Allergy Asthma Immunol Res. 2021 Jul;13(4):609-622
https://doi.org/10.4168/aair.2021.13.4.609
pISSN 2092-7355·eISSN 2092-7363

Allergy, Asthma & Immunology Research AAIR

Original Article

Check for updates

Genome-Wide Association Study of Korean Asthmatics: A Comparison With UK Asthmatics

Jin An,[1,2†] **Ah Ra Do**,[3†] **Hae Yeon Kang**,[4] **Woo Jin Kim**,[5] **Sanghun Lee**,[6] **Ji-Hyang Lee**,[1] **Woo-Jung Song**,[1] **Hyouk-Soo Kwon**,[1] **You Sook Cho**,[1] **Hee-Bom Moon**,[1] **Sile Hu**,[7] **Ian M Adcock**,[7] **Kian Fan Chung**,[7] **Sungho Won**,[3,8,9*] **Tae-Bum Kim**[1*]

[1]Department of Allergy and Clinical Immunology, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Korea
[2]Department of Pulmonary, Allergy and Critical Care Medicine, Kyung Hee University Hospital at Gangdong, College of Medicine, Kyung Hee University, Seoul, Korea
[3]Interdisciplinary Program of Bioinformatics, Seoul National University, Seoul, Korea
[4]Department of Internal Medicine, Healthcare Research Institute, Seoul National University Hospital Healthcare System Gangnam Center, Seoul, Korea
[5]Department of Internal Medicine and Environmental Health Center, School of Medicine, Kangwon National University, Chuncheon, Korea
[6]Department of Medical Consilience, Graduate School, Dankook university, Yongin, Korea
[7]Experimental Studies, Airways Disease Section, National Heart & Lung Institute, Imperial College London, London, United Kingdom
[8]Department of Public Health Sciences, School of Public Health, Seoul National University, Seoul, Korea
[9]Institute of Health and Environment, Seoul National University, Seoul, Korea

OPEN ACCESS

Received: Dec 7, 2020
Revised: Apr 19, 2021
Accepted: May 12, 2021

Correspondence to
Tae-Bum Kim, MD, PhD
Department of Allergy and Clinical Immunology, Asan Medical Center, University of Ulsan College of Medicine, 88 Olympic-ro 43-gil, Songpa-gu, Seoul 05505, Korea.
Tel: +82-2-3010-3280
Fax: +82-2-3010-6969
E-mail: tbkim@amc.seoul.kr

Sungho Won, MD, PhD
Department of Public Health Science, Seoul National University, 1 Gwanak-ro, Gwanak-gu, Seoul 08826, Korea.
Tel: +82-2-880-2714
Fax: +82-303-0942-2714
E-mail: sunghow@gmail.com

†Jin An and Ah Ra Do contributed equally to this article.

Copyright © 2021 The Korean Academy of Asthma, Allergy and Clinical Immunology · The Korean Academy of Pediatric Allergy and Respiratory Disease
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (https://creativecommons.org/licenses/by-nc/4.0/) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

ORCID iDs
Jin An
https://orcid.org/0000-0001-5416-2660

ABSTRACT

Purpose: Although genome-wide association studies (GWASs) represent the most powerful approach for identifying genes that influence asthma, to date, no studies have established genetic susceptibility to asthma in the Korean population. This study aimed to identify genetic variants associated with adult Korean asthmatics and compare them with the significant single nucleotide polymorphisms (SNPs) of UK asthmatics from the UK Biobank.
Methods: Patients were defined as having asthma if they were diagnosed by a doctor or taking medications for asthma. Controls were defined as individuals without asthma or chronic obstructive pulmonary disease. We performed quality control, genotype imputation, GWAS, and PrediXcan analyses. In the GWAS, a P value of $< 5 \times 10^{-8}$ was considered significant. We compared significant SNPs between Korean and UK patients with asthma.
Results: A total of 1,386 asthmatic patients and 5,205 controls were analyzed. The SNP rs1770, located near the human leukocyte antigen (HLA)-DQB1, was the most significant SNP ($P = 4.5 \times 10^{-10}$). In comparison with 24 SNPs in a GWAS of UK asthmatics, six SNPs were significant with the same odds ratio (OR) direction, including signals related to type 2 inflammation (*e.g.*, IL1RL1, TSLP, and GATA3) and mucus plugging (*e.g.*, MUC5AC). HLA-DQA1 showed an opposite OR direction. The HLA-DQB1 gene demonstrated significantly imputed mRNA expression in the lung tissue and whole blood.
Conclusions: The SNP rs1770 of HLA-DQB1 was the most significant in Korean asthmatics. Similarities and discrepancies were found in the genetic variants between Korean and UK asthmatics. GWAS of Korean asthmatics should be replicated and compared with those of GWAS of other ethnicities.

Generated by xmlinkpress

Ah Ra Do
https://orcid.org/0000-0002-4259-8309
Hae Yeon Kang
https://orcid.org/0000-0003-3965-7539
Woo Jin Kim
https://orcid.org/0000-0003-2927-370X
Sanghun Lee
https://orcid.org/0000-0002-0573-9555
Ji-Hyang Lee
https://orcid.org/0000-0003-4286-3114
Woo-Jung Song
https://orcid.org/0000-0002-4630-9922
Hyouk-Soo Kwon
https://orcid.org/0000-0001-7695-997X
You Sook Cho
https://orcid.org/0000-0001-8767-2667
Hee-Bom Moon
https://orcid.org/0000-0001-7542-5917
Sile Hu
https://orcid.org/0000-0002-8437-7387
Ian M Adcock
https://orcid.org/0000-0003-2101-8843
Kian Fan Chung
https://orcid.org/0000-0001-7101-1426
Sungho Won
https://orcid.org/0000-0001-5751-5089
Tae-Bum Kim
https://orcid.org/0000-0001-5663-0640

Disclosure
There are no financial or other issues that might lead to conflict of interest.

Keywords: Genome-wide association study; asthma; genetics; Korea; United Kingdom; human leukocyte antigen

INTRODUCTION

Asthma is a common chronic inflammatory disease of the airways resulting from the interaction of genetic susceptibility with environmental exposure across an individual's lifetime.[1] Both genetic and environmental factors contribute to disease risk, with genetic factors accounting for 35%–95% of susceptibility to asthma.[2] Genetic risk factors may be useful in classifying the endotypes of asthma, and genetic studies allow a structured means of understanding the causes of asthma as well as verifying targets that can be used to treat the disease.[3] Some studies on the genetic associations with asthma have been conducted, and genome-wide association studies (GWASs) represent the most powerful approach for identifying genes that influence asthma.[4]

The results of GWASs have highlighted the importance of genetic variants in or near genes implicated in asthma. Previous studies have reported 41 regions associated with asthma, providing novel insights into disease biology.[3,5,6] However, few GWASs have been performed for populations of racial and ethnic minorities. Several studies have reported significant differences in the allele frequencies of the variants and genetic architecture between European and Asian populations.[7] These suggest that genetic associations with asthma do not consistently replicate across studies because of the heterogeneity of populations. Therefore, GWASs using different ethnic populations are needed to better understand the genetic basis of asthma.

Although many recent GWASs have reported the association of multiple loci with asthma risk, the effect of genetic variation on asthma susceptibility in the Korean population has not been elucidated. For the Korean population, some studies have performed simple replication analysis and reported the results of a GWAS for aspirin-exacerbated respiratory disease (AERD).[8,9] However, no studies have yet confirmed the genetic susceptibility of asthma using a GWAS in the Korean population. Therefore, the aims of this study were to 1) identify genetic variants associated with adult asthma in the Korean population, and 2) evaluate whether significant single nucleotide polymorphisms (SNPs) in the UK Biobank are specifically correlated with those in Korean asthmatics.

MATERIALS AND METHODS

Study subjects

Subjects were obtained from the Asan Medical Center (AMC), Seoul National University Gangnam Center of Seoul National University Hospital (SNUH), and Kang-Won National University Hospital (KNUH). Of the 1,386 patients with asthma, 1,293 were diagnosed at Asan Medical Center between January 2005 and December 2018. The inclusion criteria were 1) age >18 years; 2) symptoms, such as dyspnea, wheezing, or cough, for more than three months; and 3) airway hyperresponsiveness, as indicated by a 20% reduction in forced expiratory volume in 1 second (FEV1) with a methacholine dose of 16 mg/mL (PC20) through a provocation test or airway reversibility, as indicated by an increase in FEV1 of > 12% (and at least 200 mL) following the inhalation of a short-acting β-agonist. Patients with severe

lung damage, bronchiectasis, or a history of lung resection were excluded. From the Seoul National University Gangnam Center of Seoul National University Hospital, 93 asthmatic patients provided a "yes" response to either of the following questions: "Have you ever been diagnosed with asthma?" or "Have you ever taken asthma medication at least once?" were included in this study. Of the 5,205 controls, 5,161 and 44 controls were obtained from the Seoul National University Gangnam Center of Seoul National University Hospital and Kang-Won National University Hospital, respectively, and were not diagnosed with asthma or chronic obstructive pulmonary disease. This study was approved by the Institutional Review Boards of AMC (2019-1676), SNUH (H-2005-209-1127), and KNUH (2015-08-010-021). Written informed consent was obtained from all patients.

For the replication test, the UK Biobank data were used. Only the data of British, and Irish people and those with any other white background were used if the information on height and 2 or more lung function test results were available and if they passed American Thoracic Society quality control (QC). The asthma criteria for UK Biobank include patients who have (1) a record for the 'age asthma diagnosed' item or (2) '1111' for 'self-reported non-cancer illness code'. A total of 13,475 patients with asthma and 109,871 controls were included in the replication analysis.

Genotyping and QCs

All subjects were genotyped with the Korea Biobank Array (referred to as KoreanChip).[10] KoreanChip is an Affymetrix customized SNP chip optimized for the Korean population and can demonstrate findings from the GWAS of blood biochemical traits. Variants were identified by following the Best Practices Steps provided by Affymetrix.[11] To minimize the batch effect, we applied the K-medoid algorithm.[12]

QCs of subjects and SNPs were performed with PLINK[13] and ONETOOL.[14] First, we conducted QCs for each batch. Any SNPs were excluded if any of the following conditions were satisfied: genotype call rates of less than 95% or Hardy-Weinberg equilibrium (HWE) tests[15] showing $P < 1 \times 10^{-5}$. Also, subjects were discarded if the X chromosome homozygosity was between 0.2 and 0.8, genotype call rates were < 95% or heterozygosity rates of SNPs were outside the average heterozygosity rate ± 3 standard deviations. Once QCs were performed for each batch, all batches were pooled, and QC was performed for the pooled samples. Furthermore, for each SNP, we compared the missing rates, HWE, and minor allele frequency (MAF) among batches using Pearson's χ^2 test, and any significant SNPs at the 1×10^{-5} significance level were excluded. After completing all QC processes, 8,903 subjects and 649,085 SNPs remained. The detailed procedure is presented in **Supplementary Fig. S1**.

Genotype imputation of SNP genotype data

Genotype imputation was performed using the Michigan imputation server (https://imputationserver.sph.umich.edu). Haplotype Reference Consortium release v1.1 was used and only 'non-European' or 'mixed' populations were considered.[16] Pre-phasing and imputation were performed using Eagle V 2.4[17] and Minimac4,[18] respectively. After the imputation processes, imputed SNPs were removed if the Rsq was less than 0.5, or if there were duplicated SNPs, missing genotype rates larger than 0.05, *P* values in HWE tests of less than 1×10^{-5}, or MAFs less than 0.05. Also, subjects whose identity by state was > 0.9, and whose principal component (PC) scores were outside the 5 × IQRPC were also excluded. Finally, 1,386 asthmatic patients, 5,205 controls, and 4,922,160 SNPs were used for the analysis. The detailed procedure is shown in **Supplementary Fig. S1**.

GWAS

The GWASs were conducted using logistic regression (LR). The PC scores were estimated using PLINK, and were used to adjust the population substructure. A total of 20 PC scores corresponding to the 20 largest eigenvalues, age, and sex were included as covariates. The genome-wide significance level was set at 5×10^{-8}. Genome-wide significant SNPs were annotated using ANNOVAR. LDlink[19] was used to evaluate linkage disequilibrium (LD) across the ancestral population with Phase 3 haplotype data from the 1000 Genomes Project. For the replication analysis, gene-set analysis was performed using MAGMA, which uses the whole SNP association statistics with gene size and LD as covariates, on the UK Biobank dataset.[20]

RNA expression imputation with PrediXcan and the associations with asthma

We imputed the RNA expression levels of genes on chromosome 6 for whole blood and lung tissue using PrediXcan.[21] This allows subject-level gene expression prediction for 48 human tissues with Genotype-Tissue Expression (GTEx) V7 data. The organized association data from lung tissue used in our analysis were obtained from 515 people from the lung inferior segment of the left upper lobe, 1 cm below the surface. Whole blood was obtained from a sample of 670 people, and the femoral vein, subclavian vein, and heart were other possible sites.[22] Subsequently, linear regression was used to determine whether the predicted RNA expression was associated with asthma status. For linear regression, the top 20 PC scores, age, and sex were included as covariates. For the association test, the false discovery rate and Bonferroni-corrected *P* values were calculated for all genes on chromosome 6.

Estimation of genetic heritability

We estimated SNP heritability with GCTA64,[23] which indicates the relative proportion of phenotypic variance explained according to SNPs. SNPs were pruned using PLINK with the default option, and the remaining SNPs were used to calculate the genetic relationship matrix (GRM). This GRM was used to estimate SNP heritability using the GCTA64. For ascertainment bias, the prevalence of asthma was set to 3.9%.[24]

RESULTS

Subject characteristics

A total of 1,386 asthmatic patients and 5,205 controls were used for genetic association analyses. Demographic information of the study population is presented in **Table 1**. The

Table 1. Demographic characteristics of the study population

Characteristics	Case (n = 1,386)	Control (n = 5,205)	Total (n = 6,591)	*P* value
Age (yr)	50.3 ± 15.4	49.8 ± 10.2	49.9 ± 11.5	0.3253
Female sex	766 (55.2)	2,295 (44.1)	3,061 (46.4)	< 0.001*
BMI (kg/m^2)	24.3 ± 3.6	23.1 ± 3.1	23.3 ± 3.2	< 0.001*
Smoking				0.2144
Never	761 (54.9)	1,822 (35.0)	2,583 (39.2)	
Former	384 (27.7)	982 (18.9)	1,366 (20.7)	
Current	199 (14.4)	558 (10.7)	757 (11.5)	
Unknown	42 (3.0)	1,843 (35.4)	1,885 (28.6)	
FEV1 pred pre-BD (%)	78.8 ± 20.2	101.8 ± 12.2	97.3 ± 16.8	< 0.001*
FEV1/FVC (%)	70.7 ± 0.1	82.0 ± 0.1	79.4 ± 0.1	< 0.001*

The values are shown as mean ± standard deviation or number (%) in each cell. Case group means asthmatic subjects.
BMI, body mass index; FEV1, forced expiratory volume in 1 second; BD, bronchodilator; FVC, forced vital capacity.
*$P < 0.0001$.

average ages of asthmatic patients and controls were 50.3 and 49.8 years, respectively. Although age and smoking history did not differ between the 2 groups, the proportion of females was higher in the asthmatic group than in the control group (55.2% vs. 44.1%). In addition, the body mass index was higher in the asthmatic group than in the control group (24.3 ± 3.6 vs. 23.1 ± 3.1). In comparison with the control group, the asthmatic group demonstrated lower baseline FEV1 levels (78.8 ± 20.2 vs. 101.8 ± 12.2) and FEV1/forced vital capacity (70.7 ± 0.1 vs. 82.0 ± 0.1).

Identification of SNPs on chromosome 6 by the GWAS of asthma

The GWAS was performed with LR using 1,386 asthmatic patients and 5,205 controls with 4,935,875 SNPs. **Supplementary Fig. S2** shows a quantile-quantile plot for LR, which demonstrates that the distribution of the observed *P* values was consistent with the expected distribution, except for several significant SNPs with a genomic inflation factor of 1.035. The multidimensional scaling (MDS) plots showed no evidence of population stratification (**Supplementary Fig. S3**), and the results were statistically significant. Manhattan plots of the genome demonstrated that only four SNPs had genome-wide significance (**Fig. 1**).

Table 2 shows the results for the 10 most significant SNPs with one rs1770 correlated SNP (rs9273508, $r^2 = 0.9073$). The MAF of the top 10 SNPs was compared with the MAF of the SNPs in the Korean reference dataset (Kref) and a MAF from the GnomAD.[25] If there were several genome-wide significant SNPs in the same LD block, the result for the most significant SNP was included. As shown in **Table 2**, rs1770 was the most significant SNP ($P = 4.5 \times 10^{-10}$). The genomic region of rs1770 on chromosome 6 is shown in **Fig. 2**, demonstrating that rs1770 is located near human leukocyte antigen (HLA)-DQB1. However, the MAF of rs1770 differs from that of Kref and gnomAD. Therefore, rs9273508, which also met the genome-wide significance level ($P = 1.6 \times 10^{-8}$) and is located near HLA-DQB1, was added to **Table 2** despite its correlation with rs1770. The second most significant SNP was rs9271469, which was also a genome-wide significant SNP ($P = 3.74 \times 10^{-8}$) located in the intergenic region of HLA-DRB1 (dist = 31017) and HLA-DQA1 (dist = 16553). We confirmed that the HLA-DQB1 and HLA-DQA1 genes are significant for asthma in the UK Biobank dataset from gene-based analysis (**Supplementary Table S1**).

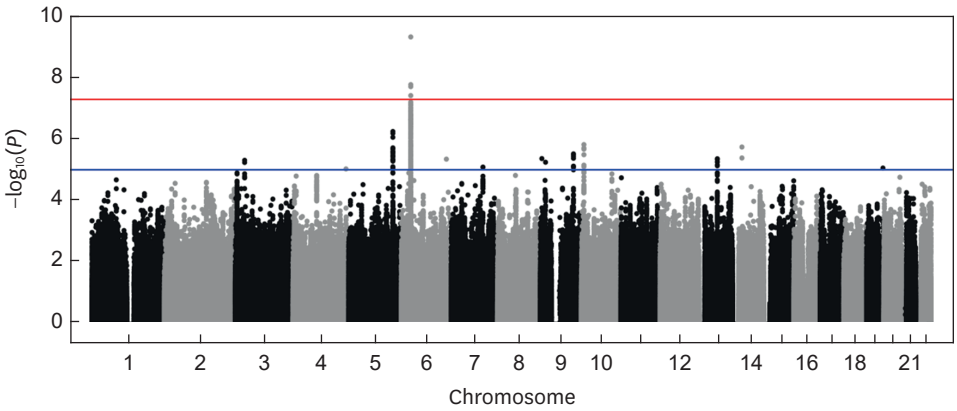


Fig. 1. Manhattan plot of *P* values from the genome-wide association study. One region on chromosome 6 met genome-wide significance ($P < 5 \times 10^{-8}$) according to logistic regression.

https://doi.org/10.4168/aair.2021.13.4.609

Table 2. Genome-wide single nucleotide polymorphisms

Chr	SNP	Nt	Minor/ major alleles	MAF					Geno type	Rsq	GENO (GG/GA/ AA/NA)	OR	SE	L95	U95	STAT	*P* value	HWE	Gene
				All	Asthma	Control	KRef	GnomAD											
6	rs1770	32627833	G/A	0.18	0.22	0.17	0.31	0.37	Imputed	0.60	203/1952/4436/0	1.41	0.06	1.27	1.58	6.24	4.5×10^{-10}	0.66	HLA-DQB1
6	rs9273508	32628439	A/G	0.32	0.36	0.31	0.38	0.38	Imputed	0.71	692/2856/3043/0	1.32	0.05	1.20	1.45	5.65	1.6×10^{-8}	0.60	HLA-DQB1
6	rs9271469	32588630	C/T	0.29	0.33	0.28	0.38	0.26	Imputed	0.72	523/2764/3304/0	1.33	0.05	1.20	1.48	5.50	3.7×10^{-8}	0.19	HLA-DRB1 (dist = 31017), HLA-DQA1 (dist = 16553)
5	rs2961757	151936718	T/G	0.24	0.28	0.23	0.24	0.61	Imputed	0.91	405/2361/3825/0	1.28	0.05	1.16	1.40	5.00	5.6×10^{-7}	0.33	NMUR2 (dist = 151878), LINC01470 (dist = 61807)
6	rs9272226	32602396	C/T	0.44	0.48	0.43	0.44	-	Imputed	0.91	1278/3299/2014/0	1.27	0.05	1.16	1.40	4.99	6.2×10^{-7}	0.08	HLA-DRB1 (dist = 44783), HLA-DQA1 (dist = 2787)
10	rs1444789	9064361	C/T	0.24	0.27	0.23	0.22	0.24	Imputed	0.94	352/2406/3833/0	1.27	0.05	1.15	1.40	4.81	1.5×10^{-6}	0.84	LOC105755953 (dist = 107802), LOC101928272 (dist = 174968)
14	rs142914081	31818569	A/G	0.06	0.07	0.05	0.06	0.003	Imputed	0.85	21/694/5876/0	1.51	0.09	1.27	1.78	4.77	1.8×10^{-6}	0.89	HEATR5A
6	rs146351421	32568678	A/C	0.14	0.12	0.15	0.14	-	Imputed	0.67	124/1598/4869/0	0.68	0.08	0.58	0.80	−4.69	2.8×10^{-6}	0.62	HLA-DRB1 (dist = 11065), HLA-DQA1 (dist = 36505)
6	rs3129879	32408907	A/G	0.33	0.30	0.34	0.34	0.28	Imputed	1.00	728/2934/2929/0	0.78	0.05	0.71	0.87	−4.67	3.0×10^{-6}	0.54	HLA-DRA
9	rs57669363	114374274	A/G	0.40	0.44	0.39	0.41	0.41	Imputed	0.97	1005/3237/2349/0	1.23	0.04	1.13	1.34	4.67	3.0×10^{-6}	0.05	LRRC37A5P
9	rs10758715	5882401	C/T	0.24	0.21	0.25	0.24	0.42	Imputed	0.93	396/2365/3830/0	0.79	0.05	0.71	0.87	−4.60	4.3×10^{-6}	0.21	ERMP1 (dist = 49320), MLANA (dist = 8508)

Logistic regression was performed with imputed SNP genotype data to identify SNPs with a significant association ($P < 5 \times 10^{-8}$) with asthma. If many SNPs in the same LD block were significant, the result for the most significant SNP was included. The *P* values and genotype information of the top 10 SNPs with one rs1770 correlated SNP (rs9273508, $r^2 = 0.9073$) are shown.
Chr, chromosome; SNP, single nucleotide polymorphism; Nt, nucleotide; MAF, minor allele frequency; KRef, Korean reference data; OR, odds ratio; SE, standard error; L95, lower bound of 95% confidence interval; U95, upper bound of 95% confidence interval; STAT, logistic regression statistic; HWE, Hardy-Weinberg equilibrium; LD, linkage disequilibrium.

https://doi.org/10.4168/aair.2021.13.4.609

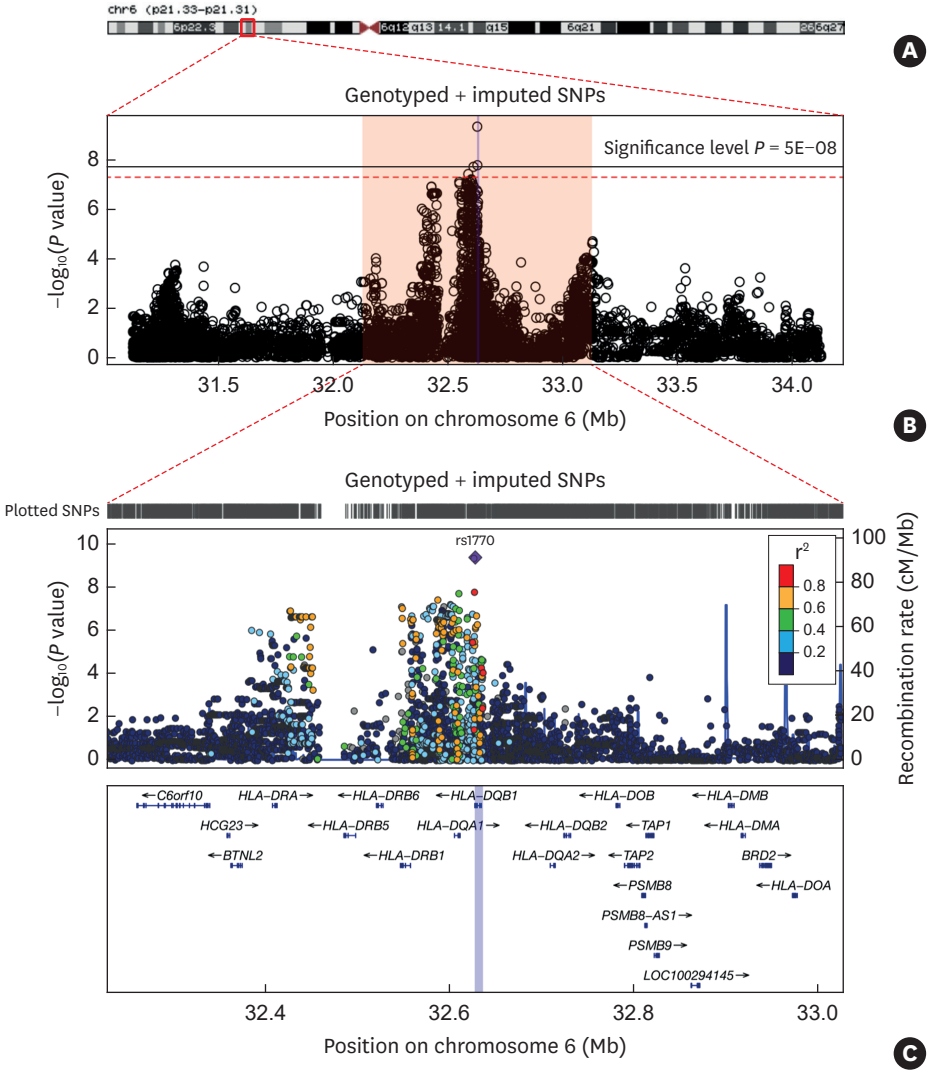


Fig. 2. The genomic region of significant SNPs on chromosome 6. (A) Ideogram of chromosome 6. (B) The 3 Mb region containing the SNPs associated with asthma. The Manhattan plot at the top shows *P* values for both genotyped and imputed SNPs in this region. (C) The expanded Manhattan plot of the 600 kb region shows both genotyped and imputed SNPs. rs1770, which was the most significant SNP, is indicated in purple, and other SNPs are colored according to their R^2 value in relation to that of rs1770. Several protein-coding genes belong to HLA class II beta chain paralogs in this range.
SNP, single nucleotide polymorphism; HLA, human leukocyte antigen.

Comparison of SNPs between Korean, UK, and Japanese asthmatics

We compared our results to those of previous studies conducted in European and Japanese populations (**Fig. 3** and **Supplementary Fig. S4**). **Fig. 3** and **Supplementary Fig. S4** show the Forest plot for 24 SNPs and five SNPs that were significant in GWASs of UK asthmatics using UK Biobank data[6] and Japanese asthmatics using the BioBank Japan, respectively.[26] When the same SNP was not available for comparison in our data, the highest correlated SNP was used. In the case of significant SNPs in our data, an asterisk was placed on the right side of the error bar. In **Fig. 3**, of the 24 SNPs, eight were significant, and six of these, namely, interleukin-1 receptor-like 1 (IL1RL1), thymic stromal lymphopoietin (TSLP), C5orf56, GATA3, MUC5AC, and CLEC16A, had the same odds ratio (OR) direction as significant SNPs in the UK Biobank data. HLA-DQA1 and C11orf30 showed the opposite ORs. In **Supplementary Fig. S4**, 3 of the 5 SNPs were significant, and one of

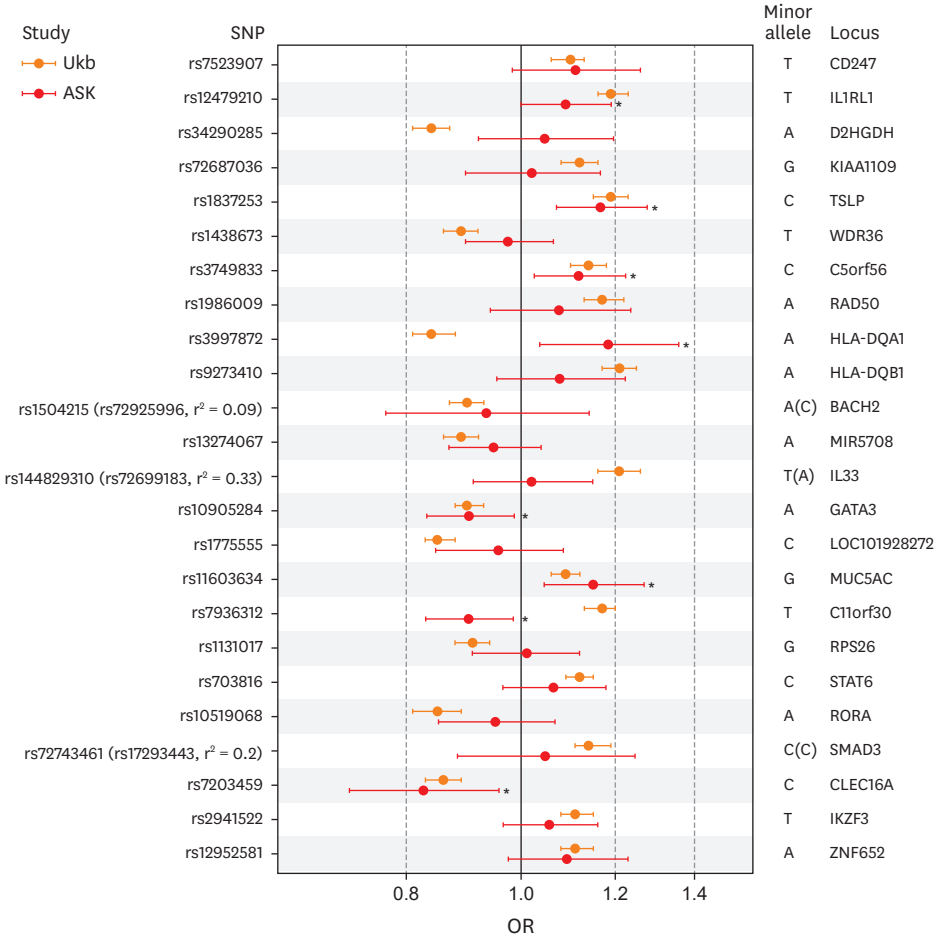


Fig. 3. Forest plot for genome-wide significant SNPs associated with asthma. ASK indicates the GWAS results using our Korean data. Ukb indicates the GWAS results using the UK Biobank data.[6]
SNP, single nucleotide polymorphism; GWAS, genome-wide association study.
*SNPs with $P < 0.05$.

these, solute carrier family 25 member 46 (SCL25A46), TSLP, LOC105755953, and LOC101928272, had the same OR direction as significant SNPs in BioBank Japan.

PrediXcan analysis of asthma on chromosome 6

We identified differentially expressed genes on chromosome 6 with imputed mRNA expression levels in the lung tissue and whole blood. At the false-discovery rate-adjusted 0.05 significance level, all significant genes are shown in **Table 3**. Significant results were observed for various HLA genes. The most significant result was found for HLA-DQA1 in the lung tissue (8.11×10^{-8}). In addition, HLA-DQB1 was significantly expressed in the lung tissue (6.12×10^{-4}) and whole blood (2.89×10^{-4}).

Estimates heritability according to the prevalence of asthma

We determined the proportion of relative phenotypic variance explained according to the pruned SNPs, h^2_{SNP}. We calculated h^2_{SNP} using 2 different scales: the liability scale and the observed 0-1 scale (**Supplementary Fig. S5**). On the observed 0–1 scale, h^2_{SNP} varied from 6% to 20% with a disease prevalence ranging from 3% to 10%. Considering a prevalence of 3.9%,[24] h^2_{SNP} was 8% (42.6% on the liability scale).

Table 3. Results of the differentially expressed gene analysis in chromosome 6 using the predicted amount with PredXcan

Tissue	Gene	Chr	Start	End	Estimate	*P* value	FDR	Bonferroni
Lung tissue	HLA-DQA1	6	32595956	32614839	−0.0678	8.11×10^{-8}	3.87×10^{-5}§	3.9×10^{-5}§
	HLA-DQA2	6	32709119	32709119	0.0972	7.61×10^{-6}	0.0018†	0.0036‡
	HLA-DRB6	6	32520490	32527799	0.0759	1.16×10^{-5}	0.0019†	0.0056‡
	GPANK1	6	31629006	31634060	−0.0115	1.85×10^{-5}	0.0022†	0.0089‡
	HLA-DQB2	6	32723875	32731330	0.0691	7.50×10^{-5}	0.0072†	0.0359†
	HLA-DRB1	6	32546546	32557625	−0.0347	4.86×10^{-4}	0.0365*	0.2327
	GAPDHP42	6	70455692	70456947	−0.0131	5.78×10^{-4}	0.0365*	0.2769
	HLA-DQB1	6	32627244	32636160	−0.0445	6.12×10^{-4}	0.0365*	0.2933
Whole blood	HLA-DQA1	6	32595956	32614839	−0.0436	3.45×10^{-7}	1.33×10^{-4}‡	0.0001‡
	PHF1	6	33378176	33384230	0.0114	7.22×10^{-7}	1.39×10^{-4}‡	0.0003‡
	HLA-DQA2	6	32709119	32709119	0.0863	1.56×10^{-6}	2.01×10^{-4}‡	0.0006‡
	HLA-DQB2	6	32723875	32731330	0.0580	5.01×10^{-5}	0.0048†	0.0194†
	HLA-DQB1	6	32627244	32636160	−0.0391	2.89×10^{-4}	0.0223*	0.1122
	HLA-DRB6	6	32520490	32527799	0.0450	3.94×10^{-4}	0.0254*	0.1530
	SYNJ2	6	158402888	159000000	0.0254	4.98×10^{-4}	0.0275*	0.1932
	HLA-DPB2	6	33080228	33102442	−0.0462	1.00×10^{-3}	0.0486*	0.3907

Chr, chromosome; FDR, false discovery rate.
*$P < 0.1$, †$P < 0.05$, ‡$P < 0.01$, §$P < 0.0001$.

DISCUSSION

For the first time, we identified a key locus that confers susceptibility to adult asthma in the Korean population. The SNP rs1770 on chromosome 6, which is located near HLA-DQB1, demonstrated a strong association with genome-wide significance. Regarding the predicted mRNA expression levels in lung tissue and whole blood, although predicted gene expression by PrediXcan has limitations for the Asian population because the reference used for Predixcan is non-Hispanic whites, we found that HLA genes were significantly associated with mRNA expression. Additionally, we compared our GWAS results with significant SNPs and genes associated with asthma identified by GWAS using the UK Biobank data. A total of 8 SNPs were significant with 6 SNPs(IL1RL1, TSLP, GATA3, MUC5AC, CLEC16A, and C5orf56)-having the same OR direction. HLA-DQA1 and C11orf30 showed opposing directions for their ORs.

GWASs of asthma have been conducted mainly in Europe and the USA, and a large-scale, consortium-based GWAS among European populations reported various loci associated with asthma.[3,27,28] In Asia, a study identified five loci that are susceptible to 1,532 Japanese asthmatics, and TSLP-WDR36 loci were found to be susceptible to asthma.[26] Next-generation sequencing studies revealed differences in allele frequencies and haplotype structures at the 17q12-21 asthma-related locus between Chinese and other populations.[29] Our GWAS of 1,386 Korean asthmatics found that rs1770 (HLA-DQB1) and rs9271469 (HLA-DRB1 and HLA-DQA1 loci) are susceptible to asthma at genome-wide significance levels. These results represent genetic heterogeneity in asthma, even within the Asian population. Thus, it is important to study populations from various ethnic backgrounds to identify shared and unique genetic predictors of asthma.

A GWAS in Korea focused on patients with AERD.[9,30-32] A study of 117 Korean asthmatics with AERD reported that HLA-DPB1 rs1042151 showed the most significant association with susceptibility to AERD.[9] A study of 80 patients with AERD identified CEP 68 as a positive risk factor for the development of aspirin intolerance in asthmatics.[33] Other GWAS verified genetic polymorphisms of CTNNA3 and CTNNA1, which could confer susceptibility to toluene diisocyanate (TDI)-induced asthma in Korea.[34] These results in the Korean

population have potential limitations due to the small sample size and low prevalence of AERD and TDI-induced asthma. Limited comprehensive GWASs have been conducted for adult asthma in the Korean population. Here, we present, to our knowledge, the largest Korean asthma GWAS available, examining 1,386 Korean asthmatics.

In the present study, the GWAS demonstrated that rs1770, which is located near HLA-DQB1, had the most significant genome-wide association ($P = 4.5 \times 10^{-10}$; OR = 1.41). In addition, the expression of HLA genes on chromosome 6 with imputed mRNA expression levels in lung tissue and whole blood were significant. These results indicate that there may be substantial genetic variants within the HLA-DQ region that can influence asthma susceptibility. HLA genes are highly polymorphic and are expressed on the surface of antigen-presenting cells, where they present peptides of exogenous origin to T-cell receptors on $CD4^+$ T cells. There are strong associations between specific HLA class II alleles, including HLA-DR, HLA-DQ, and HLA-DP alleles, and allergen-specific immunoglobulin E (IgE) responses.[35,36] HLA-DQB1 may be important in the recognition of allergen-derived peptides and plays a role in allergic sensitization and the development of asthma. These results further suggest that HLA-DQ typing could be a predictive gene marker in the Korean population as well as other populations for asthma and can be used to instigate primary prevention against asthma progression.

Studies of HLA genes in the susceptibility to asthma and related phenotypes have involved various countries and populations, *e.g.*, the USA, Chinese, UK, and Caucasian populations.[37] Recently, HLA-DQB1 region was reported to be associated with nonatopic asthma and obesity in the result of GWAS using UK Biobank data.[38] HLA-DR/DQ region is closely associated with both IL-4 and tumor necrosis factor-alpha, and HLA is associated with type 2 inflammatory responses.[39] Positive associations between HLA class II alleles and specific IgE responsiveness are consistent with the concept of strong $CD4^+$ Th2 responses of processed allergenic antigens.[35,40,41] Although the exact role of particular HLA-DQB1 remains uncertain, it might have a function to induce type2 inflammatory responses as a gene of HLA class II.

We compared our GWAS results with those from the UK Biobank and BioBank Japan data. Interestingly, our results demonstrated that the genetic signals of type 2 inflammation tended to be similar with the same OR direction to those of the UK Biobank and included IL1RL1 (a receptor for key cytokine IL-33), TSLPleading to type 2 inflammation. and GATA3 (an important transcription factor in Th2 cells).[6] The MUC5AC and CLEC16A genes also had the same OR direction. CLEC16A is a regulator of autophagy involved in airway cell homeostasis.[42] MUC5AC has been linked to mucus plugging and airway hyperresponsiveness and is a major cause of morbidity in asthma.[43] In particular, TSLP also had the same OR direction as that of BioBank Japan. Tezepelumab, a biological target of TSLP, was developed, which may result in a significant reduction in annual asthma exacerbation rates in Korean, European, and Japanese populations. Meanwhile, HLA-DQA1 and C11orf30 showed opposite OR directions compared to those of the UK Biobank. We speculated that there may be discrepancies in the genetic architecture between the European and Korean populations. Further, GWASs of multi-ethnic populations are needed to provide valuable information on genetic diversity according to racial and ethnic differences.

Our study had several limitations. First, although controls were selected to exclude subjects with asthma and other respiratory diseases using a standardized questionnaire, the choice of controls was potentially vulnerable to misclassification bias due to the inclusion of undiagnosed cases. Secondly, this study did not conduct a replication analysis; however,

we showed significant similarities in the GWAS results between Korean and UK asthmatics. Thirdly, the function of the SNP rs1770 of HLA-DQB1 was not analyzed in the present study. Functional studies of SNP rs1770 in the future could provide important insights into the genetic etiology of asthma. Finally, asthma is a heterogeneous disease with gene-environment interactions, and the contribution of environmental exposure in the development of asthma has not yet been fully assessed. In future studies, we aim to investigate gene-environment associations by performing epigenetic analysis of blood DNA methylation status in Korean asthmatics and link this to exposure to environmental pollution.

In conclusion, this is the first GWAS in Korean population. We demonstrated that the SNP rs1770 of HLA-DQB1 was the most significant SNP in adult asthmatic patients. In the comparison of genetic susceptibility between Korean and UK asthmatics from the UK Biobank, genetic variants associated with type 2 inflammation (IL1RL1, TSLP, and GATA3) and mucus plugging in the airway (MUC5AC) were significant with a similar OR direction. A discrepancy was observed for HLA-DQA1 in the opposite OR direction. Our GWAS results for Korean asthmatics should be replicated and compared with those of GWASs of other ethnicities. Moreover, further GWASs should focus on asthma phenotypes and clinical usefulness, including asthma treatment strategies beyond genetic profiling.

ACKNOWLEDGMENTS

This research was supported by the Bio & Medical Technology Development Program of the National Research Foundation (NRF) funded by the Korean government (MSIT) (2019M3E5D3073365). The Korean reference data (Kref) used in this study were provided by the CODA (http://coda.nih.go.kr), CODA accession numbers R002754, R002755, R002774, R002814, and R002854.

SUPPLEMENTARY MATERIALS

Supplementary Table S1

Gene-set analysis of the UK Biobank dataset

Click here to view

Supplementary Fig. S1

Workflow of QC for our samples. Multiple standard QCs were performed for both asthmatic subjects (Case) and control subjects (Control) to exclude outlier SNPs and subjects.

Click here to view

Supplementary Fig. S2

Quantile-quantile plot for the GWAS of asthma. The observed distributions of *P* values in the GWAS are plotted relative to the expected (null) distribution (red line) for the conditional logistic regression analysis. The grey shading indicates the 95% confidence interval for observed *P* values relative to those expected.

Click here to view

Supplementary Fig. S3

MDS plot. Gray dots represent the 1,000 Genomes Project subjects. Red and blue dots represent the subjects with asthma and control subjects, respectively. The fact that red and blue dots are not divided means that there is no population stratification between asthmatics and control subjects.

Click here to view

Supplementary Fig. S4

Forest plot for genome-wide significant SNPs associated with asthma. ASK indicates the GWAS results using our Korean data. Japanese indicates the GWAS results using the BioBank Japan data.

Click here to view

Supplementary Fig. S5

The proportion of phenotypic variances explained according to genotyped SNPs with a disease prevalence ranging from 0.03 to 0.1. The proportion of phenotypic variances explained according to genotyped SNPs was calculated with GCTA on (A) the liability scale and (B) the observed 0–1 scale. The shaded area indicates the 95% confidence interval for h^2_{SNP}.

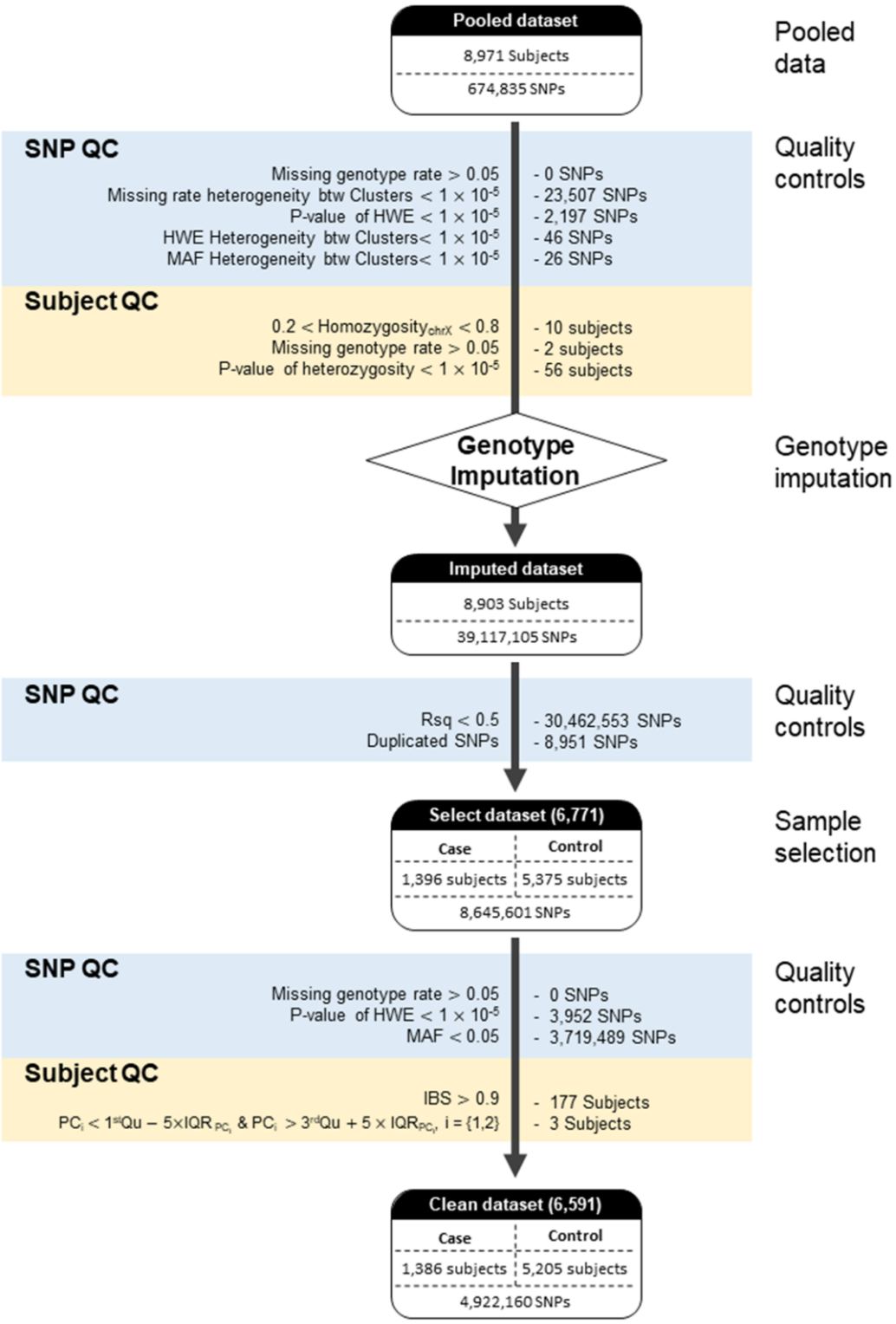
Click here to view

REFERENCES

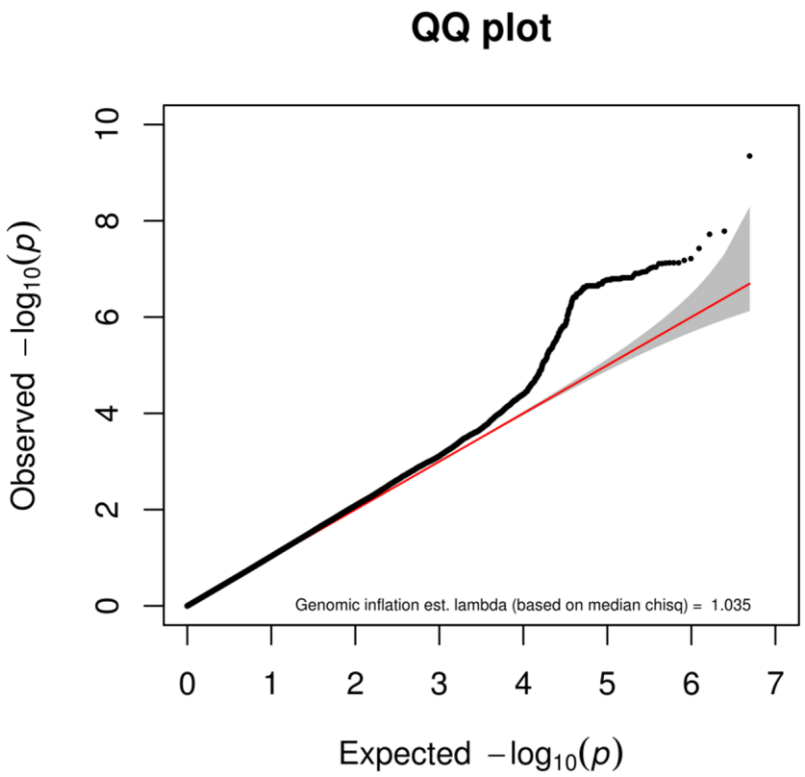
1. Ober C. Asthma genetics in the post-GWAS era. Ann Am Thorac Soc 2016;13 Suppl 1:S85-90.
PUBMED | CROSSREF
2. Ober C, Yao TC. The genetics of asthma and allergic disease: a 21st century perspective. Immunol Rev 2011;242:10-30.
PUBMED | CROSSREF
3. Moffatt MF, Gut IG, Demenais F, Strachan DP, Bouzigon E, Heath S, et al. A large-scale, consortium-based genomewide association study of asthma. N Engl J Med 2010;363:1211-21.
PUBMED | CROSSREF
4. Kontakioti E, Domvri K, Papakosta D, Daniilidis M. HLA and asthma phenotypes/endotypes: a review. Hum Immunol 2014;75:930-9.
PUBMED | CROSSREF
5. Pickrell JK, Berisa T, Liu JZ, Ségurel L, Tung JY, Hinds DA. Detection and interpretation of shared genetic influences on 42 human traits. Nat Genet 2016;48:709-17.
PUBMED | CROSSREF
6. Shrine N, Portelli MA, John C, Soler Artigas M, Bennett N, Hall R, et al. Moderate-to-severe asthma in individuals of European ancestry: a genome-wide association study. Lancet Respir Med 2019;7:20-34.
PUBMED | CROSSREF
7. Kato N, Takeuchi F, Tabara Y, Kelly TN, Go MJ, Sim X, et al. Meta-analysis of genome-wide association studies identifies common variants associated with blood pressure variation in east Asians. Nat Genet 2011;43:531-8.
PUBMED | CROSSREF
8. Yoon D, Ban HJ, Kim YJ, Kim EJ, Kim HC, Han BG, et al. Replication of genome-wide association studies on asthma and allergic diseases in Korean adult population. BMB Rep 2012;45:305-10.
PUBMED | CROSSREF
9. Park BL, Kim TH, Kim JH, Bae JS, Pasaje CF, Cheong HS, et al. Genome-wide association study of aspirin-exacerbated respiratory disease in a Korean population. Hum Genet 2013;132:313-21.
PUBMED | CROSSREF

10. Moon S, Kim YJ, Han S, Hwang MY, Shin DM, Park MY, et al. The Korea Biobank array: design and identification of coding variants associated with blood Biochemical traits. Sci Rep 2019;9:1382.
PUBMED | CROSSREF
11. Affymetrix Inc. Axiom® genotyping solution data analysis guide. Santa Clara (CA): Affymetrix Inc.; 2015.
12. Seo S, Park K, Lee JJ, Choi KY, Lee KH, Won S. SNP genotype calling and quality control for multi-batch-based studies. Genes Genomics 2019;41:927-39.
PUBMED | CROSSREF
13. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007;81:559-75.
PUBMED | CROSSREF
14. Song YE, Lee S, Park K, Elston RC, Yang HJ, Won S. ONETOOL for the analysis of family-based big data. Bioinformatics 2018;34:2851-3.
PUBMED | CROSSREF
15. Wigginton JE, Cutler DJ, Abecasis GR. A note on exact tests of Hardy-Weinberg equilibrium. Am J Hum Genet 2005;76:887-93.
PUBMED | CROSSREF
16. McCarthy S, Das S, Kretzschmar W, Delaneau O, Wood AR, Teumer A, et al. A reference panel of 64,976 haplotypes for genotype imputation. Nat Genet 2016;48:1279-83.
PUBMED | CROSSREF
17. Loh PR, Danecek P, Palamara PF, Fuchsberger C, A Reshef Y, K Finucane H, et al. Reference-based phasing using the Haplotype Reference Consortium panel. Nat Genet 2016;48:1443-8.
PUBMED | CROSSREF
18. Das S, Forer L, Schönherr S, Sidore C, Locke AE, Kwong A, et al. Next-generation genotype imputation service and methods. Nat Genet 2016;48:1284-7.
PUBMED | CROSSREF
19. Machiela MJ, Chanock SJ. LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. Bioinformatics 2015;31:3555-7.
PUBMED | CROSSREF
20. de Leeuw CA, Mooij JM, Heskes T, Posthuma D. MAGMA: generalized gene-set analysis of GWAS data. PLOS Comput Biol 2015;11:e1004219.
PUBMED | CROSSREF
21. Gamazon ER, Wheeler HE, Shah KP, Mozaffari SV, Aquino-Michaels K, Carroll RJ, et al. A gene-based association method for mapping traits using reference transcriptome data. Nat Genet 2015;47:1091-8.
PUBMED | CROSSREF
22. GTEx Consortium. The Genotype-Tissue Expression (GTEx) project. Nat Genet 2013;45:580-5.
PUBMED | CROSSREF
23. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. Am J Hum Genet 2011;88:76-82.
PUBMED | CROSSREF
24. Cho SH, Park HW, Rosenberg DM. The current status of asthma in Korea. J Korean Med Sci 2006;21:181-7.
PUBMED | CROSSREF
25. Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alföldi J, Wang Q, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. Nature 2020;581:434-43.
PUBMED | CROSSREF
26. Hirota T, Takahashi A, Kubo M, Tsunoda T, Tomita K, Doi S, et al. Genome-wide association study identifies three new susceptibility loci for adult asthma in the Japanese population. Nat Genet 2011;43:893-6.
PUBMED | CROSSREF
27. Moffatt MF, Kabesch M, Liang L, Dixon AL, Strachan D, Heath S, et al. Genetic variants regulating ORMDL3 expression contribute to the risk of childhood asthma. Nature 2007;448:470-3.
PUBMED | CROSSREF
28. Torgerson DG, Ampleford EJ, Chiu GY, Gauderman WJ, Gignoux CR, Graves PE, et al. Meta-analysis of genome-wide association studies of asthma in ethnically diverse North American populations. Nat Genet 2011;43:887-92.
PUBMED | CROSSREF
29. Leung TF, Ko FW, Sy HY, Tsui SK, Wong GW. Differences in asthma genetics between Chinese and other populations. J Allergy Clin Immunol 2014;133:42-8.
PUBMED | CROSSREF

30. Kim JH, Park BL, Cheong HS, Bae JS, Park JS, Jang AS, et al. Genome-wide and follow-up studies identify CEP68 gene variants associated with risk of aspirin-intolerant asthma. PLoS One 2010;5:e13818.
PUBMED | CROSSREF
31. Shin SW, Park J, Kim YJ, Uh ST, Choi BW, Kim MK, et al. A highly sensitive and specific genetic marker to diagnose aspirin-exacerbated respiratory disease using a genome-wide association study. DNA Cell Biol 2012;31:1604-9.
PUBMED | CROSSREF
32. Kim SH, Cho BY, Choi H, Shin ES, Ye YM, Lee JE, et al. The SNP rs3128965 of HLA-DPB1 as a genetic marker of the AERD phenotype. PLoS One 2014;9:e111220.
PUBMED | CROSSREF
33. Kim JH, Park BL, Cheong HS, Bae JS, Park JS, Jang AS, et al. Genome-wide and follow-up studies identify CEP68 gene variants associated with risk of aspirin-intolerant asthma. PLoS One 2010;5:e13818.
PUBMED | CROSSREF
34. Kim SH, Cho BY, Park CS, Shin ES, Cho EY, Yang EM, et al. Alpha-T-catenin (CTNNA3) gene was identified as a risk variant for toluene diisocyanate-induced asthma by genome-wide association analysis. Clin Exp Allergy 2009;39:203-12.
PUBMED | CROSSREF
35. Vercelli D. Discovering susceptibility genes for asthma and allergy. Nat Rev Immunol 2008;8:169-82.
PUBMED | CROSSREF
36. Shiina T, Inoko H, Kulski JK. An update of the HLA genomic region, locus information and disease associations: 2004. Tissue Antigens 2004;64:631-49.
PUBMED | CROSSREF
37. Movahedi M, Moin M, Gharagozlou M, Aghamohammadi A, Dianat S, Moradi B, et al. Association of HLA class II alleles with childhood asthma and total IgE levels. Iran J Allergy Asthma Immunol 2008;7:215-20.
PUBMED
38. Zhu Z, Guo Y, Shi H, Liu CL, Panganiban RA, Chung W, et al. Shared genetic and experimental links between obesity-related traits and asthma subtypes in UK Biobank. J Allergy Clin Immunol 2020;145:537-49.
PUBMED | CROSSREF
39. Knutsen AP, Vijay HM, Kariuki B, Santiago LA, Graff R, Wofford JD, et al. Association of IL-4RA single nucleotide polymorphisms, HLA-DR and HLA-DQ in children with Alternaria-sensitive moderate-severe asthma. Clin Mol Allergy 2010;8:5.
PUBMED | CROSSREF
40. Howell WM, Holgate ST. HLA genetics and allergic disease. Thorax 1995;50:815-8.
PUBMED | CROSSREF
41. Lasky-Su J, Himes BE, Raby BA, Klanderman BJ, Sylvia JS, Lange C, et al. HLA-DQ strikes again: genome-wide association study further confirms HLA-DQ in the diagnosis of asthma among adults. Clin Exp Allergy 2012;42:1724-33.
PUBMED | CROSSREF
42. Tam RC, Li MW, Gao YP, Pang YT, Yan S, Ge W, et al. Human CLEC16A regulates autophagy through modulating mTOR activity. Exp Cell Res 2017;352:304-12.
PUBMED | CROSSREF
43. Bonser LR, Erle DJ. Airway mucus and asthma: the role of MUC5AC and MUC5B. J Clin Med 2017;6:112.
PUBMED | CROSSREF

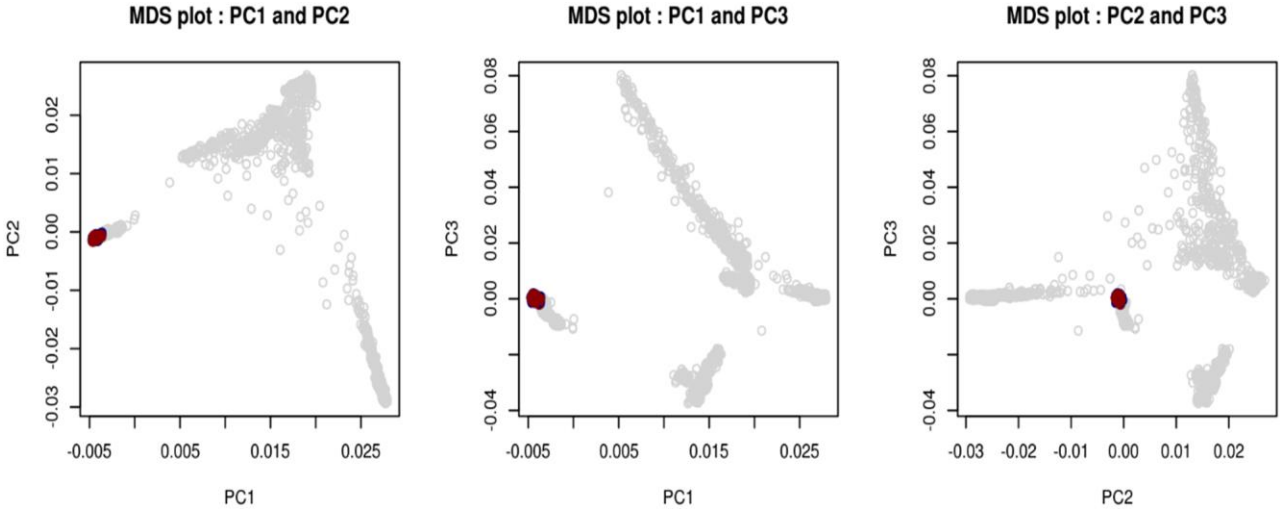


Supplementary Fig. S1. Workflow of QC for our samples. Multiple standard QCs were performed for both asthmatic subjects (Case) and control subjects (Control) to exclude outlier SNPs and subjects. HWE, Hardy-Weinberg equilibrium; MAF, minor allele frequency; QC, quality control; SNP, single nucleotide polymorphism; IBS, identity by state.

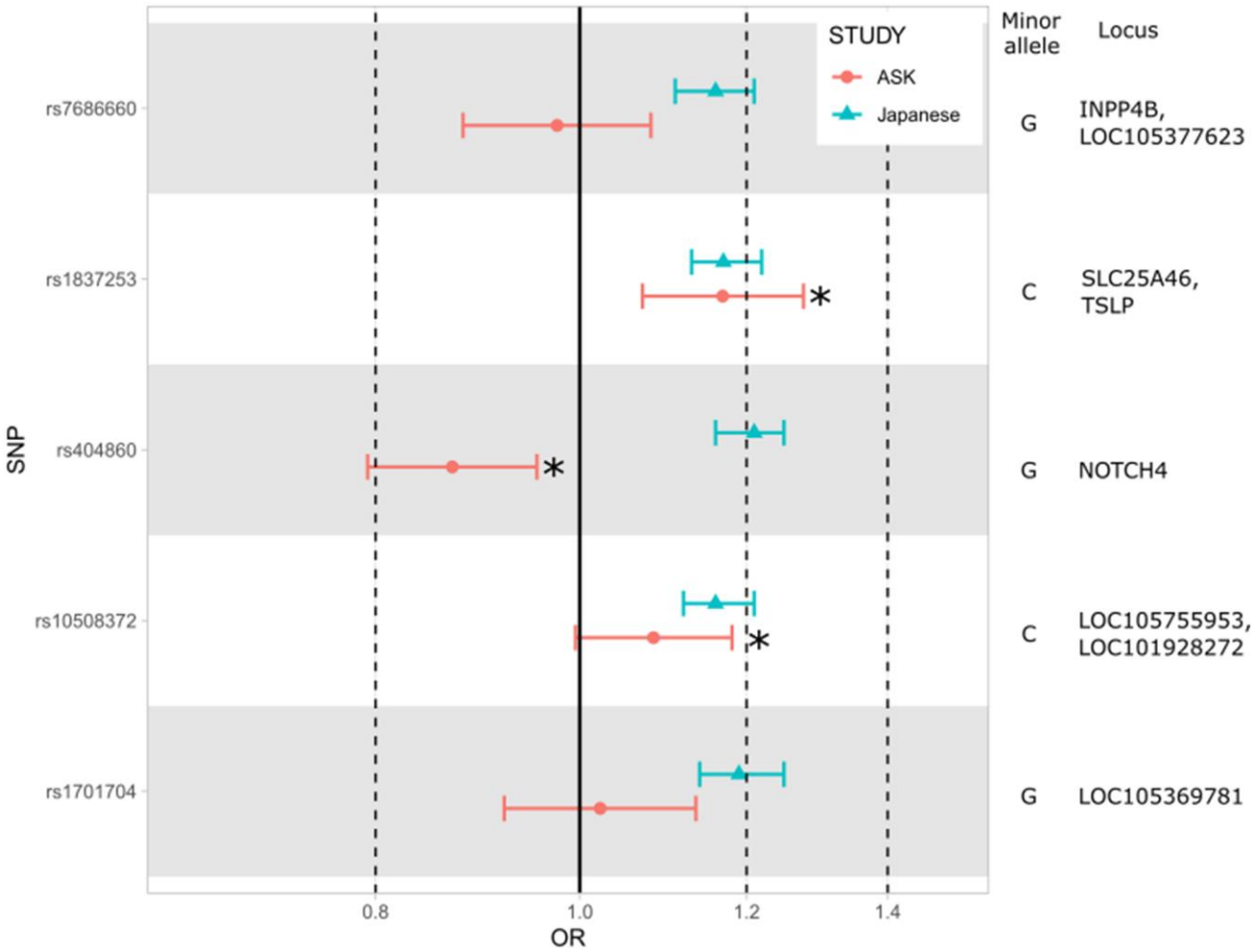


Supplementary Fig. S2. Quantile-quantile plot for the GWAS of asthma. The observed distributions of *P* values in the GWAS are plotted relative to the expected (null) distribution (red line) for the conditional logistic regression analysis. The grey shading indicates the 95% confidence interval for observed *P* values relative to those expected.

GWAS, genome-wide association study.



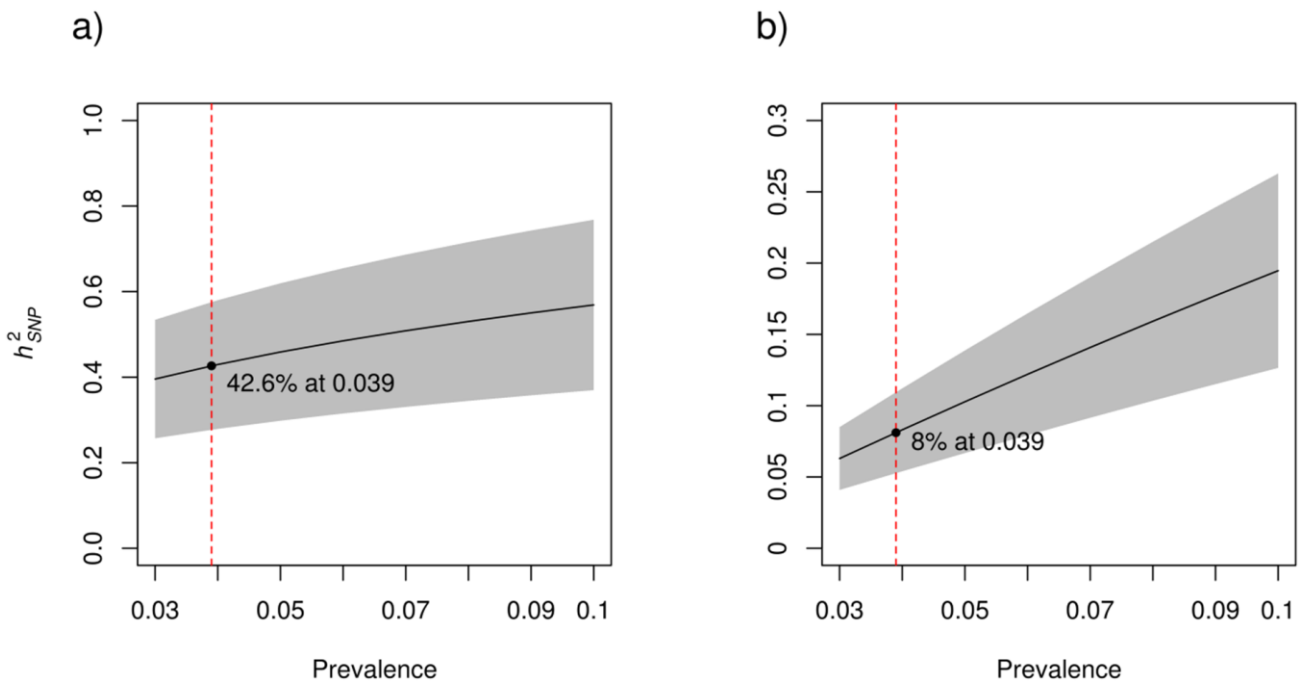
Supplementary Fig. S3. MDS plot. Gray dots represent the 1,000 Genomes Project subjects. Red and blue dots represent the subjects with asthma and control subjects, respectively. The fact that red and blue dots are not divided means that there is no population stratification between asthmatics and control subjects. MDS, multidimensional scaling.



Supplementary Fig. S4. Forest plot for genome-wide significant SNPs associated with asthma. ASK indicates the GWAS results using our Korean data. Japanese indicates the GWAS results using the BioBank Japan data.

SNP, single nucleotide polymorphism; GWAS, genome-wide association study.

*SNPs with $P < 0.05$.



Supplementary Fig. S5. The proportion of phenotypic variances explained according to genotyped SNPs with a disease prevalence ranging from 0.03 to 0.1. The proportion of phenotypic variances explained according to genotyped SNPs was calculated with GCTA on (A) the liability scale and (B) the observed 0–1 scale. The shaded area indicates the 95% confidence interval for h^2_{SNP}.

SNP, single nucleotide polymorphism.