A functional SNP upstream of the ADRB2 gene is associated with COPD

Jin-Xiu Li,1,2,8 Wei-Ping Fu,3,8 Jing Zhang,1 Xiao-Hua Zhang,1,2 Chang Sun,1,5 Lu-Ming Dai,3 Li Zhong,1,5,6 Li Yu,1,2 Ya-Ping Zhang1,7

1State Key Laboratory for Conservation and Utilization of Bio-Resource in Yunnan, Key Laboratory for Animal Genetic Diversity and Evolution of High Education in Yunnan Province, School of Life Sciences, Yunnan University, 2Department of Respiratory Critical Care Medicine, 3Department of Thoracic Surgery, The First Affiliated Hospital of Kunming Medical University, Kunming, 4College of Life Sciences, 5Provincial Demonstration Center for Experimental Biology Education, Shaanxi Normal University, Xi’an, 6State Key Laboratory of Genetic Resources and Evolution, and Yunnan Laboratory of Molecular Biology of Domestic Animals, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, China

*These authors contributed equally to this work

Correspondence: Ya-Ping Zhang
State Key Laboratory of Genetic Resources and Evolution, and Yunnan Laboratory of Molecular Biology of Domestic Animals, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming 650091, China Tel +86 871 6513 0513 Email zhangyp@mail.kiz.ac.cn

Li Zhong
State Key Laboratory for Conservation and Utilization of Bio-Resource in Yunnan, Yunnan University, Kunming 650091, China Tel +86 29 8531 0469 Email lizhong@snnu.edu.cn

Introduction

COPD, one of the most common respiratory diseases in old people, is characterized by airflow limitation, that is, a chronic persistent inflammatory process that is not fully reversible. Nowadays, COPD has become the third source of morbidity in the world. Although tobacco smoking has been suggested to be the predominant environmental factor for COPD, only ~10%~20% smokers develop airway obstruction. This phenomenon, together with the familial clustering in COPD patients, indicates that genetic factors might play an important role in the development of COPD. Recent genome-wide association studies have identified multiple COPD susceptibility genes, for example, Hedgehog interacting protein (HHIP). So far, only α1-antitrypsin (SERPINA1) has been confirmed to be a genetic risk factor for COPD. However, the mutant protease inhibitor Z homozygote of the gene, which could increase individual susceptibility to COPD, is extremely rare across worldwide populations (0.001%–4.5%), especially in Asians (~0.004%), and accounts for only 2% of COPD patients. Thus, additional genes were assumed to also play a crucial role in the predisposition to COPD and remained to be identified.

A major cause for COPD is airflow obstruction in the lung and respiratory parenchyma maintained by airway smooth muscle cells. Chronic obstructive abnormality followed by airway remodeling increases the thickness of the airway and causes airflow limitation. Chronic obstructive abnormality followed by airway remodeling increases the thickness of the airway and causes airflow limitation.
obstruction. Therefore, genes involved in the regulation of airway smooth muscle tone are good candidates for the genetic predisposition to COPD.

β₂-adrenergic receptor (ADRB2) is a G protein-coupled transmembrane receptor located on airway smooth muscle cells, and increase in ADRB2 gene expression was observed in COPD patients compared to patients with mild/moderate asthma, leading to the speculation that it is a candidate gene for COPD. Thus, many previous studies have focused on the relationship between single nucleotide polymorphisms (SNPs) in the coding region of ADRB2 and COPD in the Asian population, for example, rs1042713 (Arg16Gly) and rs1042714 (Gln27Glu), since these SNPs are thought to be affecting the construction of ADRB2 or altering the function of the translation product. However, there is still controversy in this issue, and more genotyping data and a meta-analysis are indispensable to resolve this conflict.

In addition, growing knowledge has shown that noncoding SNPs, especially the ones in the promoter region, are more important and might be functional, for example, affecting transcription factors’ (TFs) binding affinity and further influencing gene expression. Thus, investigation of the variants in the promoter of ADRB2 is valuable not only to focus on how the gene expression of ADRB2 is regulated, but also to discover the causal molecular mechanism of COPD. There are several SNPs, rs1801704 (−20T/C), rs1042711 (−477T/C), rs11959427 (−367T/C), rs11168070 (−468C/G), rs12654778 (−654G/A), rs2053044 (−1023G/A), rs2400707 (−1343A/G) and rs2895795 (−1429T/A), identified in the promoter region of ADRB2, which contain a number of putative regulatory elements. Previous studies focused on rs1042711 and found that this SNP could introduce a non-conservative amino acid change (Arg→Cys) at the 19th amino acid, further influencing the gene expression of ADRB2, since it lies within a 19 amino acid peptide (referred to as β upstream peptide) in the 5’ leader region. However, the crucial issue of whether other SNPs affect the expression of ADRB2 or the onset of COPD has never been scrutinized.

Due to the pivotal role of ADRB2 in lung function, we hypothesized that noncoding SNPs of ADRB2 could be important for its expression. To address this hypothesis, we validated whether its expression could be associated with COPD. Subsequently, a comprehensive evaluation to identify the causal variants and investigate the functional impact of this SNP on COPD pathogenesis was undertaken. This study would provide new insight into the potential molecular basis for COPD.

Materials and methods
Study population and lung tissues
A total of 422 adult subjects (200 unrelated patients with COPD and 222 healthy smokers) were recruited from the First Affiliated Hospital of Kunming Medical University (Kunming, China) for genotyping. All smokers belonged to Han nationality to minimize the potential sampling bias due to population stratification. COPD patients were diagnosed based on the results from multiple examinations, including the ratio of forced expiratory volume in 1 second/forced vital capacity (FEV₁/FVC ratio <70% and FEV₁ <80% predicted), according to the Global Initiative for Chronic Obstructive Lung Disease criteria. The healthy smokers exhibited normal pulmonary function (FEV₁/FVC ratio >70% and FEV₁ >80% predicted) and a smoking history of ≥10 pack-years. In addition, they were excluded from the possibility of COPD by chest computed tomography (CT).

To investigate whether the expression of ADRB2 could be related with COPD, human lung tissue samples from 18 COPD patients as case subjects (FEV₁ <80%) and 22 control subjects with normal lung function were also collected from the same hospital.

This study was approved by the institutional ethics committee of the First Affiliated Hospital of Kunming Medical University, and all participants were contacted by telephone to obtain verbal informed consent. Detailed information of patients and healthy smokers is presented in Table 1.

Transcription analysis
Total RNAs were isolated by Trizol (Thermo Fisher Scientific, Waltham, MA, USA) from human lung tissues stored in RNALater (Thermo Fisher Scientific) solution. cDNA was synthesized by SuperScript® III First-Strand Synthesis

Table 1 Demographic characteristics of the study subjects

<table>
<thead>
<tr>
<th>Subject Characteristics</th>
<th>DNA sample</th>
<th>RNA sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Case</td>
<td>Control</td>
</tr>
<tr>
<td>Subjects (n)</td>
<td>200</td>
<td>222</td>
</tr>
<tr>
<td>Male (%)</td>
<td>81.5</td>
<td>72.1</td>
</tr>
<tr>
<td>Age (years)</td>
<td>69.9±9.8</td>
<td>67.9±8.8</td>
</tr>
<tr>
<td>Smoking (pack-years)</td>
<td>30.6±8.6</td>
<td>28.8±7.0</td>
</tr>
<tr>
<td>FEV₁/FVC</td>
<td>49.0±16.1</td>
<td>86.0±6.1</td>
</tr>
<tr>
<td>FEV₁,% predicted</td>
<td>54.0±17.5</td>
<td>95.0±8.7</td>
</tr>
<tr>
<td>FEV₁,% &gt;50% predicted</td>
<td>112</td>
<td>0</td>
</tr>
<tr>
<td>FEV₁,% ≥50% predicted</td>
<td>88</td>
<td>0</td>
</tr>
</tbody>
</table>

Notes: Data are presented as mean ± SD. *Data available for 155 patients; †Data available for 150 subjects.

Abbreviations: FEV₁, forced expiratory volume in 1 second; FVC, forced vital capacity.
System (Thermo Fisher Scientific). Transcript levels for ADRB2 gene in the lung were measured by real-time polymerase chain reaction (PCR) with SYBR green (Kapa Biosystems, Wilmington, MA, USA) and primers (forward primer: 5′-AGGCAAGCTTCCAGAAGATTG-3′ and reverse primer: 5′-CCAGCAGAGGGTGAAAGTG-3′). All samples were tested on ABI PRISM® 7000 Sequence Detection System (Thermo Fisher Scientific) with three replications under the following cycling conditions: 2 min at 50°C, 10 min at 95°C, 40 cycles of 15 s at 95°C and 1 min at 60°C. The average cycle number at threshold (Ct) was normalized by glyceraldehyde-3-phosphate dehydrogenase (GAPDH). The expression level of ADRB2 was calculated based on the 2-ΔΔCt method.

Tag SNPs selection
SNPs within the ~8 kb region (chr5: 148203156–148211197, relative to build 37) containing the entire ADRB2 gene in East Asian population (Han Chinese in Beijing, China [CHB] and Japanese in Tokyo, Japan [JPT], n=89) were downloaded from 1,000 genomes project (www.internationalgenome.org/). 27 Tag SNPs were chosen by ldSelect software with r2>0.8. 28 Among the 32 SNPs identified in the East Asian population, eight tag SNPs were chosen for genotyping (Supplementary Figure 1).

Sample size power calculation
To assess whether our sample size for the association study was enough, genetic power calculation was used in this study. 29 with 7.3% disease prevalence in the Chinese population 30 and α=0.05, for a variant with 0.05 minor allele frequency in a dominant model. In the calculation of genetic power, ≥80% is a general threshold for acceptable power.

Genotyping of tag SNPs in ADRB2
Genomic DNA was extracted from peripheral blood by phenol–chloroform method. The genotype of each tag SNP for COPD patients and healthy smokers was screened by SNaPshot according to the manufacturer’s protocol (Thermo Fisher Scientific). In brief, multiplex PCR was performed by primers given in Supplementary Table 1 with FastStart Taq DNA polymerase (Roche, Basel, Switzerland). After alkaline phosphatase (Shrimp, Takara-Bio Inc., Kusatsu, Japan) and exonuclease I (Takara Bio Inc.) clean-up, single base extension was performed by SNaPshot Multiplex Ready Reaction Mix (Thermo Fisher Scientific) and the products were analyzed on ABI PRISM 3730 sequencer (Thermo Fisher Scientific). The genotypes of some random samples were confirmed by resequencing in 3730 sequencer.

Cell culture
Human bronchus normal epithelial cells Beas-2B (#CRL-9609; American Type Culture Collection, Manassas, VA, USA) were cultured in 1640 medium (Thermo Fisher Scientific) with 10% fetal bovine serum (Thermo Fisher Scientific) in 5% CO2 at 37°C.

Luciferase reporter assay
ADRB2 promoter ~1.38 kb region (chr5: 148205009–148206387, relative to build 37) was amplified using primers 5′-CAGTCGCTAGTTGTAAGTCACAGACGCCAG-3′ and 5′-CAGTCAAGCTTAGTCTGGCAGGTAGCGCA-3′, which introduced restriction sites for NheI and HindIII (New England Biolabs, Ipswich, MA, USA), respectively. PCR was performed by Pfu DNA polymerase (recombinant) enzyme (Thermo Fisher Scientific) to avoid artificial mutation. After digestion, the segment was cloned into the compatible sites of the pGL3-basic vector (Promega, Madison, WI, USA). A plasmid with another allele (G) for rs12654778 was generated through mutagenesis with Phusion Site-Directed Mutagenesis Kit (Thermo Fisher Scientific) and primer pair 5′-TCGTTATAAGCTTCTAGCATGCTGACCC-3′ and 5′-ACCAAGCCCATAGACCTGAGACAC-3′, according to the manufacturer’s protocol. Plasmid DNA was sequenced to exclude any PCR errors and check the orientation of the haplotypes prior to transfection.

Plasmid constructs with rs12654778-A and G (475 ng) were transfected into Beas-2B cells by Lipofectamine 2000 (Thermo Fisher Scientific) according to the manufacturer’s recommendations. After 24 hours transfection, cells were harvested and luciferase activity was measured by Dual-Luciferase Reporter Assay System (Promega) according to the manufacturer’s protocol. Plasmid pRL-TK (25 ng; Promega) was co-transfected as an internal control and the promoter activity was expressed as the ratio between firefly and Renilla luciferase. Independent transfection and reporter assays were performed six times.

Chromatin immunoprecipitation (ChIP)-PCR assay
ChIP was carried out with EZ-ChIP Assay Kit (EMD Millipore, Billerica, MA, USA) according to the manufacturer’s protocol. Briefly, Beas-2B cells were grown to reach sub-confluency. Approximately 1×107 cells were cross-linked for 10 min with formaldehyde (1% final concentration) at room temperature, which was followed by addition of glycine for 5 min to end the cross-linking. After washing twice with ice-cold phosphate-buffered saline (Thermo Fisher Scientific) containing protease
inhibitor cocktail, cells were scraped, lysed and sonicated to obtain 200–800 bp fragments in the Sonicator (Branson, Medford, NY, USA). The chromatin solution was diluted 10-fold with dilution buffer and precleared with protein G beads. After centrifuging and transferring the supernatant, 1% sample was stored as input and the remaining protein/chromatin complex was subjected to immunoprecipitation with mouse monoclonal NF-1 antibody (Santa Cruz Biotechnology, Dallas, TX, USA) or normal mouse IgG as a negative control, and precipitated by protein G beads. After washing with low salt, high salt, LiCl and Tris-EDTA buffer (twice), the immunoprecipitated protein/chromatin complex was resuspended in elution buffer and the cross-links were reversed. Protein was digested by proteinase K and DNA was recovered. The obtained DNA from ChIP preparation was quantified by real-time PCR to evaluate the enrichment using SYBR green with primers 5′-TGTGTTGGACAGGGGTGACTT-3′ and 5′-ACATTCGGAAGGAAACGAGAGT-3′. In the ChIP assay, relative enrichment was normalized by input DNA. Data are presented as the mean ± SD of triplicate experiments.

Statistical analysis
Age, smoking history and pulmonary function data are displayed as mean ± SD. Hardy–Weinberg equilibrium was evaluated by a goodness-of-fit chi-square test with one degree of freedom. The frequencies of each SNP between patients and controls were compared by two-tailed chi-square tests. To assess the independent effect of each SNP on COPD, a logistic regression analysis with tag SNPs (rs17108803, rs1432623, rs12654778, rs1042713, rs1042714, rs1042717, rs1042719 and rs1042720) as independent variables adjusted for age, sex and smoking history was also performed. For comparing the ADRB2 expression between cases and controls in the lung tissues and the luciferase activity, independent t-test was performed. All statistical tests were performed in SPSS 13.0 (SPSS Inc., Chicago, IL, USA). Odds ratios and 95% CIs were also calculated to assess the relative disease risk. In this study, the significance level was accepted when p (probability) value was <0.05.

Transcription factor-binding site prediction
For evaluating whether rs12654778 would alter the binding affinity of the TF, a putative TF-binding site was analyzed by using the web-based TRANSFAC database (http://www.gene-regulation.com/cgi-bin/pub/programs/match/bin/match.cgi).

Results
ADRB2 expression in lung tissues from cases and controls
We firstly assessed whether ADRB2 gene expression was altered in the lung tissues from COPD subjects. We measured ADRB2 expression in the lung tissues of COPD subjects and smokers who had normal lung function. By real-time PCR, we found that mRNA levels of ADRB2 were significantly reduced ~28% in COPD subjects compared with control subjects (p=0.017; Figure 1), confirming that ADRB2 is differentially expressed between pathologic and normal tissues and the decreased ADRB2 expression is associated with COPD development or onset.

Tag SNPs selection
To develop a comprehensive list of common genetic variants for ADRB2, the genotype data for East Asian population were obtained from 1,000 genomes project. In this ~8 kb region, 32 SNPs were identified, among which 12 were in the upstream region of ADRB2, six were in the 5′ untranslated region, six were in the 3′ untranslated region, five were synonymous and three were missense mutations in

---

**Figure 1** The average expression of ADRB2 in COPD patients and controls.

**Notes:** The data are presented as mean ± SD. *p<0.01.
the coding region. Subsequently, 14 blocks were observed (Supplementary Figure 1; Supplementary Table 2). Eight blocks showed minor allele frequency >5%, and thus, eight tag SNPs (rs17108803, rs1432623, rs12654778, rs1042713, rs1042714, rs1042717, rs1042719 and rs1042720) were selected from each block for further association study.

### Association study between SNPs in ADRB2 and COPD

To assess the relationship between these tag SNPs of ADRB2 and COPD, 200 COPD subjects and 222 normal controls were recruited in this study (Table 2). There was no significant difference in sex, age or smoking history between cases and controls ($p > 0.05$). In contrast, a significant difference was found in the baseline $\text{FEV}_1$ percentage predicted and $\text{FEV}_1/\text{FVC}$ between cases and controls.

Although our study sample size was moderate, our sample size provided >80% power to detect a genetic relative risk (or odds ratio) of 2.35, by using genetic power calculator,$^{29}$ with 7.3% disease prevalence in the Chinese population$^{30}$ and $\alpha=0.05$, for a variant with 0.05 minor allele frequency in a dominant model.

The tag SNPs from these eight blocks were genotyped in 422 subjects and the result is presented in Table 2. All eight SNPs were under Hardy–Weinberg equilibrium in controls ($p > 0.05$). As shown in Table 2, rs1042713 (Arg16Gly), which was detected in most previous studies,$^{13–19,31–34}$ was also investigated in our study, and the genotype (AA, AG, GG) frequency was 23%, 62% and 16% in COPD and 28%, 48% and 24% in controls, respectively, indicating that there was no statistically significant difference ($p = 0.15$) after adjusting for age, sex and smoking history. The same analysis was performed

### Table 2 Genotypes distribution in patients with COPD (case) and healthy smokers (control)

<table>
<thead>
<tr>
<th>SNP</th>
<th>Position$^a$</th>
<th>Amino substitution</th>
<th>Genotype frequency (%)</th>
<th>OR (95% CI)$^b$</th>
<th>p-value</th>
<th>Genotype</th>
<th>HWE$^*$</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs1432623</td>
<td>–2,389</td>
<td>N/A</td>
<td>Case</td>
<td>Control</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CC</td>
<td>0.13</td>
<td>0.08</td>
<td>0.10</td>
<td>0.68</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CT</td>
<td>0.42</td>
<td>0.39</td>
<td>1.92 (0.86–4.27)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TT</td>
<td>0.46</td>
<td>0.53</td>
<td>3.62 (1.11–11.86)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs17108803</td>
<td>–839</td>
<td>N/A</td>
<td>Case</td>
<td>Control</td>
<td>0.83</td>
<td></td>
<td>0.99</td>
</tr>
<tr>
<td>GG</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GT</td>
<td>0.11</td>
<td>0.13</td>
<td>0.00 (0.00–)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TT</td>
<td>0.90</td>
<td>0.87</td>
<td>0.00 (0.00–)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs12654778</td>
<td>–654</td>
<td>N/A</td>
<td>Case</td>
<td>Control</td>
<td>0.04</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>0.12</td>
<td>0.16</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AG</td>
<td>0.56</td>
<td>0.41</td>
<td>0.97 (0.36–2.61)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GG</td>
<td>0.33</td>
<td>0.43</td>
<td>2.40 (0.54–10.63)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs1042713</td>
<td>+46</td>
<td>Arg16Gly</td>
<td>Case</td>
<td>Control</td>
<td>0.15</td>
<td>0.61</td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>0.23</td>
<td>0.28</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GA</td>
<td>0.62</td>
<td>0.48</td>
<td>0.65 (0.30–1.42)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GG</td>
<td>0.16</td>
<td>0.24</td>
<td>1.17 (0.37–3.69)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs1042714</td>
<td>+79</td>
<td>Gln27Glu</td>
<td>Case</td>
<td>Control</td>
<td>0.73</td>
<td>0.98</td>
<td></td>
</tr>
<tr>
<td>GG</td>
<td>0.02</td>
<td>0.01</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GC</td>
<td>0.29</td>
<td>0.21</td>
<td>1.36 (0.27–6.98)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CC</td>
<td>0.69</td>
<td>0.78</td>
<td>1.68 (0.32–8.77)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs1042717</td>
<td>+253</td>
<td>Leu84Leu</td>
<td>Case</td>
<td>Control</td>
<td>0.19</td>
<td>0.62</td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>0.10</td>
<td>0.14</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AG</td>
<td>0.47</td>
<td>0.45</td>
<td>2.82 (0.91–8.77)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GG</td>
<td>0.44</td>
<td>0.41</td>
<td>3.82 (0.82–17.86)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs1042719</td>
<td>+1,055</td>
<td>Gly351Gly</td>
<td>Case</td>
<td>Control</td>
<td>0.69</td>
<td>0.19</td>
<td></td>
</tr>
<tr>
<td>CC</td>
<td>0.18</td>
<td>0.25</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CG</td>
<td>0.48</td>
<td>0.45</td>
<td>0.96 (0.45–2.06)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GG</td>
<td>0.35</td>
<td>0.30</td>
<td>1.29 (0.49–3.46)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs1042720</td>
<td>+1,241</td>
<td>Leu413Leu</td>
<td>Case</td>
<td>Control</td>
<td>0.13</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td>GG</td>
<td>0.22</td>
<td>0.15</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GA</td>
<td>0.46</td>
<td>0.40</td>
<td>1.55 (0.76–3.19)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>0.33</td>
<td>0.45</td>
<td>2.15 (0.86–5.35)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: *p*-value from test for HWE in controls. The position is based on the first nucleotide of the start codon being +1. The OR (95% CI) and genotype p-value have been adjusted for age, sex and smoking history.

Abbreviations: HWE, Hardy–Weinberg equilibrium; OR, odds ratio; SNP, single nucleotide polymorphism; CI, confidence interval.
on rs1042714 (Gln27Glu) and similar results were obtained ($p=0.73$; Table 2). These results were consistent with those of Brogger et al,\textsuperscript{31} while they disagree with Ho et al’s results.\textsuperscript{13} To resolve this conflict, a total of 13 previous studies on rs1042713 and rs1042714\textsuperscript{13–21,31–34} were collected and reanalyzed by meta-analysis (Supplementary Table 3). Besides the publication of Niu et al,\textsuperscript{35} four more studies (Wang et al,\textsuperscript{19,21} Vacca et al\textsuperscript{33} and our genotyping data) were involved. In total, 2,908 cases and 2,946 controls were included in this meta-analysis and no significant difference was observed in rs1042713 or rs1042714 under allele model ($p>0.05$, Supplementary Figure 2), consistent with a previous meta-analysis in 2012.\textsuperscript{35}

However, one SNP (rs12654778), located in the promoter region of \textit{ADRB2} (~654), presented a significant difference between cases (AA 12%, AG 56%, GG 33%) and controls (AA 16%, AG 41%, GG, 43%; $p=0.04$) adjusted by age, sex and smoking history. Lung function (FEV$_1$ value) of COPD patients with the different genotype of rs12654778 was also investigated. Although the result was not significant, FEV$_1$ value in individuals with AA genotype was higher than that of other genotypes (GA and GG; Supplementary Figure 3), suggesting that this SNP should be a candidate site for regulating \textit{ADRB2} gene expression and should be further associated with COPD or FEV$_1$.

Promoter activity of different alleles for rs12654778

There are two SNPs (rs12654778 and rs17778257) in this block, and rs12654778 is closer to \textit{ADRB2} translation start codon than rs17778257 (~2 kb upstream). Considering that rs17778257 is far away from the regulatory region of \textit{ADRB2} and as no TF binding near this site (Supplementary Figure 4) was found by searching the ENCODE database, rs12654778 was chosen for further functional analysis. We proposed that rs12654778 may influence the transcriptional activity of \textit{ADRB2}. To verify this hypothesis, we generated the plasmids containing different alleles of rs12654778 and transiently transfected them into Beas-2B cell lines. As shown in Figure 2, the cloned region showed ~100-fold higher luciferase expression compared with pGL3-basic plasmid, demonstrating the strong promoter activity of this region in the lung tissue. Moreover, rs12654778-A allele showed a ~21.8% reduction in promoter activity compared to the rs12654778-G allele ($p=0.0034$, Figure 2), suggesting that this SNP could regulate \textit{ADRB2} gene expression in the lung tissue.

Transcription factor neurofibromin 1 (NF1) binding rs12654778 surrounding region

Considering that rs12654778 is located within a conserved CCAAT box-like motif,\textsuperscript{36} which may function as a canonical binding site for NF1 based on TRANSFAC prediction, we hypothesized that NF1 is involved in the transcription of \textit{ADRB2} and this SNP might affect the binding affinity of NF1 to this region. To investigate whether NF1 binds the upstream region of \textit{ADRB2}, we performed ChIP assays in Beas-2B cells using an anti-NF1 antibody and quantified the enrichment in the predicted NF1 binding site by real-time PCR. As shown in Figure 3, the chromatin immunoprecipitated by NF1 antibody was significantly enriched in the region surrounding rs12654778 compared with IgG ($p=0.0021$), suggesting that NF1 binds this region in the lung tissue.

Discussion

COPD is a common respiratory disease caused by interaction of environmental risk factors with genetic background.\textsuperscript{37}
While several relevant environmental risk factors of COPD have been identified, the genetic risk factors are less well understood. ADRB2, which encodes the $\beta_2$-adrenergic receptor, is expressed in airway smooth muscle cells and is considered as an important pharmacologic target in the management of COPD, thus leading to speculation of its contributions to the onset of COPD. Previous studies have found that several genetic variants of ADRB2 are associated with COPD, but the real function of these variants is not elaborated, especially the noncoding variants, which might play a more important role in altering the noncoding variants, which might play a more important role in altering TF NF1 binding with COPD. Secondly, a noncoding SNP (rs12654778), located upstream of ADRB2, was associated with COPD. Thirdly, functional analysis indicated that rs12654778 could modulate the binding affinity of TF (NF1) and its risk allele could reduce the transcriptional activity of ADRB2 gene expression. Taking all evidence together, our results are the first to reveal that the differential NF1 binding at rs12654778 could lead to reduced ADRB2 expression level in the lung tissue and increased susceptibility to COPD.

There are several studies showing that mutations in the coding region could affect the disease, for example, SERPINA1, PiMZ heterozygote produced by α1-antitrypsin deficiency. However, most common variants identified by the association panel are located in noncoding regions, especially by the genome-wide association studies, and might be cis-regulatory elements for the nearby gene. Indeed, successful identification of functional variants in these promoter regulatory elements has been reported for $\beta$-thalassemia and pyruvate kinase deficiency. Considering that the study of regulatory elements is very important and it is difficult to identify functional genetic variants in the regulatory regions, it is worth performing intensive investigation. We contend that the identification of functional variants in such regions is an extremely important requisite for at least two reasons. On one hand, the identification of functional variants can conclusively prove which gene is actually involved in disease susceptibility. On the other hand, study of functional variants can lead to new insights into the pathophysiological mechanisms of diseases.

Since rs12654778 is located in the ADRB2 gene regulatory region, another important concern is about the potential mechanism by which this SNP in ADRB2 is associated with COPD susceptibility. The rs12654778 is located ~654 bp upstream of ADRB2, which is with the histone modification H3K27Ac, H3K4me1 and H3K4me3 (Supplementary Figure 4). Since H3K27Ac and H3K4me3 are usually correlated with activation of chromatin, it was reasonable to hypothesize that this region is pivotal for ADRB2 gene expression. Here, we firstly showed that rs12654778 could reduce ADRB2 expression level in the lung tissue by altering TF NF1 binding with ADRB2 promoter region, and contributed to the COPD onset, and this SNP has been identified with other diseases in several studies. Meanwhile, ADRB2 is also expressed in lymphoblastoid cells, and it is interesting to know whether this variant is correlated with the expression of ADRB2 in lymphoblastoid cells. To address this issue, we utilized eQTL browser (http://eqtl.uchicago.edu/cgi-bin/gbrowse/eqtl/), which collected published eQTL data of lymphoblastoid cell lines (LCL) from four HapMap populations to search for potential association. Interestingly, a significant association between rs12654778 and expression of ADRB2 in LCL was observed in Yoruban in Ibadan, Nigeria (YRI),
Utah residents with Northern and Western European ancestry from the CEPH collection and two samples which were treated as a single analysis panel of 90 Asians populations (data not shown). We further evaluated the effect of this SNP on ADRB2 expression in LCL from 726 HapMap3 individuals in GEne Expression VARiation (http://www.sanger.ac.uk/resources/software/genevar). The AA genotype of rs12654778 presented a significantly reduced expression of ADRB2, compared with the other genotypes AG or GG in most populations (CHB, Gujarati Indians in Houston, TX, USA, JPT, Mexican ancestry in Los Angeles, CA, USA, Maasai in Kinyawa, Kenya and YRI, shown in Supplementary Figure 5), which was consistent with our result.

In addition, β2-adrenoceptors, through their extracellular domain, can bind to Gs and prevent adenylyl cyclase from activating the cAMP signaling pathway, a critical pathway for embryonic lung development. Decreased ADRB2 expression leads to overactivation of the cAMP pathway in multiple types of breast tumor, which in turn contributes to uncontrolled cellular proliferation. In our study, we found that ADRB2 expression was decreased in the lung tissue from COPD cases compared with control subjects with normal lung function, indicating that lower ADRB2 expression may exacerbate COPD pathogenesis. This was inconsistent with Selivanova et al’s study, which might be due to the difference in the control group. In our study, the control group consisted of people with normal lung function, while in Selivanova et al’s study, the control group consisted of mild/middle asthma patients. Further mechanistic studies on the cAMP pathway in the context of smoking may provide novel insights into the pathogenesis of COPD.

Conclusion

Our study is the first to demonstrate that a functional SNP (rs12654778), upstream of ADRB2, was significantly associated with increased risk for COPD. These results offer valuable insights into the signaling, maintenance and regulatory mechanisms of ADRB2 in lung and its further correlation with COPD.

Acknowledgments

This work was supported by grants from the National Natural Science Foundation of China (grant no 31200940) and Yunnan University (grant no 2011ZD01). The authors wish to thank all the patients for allowing them to collect blood samples and tissues samples for genetic association studies over the years.

Disclosure

The authors report no conflicts of interest in this work.

References


