Family-Based Association Study of Pulmonary Function in a Population in Northeast Asia

Citation

Published Version

Permanent link

Terms of Use

Share Your Story

Family-Based Association Study of Pulmonary Function in a Population in Northeast Asia

Ho-Young Son[1]☯‡, Seong-Wook Sohn[2]☯‡, Sun-Hwa Im[1,3]☯‡, Hyun-Jin Kim[4], Mi Kyeong Lee[5], Bayasgalan Gombojav[4,5], Hyouk-Soo Kwon[6], Daniel S. Park[7], Hyung-Lae Kim[8], Kyung-Up Min[9], Joohon Sung[5], Jeong-Sun Seo[1,4,10]*, Jong-Il Kim[1,4,10,11]*

1 Department of Biochemistry and Molecular Biology, Seoul National University College of Medicine, Seoul, Republic of Korea, 2 Department of Internal Medicine, Dongguk University Ilsan Hospital, Goyang, Republic of Korea, 3 Department of Obstetrics and Gynecology, College of Medicine, Chung-Ang University, Seoul, Republic of Korea, 4 Genomic Medicine Institute, Medical Research Center, Seoul National University, Seoul, Republic of Korea, 5 Department of Epidemiology and Institute of Environment and Health, School of Public Health, Seoul National University, Seoul, Republic of Korea, 6 Department of Internal Medicine, Asan Medical Center, Seoul, Republic of Korea, 7 Department of Molecular and Cellular Biology, Harvard University, Cambridge, Massachusetts, United States of America, 8 Department of Biochemistry, Ewha Womans University, School of Medicine, Seoul, Republic of Korea, 9 Department of Internal Medicine, Seoul National University College of Medicine, Seoul, Republic of Korea, 10 Department of Biomedical Sciences, Seoul National University Graduate School, Seoul, Republic of Korea, 11 Cancer Research Institute, Seoul National University College of Medicine, Seoul, Republic of Korea

☯ These authors contributed equally to this work.
‡ These authors are co-first authors on this work.
* jongil@snu.ac.kr (JIK); jeongsun@snu.ac.kr (JSS)

Abstract

The spirometric measurement of pulmonary function by measuring the forced expiratory volume in one second (FEV_1) is a heritable trait that reflects the physiological condition of the lung and airways. Genome-wide linkage and association studies have identified a number of genes and genetic loci associated with pulmonary function. However, limited numbers of studies have been reported for Asian populations. In this study, we aimed to investigate genetic evidence of pulmonary function in a population in northeast Asia. We conducted a family-based association test with 706 GENDISCAN study participants from 72 Mongolian families to determine candidate genetic determinants of pulmonary function. For the replication, we chose seven candidate single nucleotide polymorphisms (SNPs) from the 5 loci, and tested 1062 SNPs for association with FEV_1 from 2,729 subjects of the Korea Healthy Twin study. We identified *TMEM132C* as a potential candidate gene at 12q24.3, which is a previously reported locus of asthma and spirometric indices. We also found two adjacent candidate genes (*UNC93A* and *TTLL2*) in the 6q27 region, which has been previously identified as a pulmonary function locus in the Framingham cohort study. Our findings suggest that novel candidate genes (*TMEM132C*, *UNC93A* and *TTLL2*) in two different regions are associated with pulmonary function in a population in northeast Asia.

Introduction

Pulmonary function, which is commonly measured by spirometry, is a good index of the physiological condition of the lung and airways [1]. Forced expiratory volume in one second (FEV_1) and its ratio to the forced vital capacity (FEV_1/FVC) are used to predict population morbidity and mortality [2–4] and are used in the clinical diagnosis of chronic obstructive pulmonary disease (COPD) and asthma [5]. Family studies have reported COPD aggregation and significant heritability of spirometry-measured pulmonary function [1, 6, 7].

To discover genetic loci related with pulmonary function and COPD, several linkage studies have been conducted using the quantitative spirometry measures FEV_1, FVC, and the FEV_1/FVC ratio [8, 9]. The Framingham Heart cohort study suggested that the linkage locus for FEV_1 was on the terminus region of chromosome 6q [10, 11].

In recent years, genome-wide association studies (GWAS) have revealed various genes showing significant associations with pulmonary function [12, 13]. Moreover, large-scale GWAS meta-analyses from CHARGE and the SpiroMeta consortium each had more than 20,000 participants and identified novel genome-wide significant loci associated with pulmonary function in the general population [14–16]. Although numerous genome-wide linkage and association studies have revealed a number of genes and genetic loci associated with pulmonary function, these genetic results explain only a small proportion of the genetic variation needed to estimate the heritability and variance of the trait [17]. In addition, most previous association studies on pulmonary function have focused on populations of European ancestry and not on those of Asian ancestry.

As part of the GENDISCAN (GENe DIScovery for Complex traits in isolated large families of Asian of Northeast) project, which was designed to investigate genetic loci associated with complex traits in extended rural families [18–20], we conducted a family based association study of pulmonary function in a Mongolian population. Subsequently, we validated the association between the candidate loci identified by our study and the pulmonary function phenotype in a population of Korean subjects to confirm the genetic evidence in two different populations.

Materials and Methods

Study design and population

From April to June 2006, a total of 2,008 participants for pulmonary function measurement were recruited in Dashbalbar, Dornod Province, Mongolia, as part of the GENDISCAN project [18, 20, 21], which was designed to identify genetic loci of complex traits in Asian populations. For this study, we selected 706 subjects from 72 large and extended pedigrees with complete genotyping. We extracted genomic DNA from peripheral blood leukocytes of all subjects according to standard protocols. Written informed consent was obtained from all participants, and our study protocols were approved by the institutional review board of Seoul National University (Approval number, H-0307-105-002). This study abided by the principles of the Declaration of Helsinki.

Phenotype measurement

Spirometry was performed with a portable spirometer (MicroPlus spirometer, Micro Medical Ltd, Rochester, Kent, England) according to American Thoracic Society criteria [22]. FEV_1 and FVC were measured, and FEV_1 was used to evaluate airway obstruction. After several practice measurements with a trained technician, measurements for data collection were taken three times. Of these measurements, the greatest FEV_1 values from acceptable tests for each subject

were selected. The subjects were instructed to take complete inspirations and expirations that lasted approximately 3 s [23]. The predicted FEV_1 and FVC values were obtained from the methods of Morris [24]. Smoking history data (current smoking, former smoking, and never smoked) was collected by questionnaire and the number of pack-year was calculated using the smoking amount (pack/day) and the duration (smoking years). Information on if the participants have any of the following respiratory diseases was also collected by questionnaire: chronic obstructive lung disease, chronic bronchitis, pneumonia, asthma or tuberculosis.

Genome-wide SNP genotyping

Details of the genotyping methods used have been previously reported [18, 25]. In brief, 706 samples from 72 extended families were genotyped using an Illumina Human 610-Quad Bead-Chip kit (San Diego, CA). To maintain the quality of the genotyping data, we checked for genotyping error at several steps before analysis. SNPs with a call rate < 99%, an error rate > 1%, minor allele frequency < 1% and a Hardy-Weinberg equilibrium $P < 1 \times 10^{-6}$ were excluded.

Statistics

Before the association test, the subject's measurements of pulmonary function were adjusted for age, age^2, sex, height, dummy variables of smoking status (current, former, never) and pack-year as covariates using the sequential oligogenic linkage analysis routines (SOLAR) package [26]. To reduce deviations from normality and the effect of outliers, we normalized the phenotypes using the inverse normal transformation option in the SOLAR package. We conducted genome-wide family-based association analysis using the additive model to investigate genetic regions that might influence pulmonary function using FBAT software version 2.0.3 [27].

Replication test

We compared the results of the GENDISCAN study with data from the Korea Healthy Twin Study. The Healthy Twin Study is an ongoing cohort study of twins and their families that was initiated in 2005. The details of the study's protocols, measurements, genotyping and imputation methods have been reported previously [28, 29, 30]. In brief, a total of 2,729 participants were recruited and genotyped using the Affymetrix Genome-wide Human SNP Array version 6.0 (Affymetrix Inc., Santa Clara, California, USA). The untyped SNPs were imputed to the HapMap3 phase 2 (JPT+CHB) and Korean HapMap (http://www.khapmap.org) reference panel using Beagle (University of Washington, Seattle, WA, USA). We selected 1062 SNPs for replication analysis that were located within 200 kb of the associated SNPs of discovery stage.

Results

The descriptive characteristics of our discovery and replication subjects are presented in Table 1. In the discovery stage, 706 individuals from the 72 families from the Mongolian population were assessed. The majority (88.7%) ethnicity of this cohort was Buryats. In the replication study, 623 families from 2,729 Korean participants were included. The mean age of the Mongolian population was 30.3 years, and the mean age of Korean population was 44.2 years. The mean (S.D.) FEV_1 values were 2.684 (0.717) and 2.915 (0.709), respectively. With respect to smoking status, Koreans were more likely to be current smokers (27.6%) and had higher mean pack-year (17.63) than the Mongolians.

The results of the family-based association between the genome-wide SNPs and FEV1 are shown in a Manhattan plot (Fig 1). Each of the approximately 510,000 SNPs is represented by

**Table 1. Descriptive characteristics of the participants.**

	Discovery stage	Replication study
Subject information		
No. of families.	72	623
Total subjects number	706	2729
Age, years	30.3 ± 14.9	44.2 ± 13.1
Female, N (%)	366 (51.8)	1667 (61.1)
Height (cm)	156.2 ± 10.9	161.6 ± 8.5
FEV_1 (L)	2.684 ± 0.717	2.915 ± 0.709
Smoking Status		
Never smokers, N (%)	565 (80)	1776 (65.1)
Current smokers, N (%)	129 (18.3)	754 (27.6)
Former smokers, N (%)	12 (1.7)	199 (7.3)
Smoking, Pack-years	8.77 ± 9.73	17.63 ± 16.99

Data represent the mean values ± standard deviation unless otherwise indicated.
FEV_1, forced expiratory volume in one second.

doi:10.1371/journal.pone.0139716.t001

single dot. Five regions were identified that contained seven suggestively associated SNPs with P values less than 1×10^{-5} (Table 2). The most significant association was observed in the intronic region of *TMEM132C* (rs12582875, $P = 2.17 \times 10^{-6}$) at 12q24.3 (Fig 2a). On chromosome 6q27, two SNPs (rs4710230 and rs3010558, $P = 2.77 \times 10^{-6}$ and $P = 8.99 \times 10^{-6}$, respectively) were located 21.7 Kb and 7.8 Kb upstream of *UNC93A* (Fig 2c). On chromosome 3p14.1, rs264676 ($P = 3.28 \times 10^{-6}$) was located in the *MAGI1* intron. We found one SNP (rs7504607) at 18q23, was associated with FEV_1 phenotype, but has no candidate gene within 200 kb. At chromosome 4q27, rs6855113 and rs6831851 (both $P = 9.48 \times 10^{-6}$) were located 17kb and 33 kb upstream of the *MAD2L1* gene, respectively.
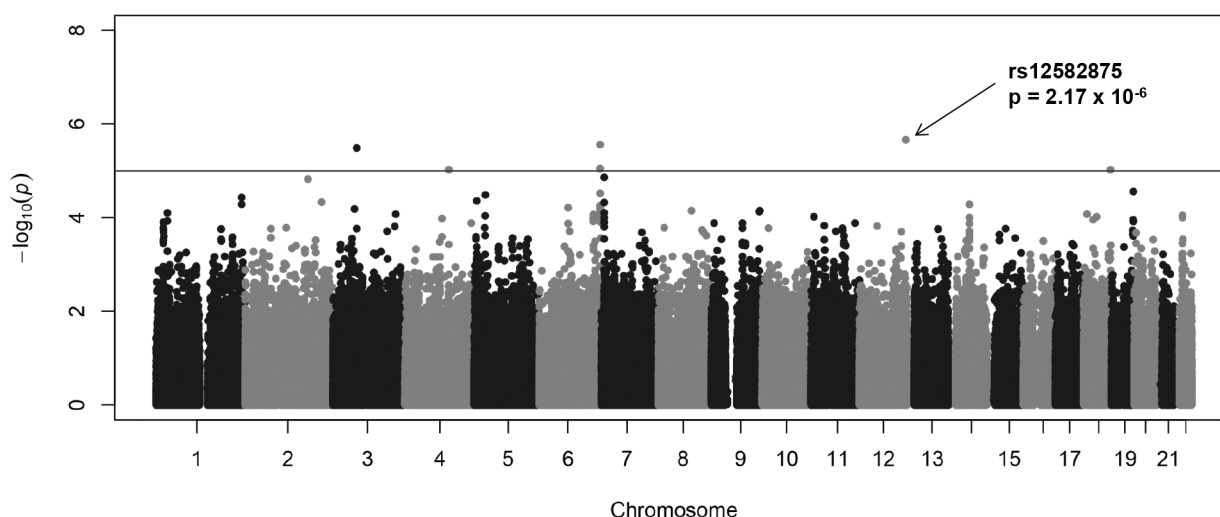
Fig 1. Manhattan plot of the genome-wide association signal with FEV_1. The x-axis represents the SNP markers on each chromosome. y-axis shows the $-\log_{10}(P$ value). The blue horizontal line represents the genome-wide suggestive threshold $P = 1.0 \times 10^{-5}$. The greatest P value ($P = 2.17 \times 10^{-6}$) was observed in rs12582875 on chromosome 12q24.3.

doi:10.1371/journal.pone.0139716.g001

Table 2. Results of the family based association test for FEV$_1$ (P value < 1.0 × 10^{-5}) in the discovery stage.

Chr	SNP	Position*	Nearest gene	Nearest gene distance	Minor Allele	MAF	P value
12	rs12582875	127598647	TMEM132C	0	A	0.161	2.17 × 10^{-6}
6	rs4710230	167603059	UNC93A	21733	T	0.387	2.77 × 10^{-6}
6	rs3010558	167616938	UNC93A	7854	T	0.383	8.99 × 10^{-6}
3	rs264676	65861112	MAGI1	0	G	0.307	3.28 × 10^{-6}
18	rs7504607	73349291	GALR1	238207	C	0.458	9.45 × 10^{-6}
4	rs6855113	121225082	MAD2L1	17621	T	0.079	9.48 × 10^{-6}
4	rs6831851	121241185	MAD2L1	33724	C	0.079	9.48 × 10^{-6}

* SNP positions are based on NCBI Build 36.

Chr, Chromosome; MAF, Minor allele frequency.

doi:10.1371/journal.pone.0139716.t002

Because it is important to take into account the effect of the respiratory disease on pulmonary function, we verified if any participants had respiratory diseases. Thirty-two participants from the discovery stage had one or more conditions among 5 respiratory-related diseases (chronic obstructive lung disease, chronic bronchitis, pneumonia, asthma and tuberculosis). To find out the extent of these participants' effect on the result, we excluded these participants with respiratory diseases and repeated the test for genetic association with FEV$_1$. The result of the seven associated SNPs is shown in S1 Table.

For the replication study, we conducted an association test between the seven SNPs from the discovery stage and lung function in 2,729 subjects from 623 families from the Korea Healthy Twin study. For the different genotyping platforms, overlapping SNPs were not sufficient. Therefore, we performed imputation, in which 5 SNPs of the 7 associated SNPs from the discovery stage were included. However, we could not find any significant association (P value < 0.05) between these SNPs and lung function in the replication study. Of the five SNPs, we found that the allele frequency of four SNPs between the discovery and replication stage was significantly different (S2 Table). These different allele frequencies suggest that the genetic architecture of two populations might be quite different. Therefore, we considered a locus specific replication.

We selected a total of 1062 SNPs from the regions surrounding the seven candidate SNPs, including 200 kb up and downstream of the SNPs. The candidate regions with the most associated SNPs that have a P value < 1 × 10^{-3} are listed in Table 3. Twelve SNPs on and near three genes were shown to have associations with FEV$_1$. Of the 225 SNPs near *TMEM132C*, rs1905942 and rs1112925 (P = 2.51 x 10^{-4}) showed the highest association (Fig 2b). There was no LD between rs1112925 and the top SNPs of the discovery stage (rs12582875) because the two SNPs were located 177 kb away from each other. On chromosome 6, ten candidate SNPs were located within two adjacent genes (*UNC93A* and *TTLL2*) (Fig 2d). The six SNPs were located in the intronic region of *UNC93A*. rs2981977, rs3010556 and rs877653 were located in the *TTLL2* intron and exon, whereas rs3010562 is 9 kb downstream of this gene. These SNPs near *UNC93* and *TTLL2* were in LD, but there was no LD between these replicated SNPs and the most significant SNP (rs4710230) of the discovery stage.

Discussion

We conducted a genome-wide association study to find genetic evidence of spirometric measures of pulmonary function in large extended families in a Mongolian population. In addition, we validated these candidate loci in 2,729 Korean individuals, and we could identify three candidate genes: *UNC93A*, *TTLL2* and *TMEC132C*.
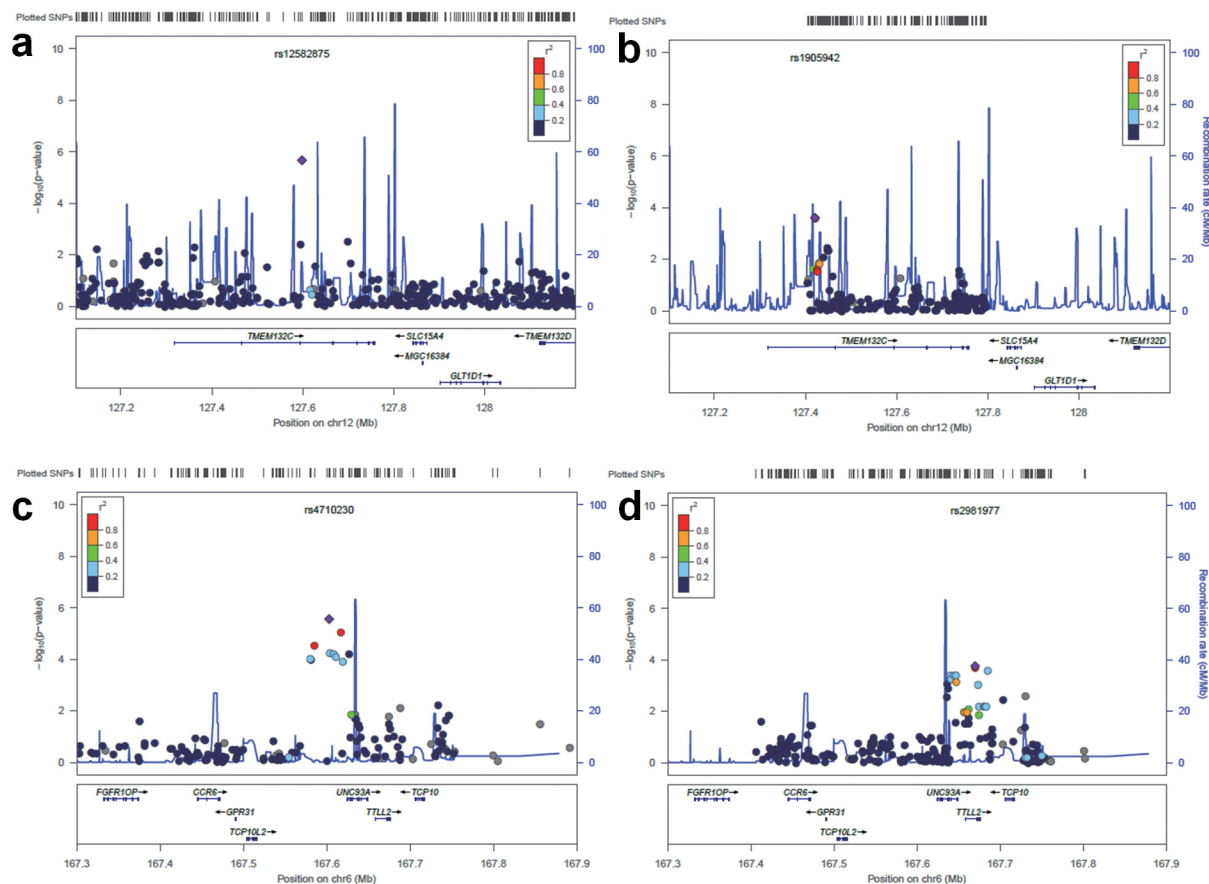
Fig 2. Regional plots for discovery and replication loci associated with FEV₁. The purple diamonds indicate the most significant SNP of each region, and nearby SNPs are color coded according to the level of LD with the top SNP. The x-axis shows chromosomal position. The left y axis shows the significance of the association, and the right y-axis shows a recombination rate across the region. Estimated recombination rates from the 1000 Genome (JPT+CHB, hg18) database are plotted with the blue line to reflect the local LD structure. Discovery (**a**) and replication (**b**) result of the *TMEM132C* region on 12q24.3. Discovery (**c**) and replication (**d**) result of the *UNC93A* and *TTLL2* region on 6q27. The regional plots were created using LocusZoom.

doi:10.1371/journal.pone.0139716.g002

UNC93A and *TTLL2* are located on chromosome 6q27, which is a region that has previously been reported to be associated with pulmonary function in many different studies. The linkage study in the Framingham cohorts revealed the greatest linkage peak at the q terminus of chromosome 6 for FEV₁ [10]. Additional markers, significant evidence of linkage and family-based association were reported at 184.5 cM of chromosome 6 in 1,115 individuals from the 182 large extended Framingham Heart Study families [11]. This linkage result was replicated again in similar regions of chromosome 6 from a different study that used 100K SNP GeneChip [13]. Subsequently, secreted modular calcium-binding protein 2 (*SMOC2*) at 6q27 was selected as a candidate gene to study, and an association with the pulmonary function phenotype was reported [31]. However, little is known about the detailed molecular and physiological functions of *SMOC2*.

Consistent with these previous findings, we identified a suggestive association with FEV1 and the 6q27 region. The SNP with the greatest association (rs4710230, $P = 2.77 \times 10^{-6}$) on chromosome 6q27 was located 21.7 Kb upstream of *UNC93A*. Moreover, *TTLL2*, *TCP10L2*, *TCP10*, *GPR31* and *CCR6* were within 200 kb. Through the replication study, ten SNPs on or near *UNC93A* and *TTLL2* were validated. To find the functional relevance of the associated SNPs, we verified regulatory chromatin states using HaploReg v3. The region of rs4710230 was

Table 3. Replication result of the associated SNPs (P value < 1.0×10^{-3}) in the Korea Healthy Twin Study.

Chr	SNP	Position*	Gene	Gene Position	Minor Allele	MAF	P value
6	rs2981977	167670140	TTLL2	Intron	G	0.307	1.76×10^{-4}
6	rs3010556	167670167	TTLL2	Intron	G	0.308	2.10×10^{-4}
12	rs1905942	127420950	TMEM132C	Intron	C	0.5	2.51×10^{-4}
12	rs1112925	127421487	TMEM132C	Intron	A	0.5	2.51×10^{-4}
6	rs3010562	167685241	TTLL2	Intergenic	C	0.477	2.64×10^{-4}
6	rs4709162	167640315	UNC93A	Intron	A	0.473	4.03×10^{-4}
6	rs4709165	167644640	UNC93A	Intron	G	0.459	4.09×10^{-4}
6	rs4709167	167646922	UNC93A	Intron	C	0.459	4.09×10^{-4}
6	rs3817767	167640832	UNC93A	Intron	C	0.464	5.74×10^{-4}
6	rs2294236	167647304	UNC93A	Intron	A	0.278	7.17×10^{-4}
6	rs3010550	167636618	UNC93A	Intron	C	0.329	8.82×10^{-4}
6	rs877653	167673681	TTLL2	Exon	C	0.458	9.57×10^{-4}

* SNP positions are based on NCBI Build 36.

Chr, Chromosome; MAF, Minor allele frequency.

doi:10.1371/journal.pone.0139716.t003

reported as a strong enhancer of *UNC93A* in HepG2 cells. The LD block of the replicated SNPs (rs4709265, rs4709167 and rs3817767) was reported as a strong and weak enhancer region. The rs2981977 and rs3010556 regions were reported as weak enhancers of *TTLL2* in HepG2 cells. A genome-wide association of methylation study reported that rs3010556 is significantly associated with the CpG (cg13033054) site in a *trans* manner [32]. rs877653 is located on third exon of *TTLL2* as a synonymous variant.

UNC93A and *UNC93B1* are human homologues to the *Caenorhabditis elegans (C. elegans) UNC93* gene. In *C. elegans*, *unc–93* is one of five genes in an interacting set (*unc–93*, *sup–9*, *sup–10*, *sup–11* and *sup–18*) involved in the contraction and coordination of muscle. Mutations in these genes produce the characteristic "rubber-band" phenotype in worms [33]. Little is known about the molecular function of human *UNC93A*. Human *UNC93B1*, however, has been identified to be related to left ventricular diastolic function, heart failure morbidity and mortality [34]. FEV_1 is determined by airway obstruction and reflects the degree of airway obstruction in obstructive lung diseases, such as asthma and COPD. Asthma is a chronic inflammatory airway disease characterized pathologically by airway smooth muscle hypertrophy and contractility dysfunction. Because of the association between *UNC93* and muscle contraction, we posited that *UNC93* gene function might influence smooth muscle function in the airway. Moreover, murine *UNC93B* has been identified as a novel component of the innate and the adaptive immune response [35]. Asthma has also been regarded as an inflammatory disease mediated by immunologic T helper type 2 (Th2) lymphocytes and *UNC93* is thought to affect this immunologic balance. Therefore, SNPs in *UNC93* could be affecting pulmonary function by negatively influencing smooth muscle contraction and/or immunologic responses in COPD and asthma. The significant SNPs identified by the discovery and replication studies were located on different LD blocks. Each SNP in *UNC93A* might also have a separate effect on the FEV1 phenotype, so this gene by itself might be influencing pulmonary function in the general population.

Another candidate gene, *TTLL2* (Tubulin tyrosine ligase-like family, member 2), is thought to be related to tubulin glutamylase, which forms polyglutamate side chains on tubulin in airway epithelial cilia [36]. Previous studies have shown that *TTLL1* depletion resulted in a loss of tubulin glutamylation and disrupted the beating of airway cilia. Moreover *TTLL1* deficiency

resulted in chronic sinusitis and abnormal development of spermatid flagella in mice [37]. From these previous results we postulated that *TTLL2* might be able to affect ciliary movement in the lung. The consequences of impaired ciliary function are abnormal mucus clearance from the airways and increased respiratory infection consistent with a condition such as COPD or asthma. *TTLL2* is located right next to the *UNC93A* gene, and the two genes are in the same LD block. Therefore, functional studies will be required to identify whether one or both of these genes influence pulmonary function.

In addition, we discovered *TMEM132C* as a novel candidate gene for FEV1 and pulmonary function. On chromosome 12, the SNP in the *TMEM132C* gene showed the highest association (rs12582875, p = 2.17×10^{-6}) with FEV1 phenotype in the Mongolian population. We checked the regulatory chromatin states using HaploReg v3. However, no functional relevance was found in the region of the *TMEM132C* SNPs.

The molecular function of *TMEM132C* has not yet been identified. *TMEM132C* is a subfamily of the *TMEM132* gene family, which has 5 subtypes (*TMEM132A—E*). *TMEM132B* is located on chromosome 12q24.31, *TMEM132C* and *TMEM132D* are tandemly located on chromosome 12q24.32 – 12q24.33. Chromosome 12 terminus region (12q24.3) is a suggestive candidate locus for asthma and spirometric indices. Studies in various populations have reported an association between 12q24.3 regions and asthma [38–40]. Recently, rs2030436 in the *TMEM132D* gene was reported to have an association with lung function decline in a mild COPD genome-wide study [41]. Suggesting the relationship between pulmonary function and the *TMEM132* family, we concluded that the *TMEM132C* gene would be a novel candidate gene for the pulmonary function.

Three other candidate regions from the discovery stage were tested for replication. However, we could not find any association in the Korean population, and we suspect this may be a result of ethnic differences between the two populations.

There were several limitations of this study. First, we performed SNP specific replication test using imputed genotype data, but discovery associated SNPs were not replicated in the Korean population. These SNPs showed a different allele frequency between the discovery and replication populations, which suggests the possibility of different genetic architecture. For this reason, we conducted a locus specific replication test. Second, FVC was not evaluated in our study. Because we allowed the expiratory maneuver to last at least 3 seconds for a more accurate FEV1 measurement, full FVC efforts lasting at least 6 seconds could not be examined [23]. Third, the sample size of the discovery stage is rather small compared to other genome-wide association studies. Therefore, we used genome-wide suggestive *P* value instead of a Bonferroni adjustment.

However, several unique values in our study design might enable us to detect candidate results. First, we used large extended families in a rural isolated population, which has several advantages in genetic studies. An isolated population is suitable for genetic research because it can minimize the environmental component of the phenotype and has a restricted genetic heterogeneity [42]. Furthermore, extended multi-generation families with a small number of founders are known to increase the genetic power [43]. Second, the replication of the results in two different ethnic groups in Asia may provide solid evidence of association with respect to pulmonary function. As described above, most of the genetic studies of pulmonary function have been focused on Caucasian populations but not on Asian populations. The only exception was very recent report on the association of the 6p21 region with pulmonary function in Korea [30]. The difference between their results and our results might be due to the different study design (general population vs. large extended family) and ethnic group (Korean vs. Mongolian) in the discovery phase. Although Mongolia and Korea are located in northeast Asia, the geographical residences of these two populations are separated by approximately 1,700 km.

Moreover, Mongolian lifestyle and dietary patterns are based on a nomadic culture, and Koreans have had an agricultural lifestyle for thousands of years.

In conclusion, our study focused on discovering a genetic determinant of pulmonary function in two northeast Asian populations. We found novel candidate genes (*UNC93A*, *TTLL2* and *TMEM132C*) in two regions that had been reported to show linkage and/or association with pulmonary function in Caucasians. Our results can be used for further functional studies of pulmonary function and provide more insight into genetic factors of pulmonary function in the general population.

Supporting Information

S1 Table. Family-based association result for FEV_1 in the discovery stage (participants without known respiratory disease only).
(DOCX)

S2 Table. The comparison of the minor allele frequency of 7 associated SNPs in the discovery and replication study.
(DOCX)

Author Contributions

Conceived and designed the experiments: JIK JSS JS KUM HLK. Performed the experiments: HYS SWS SHI BG HSK. Analyzed the data: HYS SWS SHI HJK MKL BG. Contributed reagents/materials/analysis tools: HYS SWS SHI BG HSK JSS JIK. Wrote the paper: HYS SWS SHI DSP JSS JIK.

References

1. Wilk JB, Djousse L, Arnett DK, Rich SS, Province MA, Hunt SC, et al. Evidence for major genes influencing pulmonary function in the NHLBI family heart study. Genet Epidemiol. 2000; 19(1):81–94. Epub 2000/06/22. [pii] doi: 10.1002/1098-2272(200007)19:1<81::AID-GEPI6>3.0.CO;2-8 PMID: 10861898.

2. Strachan DP. Ventilatory function, height, and mortality among lifelong non-smokers. J Epidemiol Community Health. 1992; 46(1):66–70. Epub 1992/02/01. PMID: 1573363; PubMed Central PMCID: PMC1059496.

3. Hole DJ, Watt GC, Davey-Smith G, Hart CL, Gillis CR, Hawthorne VM. Impaired lung function and mortality risk in men and women: findings from the Renfrew and Paisley prospective population study. BMJ. 1996; 313(7059):711–5; discussion 5–6. Epub 1996/09/21. PMID: 8819439; PubMed Central PMCID: PMC2352103.

4. Young RP, Hopkins R, Eaton TE. Forced expiratory volume in one second: not just a lung function test but a marker of premature death from all causes. Eur Respir J. 2007; 30(4):616–22. Epub 2007/10/02. 30/4/616 [pii] doi: 10.1183/09031936.00021707 PMID: 17906084.

5. Pauwels RA, Buist AS, Calverley PM, Jenkins CR, Hurd SS. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease. NHLBI/WHO Global Initiative for Chronic Obstructive Lung Disease (GOLD) Workshop summary. Am J Respir Crit Care Med. 2001; 163 (5):1256–76. Epub 2001/04/24. PMID: 11316667.

6. Silverman EK, Chapman HA, Drazen JM, Weiss ST, Rosner B, Campbell EJ, et al. Genetic epidemiology of severe, early-onset chronic obstructive pulmonary disease. Risk to relatives for airflow obstruction and chronic bronchitis. Am J Respir Crit Care Med. 1998;157(6 Pt 1:):1770–8. Epub 1998/06/25. PMID: 9620904.

7. Palmer LJ, Knuiman MW, Divitini ML, Burton PR, James AL, Bartholomew HC, et al. Familial aggregation and heritability of adult lung function: results from the Busselton Health Study. Eur Respir J. 2001; 17(4):696–702. Epub 2001/06/13. PMID: 11401066.

8. Silverman EK, Palmer LJ, Mosley JD, Barth M, Senter JM, Brown A, et al. Genomewide linkage analysis of quantitative spirometric phenotypes in severe early-onset chronic obstructive pulmonary disease. Am J Hum Genet. 2002; 70(5):1229–39. Epub 2002/03/27. S0002-9297(07)62515-4 [pii] doi: 10.1086/340316 PMID: 11914989; PubMed Central PMCID: PMC447597.

9. Demeo DL, Mariani TJ, Lange C, Srisuma S, Litonjua AA, Celedon JC, et al. The SERPINE2 gene is associated with chronic obstructive pulmonary disease. Am J Hum Genet. 2006; 78(2):253–64. Epub 2005/12/17. S0002-9297(07)62357-X [pii] doi: 10.1086/499828 PMID: 16358219; PubMed Central PMCID: PMC1380249.

10. Joost O, Wilk JB, Cupples LA, Harmon M, Shearman AM, Baldwin CT, et al. Genetic loci influencing lung function: a genome-wide scan in the Framingham Study. Am J Respir Crit Care Med. 2002; 165 (6):795–9. Epub 2002/03/19. PMID: 11897646.

11. Wilk JB, DeStefano AL, Joost O, Myers RH, Cupples LA, Slater K, et al. Linkage and association with pulmonary function measures on chromosome 6q27 in the Framingham Heart Study. Hum Mol Genet. 2003; 12(21):2745–51. doi: 10.1093/hmg/ddg311 PMID: 12966033.

12. Wilk JB, Chen TH, Gottlieb DJ, Walter RE, Nagle MW, Brandler BJ, et al. A genome-wide association study of pulmonary function measures in the Framingham Heart Study. PLoS Genet. 2009; 5(3): e1000429. Epub 2009/03/21. doi: 10.1371/journal.pgen.1000429 PMID: 19300500; PubMed Central PMCID: PMC2652834.

13. Wilk JB, Walter RE, Laramie JM, Gottlieb DJ, O'Connor GT. Framingham Heart Study genome-wide association: results for pulmonary function measures. BMC Med Genet. 2007; 8 Suppl 1:S8. Epub 2007/10/16. 1471-2350-8-S1-S8 [pii] doi: 10.1186/1471-2350-8-S1-S8 PMID: 17903307; PubMed Central PMCID: PMC1995616.

14. Repapi E, Sayers I, Wain LV, Burton PR, Johnson T, Obeidat M, et al. Genome-wide association study identifies five loci associated with lung function. Nat Genet. 2010; 42(1):36–44. Epub 2009/12/17. ng.501 [pii] doi: 10.1038/ng.501 PMID: 20010834; PubMed Central PMCID: PMC2862965.

15. Hancock DB, Eijgelsheim M, Wilk JB, Gharib SA, Loehr LR, Marciante KD, et al. Meta-analyses of genome-wide association studies identify multiple loci associated with pulmonary function. Nat Genet. 2010; 42(1):45–52. Epub 2009/12/17. ng.500 [pii] doi: 10.1038/ng.500 PMID: 20010835; PubMed Central PMCID: PMC2832852.

16. Soler Artigas M, Loth DW, Wain LV, Gharib SA, Obeidat M, Tang W, et al. Genome-wide association and large-scale follow up identifies 16 new loci influencing lung function. Nat Genet. 2011; 43 (11):1082–90. Epub 2011/09/29. ng.941 [pii] doi: 10.1038/ng.941 PMID: 21946350; PubMed Central PMCID: PMC3267376.

17. Soler Artigas M, Wain LV, Tobin MD. Genome-wide association studies in lung disease. Thorax. 2012; 67(3):271–3, 80. doi: 10.1136/thoraxjnl-2011-200724 PMID: 21856699.

18. Paik SH, Kim HJ, Son HY, Lee S, Im SW, Ju YS, et al. Gene mapping study for constitutive skin color in an isolated Mongolian population. Exp Mol Med. 2011; 44(3):241–9. Epub 2011/12/27. emm.2012.44.020 [pii] doi: 10.3858/emm.2012.44.3.020 PMID: 22198297; PubMed Central PMCID: PMC3317488.

19. Lee MK, Cho SI, Kim H, Song YM, Lee K, Kim JI, et al. Epidemiologic characteristics of intraocular pressure in the Korean and Mongolian populations: the Healthy Twin and the GENDISCAN study. Ophthalmology. 2012; 119(3):450–7. Epub 2012/01/17. S0161-6420(11)00872-4 [pii] doi: 10.1016/j.ophtha.2011.09.016 PMID: 22244945.

20. Park H, Kim HJ, Lee S, Yoo YJ, Ju YS, Lee JE, et al. A family-based association study after genome-wide linkage analysis identified two genetic loci for renal function in a Mongolian population. Kidney international. 2013; 83(2):285–92. doi: 10.1038/ki.2012.389 PMID: 23254893.

21. Im SW, Kim HJ, Lee MK, Yi JH, Jargal G, Sung J, et al. Genome-wide linkage analysis for ocular and nasal anthropometric traits in a Mongolian population. Exp Mol Med. 2010; 42(12):799–804. Epub 2010/12/15. emm.2010.42.080 [pii]. PMID: 21150245; PubMed Central PMCID: PMC3015153.

22. Standardization of Spirometry, 1994 Update. American Thoracic Society. Am J Respir Crit Care Med. 1995; 152(3):1107–36. Epub 1995/09/01. PMID: 7663792.

23. Crapo RO, Casaburi R, Coates AL, Enright PL, Hankinson JL, Irvin CG, et al. Guidelines for methacholine and exercise challenge testing–1999. This official statement of the American Thoracic Society was adopted by the ATS Board of Directors, July 1999. Am J Respir Crit Care Med. 2000; 161(1):309–29. doi: 10.1164/ajrccm.161.1.ats11-99 PMID: 10619836.

24. Morris JF, Koski A, Johnson LC. Spirometric standards for healthy nonsmoking adults. The American review of respiratory disease. 1971; 103(1):57–67. PMID: 5540840.

25. Park H, Lee S, Kim HJ, Ju YS, Shin JY, Hong D, et al. Comprehensive genomic analyses associate UGT8 variants with musical ability in a Mongolian population. J Med Genet. 2012; 49(12):747–52. Epub 2012/11/03. jmedgenet-2012-101209 [pii] doi: 10.1136/jmedgenet-2012-101209 PMID: 23118445; PubMed Central PMCID: PMC3512346.

26. Almasy L, Blangero J. Multipoint quantitative-trait linkage analysis in general pedigrees. Am J Hum Genet. 1998; 62(5):1198–211. Epub 1998/05/23. S0002-9297(07)61542-0 [pii] doi: 10.1086/301844 PMID: 9545414; PubMed Central PMCID: PMC1377101.

27. Horvath S, Xu X, Laird NM. The family based association test method: strategies for studying general genotype–-phenotype associations. Eur J Hum Genet. 2001; 9(4):301–6. doi: 10.1038/sj.ejhg.5200625 PMID: 11313775.

28. Sung J, Cho SI, Lee K, Ha M, Choi EY, Choi JS, et al. Healthy Twin: a twin-family study of Korea–-protocols and current status. Twin research and human genetics: the official journal of the International Society for Twin Studies. 2006; 9(6):844–8. PMID: 17254419.

29. Gombojav B, Song YM, Lee K, Yang S, Kho M, Hwang YC, et al. The Healthy Twin Study, Korea updates: resources for omics and genome epidemiology studies. Twin research and human genetics: the official journal of the International Society for Twin Studies. 2013; 16(1):241–5. PMID: 23218411.

30. Kim WJ, Lee MK, Shin C, Cho NH, Lee SD, Oh YM, et al. Genome-wide association studies identify locus on 6p21 influencing lung function in the Korean population. Respirology. 2014; 19(3):360–8. doi: 10.1111/resp.12230 PMID: 24387323.

31. Wilk JB, Herbert A, Shoemaker CM, Gottlieb DJ, Karamohamed S. Secreted modular calcium-binding protein 2 haplotypes are associated with pulmonary function. Am J Respir Crit Care Med. 2007; 175 (6):554–60. Epub 2007/01/06. 200601-110OC [pii] doi: 10.1164/rccm.200601-110OC PMID: 17204727; PubMed Central PMCID: PMC1899283.

32. Zhang D, Cheng L, Badner JA, Chen C, Chen Q, Luo W, et al. Genetic control of individual differences in gene-specific methylation in human brain. Am J Hum Genet. 2010; 86(3):411–9. doi: 10.1016/j.ajhg. 2010.02.005 PMID: 20215007; PubMed Central PMCID: PMC2833385.

33. Levin JZ, Horvitz HR. The Caenorhabditis elegans unc–93 gene encodes a putative transmembrane protein that regulates muscle contraction. The Journal of cell biology. 1992; 117(1):143–55. PMID: 1313436; PubMed Central PMCID: PMC2289394.

34. Arnlov J, Sundstrom J, Lind L, Andren B, Andersson M, Reneland R, et al. hUNC-93B1, a novel gene mainly expressed in the heart, is related to left ventricular diastolic function, heart failure morbidity and mortality in elderly men. Eur J Heart Fail. 2005; 7(6):958–65. doi: 10.1016/j.ejheart.2004.06.009 PMID: 16111919.

35. Tabeta K, Hoebe K, Janssen EM, Du X, Georgel P, Crozat K, et al. The Unc93b1 mutation 3d disrupts exogenous antigen presentation and signaling via Toll-like receptors 3, 7 and 9. Nature immunology. 2006; 7(2):156–64. doi: 10.1038/ni1297 PMID: 16415873.

36. Ikegami K, Sato S, Nakamura K, Ostrowski LE, Setou M. Tubulin polyglutamylation is essential for airway ciliary function through the regulation of beating asymmetry. Proc Natl Acad Sci U S A. 2010; 107 (23):10490–5. doi: 10.1073/pnas.1002128107 PMID: 20498047; PubMed Central PMCID: PMC2890849.

37. Vogel P, Hansen G, Fontenot G, Read R. Tubulin tyrosine ligase-like 1 deficiency results in chronic rhinosinusitis and abnormal development of spermatid flagella in mice. Veterinary pathology. 2010; 47 (4):703–12. doi: 10.1177/0300985810363485 PMID: 20442420.

38. Shao C, Suzuki Y, Kamada F, Kanno K, Tamari M, Hasegawa K, et al. Linkage and association of childhood asthma with the chromosome 12 genes. J Hum Genet. 2004; 49(3):115–22. doi: 10.1007/s10038-003-0118-z PMID: 14767694.

39. Ferreira MA, O'Gorman L, Le Souef P, Burton PR, Toelle BG, Robertson CF, et al. Robust estimation of experimentwise P values applied to a genome scan of multiple asthma traits identifies a new region of significant linkage on chromosome 20q13. Am J Hum Genet. 2005; 77(6):1075–85. doi: 10.1086/497997 PMID: 16380917; PubMed Central PMCID: PMC1285164.

40. Brasch-Andersen C, Tan Q, Borglum AD, Haagerup A, Larsen TR, Vestbo J, et al. Significant linkage to chromosome 12q24.32-q24.33 and identification of SFRS8 as a possible asthma susceptibility gene. Thorax. 2006; 61(10):874–9. doi: 10.1136/thx.2005.055475 PMID: 16738036; PubMed Central PMCID: PMC2104763.

41. Hansel NN, Ruczinski I, Rafaels N, Sin DD, Daley D, Malinina A, et al. Genome-wide study identifies two loci associated with lung function decline in mild to moderate COPD. Hum Genet. 2013; 132(1):79–90. doi: 10.1007/s00439-012-1219-6 PMID: 22986903; PubMed Central PMCID: PMC3536920.

42. Kristiansson K, Naukkarinen J, Peltonen L. Isolated populations and complex disease gene identification. Genome biology. 2008; 9(8):109. doi: 10.1186/gb-2008-9-8-109 PMID: 18771588; PubMed Central PMCID: PMC2575505.

43. Arcos-Burgos M, Muenke M. Genetics of population isolates. Clinical genetics. 2002; 61(4):233–47. PMID: 12030885.