RESEARCH ARTICLE

Polymorphisms in *VEGF-A* are associated with COPD risk in the Chinese population from Hainan province

YIPENG DING[1*], HUAN NIU[1], YIZHOU LI[2], PING HE[1], QUANNI LI[1], YANHONG OUYANG[1], MIN LI[1], ZHIGAO HU[1], YOUQING ZHONG[3], PEI SUN[1] and TIANBO JIN[4,5*]

[1]*Department of Emergency, People's Hospital of Hainan Province, Haikou, Hainan 570311, People's Republic of China*
[2]*Department of Orthopedics and Traumatology, The Second Affiliated Hospital of Inner Mongolia University, Huhehot 010030, Inner Mongolia, People's Republic of China*
[3]*Department of Respiratory Medicine, The Affiliated Hospital of Hainan Medical College, Haikou, Hainan 570102, People's Republic of China*
[4]*School of Life Sciences, Northwest University, Xi'an, Shaanxi 710069, People's Republic of China*
[5]*National Engineering Research Center for Miniaturized Detection Systems, Xi'an, Shanxi 710069, People's Republic of China*

Abstract

In this study, we examined and validated how common variants contribute to susceptibility to chronic obstructive pulmonary disease (COPD) in the Han Chinese population. Here, we genotyped 18 nucleotide polymorphisms and evaluated their association with COPD using chi-square test and genetic model analysis (246 COPD patients and 350 controls), and found three SNPs that might cause a predisposition to COPD. Both rs3025030 and rs3025033 are located on chromosome 6 in *VEGF-A*. We found one risk allele 'C' from rs3025030 and another 'G' from rs3025033 using the log-additive model (OR 1.40; 95% CI 1.05–5.96; $P = 0.022$), (OR 1.38; 95% CI 1.03–1.84; $P = 0.03$). We also found another risk allele 'A' of rs9296092 in gene region *ZBTB9-BAK1* by the allele model (OR 2.63; 95% CI 1.27–5.45; $P = 0.0078$), (adjusted OR 3.53; 95% CI 1.12–11.11; $P = 0.031$). We found a risk haplotype 'CG' associated with the risk of COPD (OR 1.39; 95% CI 1.04–1.86; $P = 0.028$). Our results when compared with previous studies showed significant association between *VEGF-A* polymorphism and COPD. We also identified rs9296092 as a risk factor for COPD.

Introduction

Chronic obstructive pulmonary disease (COPD) is one of the leading causes of morality worldwide, and is a major public health problem (Bossé 2012; Laratta and van Eeden 2014). Prevalence of the disease is estimated to be about 10% in the general population, which increases to 20–30% in the population aged above 40 years (Gagnon *et al*. 2014). COPD is a complex disease with multiple risk factors including environmental, smoking and genetic factors. COPD is defined as a preventable and treatable disease with major extrapulmonary effects. Previously, it was characterized by shortness of breath with exertion, wheezing and chronic cough (Mapel

and Roberts 2012; Naik *et al*. 2014). Currently, a decrease in forced expiratory volume in 1 s (FEV$_1$%) and the ratio of FEV$_1$ to forced vital capacity (FEV$_1$/FVC) are the diagnostic criteria of COPD (Berndt *et al*. 2012). Hypoxic vasoconstriction with advanced airflow limitation in COPD patients may result in pulmonary hypertension (Wells and Dransfield 2013). Vascular endothelial growth factor A (*VEGF-A*) is considered one of the most important regulatory factors in pulmonary hypertension. In recent years, case–control studies confirmed the relationship between single-nucleotide polymorphisms (SNPs) in *VEGF-A* and increased risk of disease development, including lung cancer, acute lung injury, etc. (Meyer *et al*. 2012; Lin *et al*. 2013; Maeda *et al*. 2013). *VEGF-A* plays an important role in the migration, proliferation and survival of endothelial cells (Gong *et al*. 2011). Previous studies report that SNPs rs10434 and rs3025028 contained within *VEGF-A* are associated with lung function

*For correspondence. E-mail: Yipeng Ding, ypding@263.net; Tianbo Jin, jintianbo@gmail.com.
Yipeng Ding and Huan Niu contributed equally to this work.

Keywords. single-nucleotide polymorphism; *VEGF-A*; chronic obstructive pulmonary disease.

Table 1. Characteristics of controls and COPD cases in this study.

Characteristic	Control ($n = 350$)	Case ($n = 246$)	P value
Female	90 (25.7)	64 (26.0)	0.934
Male	260 (74.3)	182 (74.0)	
Age (mean age $\pm$ SD)	70.92 $\pm$ 9.723	46.98 $\pm$ 11.897	<0.01
Smoking status, N. (%)			
Nonsmoking	177 (50.6)	120 (49.2)	0.739
Smoking	173 (49.4)	124 (50.8)	
Missing		2 (0.8)	

*$P < 0.05$ indicates statistical significance.

Table 2. Classification of severity of COPD.

Severity*	FEV$_1$% predicted	M MRC grate	FEV$_1$/ FVC%	Complications[#]
Mild	$\geq$80	<2	<70	No
Moderate	50–79	$\geq$2	<70	No
Severe	<50	$\geq$2	<70	Yes

*Category with the worst value should be used for severity classification.
[#]Complications include respiratory failure.
FVE, forced expiratory volume in 1 s; FEV$_1$/FVC, the ratio of FEV$_1$ to forced vital capacity.

from infancy to adulthood, and play an important role in determination of airway function (Simpson *et al.* 2012). *VEGF-A* is located on chromosome 6p21.1 and comprises a coding region separated by seven introns. To our knowledge, there is little information on SNPs in *VEGF-A* related to COPD in patients. We designed a case–control study to determine whether alterations in *VEGF-A* increase susceptibility to COPD occurrence. The purpose of our study was to identify the relationship between SNPs in *VEGF-A* and COPD.

Materials and methods

Study population

All cases and controls were restricted to Hainan province, located in the Hainan Provincial People's Hospital. All patients with COPD were diagnosed according to the criteria established by the World Health Organization Global Initiative for Chronic Obstructive Lung Disease (GOLD) (Rabe *et al.* 2007) and they did not have history of other diseases. Case individuals were chosen regardless of age, gender, smoking status or COPD classification. Normal subjects were selected randomly from a native Hainan population while attending medical checkup at their community hospital, Hainan Provincial People's Hospital. A total of 246 cases and 350 controls were recruited in this study. Basic characteristics of the participants such as gender, age and smoking status are listed in table 1 and the classification of COPD with FEV$_1$/FVC levels is shown in table 2.

Clinical data and demographic information

Personal data, including residential regions, age, smoking status, gender, education status and family history of cancer were collected using standard epidemiological questionnaire and in-person interview. The case information was collected through consultation with treating physicians or from medical chart review. All participants signed an informed consent agreement. The Human Research Committee for approval of research involving human subjects, Hainan Provincial People's Hospital approved use of human tissue in this study.

Selection of SNPs and method of genotyping

We selected SNPs from published polymorphisms associated with COPD. SNPs with minor allele frequency (MAF) >5% in a HapMap CHB population were selected resulting in 18 SNPs in different genes. Genomic DNA was extracted from whole blood using phenol–chloroform (Köchl *et al.* 2005).

Genomic DNA was extracted from whole blood using the GoldMag-Mini whole blood genomic DNA purification kit (GoldMag, Xi'an, China). We used Sequenom Mass ARRAY assay design 3.0 software (Sequenom, San Diego, USA) to design the multiplexed SNP mass EXTEND assay. SNP genotyping was performed by Sequenom Mass ARRAY RS1000 using the standard protocol recommended by the manufacturer sequenom type r4.0 software was used to perform data management and analysis.

Statistical analysis

To perform statistical analyses, we used Microsoft Excel and SPSS 17.0 statistical package (SPSS, Chicago, USA). In our study, $P \leq 0.05$ is the threshold of statistical significance. The validation of each SNP frequency in control subjects was tested for departure from Hardy–Weinberg equilibrium (HWE) using an exact test. Chi-square test was done to calculate the genotype frequencies of case and control individuals (Adamec 1964; Köchl *et al.* 2005). Unconditional logistic regression analyses with adjusted for age, gender and smoking states was performed to test. (Bland and Altman 2000). The possibility of gender difference as a source of population substructure was evaluated by a genotype test for each SNP in male and female controls, and the number of significant results at 5% level was compared with the number expected by the chi-square test. The association of certain SNPs with the risk of COPD was tested in four genetic models (allele, codominant, dominant and log-additive); unconditional logistic regression analysis with adjustment for age, gender and smoking was done to calculate ORs and 95% CIs (Bland and Altman 2000; Purcell *et al.* 2007).

Table 3. Basic information for candidate SNPs in this study.

Gene(s)	SNP no.	Chromosome	Alleles	HWE P	P value	P value adj.*	OR (95% CI)	MAF Case	MAF Control
RNF150	rs10007052	4q31.21	C/A	1	0.007	0.63	1.48 (1.11–1.97)	0.255	0.188
VEGFA	rs25648	6p21.1	T/C	1	0.286	1	1.27 (0.81–2.01)	0.077	0.061
VEGFA	rs833068	6p21.1	A/G	0.574	0.817	1	1.02 (0.81–1.30)	0.396	0.39
VEGFA	rs833070	6p21.1	A/G	0.381	0.64	1	0.93 (0.71–1.23)	0.232	0.243
VEGFA	rs3024994	6p21.1	T/C	0.39	0.703	1	1.07 (0.74–1.56)	0.11	0.103
VEGFA	rs3024997	6p21.1	A/G	0.737	0.99	1	1.00 (0.79–1.26)	0.398	0.398
VEGFA	rs3025000	6p21.1	T/C	0.822	0.861	1	0.97 (0.77–1.24)	0.382	0.387
VEGFA	rs3025030	6p21.1	C/G	0.584	0.022	1	1.39 (1.04–1.86)	0.23	0.176
VEGFA	rs3025033	6p21.1	G/A	0.584	0.03	1	1.37 (1.03–1.82)	0.23	0.179
VEGFA	rs10434	6p21.1	A/G	0.307	0.326	1	0.87 (0.66–1.14)	0.217	0.241
ZBTB9-BAK1	rs9296092	6p21.32	A/G	1	0.008	0.72	2.55 (1.24–5.24)	0.043	0.017
SCGB1A1	rs17157266	11q12.3	C/T	0.198	0.738	1	1.04 (0.79–1.38)	0.22	0.211
IREB2	rs13180	15q25.1	T/C	0.915	0.13	1	0.83 (0.66–1.05)	0.451	0.496
CHRNA5	rs667282	15q25.1	C/T	0.829	0.791	1	1.03 (0.81–1.30)	0.449	0.441
ATP2C2	rs8048576	16q24.1	A/G	0.73	0.866	1	1.02 (0.80–1.29)	0.372	0.367
TIMM21-CYB5A	rs9951925	18q22.3	A/C	0.042	0.565	1	1.07 (0.83–1.38)	0.317	0.301
EGLN2	rs7937	19q13.2	C/T	1	0.07	1	1.24 (0.98–1.57)	0.409	0.357
EGLN2	rs3733829	19q13.2	C/T	0.912	0.064	1	0.79 (0.62–1.01)	0.356	0.41

The SNPs are excluded at 5% HWE P level. *P value was adjusted by Bonferroni corrections.

Results

From the 596 participants we detected 18 SNPs. Two SNPs (rs9951925 and rs8102683) were excluded at 5% HWE statistically significant P level. The association between SNP genotypes and the susceptibility of alleles to COPD was performed by chi-square analysis. All results are shown in table 3.

We assumed that the minor allele of each SNP was a risk factor compared to the wild-type allele. Three models: codominant, dominant and log-additive were applied for analysing the association between SNPs and COPD risk by using logistic tests. We discovered one SNP (rs10007052) in the *RNF150* gene which was a risk factor for COPD in the dominant model (rs10007052, OR 1.60; 95% CI 1.13–2.26; $P = 0.0076$), codominant model (rs10007052, OR 1.55; 95% CI 1.08–2.21; $P = 0.023$) and log-additive model (rs10007052, OR 1.50; 95% CI 1.12–2.01; $P = 0.006$). We found two SNPs in *VEGF-A* which were risk factors for COPD (rs3025030, OR 1.40; 95% CI 1.05–1.87; $P = 0.022$; rs3025033, OR 1.38; 95% CI 1.03–1.84; $P = 0.03$). We also detected one SNP (rs9296092) using the allele model which had a similar relationship with the susceptibility to COPD (rs9296092, OR 2.63; 95% CI 1.27–5.45; $P = 0.078$). According to the results of the three models calculated by unconditional logistic regression analyses adjusted for age, gender and smoking status, we first found that only rs9296092 was still considered a risk factor (rs9296092, OR 3.53; 95% CI 1.12–11.11; $P = 0.031$) (table 4). Pairwise linkage disequilibrium (LD) analysis was performed for *VEGF-A* using the polymorphisms detected in this study. The pattern of LD was analysed using two parameters, r^2

Table 4. Logistic regression analysis of the association between SNPs and risk of COPD.

SNP	Model	Genotype	OR[a] (95% CI)	P[a] value	OR[b] (95% CI)	P[b] value
rs10007052	Codominant	C/A	1.55 (1.08–2.21)	0.023	1.20 (0.69–2.09)	0.26
	Dominant	C/A–C/C	1.60 (1.13–2.26)	0.0076	1.31 (0.76–2.24)	0.33
	Log-additive		1.50 (1.12–2.01)	0.0062	1.38 (0.86–2.22)	0.18
rs9296092		G/A	2.63 (1.27–5.45)	0.0078	3.53 (1.12–11.11)	0.031
rs3025030	Codominant	C/C	2.51 (1.06–5.96)	0.054	1.27 (0.30–5.32)	0.69
	Log-additive		1.40 (1.05–1.87)	0.022	1.22 (0.76–1.93)	0.41
rs3025033	Codominant	G/G	2.48 (1.05–5.87)	0.066	1.28 (0.31–5.27)	0.58
	Log-additive		1.38 (1.03–1.84)	0.03	1.25 (0.79–1.99)	0.33
rs9951925	Codominant	C/A	1.44 (1.02–2.03)	0.047	1.98 (1.13–3.46)	0.051
	dominant	C/A–A/A	1.29 (0.93–1.80)	0.12	1.88 (1.11–3.17)	0.017

OR[a] and P[a] without adjustment. OR[b] and P[b] adjusted for gender, age and smoking status. $P \leq 0.05$ indicates statistical significance. SNP, single nucleotide polymorphism; OR, odds ratio; CI, confidence interval.
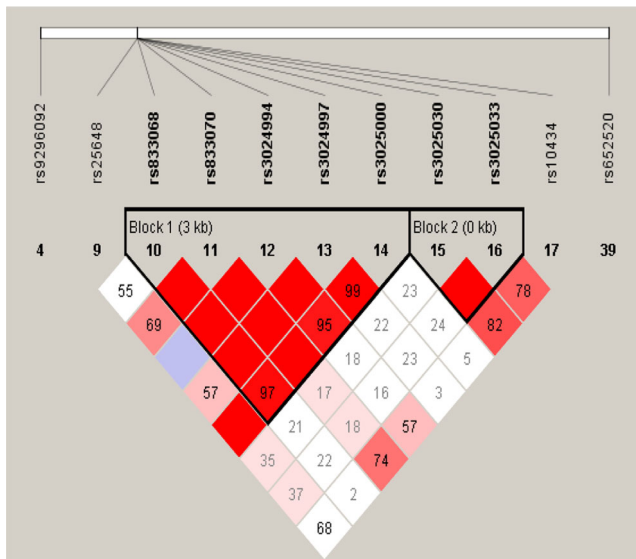
Figure 1. Linkage disequilibrium of polymorphic sites in *VEGF-A* on chromosome 6. Block 1 includes rs833068, rs833070, rs3024994, rs3024997 and rs3025000. Block 2 includes rs3025030 and rs35025033. The LD between two SNPs is standardized D' (red schemes).

and D'. One main linkage block was observed across the locus (figure 1). This block was comprised of seven closely-linked SNPs: rs833068, rs833070, rs3024994, rs3024997, rs3025000, rs3025030 and rs3025033. Then, the association between inferred haplotypes and COPD risk among the individuals was analysed. We found the haplotype 'CG' between rs3025030 and rs3025033 associated with the risk of COPD (OR 1.39; 95% CI 1.04–1.86; $P = 0.028$) (table 5).

Discussion

Environmental risk factors and effects of smoking may increase the occurrence of COPD. However, several genetic polymorphisms associated with susceptibility to COPD have been identified. Whereas each individual polymorphism may contribute to only a small relative risk of COPD, a combination of several implicated polymorphisms and other risk factors may increase the relative risk. In this study, we detected 18 SNPs from the case–control study in the Han Chinese population from Hainan city. We considered two SNPs in *VEGF-A* (rs3025030 and rs3025033) as risk factors

for COPD. Our results suggest that *VEGF-A* may play an important role in the risk of COPD in Hainan individuals.

VEGF-A which is highly polymorphic and exists with multiple common SNPs, complicating the investigation for the genetic components of COPD occurrence. Previous studies have shown that genetic variability affects *VEGF-A* activity and expression (Fang *et al.* 2009). VEGF, coded for *VEGF-A*, is a key mediator of angiogenesis and vascular permeability (Su *et al.* 2011). *VEGF-A* overexpression or suppression causes a decisive effect on the function of VEGF, which may lead to a complex change in angiogenesis of tissue. The changes of pulmonary small vessels is one of pathological changes of COPD patients (Hogg and Timens 2009), in which *VEGF-A* plays a vital role. Additionally, a previous study reported that *VEGF-A* can disrupt the pattern formation of pulmonary smooth muscle (Akeson *et al.* 2003), which may result in progression of pulmonary hypertension, further aggravating limited pulmonary ventilation. Moreover, previous literature identified several SNPs, including rs3025030 and rs3025033, as associated with colon or rectal cancer by influencing the expression of *VEGF-A* in the angiogenesis pathway (Slattery *et al.* 2014). In our case–control study, we found two *VEGF-A* polymorphisms (rs3025030, rs3025033) caused a more pronounced effect than the other single polymorphisms in *VEGF-A*. This finding suggests a novel idea that *VEGF-A* polymorphism combinations may be risk factors for COPD. SNPs may influence the expression of *VEGF-A* in the pathway, which may reduce beneficial effects of VEGF in lung tissue and increase the risk of COPD. The combination of two polymorphisms (rs3025030, rs3025033) may further increase the risk of COPD. Unfortunately, single SNP analysis of rs3025030 and rs3025033 calculated by unconditional logistic regression analyses adjusted for age, gender and smoking were nonsignificant. In our opinion, smoking may be one of the most important confounding factors and it may overshadow the genetic effect of *VEGF-A* polymorphisms in COPD. In future studies, researchers may identify the relative interaction between *VEGF-A* polymorphisms and COPD development.

A previous genomewide association study has not only proved that rs9296092 lies in the gene region between the zinc finger and BTB domain containing 9 and BCL2-antagonist/killer1 (*ZBTB9-BAK1*), but also is associated with age at smoking initiation (Siedlinski *et al.* 2011). We also found rs9296092, located near 6p21.1, predicts a 2.63-fold COPD risk by the allele model and a 3.53-fold disease susceptibility calculated by unconditional logistic regression

Table 5. Multiple genes haplotype frequencies and association with the risk of COPD.

Gene	Block	Haplotype	Frequency	OR^a (95% CI)	P^a	OR^b (95% CI)	P^b
EGLN2	1	CT	0.375	1.33 (1.02–1.73)	0.038	1.37 (0.92–2.05)	0.12
VEGF-A	2	CG	0.198	1.39 (1.04–1.86)	0.025	1.26 (0.79–1.99)	0.33

OR^a and P^a without adjustment. OR^b and P^b adjusted for gender, age and smoking status. $P \leq 0.05$ indicates statistical significance. OR, odds ratio; CI, confidence interval.

analyses adjusted for confounding factors. We hypothesize that rs9296092 may be a risk factor for COPD, although the exact genotype of the loci is not identified.

We did not find either rs7937 or rs3733829 in *EGLN2* were associated with COPD or lung function in the Chinese Han population from Hainan province. The results included strong levels of D' between rs7937 and rs3733829 observed without adjustment, which might be a relative factor with COPD patients. Overexpression of *EGLN2* at a known COPD risk locus leads to occurrence of disease (Salit *et al.* 2014)

Two limitations should be considered in the current study. First, because our sample size was relatively small, we did not conduct population stratification of smokers, therefore, we have not confirmed that this locus is significant in smokers. We will increase the sample size in subsequent experiments. Second, the heterogeneity in smoking behaviours and comorbidities was not evaluated in this study, and evaluation of heterogeneity in smoking behaviours has contributed to elucidating the pathogenesis of COPD in other studies.

We performed Bonferroni correction in our statistical analysis and found no statistical significant associations between *VEGF-A* SNPs and COPD. This may be due to the relatively small sample size, the selection criteria of *VEGF-A* SNPs (MAF 0.5%) and the weakness of Bonferroni correction itself. Adjustments for multiple tests, such as Bonferroni correction analysis, are required for medical association studies, but also create more problems than they solve. The main weakness of Bonferroni correction is that the interpretation of a finding depends on the number of other tests performed. True important differences may be deemed nonsignificant since the likelihood of type II errors are also increased. However, Bonferroni corrections are considered acceptable while performing associations without preestablished hypotheses.

In conclusion, the combined actions of different confounding factors, such as small sample size and lack of stratification of smokers resulted in a nonsignificant phenomenon with adjustment. In our opinion, both rs3025030 and rs3025033 in gene *VEGF-A* may be risk factors for COPD, and interactions between loci in *VEGF-A* may be more important than a single locus change. We also found a new link between rs9296092 and COPD risk.

References

Adamec C. 1964 Example of the use of the nonparametric test. Test X2 for comparison of 2 independent examples. *Ceskoslovenske Zdravotnictvi* **12**, 613–619.

Akeson A. L., Greenberg J. M., Cameron J. E., Thompson F. Y., Brooks S. K., Wiginton D. and Whitsett J. A. 2003 Temporal and spatial regulation of VEGF-A controls vascular patterning in the embryonic lung. *Dev. Biol.* **264**, 443–455.

Berndt A., Leme A. S. and Shapiro S. D. 2012 Emerging genetics of COPD. *EMBO Molec. Med.* **4**, 1144–1155.

Bland J. M. and Altman D. G. 2000 The odds ratio. *BMJ* **320**, 1468.

Bossé Y. 2012 Updates on the COPD gene list. *Inter. J. Chronic Obstruct. Pulmonary Dis.* **7**, 607.

Fang A. M., Lee A. Y., Kulkarni M., Osborn M. P. and Brantley Jr M. A. 2009 Polymorphisms in the VEGF-A and VEGFR-2 genes and neovascular age-related macular degeneration. *Mol. Vis.* **15**, 2710.

Gagnon P., Guenette J. A., Langer D., Laviolette L., Mainguy V., Maltais F. *et al.* 2014 Pathogenesis of hyperinflation in chronic obstructive pulmonary disease. *Inter. J. Chronic Obstruct. Pulmonary Dis.* **9**, 187.

Gong Z.-P., Qiao N.-D., Gu Y.-X., Song J.-P., Li P.-L., Qiu H.-J. *et al.* 2011 Polymorphisms of VEGF-A gene and susceptibility to hemorrhage risk of brain arteriovenous malformations in a Chinese population. *Acta Pharmacol. Sinica* **32**, 1071–1077.

Hogg J. C. and Timens W. 2009 The pathology of chronic obstructive pulmonary disease. *Ann. Rev. Patholog. Mech. Dis.* **4**, 435–459.

Köchl S., Niederstätter H. and Parson W. 2005 DNA extraction and quantitation of forensic samples using the phenol–chloroform method and real-time PCR. In *Forensic DNA typing protocols*. (ed. A. Carracedo), pp. 13–29, vol. 297. Humana Press, Totowa, USA.

Laratta C. R. and van Eeden S. 2014 Acute exacerbation of chronic obstructive pulmonary disease: cardiovascular links. *BioMed. Res. Int.* **2014**, 528789.

Lin L., Cao K., Chen W., Pan X. and Zhao H. 2013 Four common vascular endothelial growth factor polymorphisms ($-2578C>A$, $-460C>T$, $+936C>T$, and $+405G>C$) in susceptibility to lung cancer: a meta-analysis. *PLoS One* **8**, e75123.

Maeda A., Nakata M., Yasuda K., Yukawa T., Saisho S., Okita R. *et al.* 2013 Influence of vascular endothelial growth factor single nucleotide polymorphisms on non-small cell lung cancer tumor angiogenesis. *Oncol. Rep.* **29**, 39–44.

Mapel D. W. and Roberts M. H. 2012 New clinical insights into chronic obstructive pulmonary disease and their implications for pharmacoeconomic analyses. *Pharmacoeconomics* **30**, 869–885.

Meyer N. J., Daye Z. J., Rushefski M., Aplenc R., Lanken P. N., Shashaty M. G. *et al.* 2012 SNP-set analysis replicates acute lung injury genetic risk factors. *BMC Med. Genet.* **13**, 52.

Naik D., Joshi A., Paul T. V. and Thomas N. 2014 Chronic obstructive pulmonary disease and the metabolic syndrome: consequences of a dual threat. *Ind. J. Endocrin. Metab.* **18**, 608.

Purcell S., Neale B., Todd-Brown K., Thomas L., Ferreira M. A., Bender D. *et al.* 2007 PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575.

Rabe K. F., Hurd S., Anzueto A., Barnes P. J., Buist S. A., Calverley, P. *et al.* 2007 A chronic obstructive pulmonary disease—Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease: GOLD executive summary. *Am. J. Resp. Crit. Care Med.* **176**, 532–555.

Salit J., Walters M., Vincent T., Agosto-Perez F., Mezey J., Crystal R. and Ryan D. 2014 Smoking-induced alternative splicing of transcript variants in human airway basal cells at COPD risk locus 19q13. 2. *Locus* **19**, 2.

Siedlinski M., Cho M. H., Bakke P., Gulsvik A., Lomas D. A., Anderson W. *et al.* 2011 Genome-wide association study of smoking behaviours in patients with COPD. *Thorax* **66**, 894–902.

Simpson A., Custovic A., Tepper R., Graves P., Stern D. A., Jones M. *et al.* 2012 Genetic variation in vascular endothelial growth factor-a and lung function. *Am. J. Resp. Crit. Care Med.* **185**, 1197–1204.

Slattery M. L., Lundgreen A., Mullany L. E., Penney R. B. and Wolff R. K. 2014 Influence of CHIEF pathway genes on gene expression: a pathway approach to functionality. *Inter. J. Molec. Epidemiol. Genet.* **5**, 100.

Su M.-T., Lee I.-W., Chen Y.-C. and Kuo P.-L. 2011 Association of progesterone receptor polymorphism with idiopathic recurrent pregnancy loss in Taiwanese Han population. *J. Assis. Reproduc. Genet.* **28**, 239–243.

Wells J. M. and Dransfield M. T. 2013 Pathophysiology and clinical implications of pulmonary arterial enlargement in COPD. *Inter. J. Chronic Obstruct. Pulmonary Dis.* **8**, 509.