Population-based frequency of surfactant dysfunction mutations in a native Chinese cohort

Yu-Jun Chen, Jennifer Anne Wambach, Kelcey DePass, Daniel James Wegner, Shao-Ke Chen, Qun-Yuan Zhang, Hillary Heins, Francis Sessions Cole, Aaron Hamvas

St. Louis, USA

Original article

Background: **Rare mutations in surfactant-associated genes contribute to neonatal respiratory distress syndrome. The frequency of mutations in these genes in the Chinese population is unknown.**

Methods: **We obtained blood spots from the Guangxi Neonatal Screening Center in Nanning, China that included Han (*n*=443) and Zhuang (*n*=313) ethnic groups. We resequenced all exons of the surfactant proteins-B (*SFTPB*), -C (*SFTPC*), and the ATP-binding cassette member A3 (*ABCA3*) genes and compared the frequencies of 5 common and all rare variants.**

Results: **We found minor differences in the frequencies of the common variants in the Han and Zhuang cohorts. We did not find any rare mutations in *SFTPB* or *SFTPC*, but we found three *ABCA3* mutations in the Han [minor allele frequency (MAF)=0.003] and 7 in the Zhuang (MAF=0.011) cohorts (*P*=0.10). The *ABCA3* mutations were unique to each cohort; five were novel. The collapsed carrier rate of rare *ABCA3* mutations in the Han and Zhuang populations combined was 1.3%, which is significantly lower than that in the United States (*P*<0.001).**

Conclusions: **The population-based frequency of mutations in *ABCA3* in south China newborns is significantly lower than that in United States. The contribution of these rare *ABCA3* mutations to disease burden in the south China population is still unknown.**

World J Pediatr 2016;12(2):190-195

Key words: **genetic epidemiology; human population genetics; neonatal respiratory distress syndrome; pulmonary surfactant**

Introduction

A developmentally regulated, quantitative deficiency of pulmonary surfactant is the underlying cause of neonatal respiratory distress syndrom (RDS), the most common respiratory cause of mortality and morbidity among infants less than one year of age in the United States.[1,2] Even after controlling for prematurity, differences in the risk for RDS among United States newborns of African, European, Asian, and Latino descent persist.[3] Mutations in surfactant-associated genes, including surfactant protein-B (*SFTPB*), -C (*SFTPC*), and the ATP-binding cassette transporter A3 (*ABCA3*), demonstrate that a proportion of RDS in infants of European and African descent likely has a heritable component.[4-15] Population-based studies of the epidemiology of disease-associated mutations in these genes in the US demonstrate that mutations in *SFTPB* and *SFTPC* are very rare (<0.1%) and contribute very little to the heritability of RDS.[5,16-20] In contrast, mutations in *ABCA3* are present in 3%-5% of the general population in the US and account for approximately 10% of the attributable risk of RDS in infants of European descent, but not in infants of African descent.[21] Race-specific differences in prevalence of variants in these genes, though, suggest these differences may explain the differences in the risk for RDS.

In China, even though there are no data about the population-based epidemiology of RDS, institutional reports suggest that RDS is the most common cause of respiratory disease in term and late preterm infants and the most common reason for mechanical ventilation

Author Affiliations: Division of Neonatology, Department of Pediatrics, the First Affiliated Hospital of Guangxi Medical University, Nanning, China (Chen YJ); Division of Newborn Medicine, Edward Mallinckrodt Department of Pediatrics, Washington University School of Medicine, St. Louis, USA (Chen YJ, Wambach JA, DePass K, Wegner DJ, Heins H, Cole FS, Hamvas A); Department of Pediatrics, Guangxi Maternal and Child Health Hospital, Nanning, China (Chen SK); Division of Statistical Genomics, Department of Genetics, Washington University School of Medicine, St. Louis, USA (Zhang QY)

Corresponding Author: Aaron Hamvas, MD, Division of Neonatology, Ann and Robert H. Lurie Children's Hospital, 225 E. Chicago Ave, Box No. 45, Chicago, IL 60611, USA (Tel: 312-227-4190; Fax: 312-227-9758; Email: ahamvas@luriechildrens.org)

doi: 10.1007/s12519-015-0047-x
Online First November 2015

among all infants with respiratory failure.[22,23] Furthermore, reports that a common non-synonymous variant in *SFTPB*, p.T131I, and a synonymous variant in *ABCA3*, p.P585P, are over-represented in newborns with RDS suggest an underlying genetic susceptibility to RDS in Chinese newborns.[24,25] However, there are 56 different ethnic groups in China and thus, any investigation into the genetic basis of disease must account for genetic diversity among these different groups.

Han is the largest ethnic group in China overall and common variants across the genome have been described and cataloged for this population (www.HapMap.org).[24] Zhuang is the largest minority ethnic group in China centered in the Guangxi Zhuang Autonomous Region in south China. Although there has been less genetic characterization of the Zhuang genome, differences in the apolipoprotein B and sterol regulatory element-binding protein-2 genes have been identified between the Han and Zhuang population, however, even the most remote Zhuang subpopulation (Hei-Yi Zhuang) is not genetically isolated.[26-28]

Because RDS is the primary cause of moderate and severe respiratory failure in late preterm and term infants in China, we hypothesize that some of these cases are attributed to inherited variants in surfactant-associated genes and mutations in surfactant-associated genes are highly prevalent in the general population. However, before embarking on the significantly more complex case-control studies necessary to determine the genetic risk for RDS, we wanted to determine the population-based frequency of rare and common genetic variation in *SFTPB*, *SFTPC*, and *ABCA3* in the Han and Zhuang populations.

Methods

Population-based study cohort

The study cohort was derived from the native population of Nanning, China. Nanning, the capital of the Guangxi Zhuang autonomous region, has a population of more than 7 million and an ethnic profile of 58.02% of minority (http://english.nanning.gov.cn/about_nanning/Overview/2014 10/t20141023_190651.html). Zhuang, the native minority, has always accounted for more than 50% of total population (http://english.nanning.gov.cn/about_nanning/Overview/200809/t200809 20_148567.html). We obtained 756 randomly chosen blood spots from the Guangxi Neonatal Screening Center that were de-identified except for sex, birth weight, gestational age, and self-identified ancestry group, either Han (n=443) or Zhuang (n=313). No other clinical characteristics were linked with the blood spot. The demographics of the population are described in Table 1. These studies were performed with a waiver of consent from the Washington University Human Research Protection Office.

DNA isolation and sequencing

DNA was extracted with chelex-100 and quantified from individual blood spots, as described previously.[29] We divided equimolar amounts of DNA into 4 pools: 245 Han males, 198 Han females, 160 Zhuang males and 153 Zhuang females, respectively. Using the pooled DNA as a template, we performed polymerase chain reaction amplification, library preparation, and sequencing of all coding exons and flanking regions of *SFTPB* [NM_000542.3, HUGO Gene Nomenclature Committee (HGNC): 10801, Gene ID 6439], *SFTPC* (NM_003018.3, HGNC: 10802, Gene ID 6440), and *ABCA3* (NM_001089.2, HGNC: 33, Gene ID 21), for a total of 20703 base pairs per individual, as previously described (Supplemental Table 1).[19,21,30]

Variant detection and validation

After sequencing, we used the short indel prediction by large deviation inference and nonlinear true frequency estimation by recursion, the in-house variant identification algorithm at Washington University, to capture all variants in each pool.[31,32] After annotation using Genome Reference Consortium Human Build 37, we identified all rare non-synonymous variants [minor allele frequency (MAF) <0.05] and tested for predicted functionality using sorting intolerant from tolerant (SIFT) and polymorphism phenotyping (PolyPhen) (http://sift.cchmc.org and http://genetics.bwh.harvard.edu/pph). We defined a mutation as one that was predicted to be "not tolerated" by SIFT and "damaging" by PolyPhen. We also selected 5 variants with MAF >0.20: the common 4 coding region variants

Table 1. Characteristics of study population

Variables	Han (n=443)		Zhuang (n=313)		P value
	Male	Female	Male	Female	
Number (%)	245 (55)	198 (45)	160 (51)	153 (49)	0.26, H *vs*. Z
Birth weight (kg)	3.20±0.40	3.16±0.40	3.21±0.40	3.06±0.40	0.95, HM *vs*. ZM 0.03, HF *vs*. ZF
Gestational age (wk)	38.9±1.3	39.0±1.4	38.9±1.2	38.9±1.5	0.63, HM *vs*. ZM 0.76, HF *vs*. ZF

H: Han; Z: Zhuang; HM: Han male; HF: Han female; ZM: Zhuang male; ZF: Zhuang female.

identified in *SFTPB*, *SFTPC*, and *ABCA3* and one non-coding region variant near exon 5 in *SFTPC* that directs alternative splicing.

We used Taqman® genotyping methods to validate the 23 non-synonymous variants that were predicted by both SIFT and Polyphen to be deleterious and that had not been previously detected in any of our previous studies or in the Exome Sequencing Project database (http://evs.gs.washington.edu/EVS/, accessed June 4, 2013).[33] Seven of these mutations were confirmed (Table 2). We did not perform further validation on the 5 selected common variants because they were found in all pools and with similar frequencies between males and females within the two ethnic groups (data not shown). All novel single nucleotide polymorphisms (SNPs) have been submitted to database of SNPs.

Statistical methods

Because there were no significant differences in variant frequency between the male and female pools (data not shown), we combined the Han male and female pools and compared allele frequencies with the combined Zhuang male and female pool. For common variants, we performed Chi-square analyses for each variant to compare the frequencies between the Han and Zhuang groups. For rare mutations, we collapsed the variants across each gene and performed Fisher's exact analysis to compare the frequency of mutations in each gene between the ethnic groups.[32]

Results

Frequency of common variants

Some minor differences in frequencies of the 5 common variants were noted between the Han and Zhuang groups (Table 3) as has been reported previously for other genes.[26,27] The Han and Zhuang cohort shared 21 of 26 common non-coding variants in *ABCA3*, *SFTPB* and *SFTPC* combined (Supplemental Tables 2&3).

Frequency of rare variants

In the combined Han and Zhuang cohorts, there were no

Table 2. Rare mutations in *ABCA3* identified in Chinese population

Mutation	SIFT score	PolyPhen score	Han (*n*=886 alleles)		Zhuang (*n*=626 alleles)		ESP-ED	ESP-AD
			No. of alleles	Frequency	No. of alleles	Frequency	Frequency	Frequency
p.G205R	0.00	0.999	2	0.002				
p.E292V	0.00	0.999	1	0.001			0.005	0.001
p.L654V	0.01	0.999			1	0.002	0.0003	
p.G668D	0.00	0.992			2	0.003		
p.A823P	0.01	0.992			1	0.002		
p.F144IS	0.00	0.996			1	0.002		
p.G1608C	0.01	1.000			2	0.003		
Collapsed MAF			3	0.003*	7	0.011*		
Collapsed carrier frequency Han and Zhuang combined, %					1.3†			

ABCA3: ATP-binding cassette transporter A3; SIFT: sorting intolerant from tolerant; PolyPhen: polymorphism phenotyping; ESP: exome variant server; ED: European descent; AD: African descent; MAF: minor allele frequency. *: *P*=0.13 for the collapsed frequency of Han *vs.* Zhuang; †: collapsed carrier frequency of rare mutations per individual.

Table 3. Minor allele frequency of common variants in *SFTPC*, *SFTPB*, and *ABCA3*

Variables	Han (*n*=443)	Zhuang (*n*=313)	*P* value, H *vs.* Z*	HAPMAP CHB†	ESP-ED‡ (MAF)	ESP-AD§	*P* value, combined H-Z *vs.* ED‖
SFTPC							
p.T138N rs4715	0.29 (253/443)	0.22 (137/313)	0.004	0.28 (76/272)	0.27 (2277/8296)	0.06	0.180
p.S186N rs1124	0.31 (279/443)	0.28 (176/313)	0.160	0.33 (88/270)	0.33 (2800/8360)	0.11	0.010
c.436-8C>G rs2070687	0.20 (174/443)	0.28 (176/313)	0.001	0.38 (34/90)	0.24 (1965/8338)	0.30	0.720
SFTPB							
p.T143I (formerly p.T131I) rs1130866	0.26 (231/443)	0.24 (153/313)	0.470	0.30 (81/274)	0.48 (4086/8600)	0.26	<0.001
ABCA3							
p.F353F rs13332514	0.41 (367/443)	0.39 (247/313)	0.440	0.38 (103/274)	0.10 (847/8600)	0.09	<0.001

SFTPC: surfactant protein-C; *SFTPB:* surfactant protein-B; *ABCA3:* ATP-binding cassette transporter A3; HAPMAP: haplotype map; CHB: Han Chinese in Beijing, China; ED: European descent; ESP: exome variant server; MAF: minor allele frequency; AD: African descent; H: Han; Z: Zhuang. *: minor allele frequency comparison: Han *vs.* Zhuang; †: minor allele frequency and primary numbers of the CHB in HAPMAP release #28; ‡: minor allele frequency and primary numbers of the European descent population in ESP accessed 09/2013; §: minor allele frequency of the African descent population in ESP accessed 09/2013; ‖: minor allele frequency comparison of the combined H-Z cohort *vs.* the European-descent cohort in ESP.

mutations in *SFTPB* or *SFTPC*. We validated 7 mutations in *ABCA3* (Table 2), 2 of which were present in the exome variant server, including the common disease-associated mutation p.E292V[9,34] and 5 of which were novel. Each of these mutations was distinct to either the Han or Zhuang cohorts. No individual carried more than a single *ABCA3* mutation. The combined, collapsed frequency of rare mutations in *ABCA3* was 0.007, which translates to an *ABCA3* mutation carrier rate of 1.3% and is significantly less than the frequency of 3%-5% reported previously in cohorts of European or African descent ($P<0.001$).[18]

Discussion

Rare mutations in *SFTPB*, *SFTPC*, and *ABCA3* result in lethal neonatal RDS or later-onset interstitial lung disease in term newborns and young children of European or African descent.[7] Furthermore, single mutations in *ABCA3* are over-represented in late preterm infants of European descent with RDS.[21] Therefore, since RDS, especially in term newborns, is a prominent cause for admission to Chinese neonatal intensive care units, we wanted to determine if the carrier rate for mutations in these genes in the Chinese population was similar to or greater than that of the European descent population. Contrary to our initial hypothesis, we found very few mutations in any of these surfactant-associated genes, which suggests that, if there is a genetic susceptibility to RDS in the Han or Zhuang populations, genes involved in other aspects of lung functional development are more likely to be contributing.

This is the largest study to date of multiple surfactant-associated genes that were simultaneously sequenced in the Chinese population. Our study cohort consisted of local residents and was representative of the Han and Zhuang populations of the Nanning District of the Guangxi Zhuang Autonomous Region. The slight differences in the frequencies of the common variants in all 3 genes and the distinct distribution of the rare *ABCA3* mutations are consistent with previous studies that have demonstrated differences in the numbers of variable number tandem repeat elements in the apolipoprotein B gene, on the Y chromosome, and in mitochondrial DNA, and further suggest similar, yet distinct genetic backgrounds of the 2 ethnic groups.[26-28]

The frequency of the common *SFTPB* variant p.T143I, also known as p.T131I (g.1580) in earlier literature, in our cohort was 0.24-0.26, which is approximately half that of the European-descent population, but is similar to the frequency of 0.28 in the Han Chinese population from Beijing in the haplotype map (HAPMAP) database (http://www.ncbi.nlm.nih.gov/projects/SNP/snp_viewTable.cgi?pop=12157). This variant encodes differential glycosylation of the surfactant protein B pro-protein and has been variably associated with multiple respiratory diseases, including RDS, chronic obstructive pulmonary disease, and others.[35-37] A recent study of 80 Han Chinese newborns with and without RDS suggested that this variant was over-represented in those with RDS.[24] However, the 0.12 minor allele frequency in the control population in that study was significantly less than our and previously published cohorts. Combined with our previous work in which we found no differences in frequencies of p.T143I in newborns with and without RDS,[20] the relatively high prevalence of this variant in the general population, the lack of functionality based on computational prediction algorithms, and conflicting results in the literature, we think it is unlikely that this variant provides significant genetic susceptibility to neonatal RDS.

Finally, the common synonymous *ABCA3* variant, p.F353F, is significantly more prevalent in the Chinese population, both in our study and in the HAPMAP (0.39) than in either European (MAF=0.098) or African-descent (MAF=0.087) cohorts in the United States. This variant and another less common synonymous variant, p.P585P, have been associated with RDS in separate studies of a Finnish and a Chinese cohort of premature infants,[25,38] but we did not find this association in late preterm or term infants with RDS.[39] Thus, the contribution of these synonymous *ABCA3* variants to RDS is difficult to determine and will require further study.

As anticipated, we did not find any mutations in *SFTPB* and *SFTPC*, similar to previous studies in the European and African descent populations in the United States.[18,21] However, we were surprised to find that the frequency of deleterious variants in *ABCA3* in both the Han and Zhuang cohorts was lower than observed in the United States. In that study of newborns >34 weeks' gestation, the attributable risk of RDS due to *ABCA3* mutations was approximately 10%.[21] Although we do not have a comparable study in the Chinese population, the significantly lower population-based frequency of *ABCA3* mutations suggests that this gene probably does not account for a significant proportion of RDS.

We were limited in sample size, which also limits the number of rare variants that would be detected.[40] However, the total Chinese cohort size of 756 individuals is similar in magnitude to the 871 in the ED Missouri cohort in which the about 4% frequency of deleterious *ABCA3* mutations using the same methodology was ascertained.[21] Furthermore, as we did not have phenotype information, we can only infer that the contributions of variants in *SFTPB*, *SFTPC*, or *ABCA3* to RDS, based on the population-based frequencies,

is low. We also cannot rule out that variants in non-coding regions of these genes might be influencing the risk of RDS.[15] Further study using a case-control cohort design to determine the contribution of these genes to RDS will be necessary to identify a substantial correlation.

We conclude that deleterious variation in *ABCA3* in the native Chinese Han and Zhuang populations is significantly less than that in populations of European or African descent, and thus, emphasis on genetic impact of *SFTPB*, *SFTPC*, or *ABCA3* for RDS should be conserved. To determine whether there is a significant genetic component to the high rate of RDS in the Chinese population, genes that are important in other aspects of alveolar type II cell metabolism or lung development along with other epigenetic, environmental, or developmental factors should also be investigated.

Acknowledgements

The authors would also like to thank the National Heart, Lung, and Blood Institute GO Exome Sequencing Project and its ongoing studies which produced and provided exome variant calls for comparison: the Lung GO Sequencing Project (HL-102923), the Women's Health Initiative Sequencing Project (HL-102924), the Broad GO Sequencing Project (HL-102925), the Seattle GO Sequencing Project (HL-102926) and the Heart GO Sequencing Project (HL-103010).

Funding: This study was supported by the National Natural Science Foundation of China 81260094 (Chen YJ), National Institutes of Health R01 HL065174 (Cole FS, Hamvas A), R01 HL082747 (Cole FS, Hamvas A), K12 HL089968 (Wambach JA, Cole FS), K08 HL105891 (Wambach JA), Foundation of Health Department of Guangxi Zhuang Autonomous Region 2012059 (Chen YJ).

Ethical approval: These studies were performed with a waiver of consent from the Washington University Human Research Protection Office and approved by the Ethics Committee of The First Affiliated Hospital of Guangxi Medical University.

Competing interest: None of the authors has any financial, personal, or professional interests that could be construed to have influenced the paper.

Contributors: Chen YJ, Hamvas A and Wambach JA contributed to the concept and design. Wegner DJ and Zhang QY contributed to the analysis and interpretation of data. Chen YJ and Hamvas A drafted the article. Hamvas A revised the manuscript. All authors approved the final version.

References

1 Avery ME, Mead J. Surface properties in relation to atelectasis and hyaline membrane disease. AMA J Dis Child 1959;97:517-523.
2 Barber M, Blaisdell CJ. Respiratory causes of infant mortality: progress and challenges. Am J Perinatol 2010;27:549-558.
3 Anadkat JS, Kuzniewicz MW, Chaudhari BP, Cole FS, Hamvas A. Increased risk for respiratory distress among white, male, late preterm and term infants. J Perinatol 2012;32:780-785.
4 Nogee LM, de Mello DE, Dehner LP, Colten HR. Brief report: deficiency of pulmonary surfactant protein B in congenital alveolar proteinosis. N Engl J Med 1993;328:406-410.
5 Nogee LM, Garnier G, Dietz HC, Singer L, Murphy AM, deMello DE, et al. A mutation in the surfactant protein B gene responsible for fatal neonatal respiratory disease in multiple kindreds. J Clin Invest 1994;93:1860-1863.
6 Cole FS, Nogee LM, Hamvas A. Defects in surfactant synthesis: clinical implications. Pediatr Clin North Am 2006;53:911-927, ix.
7 Hamvas A. Evaluation and management of inherited disorders of surfactant metabolism. Chin Med J (Engl) 2010;123:2943-2947.
8 Nogee LM, Wert SE, Proffit SA, Hull WM, Whitsett JA. Allelic heterogeneity in hereditary surfactant protein B (SP-B) deficiency. Am J Respir Crit Care Med 2000;161:973-981.
9 Shulenin S, Nogee LM, Annilo T, Wert SE, Whitsett JA, Dean M. *ABCA3* gene mutations in newborns with fatal surfactant deficiency. N Engl J Med 2004;350:1296-1303.
10 Cameron HS, Somaschini M, Carrera P, Hamvas A, Whitsett JA, Wert SE, et al. A common mutation in the surfactant protein C gene associated with lung disease. J Pediatr 2005;146:370-375.
11 Brasch F, Schimanski S, Mühlfeld C, Barlage S, Langmann T, Aslanidis C, et al. Alteration of the pulmonary surfactant system in full-term infants with hereditary *ABCA3* deficiency. Am J Respir Crit Care Med 2006;174:571-580.
12 Wegner DJ, Hertzberg T, Heins HB, Elmberger G, MacCoss MJ, Carlson CS, et al. A major deletion in the surfactant protein-B gene causing lethal respiratory distress. Acta Paediatr 2007;96:516-520.
13 McBee AD, Wegner DJ, Carlson CS, Wambach JA, Yang P, Heins HB, et al. Recombination as a mechanism for sporadic mutation in the surfactant protein-C gene. Pediatr Pulmonol 2008;43:443-450.
14 Wambach JA, Yang P, Wegner DJ, An P, Hackett BP, Cole FS, et al. Surfactant protein-C promoter variants associated with neonatal respiratory distress syndrome reduce transcription. Pediatr Res 2010;68:216-220.
15 Agrawal A, Hamvas A, Cole FS, Wambach JA, Wegner D, Coghill C, et al. An intronic *ABCA3* mutation that is responsible for respiratory disease. Pediatr Res 2012;71:633-637.
16 Hamvas A. Inherited surfactant protein-B deficiency. Adv Pediatr 1997;44:369-388.
17 Cole FS, Hamvas A, Rubinstein P, King E, Trusgnich M, Nogee LM, et al. Population-based estimates of surfactant protein B deficiency. Pediatrics 2000;105:538-541.
18 Garmany TH, Wambach JA, Heins HB, Watkins-Torry JM, Wegner DJ, Bennet K, et al. Population and disease-based prevalence of the common mutations associated with surfactant deficiency. Pediatr Res 2008;63:645-649.
19 Hamvas A, Wegner DJ, Carlson CS, Bergmann KR, Trusgnich MA, Fulton L, et al. Comprehensive genetic variant discovery in the surfactant protein B gene. Pediatr Res 2007;62:170-175.
20 Hamvas A, Heins HB, Guttentag SH, Wegner DJ, Trusgnich MA, Bennet KW, et al. Developmental and genetic regulation of human surfactant protein B *in vivo*. Neonatology 2009;95:117-124.
21 Wambach JA, Wegner DJ, Depass K, Heins H, Druley TE, Mitra RD, et al. Single *ABCA3* mutations increase risk for neonatal respiratory distress syndrome. Pediatrics 2012;130:e1575-e1582.
22 Qian L, Liu C, Zhuang W, Guo Y, Yu J, Chen H, et al. Neonatal respiratory failure: a 12-month clinical epidemiologic study from 2004 to 2005 in China. Pediatrics 2008;121:e1115-e1124.

23 Ma XL, Xu XF, Chen C, Yan CY, Liu YM, Liu L, et al. Epidemiology of respiratory distress and the illness severity in late preterm or term infants: a prospective multi-center study. Chin Med J (Engl) 2010;123:2776-2780.
24 Yin X, Meng F, Wang Y, Xie L, Kong X, Feng Z. Surfactant protein B deficiency and gene mutations for neonatal respiratory distress syndrome in China Han ethnic population. Int J Clin Exp Pathol 2013;6:267-272.
25 Jiang L, Wu YD, Xu XF, Du LZ. Polymorphism analysis of the *ABCA3* gene: association with neonatal respiratory distress syndrome in preterm infants. Chin Med J (Engl) 2012;125:1594-1598.
26 Yin RX, Chen GQ, Wang Y, Lin WX, Yang DZ, Pan SL. Effect of the 3'APOB-VNTR polymorphism on the lipid profiles in the Guangxi Hei Yi Zhuang and Han populations. BMC Med Genet 2007;8:45.
27 Deng YJ, Yin RX, Li YY, Zhou YJ, Lin WX, Pan SL, et al. Polymorphism of the sterol regulatory element-binding protein-2 gene and its association with serum lipid levels in the Guangxi Hei Yi Zhuang and Han populations. Am J Med Sci 2009;337:14-22.
28 Zhao Q, Pan S, Qin Z, Cai X, Lu Y, Farina SE, et al. Gene flow between Zhuang and Han populations in the China-Vietnam borderland. J Hum Genet 2010;55:774-776.
29 Hamvas A, Trusgnich M, Brice H, Baumgartner J, Hong Y, Nogee LM, et al. Population-based screening for rare mutations: high-throughput DNA extraction and molecular amplification from Guthrie cards. Pediatr Res 2001;50:666-668.
30 Metzker ML. Sequencing technologies-the next generation. Nat Rev Genet 2010;11:31-46.
31 Druley TE, Vallania FL, Wegner DJ, Varley KE, Knowles OL, Bonds JA, et al. Quantification of rare allelic variants from pooled genomic DNA. Nat Methods 2009;6:263-265.
32 Li B, Leal SM. Methods for detecting associations with rare variants for common diseases: application to analysis of sequence data. Am J Hum Genet 2008;83:311-321.
33 Shen GQ, Abdullah KG, Wang QK. The TaqMan method for SNP genotyping. Methods Mol Biol 2009;578:293-306.
34 Garmany TH, Moxley MA, White FV, Dean M, Hull WM, Whitsett JA, et al. Surfactant composition and function in patients with ABCA3 mutations. Pediatr Res 2006;59:801-805.
35 Lin Z, Pearson C, Chinchilli V, Pietschmann SM, Luo J, Pison U, et al. Polymorphisms of human *SP-A*, *SP-B*, and *SP-D* genes: association of *SP-B* Thr131Ile with ARDS. Clin Genet 2000;58:181-191.
36 Floros J, Fan R, Diangelo S, Guo X, Wert J, Luo J. Surfactant protein (SP) B associations and interactions with SP-A in white and black subjects with respiratory distress syndrome. Pediatr Int 2001;43:567-576.
37 Guo X, Lin HM, Lin Z, Montaño M, Sansores R, Wang G, et al. Surfactant protein gene A, B, and D marker alleles in chronic obstructive pulmonary disease of a Mexican population. Eur Respir J 2001;18:482-490.
38 Karjalainen MK, Haataja R, Hallman M. Haplotype analysis of *ABCA3*: association with respiratory distress in very premature infants. Ann Med 2008;40:56-65.
39 Wambach JA, Wegner DJ, Heins HB, Druley TE, Mitra RD, Hamvas A, et al. Synonymous *ABCA3* variants do not increase risk for neonatal respiratory distress syndrome. J Pediatr 2014;164:1316-1321.e3.
40 Kruglyak L, Nickerson DA. Variation is the spice of life. Nat Genet 2001;27:234-236.

Received April 19, 2015
Accepted after revision July 14, 2015

(Supplementary information is linked to the online version of the paper on the *World Jounal of Pediatrics* website)

Supplementary information

Table 1. PCR primers

Amplicon (48)	Size (bp)	Forward primer	Reverse primer
*ABCA3*_ex2	313	CACTCAAACACCTTCCATCTGTCCAA	CAGGGCTGGGAGAGAAGGTCAGAAA
*ABCA3*_ex3	419	CGTGCATCTTAACCTGGCTGATGGA	AAGGAACACAGACACTGAACCCAGA
*ABCA3*_ex4-5	672	CCAAATCCCCACTCTG	AGGCCAAGTCTGCACAGGGTGAACT
*ABCA3*_ex6	345	CCGTCTTTCATCTGCCAGTGACCTG	TGACTTGCAGGCAGGCAGAGGTTTA
*ABCA3*_ex7	367	AGGGACCACTCAGTGTGACATTCCG	GGCTGGTAACACGAACCCTAACCGA
*ABCA3*_ex8	426	TGAGCTGAAGTCACTCTGTTGCCCC	ACAGCGCGGTTTCTAGAGTGTTGGG
*ABCA3*_ex9	373	CTGCTGGGACAGTCGGACTCAGG	CACCGAGAGGAGTGGGACATTGACA
*ABCA3*_ex10	468	GGGCCCTCTTGGGAAGAACTTTGTG	CGCTGACTTTCCTCCTTCCAGTCCA
*ABCA3*_ex11	353	GTGCTGGAGCTTGTGTCCCGTGTAG	ACAGGCTGGACAAGGCAAACACTCA
*ABCA3*_ex12	272	GGGCCACTTTCCTGATGTGTCTTCC	GGTACTGGGGACACCTCTGCACTCA
*ABCA3*_ex13	391	AAGTTGGGACTCTCTGGGGCTCTCC	TATGAGGTCTCACTGCCGTGCTGGT
*ABCA3*_ex14	505	CTCAGGAAATGCCAGACTCAGCCGT	GAAAGCCCCATTGAGGGAGTGAGG
*ABCA3*_ex15	446	GTGTCGTGGGTTTCTCCTCCCTGAC	GAGCACATCAGTGGAAACACCCCTG
*ABCA3*_ex16	314	ATCTCCCTGCGTCCCCTGT	GGCTTG AGTCCTCCAAGGATGGTGA
*ABCA3*_ex17	453	GACAAGGCCATCACCATCCTTGGAG	CTAGAAAAGGCCACCCCTGCCTCAT
*ABCA3*_ex18	322	TTGCCTGGTGTGAGCCCTAGAGACC	CCTCTGAGCACAAAGCCCTCATGG
*ABCA3*_ex19	230	ACTGTGCCTGGCCGAGGGG	AGCCCAGTCCTAGGTGGACGGAAAT
*ABCA3*_ex20	351	TGCTCCCCTCAGTGCCTCTAACCAT	CTGCATGGGCTTACATGAGGCGTTT
*ABCA3*_ex21	567	AGAGTCTGCACAGGTGACCCTGCC	GGCGAACTCTGGCTGCAGGACT
*ABCA3*_ex22	593	CCGGTTCCTGACTGGCATTCAAAGT	TTGGGAGGGCAGACACAATGCTCTA
*ABCA3*_ex23	511	GTGCTCCGTCCCTGACCTTCTCTGT	ACCCTCTCTGTCCCAGTTTCCCCTC
*ABCA3*_ex24	386	AGGGGTCTGAGGACCTCCAAATGCT	CTCCCTGTCTGGGCGGAGTGG
*ABCA3*_ex25	405	CCCTCACTCCACACAGCACGGATAA	AAGGCGGTACAGAGGAACGCACCAG
*ABCA3*_ex26	473	TCGAGAGGCAGCTGTGACCTACTGG	CTGAGGCCGTACAGTGGGAGACCAT
*ABCA3*_ex27-28	662	CATGCGGTCTTTGTCCTGGTCAATC	CTTGTCTCGCTGTCCAGAGGCATGT
*ABCA3*_ex29	467	TGTGTCCCTGTTCCAAGAGCTTCCA	GAGCGGTCACTCCCAGCTCTATGCT
*ABCA3*_ex30	421	TTTCCAGGTGCACACACAGCTCCTT	CTCTGCACCAGATGCTGATGGGTCT
*ABCA3*_ex31-32	725	ATCAGGAACAGCCTGATCGGAGAGC	AAAACCCCCAAACCAGCACGTATCA
*ABCA3*_ex33a	577	GGCTCAGAAAGGGAACATCACTGGC	GCTGCACTCGTCCATTCTGTGCATAC
*ABCA3*_ex33b	863	CAAGCAGGGCCCATCTTACATCCTC	CAACGTCCCACGTTTGTGTGATTGA
*SFTPB*_ex1	242	CCTGGAGGGCTCTTCAGAGCAAAGA	GCTCAGTGAGTGGTGGAGCTGCCTA
*SFTPB*_ex2	401	AAAGACAAGGCAGCTGGGGTTCAGA	CACCCAGCACCCTTCATTTCAGACC
*SFTPB*_ex3	494	GGGATGGGATGGGATGACACAGAAT	TCAGGGAAGACCATCTCTGGCTGTG
*SFTPB*_ex4	406	GTCATGGCCCTGAGCTCAATAGCAC	CTCCCCATGGGTGGGCACAG
*SFTPB*_ex5	649	ACGCTCACACACCCTTACACCCTCA	CAGGCTCTCCTCCCCTCTCTCTTCC
*SFTPB*_ex6	213	GGAGAGCCTGGAGGACTCTTCTCCC	GCTGCAGGGAGCTACAGGTATGCGT
*SFTPB*_ex7	319	CAGAGAGTGGAGGCTTGCCAAGTGA	AGAGGGTCATGCAGTGGCAGGGTAG
*SFTPB*_ex8	430	CACTCCTTAGCCCAATGCCTGCTCT	CCTGTCTGCCTGTCTGTGCTCCATT
*SFTPB*_ex9	560	TGAGATTCCACCCCTCTGCCTGAGT	ATAATGGACATCCAGCCGCACTCCT
*SFTPB*_ex10	347	GTTCTTTCCCGGAAGAGCTGGGTGT	GTCCGGAAAGGGTGTTTTCCTGATG
*SFTPB*_promoter	729	CCCTCAGCCTGTGAGCTTTTTCTCC	TCTGTAGGAGTGGCAGCGACCTCAG
*SFTPC*_ex1	220	CTGTCCCCTCTCCCTACGGACACAT	GGATAAGGAAACAGGCCAGGGAGGA
*SFTPC*_ex2	332	CCGTGGGAGGGTGTTCAGCTTGTAT	TTGGACAGTTTCCTATCGCCCATCC
*SFTPC*_ex3	216	GTAGGAAAGGGGAAGACCAGGTGGC	AGCAGAGCCTTGTCATTGGTCCCAG
*SFTPC*_ex4	300	GTATGACTCCCGTGCCCAACCTAGA	ACAGGTGACAAATAGCCACCATTCGG
*SFTPC*_ex5	567	CCGAATGGTGGCTATTTGTCACCTG	GATGACCCCGCTTCAGTGGACG
*SFTPC*_upstream	608	TCTGGCTGCTGG AGTCTTAGGCAAA	CTGGGCTCGCTCCCTTAACCTCCTA

ABCA3: ATP-binding cassette transporter A3; *SFTPB*: surfactant protein-B; *SFTPC*: surfactant protein-C; PCR: polymerase chain reaction; ex: exon.

Table 2. Number of common SNPs per gene in each cohort

Genes	Subgroup	Number of SNPs	Total SNPs in both Han and Zhuang	Identical SNPs in Han and Zhuang
ABCA3	Han	7	8	7
	Zhuang	8		
SFTPB	Han	6	8	4
	Zhuang	6		
SFTPC	Han	10	10	10
	Zhuang	10		

ABCA3: ATP-binding cassette transporter A3; *SFTPB*: surfactant protein-B; *SFTPC*: surfactant protein-C; SNPs: single nucleotide polymorphisms.

Table 3. rs numbers of SNPs that were similar and different between cohorts

	ABCA3	*SFTPB*	*SFPTC*
Variants common to the Han and Zhuang cohorts	rs13332514 rs13332547 rs4787273 rs170447 rs2240523 rs313909 rs75808174	rs3024798 rs1130866 rs893159 rs2077079	rs13248346 rs2070684 rs2070685 rs2070686 rs2070687 rs1124 rs8192337 rs6557857 rs78177348 rs4715
Variants different between the Han and Zhuang cohorts	rs2302035	rs762548 rs34024265 rs2304566 rs3024791	None

ABCA3: ATP-binding cassette transporter A3; *SFTPB*: surfactant protein-B; *SFTPC*: surfactant protein-C; SNPs: single nucleotide polymorphisms.