 healthy subjects from the same geographic area with normal lung volume according to the GOLD standard, including 52 smokers and 118 non-smokers. Subjects with other lung diseases, infections, inflammation, endocrine, gastrointestinal, liver, kidney and tumor diseases, and excessive alcohol consumption (≥ 40 g /day) were excluded. Peripheral venous blood (5 ml) was collected in a test tube containing EDTA, centrifuged, and stored at -80 °C for subsequent use.

SNP selection and genotyping

To verify the effectiveness of the genotyping of the selected population, the SNPs on chromosomes 11 and 7 were based on the smallest level according to the CHB (Beijing Han

Table 3 Relationships of the SNPs and the risk of COPD.

SNP_ID	Model	Genotypes	Case	Controls	With adjustment	
					OR (95% CI)	<i>p</i>
rs4719841	Codominant	G/G	145	100	1.00	
		G/A	146	160	0.62(0.44–0.87)	0.0060
		A/A	24	54	0.31(0.18–0.53)	<0.0001
	Dominant	G/G	145	100	1.00	
		G/A-A/A	170	214	0.54(0.39–0.75)	0.0002
	Recessive	G/G-G/A	291	260	1.00	
	A/A	24	54	0.4(0.24–0.66)	0.0004	
	Log-additive	–	–	0.57(0.45–0.73)	<0.0001	
rs7934083	Codominant	C/C	249	206	1.00	
		C/T	60	89	0.55(0.38–0.81)	0.0022
		T/T	3	18	0.13(0.04–0.46)	0.0015
	Dominant	C/C	249	206	1.00	
		C/T-T/T	63	107	0.48(0.34–0.7)	0.0001
	Recessive	C/C–C/T	309	295	1.00	
		T/T	3	18	0.16(0.05–0.54)	0.0032
	Log-additive	–	–	–	0.49(0.36–0.67)	<0.0001

SNP: Single nucleotide polymorphism; OR: Odds ratio; 95% CI: 95% confidence interval; COPD: chronic obstructive pulmonary disease.

p-values were calculated by unconditional logistic regression analysis with the comparison between COPD and healthy controls with adjustments for age and gender.

Bonferroni's multiple adjustment was applied, with $p < 0.0017(0.05/30)$.

Bold indicates a statistically significant SNP ($p < 0.0017$).

nationality) data in the Thousand Genome Project (<http://www.internationalgenome.org/>) SNPs with locus frequency less than 5% were initially screened. Subsequently, the HaploReg v4.1 (<https://pubs.broadinstitute.org/mammals/haploreg/haploreg.php>) was used to predict the potential functions of selected SNPs, which are related to promoter histone marks, enhancer histone marks, DNase, proteins bound and selected eQTL hits, indicating the mechanisms by which these SNPs may affect the disease occurrence and development. Finally, we used a few previously reported SNPs located adjacent to *LOC105375199* (rs10245353 and rs2290263), *Loc105375199* (rs4719841 and rs7780562) and *MIR4300HG* (rs7934083 and rs78750958) as candidate loci for this study. Genomic DNA was extracted according to the GoldMag® nanoparticles method and its concentration was measured using the NanoDrop 2000. We conducted Multiplexed SNP MassEXTEND experiment, SNP genotyping and data management and analysis using the Sequenom MassARRAY Assay Design 3.0 software, Sequenom MassARRAY RS1000, and Sequenom Typer 4.0 software, respectively, as reported previously.^{22,23}

Data analysis

To analyze the mass spectrometry results, the undetected, repeated data and the success rate of typing 90% of the sites or samples were deleted. This study used Welch's *t*-test to analyze the average age between the cases and control groups. Hardy Weinberg equilibrium (HWE) analysis is a method of quality control of the genetic data of the control group. Theoretically, the control group should follow the Hardy Weinberg equilibrium, $p > 0.05$, indicating that the control group reaches genetic balance, ensuring

that the data of the control group is credible. Fisher precision test and Wald test were used to compare the allele/genotype frequency of each locus between the cases and the controls, and genetic model analysis was performed using SNPStats software (<http://bioinfo.iconco-logia.net/snpstats/start.htm>) to assess the association between the SNPs and COPD risk. Multi-factor dimensionality reduction (MDR) analysis is a nonparametric, nongenetic approach suitable for case-control studies that only require individual genetic data (such as SNPs) and no other special conditions to analyze gene-gene interactions are needed.^{24,25} One-way analysis and ANOVA were used to analyze the differences in clinical characteristics between the different genetic models. Multiple comparisons were Bonferroni-corrected.²⁶

Results

Study subjects

The basic information of the participants in this study is shown in Table 1. A total of 629 subjects, 315 and 314 patients in the COPD and control groups, respectively, were included in this study. There were 239 men and 76 women in the COPD group and 237 men and 77 women in the control group. There were no significant differences in age, gender, smoking, complication, wheezing, chest tightness, recent respiratory infection status and BMI between the two groups. Detailed information such as the average value and standard deviation value of patients' respiratory rate (per minute), pulse rate (per minute), forced vital capacity (L), quantity per second (L), and rate per second (s) are shown in Table 1.

Table 4 Stratified analysis of the age and gender on association between the SNPs and COPD risk.

SNP	Model	Genotype	> 70		≤ 70		Male		Female	
			OR (95%CI)	<i>p</i> -value						
rs4719841	Allele	G	1		1		1		1	
		A	0.65(0.48–0.88)	0.0055	0.52(0.36–0.74)	0.0004	0.6(0.46–0.78)	0.0001	0.6(0.37–0.95)	0.0305
	Codominant	G/G	1		1		1		1	
		G/A	0.61(0.39–0.98)	0.0411	0.71(0.42–1.22)	0.2136	0.64(0.43–0.95)	0.0272	0.55(0.27–1.1)	0.0884
		A/A	0.47(0.23–0.96)	0.0377	0.18(0.07–0.48)	0.0006	0.3(0.16–0.55)	0.0001	0.33(0.11–1)	0.0498
	Dominant	G/G	1		1		1		1	
		G/A-A/A	0.58(0.37–0.91)	0.0184	0.56(0.33–0.92)	0.0233	0.55(0.38–0.8)	0.0018	0.5(0.26–0.97)	0.0403
	Recessive	G/G-G/A	1		1		1		1	
A/A		0.63(0.33–1.21)	0.1619	0.21(0.08–0.55)	0.0014	0.38(0.21–0.68)	0.0011	0.46(0.16–1.31)	0.1457	
rs7934083	Log-additive	–	0.67(0.48–0.93)	0.0154	0.52(0.35–0.77)	0.0010	0.57(0.43–0.76)	0.0001	0.56(0.34–0.93)	0.0259
		Allele	C	1		1		1		1
	Codominant	T	0.45(0.29–0.68)	0.0001	0.51(0.31–0.85)	0.0080	0.55(0.38–0.78)	0.0007	0.23(0.1–0.54)	0.0003
		C/C	1		1		1		1	
		C/T	0.49(0.29–0.83)	0.0077	0.58(0.32–1.05)	0.0712	0.66(0.43–1)	0.0501	0.25(0.1–0.64)	0.0036
	Dominant	T/T	0.11(0.02–0.54)	0.0063	0.13(0.02–1.1)	0.0606	0.16(0.05–0.58)	0.0050	–	0.9988
		C/C	1		1		1		1	
	Recessive	C/T-T/T	0.42(0.25–0.69)	0.0007	0.51(0.28–0.9)	0.0207	0.57(0.38–0.85)	0.0058	0.22(0.09–0.55)	0.0012
C/C–C/T		1		1		1		1		
Log-additive	T/T	0.14(0.03–0.65)	0.0126	0.15(0.02–1.24)	0.0788	0.18(0.05–0.65)	0.0082	–	0.9988	
	–	0.43(0.28–0.67)	0.0002	0.51(0.3–0.85)	0.0094	0.56(0.39–0.79)	0.0010	0.23(0.09–0.56)	0.0012	

SNP: single-nucleotide polymorphism; OR, odds ratio; CI, confidence interval; COPD: chronic obstructive pulmonary disease.

Bonferroni's multiple adjustment was applied, with $p < 0.0017(0.05/30)$.

Bold indicates a statistically significant SNP ($p < 0.0017$).

Table 5 Stratified analysis of the smoking on association between the SNPs and COPD risk.

SNP	Model	Genotype	Yes		No	
			OR (95%CI)	p-value	OR (95%CI)	p-value
rs4719841	Allele	G	1		1	
		A	0.65(0.41–1.03)	0.0631	0.47(0.33–0.66)	<0.0001
	Codominant	G/G	1		1	
		G/A	0.86(0.42–1.77)	0.6893	0.39(0.23–0.67)	0.0006
		A/A	0.31(0.11–0.84)	0.0218	0.20(0.09–0.46)	0.0001
	Dominant	G/G	1		1	
		G/A-A/A	0.69(0.35–1.37)	0.2887	0.34(0.2–0.57)	<0.0001
	Recessive	G/G-G/A	1		1	
		A/A	0.34(0.14–0.84)	0.0190	0.34(0.16–0.73)	0.0050
	Log-additive	–	0.61(0.37–1)	0.0507	0.43(0.29–0.63)	<0.0001

SNP: single-nucleotide polymorphism; OR, odds ratio; CI, confidence interval; COPD:chronic obstructive pulmonary disease. Bonferroni's multiple adjustment was applied, with $p < 0.0017(0.05/30)$. Bold indicates a statistically significant SNP ($p < 0.0017$).

Association between loc105375199/ mir4300hg SNPs and COPD risk

The basic information of all the SNPs is shown in Table 2, including SNP-ID, gene, chromosomes, position, alleles, minor allele frequency (MAF), p-HWE, OR (95%CI) and p-value. After Bonferroni correction, the results showed that individuals with the A allele of rs4719841 had a lower risk of COPD than those with the G allele ($p < 0.0001$). Individuals with the G allele of rs78750958 had a lower risk of COPD than those with the A allele ($p < 0.0001$). HaploReg software function prediction showed that the selected SNPs (rs10245353 rs2290263, rs4719841, rs7780562, rs7934083 and rs78750958) are related to siPhy cons, DNase, motif changes, selected eQTL hits, promoter histone marks, enhancer histone marks, and proteins, suggesting that these SNPs may affect the occurrence and development of the disease. A multi-genetic model was used to analyze the correlation between the selected SNPs and the risk of COPD, and

the results showed that rs4719841 and rs7934083 were related to the risk of COPD ($p < 0.0017$) (Table 3). Specifically, the A/A genotype of rs4719841 was observed to reduce the risk of COPD. In the co-dominant, dominant and additive models, rs7934083 reduced the risk of COPD.

Stratified analysis of the impact of selected SNPs on the risk of COPD

To further clarify the correlation between the six selected SNPs and the risk of COPD, we conducted 10 stratified analyses, including stratification by age, gender, complications, BMI, asthma, recent respiratory infection, smoking, wheezing, chest tightness, and GOLD spirometric grade (Supplementary Table 1). The results showed that COPD risk due to some of the selected SNPs was associated with age, gender, and non-smoking categories, whereas in case of other selected SNPs, COPD risk was not associated with stratified analyses.

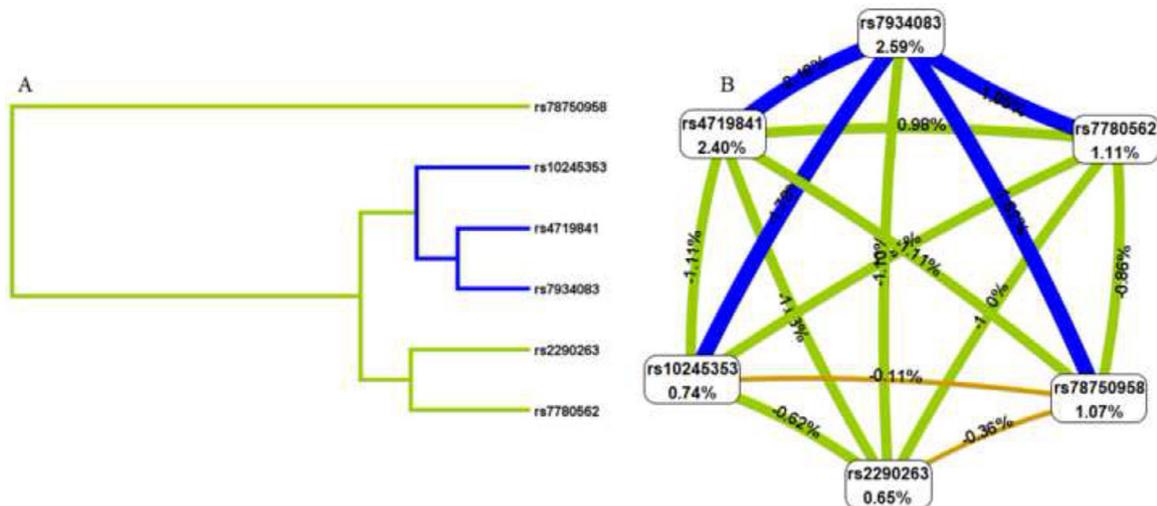


Figure 1 The dendrogram (A) and fruchterman Rheingold (B) of SNP-SNP interaction for COPD risk. (A) Short connections among nodes represent stronger redundant interactions. (B) Negative percent entropy indicates redundancy.

Table 6 Clinical characteristics of patients based on the genetic models of the SNPs.

SNP_ID	Genotypes	Characteristics									
		Respiratory rate (per minute)	p	Pulse rate (per minute)	p	Forced vital capacity (L)	p	Quantity per second (L)	p	Rate per second (s)	p
rs4719841	G/G	22.11±2.26	0.324	85.75±10.66	0.252	1.92±0.65	<0.001	1.20±0.54	0.289	40.64±29.92	0.322
	G/A	22.41±2.57		87±12.06		2.05±0.72		1.26±0.66		39.84±28.67	
	A/A	22.85±2.91		85.55±15.26		6.1 ± 14.07		1.09±0.49		35.98±32.22	
rs7934083	C/C	21.67±2.08	0.889	82.00±13.11	0.546	1.83±0.00	0.628	0.80±0.00	0.725	0.44±0.00	0.001
	C/T	22.57±2.47		86.31±9.72		2.04±0.80		1.15±0.54		48.22±24.04	
	T/T	22.25±2.47		86.38±12.13		2.33±4.13		1.24±0.61		38.01±30.11	
rs7780562	C/C	22.07±2.23	0.002	85.77±10.91	0.668	1.93±0.68	<0.001	1.16±0.51	0.130	40.68±30.07	0.737
	C/A	22.65±2.80		87.06±12.44		2.08±0.68		1.35±0.69		38.41±29.21	
	A/A	21.83±1.50		84.28±12.6		6.52±14.76		1.02±0.44		43.27±29.09	

SNP: single-nucleotide polymorphism; COPD: chronic obstructive pulmonary disease; “-” indicates data missing; Bonferroni’s multiple adjustment was applied, with $p < 0.05$. Bold indicates a statistically significant SNP ($p < 0.05$).

In subjects aged > 70 years, we found that rs7934083 was associated with reducing COPD risk ($p < 0.0017$), while in subjects aged ≤ 70 years, rs4719841 was associated with a reduction in the risk of COPD ($p < 0.0017$) (Table 4). The results of the male stratification analysis showed that rs4719841 reduced the risk of COPD in the allele, collinearity model, additive model and recessive model ($p < 0.0017$). The results of the male stratification analysis also showed that individuals with the T allele rs7934083 had a lower risk of COPD than individuals with the C allele ($p = 0.0007$). In the female stratification analysis, we only observed that rs7934083 was associated with a reduced risk of COPD ($p < 0.0017$). In the stratified analysis of non-smokers, we observed that the other five selected SNPs were not related to the risk of COPD except for rs4719841, which is related to a reduced risk of COPD (Table 5).

MDR analysis

To better understand the impact of selected SNP interactions on COPD risk, we conducted MDR analysis (Supplementary Table 2 and Fig. 1). The dendrogram and frauctman-reiringold represent the interactions between these SNPs. In Fig. 1(A), the connection between nodes is shorter, indicating that the redundant interaction is stronger. A negative double-track entropy in Fig. 1(B) indicates that it is an antagonistic effect. A positive value indicates a synergistic effect. The results showed that there was a specific antagonism between the selected genes. Supplementary Table 2 shows the model combination of candidate SNP interactions for the COPD risk assessment. Among the many locus models, the best combination is a three-locus model consisting of rs4719841, rs7934083 and rs78750958, which increases the risk of COPD (test accuracy=0.618, CVC=10/10, OR=2.93, 95% CI: 2.11 – 4.05, $p < 0.001$). **3.4 3.5 Genetic models and clinical characteristics.**

In addition, we explored the relationship between different SNP models and the clinical indices of patients, including respiratory rate (times per minute), pulse rate (times per minute), forced vital capacity (L), one-second volume (L), and one-second rate (S). There was no significant correlation between the genetic models of rs10245353, rs2290263, and rs78750958 and the above clinical parameters ($p > 0.05$). However, the genetic models of rs4719841, rs7934083, and rs7780562 are related to forced vital capacity, rate per second, and respiratory rate/forced vital capacity of COPD patients ($p < 0.05$), as shown in Table 6.

Discussion

Based on a case-control study, we investigated the relationship between the six SNPs (rs10245353, rs2290263, rs78750958, rs4719841, rs7934083, and rs7780562) and susceptibility to COPD in the Chinese population. In the overall analysis, only two SNPs (rs4719841, rs7934083) were found to be associated with susceptibility to COPD. In multiple stratified analyses, we observed that only some selected SNPs were associated with COPD risk in the stratified analysis of age, gender, and non-smoking. The results of the MDR analysis showed that the three-locus model comprising rs4719841, rs7934083 and rs78750958 is the best model for COPD risk

assessment. In the analysis of clinical indicators, the genetic models of rs4719841, rs7934083, and rs7780562 were associated with forced vital capacity, respiratory rate per second, and respiratory rate/forced vital capacity, respectively.

COPD is chronic bronchitis and emphysema with airflow obstruction characteristics. It is currently one of the leading causes of chronic morbidity and mortality.²⁷ Studies have shown that genetic factors play a key role in the occurrence and development of COPD.²⁸ Genetic association research is an important way to understand the pathogenesis and epidemiology of COPD. A study by Almira Akparova et al. showed that the *EPHX1 Y113H* polymorphism is associated with an increased risk of COPD in the Kazakhstan population.²⁹ Studies have also shown that vitamin D binding protein (*VDBP*), promoter regions of matrix metalloproteinases, cyclooxygenase-2 (*COX-2*), and angiotensin converting enzyme gene (*ACE*) polymorphisms are associated with COPD susceptibility.^{30–33}

Non-coding RNAs are associated with various diseases.^{7–9} However, even searching many databases, no report of *Loc105375199*, a non-coding RNA, was found. Our study showed that the polymorphism of *Loc105375199* is associated with susceptibility to COPD; that is, rs4719841 in the *Loc105375199* gene is associated with a reduction in susceptibility to COPD. This is the first report on *Loc105375199* and COPD, which provides a unique perspective for future studies on *Loc105375199* and COPD and a new direction for follow-up studies.

A previous study only involved rs10245353, located at the locus adjacent to *Loc105375199*, which indicated that rs10245353 mutation may increase the risk of gestational diabetes mellitus (GDM) in the Chinese population.³⁴ There was no statistically significant relationship between rs10245353 mutation and COPD risk in our results, which may also be because our sample size was not large enough; this should be verified by further expanding sample size and mechanism studies.

MIR4300HG is a host gene for miRNAs. There is only one report on *MIR4300HG*, indicating that the functional variation of *MIR4300HG* is related to the progression of adolescent idiopathic scoliosis (AIS).¹⁵ Our study showed that rs7934083 on *MIR4300HG* was associated with a reduced risk of COPD, which also verified our hypothesis.

This study has several limitations. First, the sample size of this study was not large enough. Therefore, we still need to expand the sample size to verify the results further. Second, this study only involved related studies, and cell and animal experiments are needed to clarify the specific mechanism further.

Conclusion

This study is the first to report that the polymorphisms of *Loc105375199* and *MIR4300HG* are associated with COPD susceptibility, which provides a new direction for the study of COPD and is more conducive to the promotion of personalized medicine and pharmacogenomics.

Declaration of Competing Interest

None.

Acknowledgments

We are very grateful to all the participants in this study.

Funding

This study was supported by Special project for University Research Program of Xinjiang (XJEDU2019I017), Business Construction of National TCM Clinical Research Base of State Administration of Traditional Chinese Medicine (*JDZX2015233*), National Natural Science Foundation of China (81960848).

Supplementary materials

Supplementary material associated with this article can be found in the online version at [doi:10.1016/j.pulmoe.2022.02.004](https://doi.org/10.1016/j.pulmoe.2022.02.004).

References

- Hattab Y, Alhassan S, Balaan M, Lega M, Singh AC. Chronic obstructive pulmonary disease. *Crit Care Nurs Q*. 2016;39(2):124–30. <https://doi.org/10.1097/cnq.000000000000105>.
- Vos T, Flaxman AD, Naghavi M, Lozano R, Michaud C, Ezzati M. Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990–2010: a systematic analysis for the global burden of disease study 2010. *Lancet*. 2012;380(9859):2163–96. [https://doi.org/10.1016/s0140-6736\(12\)61729-2](https://doi.org/10.1016/s0140-6736(12)61729-2).
- Rabe KF, Watz H. Chronic obstructive pulmonary disease. *Lancet*. 2017;389(10082):1931–40. [https://doi.org/10.1016/s0140-6736\(17\)31222-9](https://doi.org/10.1016/s0140-6736(17)31222-9).
- Decramer M, Janssens W, Miravittles M. Chronic obstructive pulmonary disease. *Lancet*. 2012;379(9823):1341–51. [https://doi.org/10.1016/s0140-6736\(11\)60968-9](https://doi.org/10.1016/s0140-6736(11)60968-9).
- Mercer TR, Dinger ME, Mattick JS. Long non-coding RNAs: insights into functions. *Nat Rev Genet*. 2009;10(3):155–9. <https://doi.org/10.1038/nrg2521>.
- Zhu S, Wang J, He Y, Meng N, Yan GR. Peptides/proteins encoded by non-coding rna: a novel resource bank for drug targets and biomarkers. *Front Pharmacol*. 2018;9:1295. <https://doi.org/10.3389/fphar.2018.01295>.
- Bai J, Huang G. Role of long non-coding RNA NEAT1 in the prognosis of prostate cancer patients. *Medicine*. 2020;99(22). <https://doi.org/10.1097/md.00000000000020204>. e20204.
- Costa SEMC, Kurc S, Drożdż A, Cortez-Dias N, Enguita FJ. The circulating non-coding RNA landscape for biomarker research: lessons and prospects from cardiovascular diseases. *Acta Pharmacol Sin*. 2018;39(7):1085–99. <https://doi.org/10.1038/aps.2018.35>.
- Paloschi V, Winter H, Viola J, Soehnlein O, Maegdefessel L. Mechanistic links between non-coding RNAs and myeloid cell inflammation in atherosclerosis. *Thromb Haemost*. 2019;119(8):1205–11. <https://doi.org/10.1055/s-0039-1687874>.
- Fernandez Fernandez E, De Santi C, De Rose V, Greene CM. CFTR dysfunction in cystic fibrosis and chronic obstructive pulmonary disease. *Expert Rev Respir Med*. 2018;12(6):483–92. <https://doi.org/10.1080/17476348.2018.1475235>.
- Siganaki M, Koutsopoulos AV, Neofytou E, Vlachaki E, Psarrou M, Soultz N. Deregulation of apoptosis mediators' p53 and bcl2 in lung tissue of COPD patients. *Respir Res*. 2010;11(1). <https://doi.org/10.1186/1465-9921-11-46>. 46.

12. Stolarczyk M, Scholte BJ. The EGFR-ADAM17 axis in chronic obstructive pulmonary disease and cystic fibrosis lung pathology. *Mediators Inflamm.* 2018;1067134. <https://doi.org/10.1155/2018/1067134>. 2018.
13. Wang XW, Zhang YL, Xiong Y. Impact of ABCB1 single-nucleotide polymorphisms on treatment outcomes with salmeterol/fluticasone combination therapy for stable chronic obstructive pulmonary disease. *Genet Test Mol Biomarkers.* 2015;19(10):566–72. <https://doi.org/10.1089/gtmb.2015.0108>.
14. Xu X, Wang H, Li H, Cui X, Zhang H. SERPINE1 -844 and -675 polymorphisms and chronic obstructive pulmonary disease in a Chinese Han population. *J Int Med Res.* 2016;44(6):1292–301. <https://doi.org/10.1177/0300060516664270>.
15. Ogura Y, Kou I, Takahashi Y, Takeda K, Minami S, Kawakami N. A functional variant in MIR4300HG, the host gene of microRNA MIR4300 is associated with progression of adolescent idiopathic scoliosis. *Hum Mol Genet.* 2017;26(20):4086–92. <https://doi.org/10.1093/hmg/ddx291>.
16. Ishii T, Matsuse T, Teramoto S, Matsui H, Miyao M, Hosoi T. Glutathione S-transferase P1 (GSTP1) polymorphism in patients with chronic obstructive pulmonary disease. *Thorax.* 1999;54(8):693–6. <https://doi.org/10.1136/thx.54.8.693>.
17. Stoll P, Wuertemberger U, Bratke K, Zingler C, Virchow JC, Lommatzsch M. Stage-dependent association of BDNF and TGF- β 1 with lung function in stable COPD. *Respir Res.* 2012;13(1):116. <https://doi.org/10.1186/1465-9921-13-116>.
18. De Falco G, Colarusso C, Terlizzi M, Popolo A, Pecoraro M, Commodo M. Chronic obstructive pulmonary disease-derived circulating cells release IL-18 and IL-33 under ultrafine particulate matter exposure in a caspase-1/8-independent manner. *Front Immunol.* 2017;8:1415. <https://doi.org/10.3389/fimmu.2017.01415>.
19. Dima E, Koltsida O, Katsounou P, Vakali S, Koutsoukou A, Koulouris NG. Implication of Interleukin (IL)-18 in the pathogenesis of chronic obstructive pulmonary disease (COPD). *Cytokine.* 2015;74(2):313–7. <https://doi.org/10.1016/j.cyto.2015.04.008>.
20. Sayan M, Mossman BT. The NLRP3 inflammasome in pathogenic particle and fibre-associated lung inflammation and diseases. *Part Fibre Toxicol.* 2016;13(1):51. <https://doi.org/10.1186/s12989-016-0162-4>.
21. Mirza S, Clay RD, Koslow MA, Scanlon PD. COPD guidelines: a review of the 2018 GOLD report. *Mayo Clin Proc.* 2018;93(10):1488–502. <https://doi.org/10.1016/j.mayocp.2018.05.026>.
22. Gabriel S, Ziaugra L, Tabbaa D. SNP genotyping using the Sequenom MassARRAY iPLEX platform. *Curr Protoc Hum Genet.* 2009;2. <https://doi.org/10.1002/0471142905.hg0212s60>. Unit 2.12.
23. Thomas RK, Baker AC, Debiasi RM, Winckler W, Laframboise T, Lin WM. High-throughput oncogene mutation profiling in human cancer. *Nat Genet.* 2007;39(3):347–51. <https://doi.org/10.1038/ng1975>.
24. Gola D, Mahachie John JM, van Steen K, König IR. A roadmap to multifactor dimensionality reduction methods. *Brief Bioinform.* 2016;17(2):293–308. <https://doi.org/10.1093/bib/bbv038>.
25. Ritchie MD, Hahn LW, Roodi N, Bailey LR, Dupont WD, Parl FF. Multifactor-dimensionality reduction reveals high-order interactions among estrogen-metabolism genes in sporadic breast cancer. *Am J Hum Genet.* 2001;69(1):138–47. <https://doi.org/10.1086/321276>.
26. Sjölander A, Vansteelandt S. Frequentist versus Bayesian approaches to multiple testing. *Eur J Epidemiol.* 2019;34(9):809–21. <https://doi.org/10.1007/s10654-019-00517-2>.
27. Stanojkovic I, Kotur-Stevuljevic J, Milenkovic B, Spasic S, Vujic T, Stefanovic A. Pulmonary function, oxidative stress and inflammatory markers in severe COPD exacerbation. *Respir Med.* 2011;105. [https://doi.org/10.1016/s0954-6111\(11\)70008-7](https://doi.org/10.1016/s0954-6111(11)70008-7). Suppl 1531-7.
28. Marciniak SJ, Lomas DA. What can naturally occurring mutations tell us about the pathogenesis of COPD? *Thorax.* 2009;64(4):359–64. <https://doi.org/10.1136/thx.2008.099408>.
29. Akparova A, Abdrakhmanova B, Banerjee N, Bersimbaev R. EPHX1 Y113H polymorphism is associated with increased risk of chronic obstructive pulmonary disease in Kazakhstan population. *Mutat Res Genet Toxicol Environ Mutagen.* 2017; 1–6. <https://doi.org/10.1016/j.mrgentox.2017.02.004>. 816-817.
30. Chen LJ, Xu W, Taooka Y, Ohe M, Takahashi H, Sutani A. Cyclooxygenase 2 1195 G >A polymorphism is associated with chronic obstructive pulmonary disease in Japanese and Chinese patients. *Chin Med J.* 2013;126(12):2215–21.
31. Jung JY, Choi DP, Won S, Lee Y, Shin JH, Kim YS. Relationship of vitamin D binding protein polymorphisms and lung function in Korean chronic obstructive pulmonary disease. *Yonsei Med J.* 2014;55(5):1318–25. <https://doi.org/10.3349/ymj.2014.55.5.1318>.
32. Kanazawa H, Okamoto T, Hirata K, Yoshikawa J. Deletion polymorphisms in the angiotensin converting enzyme gene are associated with pulmonary hypertension evoked by exercise challenge in patients with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med.* 2000;162:1235–8. <https://doi.org/10.1164/ajrccm.162.4.9909120>. 4 Pt 1.
33. Korytina GF, Akhmadishina LZ, Ianbaeva DG, Viktorova TV. Polymorphism in promoter regions of matrix metalloproteinases (MMP1, MMP9, and MMP12) in chronic obstructive pulmonary disease patients. *Genetika.* 2008;44(2):242–9.
34. Yan J, Su R, Ao D, Wang Y, Wang H, Yang H. Genetic variants and clinical relevance associated with gestational diabetes mellitus in Chinese women: a case-control study. *J Matern Fetal Neonatal Med.* 2018;31(16):2115–21. <https://doi.org/10.1080/14767058.2017.1336225>.