The Genome-wide Association Study of Serum IgE Levels Demonstrated the Shared Genetic Background in Allergic Diseases

Hsing-Fang Lu¹, Chen-Hsing Chou², Ying-Ju Lin¹, Shunsuke Uchiyama³, Chikashi Terao³, Yu-Wen Wang¹, Jai-Sing Yang¹, Ting-Yuan Liu¹, Henry Sung-Ching Wong⁴, Sean Chun-Chang Chen⁴, and Fuu-Jen Tsai¹

¹China Medical University Hospital
²China Medical University
³RIKEN Center for Integrative Medical Sciences
⁴Taipei Medical University

August 8, 2023

Abstract

Background Immunoglobulin E (IgE) is highly related to a variety of atopic diseases, and several genome-wide correlation studies (GWASs) have demonstrated the association between genes and IgE. In this study, we conducted the largest genome-wide association study of IgE in a Taiwanese Han population and aimed to elucidate the genetic architecture of IgE. Methods Genome-wide association study was used to discover the association between variants and IgE. Through HLA imputation, we explored the association between HLA alleles and IgE. In order to explore the pleiotropy relationship between IgE and atopic diseases, we performed both global and local genetic correlation analysis. Moreover, we divided our cohort into a training group and a validation group to construct the polygenic risk score (PRS) of IgE and applied it to test the risk of asthma and atopic dermatitis. Results A total of 8 independent variants showed genome-wide significance, and rs147642819 at 6p21.32 was the most significant signal ($p=1.8 \times 10^{-19}$). Seven of the loci were replicated successfully after a meta-analysis of the Japanese population. Among all the HLA alleles, HLA-DQB1*03:03 is the most significant allele. The global genetic correlation showed significance between IgE and asthma. The IgE PRS significantly correlated with the total IgE level. Furthermore, the top 10 quantile IgE PRS group had a potentially higher risk for asthma, which was replicated in the Japanese population as well. Conclusions Our study provided a more comprehensive understanding of the impact of genomic variants including complex HLA alleles on serum IgE.
5. Genetic Center, Department of Medical Research, China Medical University Hospital, Taichung, Taiwan
6. School of Chinese Medicine, China Medical University, Taichung, Taiwan
7. Genetic Center, Department of Medical Research, China Medical University Hospital, Taichung, Taiwan
8. Clinical Research Center, Shizuoka General Hospital, Shizuoka, Japan
9. The Department of Applied Genetics, The School of Pharmaceutical Sciences, University of Shizuoka, Shizuoka, Japan
10. Graduate Institute of Biomedical Informatics, College of Medical Science and Technology, Taipei Medical University, Taipei, Taiwan.
11. School of Chinese Medicine, China Medical University, Taichung, Taiwan
12. Department of Biotechnology and Bioinformatics, Asia University, Taichung, Taiwan

+ Hsing-Fang Lu and Chen-Hsing Chou contributed equally to this work.

*Corresponding Author:
Fuu-Jen Tsai M.D. Ph.D.
School of Chinese Medicine, China Medical University, Taiwan
No.91, Hsueh-Shih Road, Taichung 40402, Taiwan
Phone: 886-4-22053366
E-mail: d0704@mail.cmuh.org.tw

Competing interests
All authors declare that they have no other relevant conflicts of interest.

Acknowledgements
We thank the members of the BioBank Japan Project, headquartered in the University of Tokyo Institute of Medical Science, for supporting this project.

Abstract

Background
Immunoglobulin E (IgE) is highly related to a variety of atopic diseases, and several genome-wide correlation studies (GWASs) have demonstrated the association between genes and IgE. In this study, we conducted the largest genome-wide association study of IgE in a Taiwanese Han population and aimed to elucidate the genetic architecture of IgE.

Methods
Genome-wide association study was used to discover the association between variants and IgE. Through HLA imputation, we explored the association between HLA alleles and IgE. In order to explore the pleiotropy relationship between IgE and atopic diseases, we performed both global and local genetic correlation analysis. Moreover, we divided our cohort into a training group and a validation group to construct the polygenic risk score (PRS) of IgE and applied it to test the risk of asthma and atopic dermatitis.

Results
A total of 8 independent variants showed genome-wide significance, and rs147642819 at 6p21.32 was the most significant signal ($p=1.8\times10^{-19}$). Seven of the loci were replicated successfully after a meta-analysis of the Japanese population. Among all the HLA alleles, \textit{HLA-DQB1}\textsuperscript{*}03:03 is the most significant allele. The global genetic correlation showed significance between IgE and asthma. The IgE PRS significantly correlated with the total IgE level. Furthermore, the top 10 quantile IgE PRS group had a potentially higher risk for asthma, which was replicated in the Japanese population as well.

Conclusions
Our study provided a more comprehensive understanding of the impact of genomic variants including complex HLA alleles on serum IgE.

**Keywords**
Asthma, GWAS, HLA, IgE

**Introduction**
Allergic sensitization is the process of stimulating B cells and plasma cells with environmental antigens and releasing immunoglobulin E (IgE) antibodies in serum. IgE is highly related to a variety of atopic diseases, such as asthma, atopic dermatitis, and rhinitis.

The heritability of serum IgE levels ranges from 30% to 80% according to twin-based studies, which underlies the importance of the genetic impact of IgE. Several genome-wide association studies (GWAS) demonstrated the correlation between genes and IgE. The findings of significant regions varied between populations. Moffatt et al. demonstrated that HLA-DR is significantly associated with IgE concentration in Caucasian ancestry. A multi-ancestry meta-analysis reveals that HLA-DQB1 *03:02 was reported as a genome-wide significant signal for IgE in Latino and Hispanic populations. More recently, the MHC region and IL4r showed significance with IgE in the Japanese population. In contrast, no GWAS-significant signals of IgE were identified in a Han Chinese population-based study with a relatively small sample size.

Earlier studies implicated the shared genetic components between IgE and atopic diseases. The largest meta-analysis of the Caucasian population revealed that both HLA-DRB1 and IL4r/IL2r were related to IgE but had no association with asthma. Kim et al. pointed out that there may be a genetic basis shared between IgE and atopic dermatitis. However, the global genetic correlation and the local genetic correlation between these traits is still lacking.

Our study aimed to improve the understanding of the genetic architecture that affect IgE in a Taiwanese Han population and explore the pleiotropy effect between IgE and atopic diseases including asthma and atopic dermatitis. Finally, we constructed the PRS of IgE and explored the relationship between IgE PRS and allergic diseases.

**Materials and Methods**

**Sample Collection**
We included study samples from China Medical University Hospital Precision Medicine Project cohort (CM-UH, Taiwan). The electronic medical record (EMR) of the CMUH was used to collect clinical information from each individual. The first serum IgE value of each patient in the EMR was collected. We collected diseases that may influence the level of IgE, including asthma (ICD9: 493), eczema/dermatitis (ICD 9: 691, 692), allergic rhinitis (ICD 9: 477), anaphylaxis (ICD 9: 995.0, 999.4), conjunctivitis (ICD 9: 372.1, 372.3), food allergy (ICD 9: 693.1), urticaria/angioedema (ICD 9: 708, 995.1), other allergy (ICD9: 995.3, 998.5), rheumatoid arthritis (ICD 9: 714.0, 714.4, 714.89, 714.9), systemic lupus erythematosus (ICD 9: 710.0), nephrotic syndrome (ICD 9: 581.3, 581.1, 581.2, 581.0, 581.89, 581.9), anaphylactoid purpura (ICD 9: 287.0), and cancers (ICD9: 140—172, 174—195.8, and 200—208). We also collected the drug usage before the examination of serum IgE level, including antihistamines (ATC code: R06), omalizumab (anti-IgE therapy; ATC code: R03DX05), adrenaline (ATC code: A01AD01, B02BC09, C01CA24, R01AA14, R03AA01, S01EA01), inhaled corticosteroid (ATC code: R03BA), short- and long-acting Beta2 agonist (ATC code: R03AC), leukotrien receptor antagonist (ATC code: R03DC), kombinationspræparater (ATC code: R03AK), and omalizumab (anti-IgE therapy; ATC code: R03DX05). A total of 21,662 samples contained serum IgE level records. Cancer patients, omalizumab users and any disease with samples size < 20% of total samples (Figure S1) were excluded from the following analysis which remained N=17,884 as the final sample size. This study was approved by the institutional review board and the ethics committee of the Human Studies Committee of China Medical University Hospital (CMUH111-REC1-176).

**Genotyping Quality Control & Imputation**
Detailed information of genotyping data and imputation method are described previously\textsuperscript{14}. In summary, PLINK1.9 was used to perform quality control for array data\textsuperscript{15}. We excluded SNPs with call rate < 0.98, minor allele frequency (MAF) < 0.001, p value of Hardy–Weinberg Equilibrium (HWE) p value < 1x10\textsuperscript{-6}, and polymorphic variants. For sample-level quality control, we excluded individuals call rate < 0.98, heterozygosity > 5 SD, and remove duplicate samples. Besides, the outliers from 1000 genome EAS population (>3 IQR) were removed from the analysis. For data after imputation, we kept variants with INFO score \textsuperscript{\textsuperscript{\textsuperscript{0.3}}}.

Association study

If the sample is affected by diseases that may influence IgE levels, we will include it in the category of ”any disease.” The same applies to medication usage, which will be consolidated under ”any medication.” The covariates included age, sex, PC1 to PC10, any disease status, and any medication. Bolt-LMM was used to perform the association analysis of IgE\textsuperscript{16}. We regressed IgE by covariates and extracted the rank-based inverse normalize residuals as the input phenotype. No covariates were included during the association test. Multi-SNP-based conditional & joint association analysis using GWAS summary data (GCTA-COJO) was used to conduct conditional analysis and extract the independent signals\textsuperscript{17}. We applied ANNOVAR to conduct cytoband and functional annotation. In order to replicate our results, we conducted meta-analysis of lead SNP +- 250kb region with IgE summary statistics from Tohoku Medical Megabank Organization (ToMMo)\textsuperscript{9}. Meta-analysis was conducted via METAL.

HLA imputation & association analysis

We applied HLA genotype imputation with attribute bagging (HIBAG)\textsuperscript{18} to do the imputation of HLA with 4-digit resolution, including the HLA class I genes (\textit{HLA -A}, \textit{HLA-B}, and \textit{HLA-C}) and class II genes (\textit{HLA-DPB1}, \textit{HLA-DQA1}, \textit{HLA-DQB1} and \textit{HLA-DRB1}). The detailed methods have been described previously\textsuperscript{19}. We removed the 2-degree relative (n=4,071) before the association analysis. The R package - MiDasHLA\textsuperscript{20} was used to conduct the association analysis, and the covariates were the same as previously described. The \textit{HLA} allele with the possibility < 0.8, frequency less than 0.01 and the p-value of HWE < 1x10\textsuperscript{5} were removed from the analysis. We further tested the haplotype association of \textit{HLA-DQA1} and \textit{HLA-DQB1} and serum IgE level.

Pathway analysis

To identify genes that could be involved in the regulation of the serum IgE level, we mapped 1,384 variants that passed a threshold of suggested significance (p < 1x 10\textsuperscript{-6}) to the genes. We further used Ingenuity Pathway Analysis (IPA) to conduct the canonical pathway analysis. To eliminate multiple test errors, the false discovery rate (FDR) q value < 0.05 was defined as significant.

Heritability estimation & Genetic correlation

The average heritability was estimated using the LD score (LDSC)\textsuperscript{21}. We downloaded the summary statistics from Biobank Japan (BBJ)\textsuperscript{22}. In order to update the genome build from Hg19 to GRCh38, we applied liftover to the summary statistics of BBJ. Furthermore, we used LDSC to conduct the genetic correlation (r\textsubscript{g}) between IgE and BBJ phenotypes\textsuperscript{21}. The significant threshold was defined as \(P < 0.002\), according to the Bonferroni correction (0.05/165).

Local genetic correlation

Local Analysis of [co]Variant Association (LAVA)\textsuperscript{23} was used to find the potential loci of genetic correlation between two diseases. Since global SNP genetic correlation (r\textsubscript{g}) represented the mean of local correlations in the genome, there is the possibility that opposing local genetic correlations are masked. To explore the local genetic correlation between IgE and atopic disease, we applied LAVA to conduct the analysis. The regions with p < 0.05 in univariate local genetic analysis was submitted to bivariate analysis. We applied 1000 genome EAS as the reference panel. We defined FDR < 0.05 as a significant threshold.

PRS Construction
We separated the 70% sample size of our cohort to construct the training GWAS summary statistics. PRS-CS\textsuperscript{24} was used to infer the posterior effect size of SNPs using the IgE summary statistics of the training group. The 1000 genome EAS phase3 samples were used as LD reference panel. We further adopted the effect size in the rest of 30% samples and used PLINK2.0\textsuperscript{25} to sum up the effect allele dosages of each individual and generate the IgE PRS. We checked the correlation between IgE PRS and IgE using Pearson correlation coefficient test. In order to investigate the association between different quantiles of IgE PRS groups and IgE, we divided the IgE PRS into 11 parts according to the IgE PRS quantiles (0-10th, 10-20th, 20-30th, 30-40th, 40-50th, 50-55th, 55-60th, 60-70th, 70-80th, 80-90th, and 90-100th). The population within the 50th to 55th quantile range was selected as the reference group. The logistic regression was performed compared with each group with the reference group and adjust with age, sex, and PC1 to PC10.

**Evaluation of IgE PRS for the Risk of Asthma & Atopic Dermatitis**

To reduce potential bias, the samples within the IgE GWAS training set were excluded from the following analysis. We test the association between the case-control status of atopic diseases (asthma and atopic dermatitis) in the top decile (10%) IgE PRS with the rest of 9 deciles (bottom 90%) IgE PRS population under a logistic regression model. Age, sex, and PC1 to PC10 were included as covariant. To explore the application of our IgE PRS, we further applied our IgE PRS in the BBJ population. We applied the posterior obtained from PRS-CS to construct the PRS of individuals in BBJ and replicate the association test in asthma & atopic dermatitis. We used 748,819 SNP which is common between BBJ and Taiwan SNP list as PRS input. A total 748,370 variants were processed for the calculation of IgE PRS. The age, sex, 46 disease status at registry, and PC1 to PC 10 were included as covariates.

**Results**

**Sample Collection**

The study design was shown in supplementary figure 1. Overall, 17,884 individuals were included in this study and 7,468 (41.8%) were men (Table 1). The mean age is 42.1 +- 20.7, and the mean age of male and female was 43.03 +- 19.49 and 40.72 +- 22.17, respectively. Most of the included individuals (76.4%) have at least one of the following diseases: allergic rhinitis, asthma, dermatitis, or urticaria. Some of them received at least one therapy that may influence the serum IgE level (n = 3,161).

**Novel Signal was identified in the IgE GWAS**

The Manhattan plot and QQ plot ($\lambda_{GC}$:1.05) were shown in Figure 1. The LDSC demonstrated an intercept of 1.01 (SE, 0.009), and heritability of IgE was 10.7% (SE, 4.1%), which demonstrated the polygenic effect of IgE. After conditional analysis, a total of 8 independent significant loci were identified (Table 2). The most significant SNP- rs147642819 (6p21.32) was within the intronic region of HLA-DRA. The intronic SNP of CD28 rs1181388 ($p = 2.5 \times 10^{-12}$) on 2q33.2 and an intergenic SNP rs1002957030 ($p = 4.7 \times 10^{-8}$) on 11q23.2 were novel signal for IgE. Six independent signals within chr6 were found included 6p22.1, 6p21.32, and 6q23.3 region (S Figure 2). We also identified an intergenic SNP rs112731607 (IL4R; IL21R) at 16p12.1 as GWAS significance. This region has been reported associated with total serum IgE in a Japanese population as well\textsuperscript{9}. After conducting a meta-analysis with Japanese population (ToMMo), six of the SNPs demonstrated a significant level of association. Among them, rs369358206 within the intergenic region of HLA-DQA1/HLA-DQB1 is the most significant SNP with $p_{\text{meta}} = 1.7 \times 10^{-23}$. No increased significance was observed for on chr11 due to lack of SNP results. Although there were no results of lead SNP (rs112731607) in ToMMo for 16q12.1, but another SNP nearby named rs144651842 showed genome-wide significance after meta-analysis ($\text{beta}=0.18, \text{SE}=0.02, \; p_{\text{meta}} = 9.75 \times 10^{-17}$).

**HLA fine mapping**

Multiple signals were identified in the HLA region according to GWAS; therefore, we further discover the association between the HLA allele and IgE. The HLA star allele with FDR adjusted p values less than 0.05 were shown in Table 3. The most significant signal, HLA-DQB1*03:03 ($p=0.5$, SE=0.05, $p = 3.70 \times 10^{-14}$) is associated with IgE under dominant genetic model. After stepwise conditional analysis under
the dominant model, no other signals remained genome-wide significant (S Table 1). We further select two most significant HLA genes, HLA-DQB1 and HLA-DQA1 to conduct haplotype analysis. The haplotype HLA-DQA1*03:02 - HLA-DQB1*03:03 with frequency 0.15 became more significant positively associated with serum IgE (OR = 1.25, SE = 0.02, p-value = 8.88×10^{-16}) compared with the most protective haplotype (S Table 2).

Pathway analysis

We mapped SNPs with p < 1×10^{-6} to the closest genes and generate a list contained 128 genes (S Table 3). These genes were used to perform the pathway analysis. The top 20 significant pathways were shown in Figure 2a (S Table 4). The most significantly enriched pathway is the antigen presentation pathway (q value = 2.55×10^{-19}), which included HLA-A, HLA-B, HLA-C, HLA-DOB, HLA-DQA1, HLA-DQA2, HLA-DQB1, HLA-DQB2, HLA-DRA, HLA-DRB1, HLA-F, HLA-G, and TAP2. The PD-1, PD-L1 cancer immunotherapy pathway and the type 2 T helper pathway (Th2) were also significantly enriched.

Genetic correlation

A total of 165 traits from the Biobank Japan project were included in the genetic correlation analysis (S Table 5). The p-value less than 0.05 were shown in Figure 2b. Asthma showed a significant positive genetic correlation with IgE level (r_{g} = 0.5, SE = 0.14, p-value = 5×10^{-4}) even after Bonferroni correction (p < 3.7×10^{-4}). On the other hand, the atopic dermatitis did not show significant result (r_{g} = 0.27, p = 0.09). Except for the genetic correlation on a genome-wide scale, we further explore the local pleiotropy among IgE and atopic diseases (S Table 6). A total of seven significant local genetic correlation signals with q-value < 0.05 were founded between IgE and asthma (Figure 2c), among them, 5 were within chr6 including 6p21.32, 6p21.33, 6q23.3-q23.2, and 6p24.1. The cytoband-16p12.1 showed the strongest local genetic correlation between IgE and Asthma (r_{g}=0.81, q-value = 2×10^{-4}). We also observed one local genetic correlation of 5q31.1 between IgE and AD (r_{g} = 0.79, q-value = 0.04) even though the global genetic correlation did not show any significance.

PRS Construction

We used the 70% (n = 12,520) samples to calculate training GWAS (S Figure 3). After utilized PRS-CS to generate the posterior of variants, a total of 766,640 SNPs were included to generate the PRS. We further utilized the remaining 30% (n = 5,364) target data to generate IgE PRS for each individual. The IgE PRS was positively correlated with serum IgE level (Spearman’s rank correlation coefficient = 0.08, p = 1.59×10^{-8}). From the strata plot, there was a positive trend that higher percentile revealed potentially higher OR of IgE level compared with the reference quantile (Figure 3a).

Application of IgE PRS

Since IgE demonstrated genetic correlation between asthma and atopic dermatitis, indicated the pleiotropy effect of these regions and the highly shared genetic components between IgE and these two diseases. To reveal how the risk of higher IgE may relate to asthma and atopic dermatitis, we evaluated the association between IgE PRS and the risk status of these diseases (Figure 3b) in another testing group. The top decile showed significant higher risk in asthma (OR = 1.11, 95% CI = 1.03-1.20, p = 0.006) and atopic dermatitis (OR = 1.06, 95%CI = 1.01-1.11, p = 0.027) compared with the bottom 9 deciles. In BBJ, we observed the consistent association between IgE PRS with asthma (OR = 1.10, 95% CI = 1.03-1.18, p = 0.006) and atopic dermatitis, the trend remained positive (OR = 1.02) but no statistics significance (p = 0.832).

Discussion

Our study was the largest number of East Asian samples to analyze IgE GWAS to date. By imputing the HLA allele, we obtained a higher resolution of the association between HLA and serum IgE level. Furthermore, we explore the pleiotropy relationship between IgE and atopic diseases. Finally, we construct the PRS of IgE and found the positive relationship between IgE PRS and asthma and atopic dermatitis.
The most significant SNP rs147642819 on chr6 has been reported to be associated with several blood cell including white blood cell count, neutrophil count, eosinophil counts, and type 1 diabetes. Besides, this SNP is an eQTL for HLA-DQB1 in several tissues (including venous blood, thyroid gland, and esophagus squamous epithelium), and locates within active enhancer region in CD14 positive monocyte with the RegulomeDB score of 1f. This evidence suggested that the variant may be biological relevance and participated in gene expression via transcriptional regulation.

Previous family study points out the linkage of 2q33 with total serum IgE in asthma families. This cytoband harbors two candidate genes - CTLA4 and CD28, and both of them are involved in regulation of T-cell activation. Howard et al. demonstrates that the polymorphisms of CTLA4 associate with the total serum IgE levels, but no association is observed of the variants from CD28. Our study firstly demonstrated an intron variant of CD28 - rs1181388 in cytoband 2q33.2 is significantly associated with IgE. Interestingly, this SNP is in DHS promoter of CTLA4, and is an eQTL of CTLA4 in blood as well. This variant hasn’t been reported as genome wide significant signals with other traits but being nominal significance (p=2.5 × 10^-4) in age hay fever, rhinitis or eczema diagnosed in UKBB. Besides, the SNP is in highly LD (R^2=0.92) with the asthma lead SNP - rs55730955, which is previous reported in the Japanese population. Further studies are still needed to deeper elucidate the relationship and impact of variants within CD28 and CTLA4 to IgE level.

Multiple signals were identified in the MHC regions after conditional analysis, suggesting the importance and high heterogeneity of HLA genes. Several studies mentioned the association between HLA genes and serum IgE, but the identified gene and HLA allele vary across different populations. In Latino and Hispanic populations, they report HLA-DQB1 *03:02 as the most significant signal for IgE. Our study identified HLA-DQB*03:03 as the most significant star allele in the Taiwanese Han population, while HLA-DQB1 *03:02 did not show significant association. The frequency of HLA-DQB1 *03:02 is higher in Hispanic (0.194) compared with HLA-DQB*03:02 (0.02), while the frequency of HLA-DQB*03:03 (0.14) is higher in our cohort compared with HLA-DQB1 *03:02 (0.08). The difference distribution of HLA allele frequencies across ancestries emphasizes the value of using diverse ethnic groups to identify the shared and unique population-specific association signals.

For atopic dermatitis, the global genetic correlation did not show significance, but the local genetic correlation revealed borderline significant signals. Our study revealed that IgE and atopic dermatitis shared genetic components in 5q31.1. Previous study has demonstrated the association between 5q31.1 and IgE, and the linkage association of 5q31-33 with AD has been reported as well. The local genetic correlation enabled us having a deeper understanding of the shared genetic components of both traits.

The IgE PRS was significantly correlated with the serum IgE level and the risk of asthma and atopic dermatitis in testing group. Furthermore, the association between IgE PRS and asthma was confirmed in a Japanese population. This led to the high possibility of the future application of IgE PRS. However, the AD did not show significance signal in Japanese population. This may be due to relatively smaller sample size in Japanese cohort (case = 4,288; control = 124,753) compared with our testing group (case = 17,327; control = 247,433). Currently in clinical practice, patients undergo testing for total IgE levels only when allergic symptoms appear. Therefore, our study aims to identify individuals with high-risk to allergic disease before their IgE levels rise even in newborn babies. With the decreasing cost and increasing accessibility of DNA microarrays, genotyping screening may become easier to implement in hospitals. By applying IgE PRS, each subject can obtain their risk of allergic diseases before symptoms manifest. This information can be particularly beneficial for high-risk patients, as it can aid in preventing antigen exposure and reducing the risk of allergies through early prevention measures.

In summary, our study identified 8 variants associated with IgE in the Taiwanese Han population, successfully replicated six of them. Additionally, the IgE PRS showcased its potential for future clinical applications. These findings contribute to a deeper understanding of the genetic architecture of serum IgE and its relationship with atopic diseases.
Funding

H-FL was supported by China Medical University Hospital, Taiwan (DMR-111-190). C.T. was supported by Japan Agency for Medical Research and Development (AMED) grants 22ek0109555, 21tm0424220, 21ck0106642, 23ek0410114 and 23tm0424225, Japan Society for the Promotion of Science (JSPS) KAKENHI grant JP20H00462. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author contribution

HFL, CHC, SU, YWW and JSY analyzed the data. HFL, CHC, YWW, and TYL collect the data. FJT, HFL, YJL and CT planned the study design. HFL wrote the manuscript. FJT, YJL and CT reviewed the manuscript. FJT organized the database.

Table 1. Demographic and baseline characteristics of included samples

<table>
<thead>
<tr>
<th>Demographic characteristics</th>
<th>Total</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, year, mean (SD)</td>
<td>42.07 (20.68)</td>
<td>40.72 (22.17)</td>
<td>43.03 (19.49)</td>
</tr>
<tr>
<td>Male, n (%)</td>
<td>7,468 (41.76)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Serum IgE, mean (SD)</td>
<td>268.04 (544.08)</td>
<td>327.72 (616.81)</td>
<td>225.25 (480.76)</td>
</tr>
<tr>
<td>Diseases</td>
<td></td>
<td>5,470 (30.59)</td>
<td>2,492 (13.93)</td>
</tr>
<tr>
<td>Allergic Rhinitis, n (%)</td>
<td></td>
<td>3,613 (20.2)</td>
<td>1,633 (9.13)</td>
</tr>
<tr>
<td>Asthma, n (%)</td>
<td></td>
<td>5,093 (28.48)</td>
<td>3,093 (17.29)</td>
</tr>
<tr>
<td>Dermatitis, n (%)</td>
<td></td>
<td>13,664 (76.4)</td>
<td>8,037 (44.94)</td>
</tr>
<tr>
<td>Urticaria, n (%)</td>
<td></td>
<td>4,760 (26.62)</td>
<td>3,152 (17.62)</td>
</tr>
<tr>
<td>Any diseases, n (%)</td>
<td></td>
<td>4,760 (26.62)</td>
<td>3,152 (17.62)</td>
</tr>
</tbody>
</table>

Drug user

<table>
<thead>
<tr>
<th>Drug user</th>
<th>Total</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adrenaline, n (%)</td>
<td>1,207 (6.75)</td>
<td>609 (3.41)</td>
<td>608 (3.34)</td>
</tr>
<tr>
<td>Antihistamines, n (%)</td>
<td>2,111 (11.8)</td>
<td>938 (5.24)</td>
<td>1,173 (6.56)</td>
</tr>
<tr>
<td>Beta2 Agonist, n (%)</td>
<td>1,411 (7.89)</td>
<td>666 (3.72)</td>
<td>745 (4.17)</td>
</tr>
<tr>
<td>Corticosteroid, n (%)</td>
<td>459 (2.57)</td>
<td>173 (0.97)</td>
<td>286 (1.6)</td>
</tr>
<tr>
<td>Leukotrien Receptor Antagonist, n (%)</td>
<td>407 (2.28)</td>
<td>171 (0.96)</td>
<td>236 (1.32)</td>
</tr>
<tr>
<td>Any drugs, n (%)</td>
<td>3,161 (17.68)</td>
<td>1,417 (7.92)</td>
<td>1,744 (9.75)</td>
</tr>
</tbody>
</table>

IgE, immunoglobulin E; SD, standard deviation.

Table 2. Genome-wide significant variants of IgE after conditional analysis.

<table>
<thead>
<tr>
<th>Chr</th>
<th>POS</th>
<th>SNP</th>
<th>cytoband</th>
<th>Ref</th>
<th>Alt</th>
<th>CMUH</th>
<th>CMUH</th>
<th>CMUH</th>
<th>CMUH</th>
<th>ToMMo</th>
<th>ToMMo</th>
<th>ToMMo</th>
<th>ToMMo</th>
<th>Meta</th>
<th>Meta</th>
<th>Meta</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>32460320</td>
<td>rs147642819</td>
<td>6p21.32</td>
<td>T</td>
<td>TG</td>
<td>0.229</td>
<td>0.112</td>
<td>0.012</td>
<td>1.80E-19</td>
<td>0.275</td>
<td>0.061</td>
<td>0.019</td>
<td>0.002</td>
<td>0.243</td>
<td>0.097</td>
<td>0.011</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>29680621</td>
<td>rs365052</td>
<td>6p22.1</td>
<td>G</td>
<td>C</td>
<td>0.535</td>
<td>0.095</td>
<td>0.011</td>
<td>2.40E-19</td>
<td>0.427</td>
<td>0.001</td>
<td>0.017</td>
<td>0.91</td>
<td>0.494</td>
<td>-0.07</td>
<td>0.009</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>32652743</td>
<td>rs369358206</td>
<td>6p21.32</td>
<td>CT</td>
<td>C</td>
<td>0.542</td>
<td>0.088</td>
<td>0.011</td>
<td>1.50E-16</td>
<td>0.543</td>
<td>0.096</td>
<td>0.017</td>
<td>2.70E-08</td>
<td>0.542</td>
<td>0.090</td>
<td>0.009</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>135204258</td>
<td>rs3819409</td>
<td>6q23.3</td>
<td>G</td>
<td>T</td>
<td>0.435</td>
<td>0.066</td>
<td>0.011</td>
<td>8.60E-10</td>
<td>0.327</td>
<td>0.060</td>
<td>0.018</td>
<td>0.001</td>
<td>0.593</td>
<td>-0.064</td>
<td>0.009</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>203711228</td>
<td>rs1181388</td>
<td>2q33.2</td>
<td>A</td>
<td>G</td>
<td>0.47</td>
<td>0.072</td>
<td>0.011</td>
<td>2.50E-12</td>
<td>0.485</td>
<td>0.070</td>
<td>0.017</td>
<td>3.40E-05</td>
<td>0.474</td>
<td>0.072</td>
<td>0.009</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>32979175</td>
<td>rs16901848</td>
<td>6p22.1</td>
<td>G</td>
<td>A</td>
<td>0.127</td>
<td>-0.108</td>
<td>0.016</td>
<td>5.40E-12</td>
<td>0.163</td>
<td>-0.071</td>
<td>0.023</td>
<td>0.003</td>
<td>0.862</td>
<td>0.096</td>
<td>0.013</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>135204258</td>
<td>rs3819409</td>
<td>6q23.3</td>
<td>G</td>
<td>T</td>
<td>0.435</td>
<td>0.066</td>
<td>0.011</td>
<td>8.60E-10</td>
<td>0.327</td>
<td>0.060</td>
<td>0.018</td>
<td>0.001</td>
<td>0.593</td>
<td>-0.064</td>
<td>0.009</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>27378605</td>
<td>rs112731607</td>
<td>16p12.1</td>
<td>C</td>
<td>T</td>
<td>0.101</td>
<td>-0.109</td>
<td>0.019</td>
<td>7.20E-09</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>0.899</td>
<td>0.109</td>
<td>0.019</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>112926911</td>
<td>rs1002957030</td>
<td>11q23.2</td>
<td>A</td>
<td>G</td>
<td>0.004</td>
<td>0.532</td>
<td>0.097</td>
<td>4.70E-08</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>0.004</td>
<td>0.532</td>
<td>0.097</td>
</tr>
</tbody>
</table>

Ref, reference allele; Alt, alternative allele; Freq, allele frequency of alternative allele; SE, standard error; P, p-value; ToMMo, the results from Tohoku Medical Megabank Organization.
Table 3. HLA alleles significantly associated with IgE levels after conditional analysis

<table>
<thead>
<tr>
<th>HLA allele</th>
<th>N</th>
<th>%</th>
<th>p value</th>
<th>Beta</th>
<th>SE</th>
<th>covariates</th>
</tr>
</thead>
<tbody>
<tr>
<td>DQB1*03:03</td>
<td>3788</td>
<td>13.71%</td>
<td>3.71E-14</td>
<td>1.157</td>
<td>0.019</td>
<td></td>
</tr>
<tr>
<td>DQB1*05:02</td>
<td>2600</td>
<td>9.41%</td>
<td>1.98E-05</td>
<td>0.910</td>
<td>0.022</td>
<td>DQB1*03:03</td>
</tr>
<tr>
<td>A*02:01</td>
<td>1463</td>
<td>5.30%</td>
<td>2.03E-05</td>
<td>0.888</td>
<td>0.028</td>
<td>DQB1<em>03:03 + DQB1</em>05:02</td>
</tr>
<tr>
<td>DRB345*03:01</td>
<td>2413</td>
<td>8.73%</td>
<td>8.81E-05</td>
<td>0.915</td>
<td>0.023</td>
<td>DQB1<em>03:03 + DQB1</em>05:02 + A*02:01</td>
</tr>
</tbody>
</table>

HLA, human leukocyte antigen alleles; SE, standard error; The P adjusted value was according to the FDR test.

References:


**Figure legends:**

**Figure 1.** Genome-wide association study for serum IgE level in a Taiwanese Han population. The Manhattan plot (a) and the quantile-quantile plot (b) of the IgE GWAS.

**Figure 2.** Pathway analysis and the genetic correlation results. (a) The top 20 enriched pathways were shown. (b) The global genetic correlation between IgE and BBJ. Only the results with $p < 0.05$ were shown (c) The $q$ value of the local genetic correlation between IgE and allergic diseases (asthma and atopic dermatitis) were shown. The red dashed line represented $q$ value $=0.05$.

**Figure 3.** Comparison the risk among different strata of IgE polygenic risk score (PRS) for IgE and allergic disease. The error bar represented standard deviation. (a) The quantile plot showed the odds ratio (OR) for each stratum of IgE PRS percentile compared with the median IgE PRS group. (b) The OR of allergic disease among top 10% compared with the bottom 90% of the IgE PRS group. The top 10% IgE PRS group showed significant higher risk for asthma in both CMUH and BBJ. CMUH, China Medical University Hospital; BBJ, Biobank Japan. AD, atopic dermatitis.

**Hosted file**