BMB
Reports

Replication of genome-wide association studies on asthma and allergic diseases in Korean adult population

Dankyu Yoon[1,2,#], Hyo-Jeong Ban[3,#], Young Jin Kim[4], Eun-Jin Kim[1,2], Hyung-Cheol Kim[4], Bok-Ghee Han[5], Jung-Won Park[1,6], Soo-Jong Hong[1,7], Sang-Heon Cho[1,8], Kiejung Park[3,] & Joo-Shil Lee[1,2,*]*

[1]Research Center for Standardization of Allergic Disease, [2]Allergy TF, Department of Immunology and Pathology, Korea National Institute of Health, [3]Division of Bio-Medical Informatics, Center for Genome Science, Korea National Institute of Health, [4]Division of Structural and Functional Genomics, Center for Genome Science, Korea National Institute of Health, [5]Center for Genome Science, Korea National Institute of Health, Osong 363-951, [6]Department of Internal Medicine and Institute of Allergy, Yonsei University College of Medicine, Seoul 120-752, [7]Childhood Asthma Atopy Center, Department of Pediatrics, Asan Medical Center, University of Ulsan College of Medicine, Seoul 138-736, [8]Department of Internal Medicine, Seoul National University College of Medicine, Seoul 110-799, Korea

Allergic diseases such as asthma, allergic rhinitis, and atopic dermatitis are heterogeneous diseases characterized by multiple symptoms and phenotypes. Recent advancements in genetic study enabled us to identify disease associated genetic factors. Numerous genome-wide association studies (GWAS) have revealed multiple associated loci for allergic diseases. However, the majority of previous studies have been conducted in populations of European ancestry. Moreover, the associations of single nucleotide polymorphisms (SNPs) with allergic diseases have not been studied amongst the large-scale general Korean population. Herein, we performed the replication study to validate the previous variants, known to be associated with allergic diseases, in the Korean population. In this study, we categorized three allergic related phenotypes, one allergy and two asthma related phenotypes, based on self-reports of physician diagnosis and their symptoms from 8,842 samples. As a result, we found nominally significant associations of 6 SNPs with at least one allergic related phenotype in the Korean population. [BMB reports 2012; 45(5): 305-310]

INTRODUCTION

Allergic diseases are complex ailments caused by crosstalk between multiple genes and environmental factors (1). According to a recent report from the World Allergy Organization (WAO), the prevalence of allergic diseases is rising dramatically on a global scale (2). Allergic diseases brought the significant burden of global morbidity and mortality in the world and are regarded as one of the major contributors to the considerable drain on health budgets of developed and emerging economies (2). Given the reports from the Korean National Health and Nutrition Examination Survey (3), the prevalence of allergic diseases is increasing in Korea, and the prevalence of asthma, allergic rhinitis and atopic dermatitis in adult population are 7.6%, 11.9% and 2.95%, respectively. Therefore, it is valuable to identify genes or loci associated with allergic diseases for understanding underlying complex mechanisms of diseases.

Numerous previous genetic studies have been conducted using linkage designs and candidate gene association studies to elucidate etiology of allergic diseases (4). Although these cases provide strong evidence for the involvement of genetic factors in allergic diseases, identifying these susceptible genes has been a challenge. A number of genome-wide association studies (GWAS) have been conducted for various allergic disease/phenotypes, including asthma (5-12), atopy (13), atopic dermatitis (14), serum IgE levels (15), eosinophilic esophagitis (16), and plasma eosinophil count (17). However, most of these GWASs have been focused primarily on samples of European origin. Currently, more than two dozen GWAS for asthma and allergic disease have been performed (18). However, few GWAS have been conducted in populations of racial and ethnic minorities. Moreover, several prior studies reported significant discrepancies in allele frequencies of the variants and differences in genetic architecture between European and Asian (19, 20). In this context, we conducted the replication study on allergic disease amongst an Asian population.

RESULTS

Description of KoGES for KARE project in its relation allergic phenotypes

The Korean Association REsource (KARE) was established as a part of the Korean Genome and Epidemiology study (KoGES). In

*Corresponding author. Joo-Shil Lee, Tel: +82-43-719-8400; Fax: +82-43-719-8519; E-mail: jooshil@korea.kr, Kiejung Park, Tel: +82-43-719-8850; Fax: +82-43-719-8869; E-mail: kjpark63@gmail.com
*These authors jointly directed this work.
#These authors contributed equally to this work.
http://dx.doi.org/10.5483/BMBRep.2012.45.5.305

Received 27 September 2011, Revised 3 November 2011, Accepted 27 January 2012

Keyword: Allergic diseases, Asthma, GWAS replication, Korean

Copyright ⓒ 2012 by the The Korean Society for Biochemistry and Molecular Biology
This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/3.0) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

KARE, 8,842 subjects with 352,228 SNPs were analyzed. Average ages of male and female were 51.78 ± 8.78 (SD) and 52.61 ± 9.01 (SD), respectively. Female participants were 52.7% (N=4,659) of all subjects, 5.5% and 2.2% of these had allergy and asthma diagnostic history, respectively. The characteristics of the study are shown in Table 1.

Comparison analysis between previous GWAS results and KARE

To gain insight into the genetic influence of previously allergic diseases associated loci in an Asian population, we conducted a replication study for those known candidate loci in the Korean population. From GWAS catalog, 46 significant associations of 12 independent GWA studies on allergic diseases were retrieved. Supplementary Table 1 shows the summarized information of all 46 known associations. For comparison analysis, we retrieved the corresponding SNPs comprising directly genotyped, imputed genotype and proxy SNPs in LD ($r^2 > 0.8$). As a result, 32 SNPs were available in KARE (Supplementary Table 1). As expected, difference in allele frequencies across populations was observed. For example, the SNPs including rs1342326, rs17525472, rs3184504 and rs4815617 are monomorphic in Asian populations. Due to this disparity in allele frequencies, those loci were not able to directly compare to those of SNPs in our dataset. We defined SNPs with P values less than 0.05 as a replicated result. Given these criteria, we observed 6 replicated SNPs in total. We summarized the comparison analysis (Supplementary Table 2) and the replicated results (Table 2). We successfully replicated three allergy phenotype associated SNPs comprising rs12619285 (near *IKZF2*; IKAROS family zinc finger 2), rs1295686 (at *IL13*; Interleukin 13) and rs2073643 (at *SLC22A5*; solute carrier family 22 (organic cation/carnitine transporter), member 5). In asthma phenotypes, rs13106227 (at *SHROOM3*; shroom family member 3) for both asthma I and asthma II, rs2416257 (at *WDR36*; WD repeat domain 36 and near *TSLP*; Thymic stromal lymphopoietin) for asthma II and rs3806932 (near WDR36) for asthma I were replicated in our study.

Table 1. Characteristics of the study population

Characteristic	Total	Male	Female
Sample size (N, %)	8,842	4,183 (47.31)	4,659 (52.69)
Age (year) (mean ± SE)	52.22 ± 8.91	51.78 ± 8.78	52.61 ± 9.01
BMI	24.60 ± 3.12	24.25± 2.92	24.90±3.25
FEV_1 (%)	111.7 ± 17.71	106.3 ± 16.25	116.6 ± 17.58
FEV_1/FVC (%)	79.88 ± 7.76	77.91 ± 8.39	81.66 ± 6.65
Allergic disease phenotype			
Allergy diagnosis (N)			
Control	8,348	4,008	4,340
Case	490	173	317
Asthma I (N)			
Control	8,645	4,111	4,534
Case	193	70	123
Asthma II (N, %)			
Control	7,450	3,632	3,818
Case	44	14	30

BMI: Body Mass Index (kg/m^2), FEV_1(%): % predicted value of forced expiratory volume in 1 second, FEV_1/FVC (%): % predicted value of FEV_1 and forced vital capacity (FVC). Asthma I: asthma diagnosis, Asthma II: asthma diagnosis and self-reported symptom.

In order to scrutinize the reason for replication failure of 26 SNPs, the statistical power was calculated. The power analysis revealed that 26 non-replicated SNPs in our study would be detected due to the lack of statistical power (Supplementary Table 2). All 6 replicated loci have relatively high statistical power (0.56-0.86) for the disease association while the power of the remaining variants was found to be relatively low (< 0.47).

Since the statistical power is the main reason for non-replication, we performed an alternative association analysis. Gene-based association analysis is a well-known approach for testing the association between multiple markers within gene region and a trait. It is known to be more powerful than the classical single marker association test under the condition of genetic architecture (21). We performed a gene-based association test for the candidate gene regions near previously reported 46 variants. We observed significant results (P-value < 0.05) at the loci near replicated variants from single SNP analysis in one or more allergic disease related phenotypes (Supplementary Table 3). Unfortunately, however, we did not observe any newly replicated locus that is not listed in the single SNP analysis.

DISCUSSION

We performed the replication study on previously known allergic diseases associated loci in the Korean population. Previously, Kim *et al.* (8) reported GWA results of toluene-induced asthma. However, the investigation was mainly conducted in a population with specific material induced disease. Thus, to the best of our knowledge, we present the first replication study focused on the general Korean population.

From 46 previously reported allergic disease related variants, 32 SNPs were available for the single SNP association test in the current study excluding 4 monomorphic SNPs and 9 SNPs with low imputation quality. Among 32 SNPs, 6 SNPs were successfully replicated in this study (P-value < 0.05). The result from gene-based association test also confirmed the associations of 4 loci that were identified by single marker analysis. Four of six SNPs were found to be located at chromosome 5 including rs1295686 at *IL13*, rs2073643 at *SLC22A5*, rs2416257 at *WDR36* and rs3806932 near *WDR36*. Both *IL13* and *SLC22A5* have been reported to be associated with asthma (6). *WDR36* was previously associated with eosinophilic esophagitis (16), plasma eosinophil count, as well as with atopic asthma (17). Interestingly, Gudbjartsson *et al.* (17) confirmed this signal in both Europeans and Asians. *SHROOM3* and *IKZF2* were reported to be associated with eosinophilic esophagitis (16) and plasma eosinophil count (17), respectively. Here we excluded 9 SNPs due to low im-

Table 2. Significant results of replication analysis on allergic disease phenotype in KARE

SNP	chr	Position	Gene	Allele	Allergy			Asthma			Asthma& self reported symptoms		
					OR	SE	P value	OR	SE	P value	OR	SE	P value
rs12619285	2q34	213532290	IKZF2	A	0.865	0.07	0.038	0.907	0.109	0.371	1.028	0.22	0.899
rs13106227	4q21.1	77637705	SHROOM3	G	0.901	0.067	0.123	0.733	0.107	0.004	0.462	0.241	0.001
rs1295686	5q31.1	132023742	IL13	T	1.237	0.069	0.002	1.042	0.11	0.707	1.19	0.222	0.433
rs2073643	5q31.1	131751187	SLC22A5	C	0.84	0.082	0.033	0.785	0.129	0.059	0.763	0.271	0.319
rs2416257	5q22.1	110463389	WDR36, TSLP	T	1.007	0.151	0.962	1.079	0.227	0.738	2.723	0.327	0.002
rs3806932	5q22.1	110433574	WDR36	G	0.962	0.07	0.586	1.271	0.107	0.025	1.342	0.22	0.182

putation quality prior to comparison analysis based on single SNP analysis. Despite the unavailability for directly matched or proxy makers of these 9 SNPs, we were able to perform the gene-based association test on 7 of 9 excluded loci in the single SNP analysis while 2 gene desert loci were excluded from the analysis. However, these 7 loci were not replicated in the gene-based association analysis (P-value $>$ 0.05). The remainder of the unreplicated 26 SNPs was due to lack of statistical power. In the pool of the failed SNPs, all SNPs were below 50% of statistical power. These results together indicate that further analysis including more sampling and genotyping is required for exploring the possible association of non-replicated loci and SNPs with low imputation quality.

As mentioned above, some SNPs (such as rs1342326, rs3184504, rs4815617, and rs17525472) were not polymorphic in Asian populations. All 4 loci showed distinct LD patterns and some evidence of genetic differentiation across populations (Supplementary Fig. 1 and Supplementary Fig. 2). Given iHS (Integrated Haplotype Score) values within the loci, little evidence of population difference were shown in terms of positive selection. However, F_{st} score indicates that genetic diversity between populations exists around the loci. For instance, the regions around the 4 loci show numerous points of SNPs in the top 1% difference between HapMap CEU and CHB/JPT (Supplementary Fig. 2). This implies that genetic diversity may cause population specific effects of some genomic regions. For example, Soranzo *et al.* (22) reported the hematological trait associated loci at 12q24 and showed this region is positively selected in populations with European ancestry. However, Kato *et al.* (20) reported the blood pressure associated loci at 12q24 in Asian populations and showed underlying complexity in the genetic architecture among populations. Reports have indicated that the phenotype associated loci in a certain population may be varied by genetic diversity in terms of population specificity. Interestingly, previously known allergic disease related rs3184504, monomorphic in Asians, is also located at the 12q24 region. Therefore, further study concentrating on fine mapping and targeted resequencing analysis would be required for understanding the effect of variants in a population specific manner.

In conclusion, our study may provide a deeper understanding of the unraveled mechanisms of allergic diseases. Moreover, the comparison results among populations may play a role as a guide to personal genomics. As shown in this study, previously allergic disease associated loci may have an effect in a population specific manner. Therefore, the association study in multi-ethnic populations would provide valuable information to discover the underlying mechanism of allergic diseases. Moreover, further investigation is needed with a large number of disease samples and accurately defined phenotype information in order to understand the disease mechanism and genetic diversity in the context of disease associated loci.

MATERIALS AND METHODS

Subjects and genotyping

All participants provided informed consent using a form that was approved by all participating Institutional Review Boards. The relevant information for the subjects used in the current study has been described in the previous study (23). In brief, the epidemiologic and clinical data for the study were collected from 10,038 participants within a large-scale Ansan and Ansung general population-based cohort, which was established as part of the KoGES by Korea Center for Diseases Control and Prevention (KCDC). The genomic DNAs were isolated from peripheral blood drawn from the participants and were genotyped on the Affymetrix Genome-Wide Human SNP array 5.0 containing 500,568 SNPs. After quality control on samples and SNPs, we selected 352,228 SNPs with 8,842 individuals for further analysis. The genotype data of the participants were provided from the Korean Genome Analysis project and GWAS was conducted as part of the KARE project.

Definition of Allergic related phenotype

Allergic disease related phenotypes are defined based on information from self-report of physician diagnosis and their symptoms. On the basis of the survey questionnaire, three allergic disease phenotypes were defined: one of them for allergy and two of them for asthma related phenotypes. Allergy diagnosis phenotype was defined based on the history of physician diagnosis on allergic rhinitis and atopic dermatitis. Asthma phenotypes were defined according to the history of physician diagnosis on asthma and self-reported symptoms. Asthma related

phenotypes were classified in detail by the presence of self-reported symptoms. Asthma I was defined based on the history of physician diagnosis on asthma. Asthma II was defined by both the history of physician diagnosis as well as awakening with chest tightness, cough, wheeze, or night time awakening due to shortness of breath during the last 12 months. The control samples were retrieved from 8,842 samples based on no history of asthma, wheezing and allergy.

Comparison between previous GWAS result and KARE

Twelve GWAS on allergic diseases have been reported and 46 statistically significant results from each study (P-value $< 1 \times 10^{-5}$) were retrieved from GWAS catalog (24). Due to the limited amount of directly genotyped data, we used imputed genotype data for the comparison analysis. Those SNPs that were not available in our directly genotyped or imputed data were compared to proxy markers (SNPs in linkage disequilibrium, $r^2 >$ 0.8 in CHB/JPT HapMap Phase II samples) within our dataset. Among 46 previously reported SNPs, 13 SNPs were excluded for further analysis due to monomorphism in Asians (4 SNPs) and low imputation quality (9 SNPs, information score < 0.5).

Statistical analysis

Statistical analysis was performed using PLINK (25) and R software. We performed logistic regression analysis for the binary phenotypes, adjusting age, sex, and recruiting geographic area under additive models. Imputation was carried out by using the IMPUTE program on the basis of NCBI build 36 and dbSNP build 126, we initially used 90 individuals from JPT and CHB founders in HapMap as a reference panel (release 22) (26). Before the association analysis on the imputed genotypes, we removed imputed SNP markers with low genotype information content (info < 0.5), posterior probability score < 0.90, low call rate < 0.90, MAF < 0.01 and HWE (P-value $< 1 \times 10^{-7}$). Among the markers, we selected 32 previously known variants for further analysis. To calculate the statistical power, we used the R statistics software and its package "GeneticsDesign". The power analysis was performed under the additive model and with specific parameters such as odds ratio as estimated in this study, a disease prevalence of 0.07 (3), marker allele frequency in the current study, and alpha as 0.05.

Gene-based associations

A gene-based approach tests for the association between a phenotype and all variants within a gene region. We used set-based tests implemented in PLINK software for the gene-based association test. Gene information near 46 previously allergic diseases associated SNPs was retrieved from UCSC genome browser (hg18). In the association test, we selected options for analysis as r-square = 0.5, 10,000 permutations, and no threshold for p-value of the variants within gene region. After excluding the gene regions with no available SNPs, we performed the set-based test on 34 genes near previously reported variants from asthma and allergic GWAS.

Assessment of linkage disequilibrium (LD) and population structure

To view the LD patterns of genomic regions across populations, we used Haploview software (27) to draw LD plots based on HapMap phase II CEU and CHB/JPT populations. In addition to the graphical view of difference in genetic structure, we also assessed the differential levels of genetic structure using iHS (Integrated Haplotype Score) and F_{st} score. Haplotter is used for retrieving pre-calculated iHS and F_{st} score (28). iHS is a statistic representing the degree of recent positive selection at a locus. F_{st} is a measure of the magnitude of population differentiation between two populations.

Acknowledgements

This study was supported by a grant from the Korea Healthcare Technology R&D Project, Ministry of Health & Welfare, Republic of Korea (A092076). This work was also supported by grants from the Korea Centers for Disease Control and Prevention (The Korean Genome Analysis Project (4845-301), KoGES (4851-302), and Korea Biobank Project (4851-307)) and an intramural grant from the Korea National Institute of Health (2001-N73002-00, 2009-N72001-00), the Republic of Korea. We thank Nam H. Cho (Department of Preventive Medicine, Ajou University, South Korea) and Chol Shin (Department of Internal Medicine, Korea University Ansan Hospital, South Korea) for their great efforts in generating and providing this valuable data.

REFERENCES

1. Vercelli, D. (2008) Discovering susceptibility genes for asthma and allergy. *Nat. Rev. Immunol.* **8**, 169-182.
2. Pawankar, R., Canonica, G., Holgate, S. and Lockey, R. (2011) WAO White Book on Allergy, World Allergy Organisation, Milwaukee, Wisconsin, USA.
3. KCDC (2010) Korea Health Statistics 2009: Korea National Health and Nutrition Examination Survey (KNHANES IV-3).
4. Akhabir, L. and Sandford, A. J. (2011) Genome-wide association studies for discovery of genes involved in asthma. *Respirology* **16**, 396-406.
5. Moffatt, M. F., Kabesch, M., Liang, L., Dixon, A. L., Strachan, D., Heath, S., Depner, M., von Berg, A., Bufe, A., Rietschel, E., Heinzmann, A., Simma, B., Frischer, T., Willis-Owen, S. A., Wong, K. C., Illig, T., Vogelberg, C., Weiland, S. K., von Mutius, E., Abecasis, G. R., Farrall, M., Gut, I. G., Lathrop, G. M. and Cookson, W. O. (2007) Genetic variants regulating ORMDL3 expression contribute to the risk of childhood asthma. *Nature* **448,** 470-473.
6. Moffatt, M. F., Gut, I. G., Demenais, F., Strachan, D. P., Bouzigon, E., Heath, S., von Mutius, E., Farrall, M., Lathrop, M. and Cookson, W. O. (2010) A large-scale, consortium-based genomewide association study of asthma. *N. Engl. J. Med.* **363,** 1211-1221.
7. Hancock, D. B., Romieu, I., Shi, M., Sienra-Monge, J. J., Wu, H., Chiu, G. Y., Li, H., del Rio-Navarro, B. E., Willis-Owen, S. A., Weiss, S. T., Raby, B. A., Gao, H., Eng, C., Chapela, R., Burchard, E. G., Tang, H., Sullivan, P. F. and London, S. J.

(2009) Genome-wide association study implicates chromosome 9q21.31 as a susceptibility locus for asthma in mexican children. *PLoS Genet*. **5,** e1000623.
8. Kim, S. H., Cho, B. Y., Park, C. S., Shin, E. S., Cho, E. Y., Yang, E. M., Kim, C. W., Hong, C. S., Lee, J. E. and Park, H. S. (2009) Alpha-T-catenin (CTNNA3) gene was identified as a risk variant for toluene diisocyanate-induced asthma by genome-wide association analysis. *Clin. Exp. Allergy*. **39,** 203-212.
9. Himes, B. E., Hunninghake, G. M., Baurley, J. W., Rafaels, N. M., Sleiman, P., Strachan, D. P., Wilk, J. B., Willis-Owen, S. A., Klanderman, B., Lasky-Su, J., Lazarus, R., Murphy, A. J., Soto-Quiros, M. E., Avila, L., Beaty, T., Mathias, R. A., Ruczinski, I., Barnes, K. C., Celedon, J. C., Cookson, W. O., Gauderman, W. J., Gilliland, F. D., Hakonarson, H., Lange, C., Moffatt, M. F., O'Connor, G. T., Raby, B. A., Silverman, E. K. and Weiss, S. T. (2009) Genome-wide association analysis identifies PDE4D as an asthma-susceptibility gene. *Am. J. Hum. Genet*. **84,** 581-593.
10. Li, X., Howard, T. D., Zheng, S. L., Haselkorn, T., Peters, S. P., Meyers, D. A. and Bleecker, E. R. (2010) Genome-wide association study of asthma identifies RAD50-IL13 and HLA-DR/DQ regions. *J. Allergy Clin. Immunol*. **125**, 328-335, e311.
11. Sleiman, P. M., Flory, J., Imielinski, M., Bradfield, J. P., Annaiah, K., Willis-Owen, S. A., Wang, K., Rafaels, N. M., Michel, S., Bonnelykke, K., Zhang, H., Kim, C. E., Frackelton, E. C., Glessner, J. T., Hou, C., Otieno, F. G., Santa, E., Thomas, K., Smith, R. M., Glaberson, W. R., Garris, M., Chiavacci, R. M., Beaty, T. H., Ruczinski, I., Orange, J. M., Allen, J., Spergel, J. M., Grundmeier, R., Mathias, R. A., Christie, J. D., von Mutius, E., Cookson, W. O., Kabesch, M., Moffatt, M. F., Grunstein, M. M., Barnes, K. C., Devoto, M., Magnusson, M., Li, H., Grant, S. F., Bisgaard, H. and Hakonarson, H. (2010) Variants of DENND1B associated with asthma in children. *N. Engl. J. Med*. **362,** 36-44.
12. Ferreira, M. A., McRae, A. F., Medland, S. E., Nyholt, D. R., Gordon, S. D., Wright, M. J., Henders, A. K., Madden, P. A., Visscher, P. M., Wray, N. R., Heath, A. C., Montgomery, G. W., Duffy, D. L. and Martin, N. G. (2011) Association between ORMDL3, IL1RL1 and a deletion on chromosome 17q21 with asthma risk in Australia. *Eur. J. Hum. Genet*. **19,** 458-464.
13. Castro-Giner, F., Bustamante, M., Ramon Gonzalez, J., Kogevinas, M., Jarvis, D., Heinrich, J., Anto, J. M., Wjst, M., Estivill, X. and de Cid, R. (2009) A pooling-based genome-wide analysis identifies new potential candidate genes for atopy in the European Community Respiratory Health Survey (ECRHS). *BMC Med. Genet*. **10,** 128.
14. Esparza-Gordillo, J., Weidinger, S., Folster-Holst, R., Bauerfeind, A., Ruschendorf, F., Patone, G., Rohde, K., Marenholz, I., Schulz, F., Kerscher, T., Hubner, N., Wahn, U., Schreiber, S., Franke, A., Vogler, R., Heath, S., Baurecht, H., Novak, N., Rodriguez, E., Illig, T., Lee-Kirsch, M. A., Ciechanowicz, A., Kurek, M., Piskackova, T., Macek, M., Lee, Y. A. and Ruether, A. (2009) A common variant on chromosome 11q13 is associated with atopic dermatitis. *Nat. Genet*. **41,** 596-601.
15. Weidinger, S., Gieger, C., Rodriguez, E., Baurecht, H., Mempel, M., Klopp, N., Gohlke, H., Wagenpfeil, S., Ollert, M., Ring, J., Behrendt, H., Heinrich, J., Novak, N., Bieber, T., Kramer, U., Berdel, D., von Berg, A., Bauer, C. P., Herbarth, O., Koletzko, S., Prokisch, H., Mehta, D., Meitinger, T., Depner, M., von Mutius, E., Liang, L., Moffatt, M., Cookson, W., Kabesch, M., Wichmann, H. E. and Illig, T. (2008) Genome-wide scan on total serum IgE levels identifies FCER1A as novel susceptibility locus. *PLoS Genet*. **4,** e1000166.
16. Rothenberg, M. E., Spergel, J. M., Sherrill, J. D., Annaiah, K., Martin, L. J., Cianferoni, A., Gober, L., Kim, C., Glessner, J., Frackelton, E., Thomas, K., Blanchard, C., Liacouras, C., Verma, R., Aceves, S., Collins, M. H., Brown-Whitehorn, T., Putnam, P. E., Franciosi, J. P., Chiavacci, R. M., Grant, S. F., Abonia, J. P., Sleiman, P. M. and Hakonarson, H. (2010) Common variants at 5q22 associate with pediatric eosinophilic esophagitis. *Nat. Genet*. **42,** 289-291.
17. Gudbjartsson, D. F., Bjornsdottir, U. S., Halapi, E., Helgadottir, A., Sulem, P., Jonsdottir, G. M., Thorleifsson, G., Helgadottir, H., Steinthorsdottir, V., Stefansson, H., Williams, C., Hui, J., Beilby, J., Warrington, N. M., James, A., Palmer, L. J., Koppelman, G. H., Heinzmann, A., Krueger, M., Boezen, H. M., Wheatley, A., Altmuller, J., Shin, H. D., Uh, S. T., Cheong, H. S., Jonsdottir, B., Gislason, D., Park, C. S., Rasmussen, L. M., Porsbjerg, C., Hansen, J. W., Backer, V., Werge, T., Janson, C., Jonsson, U. B., Ng, M. C., Chan, J., So, W. Y., Ma, R., Shah, S. H., Granger, C. B., Quyyumi, A. A., Levey, A. I., Vaccarino, V., Reilly, M. P., Rader, D. J., Williams, M. J., van Rij, A. M., Jones, G. T., Trabetti, E., Malerba, G., Pignatti, P. F., Boner, A., Pescollderungg, L., Girelli, D., Olivieri, O., Martinelli, N., Ludviksson, B. R., Ludviksdottir, D., Eyjolfsson, G. I., Arnar, D., Thorgeirsson, G., Deichmann, K., Thompson, P. J., Wjst, M., Hall, I. P., Postma, D. S., Gislason, T., Gulcher, J., Kong, A., Jonsdottir, I., Thorsteinsdottir, U. and Stefansson, K. (2009) Sequence variants affecting eosinophil numbers associate with asthma and myocardial infarction. *Nat. Genet*. **41,** 342-347.
18. Barnes, K. C. (2010) Genomewide association studies in allergy and the influence of ethnicity. *Curr. Opin. Allergy Clin. Immunol*. **10,** 427-433.
19. Kamatani, Y., Matsuda, K., Okada, Y., Kubo, M., Hosono, N., Daigo, Y., Nakamura, Y. and Kamatani, N. (2010) Genome-wide association study of hematological and biochemical traits in a Japanese population. *Nat. Genet*. **42,** 210-215.
20. Kato, N., Takeuchi, F., Tabara, Y., Kelly, T. N., Go, M. J., Sim, X., Tay, W. T., Chen, C. H., Zhang, Y., Yamamoto, K., Katsuya, T., Yokota, M., Kim, Y. J., Ong, R. T., Nabika, T., Gu, D., Chang, L. C., Kokubo, Y., Huang, W., Ohnaka, K., Yamori, Y., Nakashima, E., Jaquish, C. E., Lee, J. Y., Seielstad, M., Isono, M., Hixson, J. E., Chen, Y. T., Miki, T., Zhou, X., Sugiyama, T., Jeon, J. P., Liu, J. J., Takayanagi, R., Kim, S. S., Aung, T., Sung, Y. J., Zhang, X., Wong, T. Y., Han, B. G., Kobayashi, S., Ogihara, T., Zhu, D., Iwai, N., Wu, J. Y., Teo, Y. Y., Tai, E. S., Cho, Y. S. and He, J. (2011) Meta-analysis of genome-wide association studies identifies common variants associated with blood pressure variation in east Asians. *Nat. Genet*. **43**, 531-538.
21. Liu, J. Z., McRae, A. F., Nyholt, D. R., Medland, S. E., Wray, N. R., Brown, K. M., Hayward, N. K., Montgomery, G. W., Visscher, P. M., Martin, N. G. and Macgregor, S. (2010) A versatile gene-based test for genome-wide association studies. *Am. J. Hum. Genet*. **87**, 139-145.
22. Soranzo, N., Spector, T. D., Mangino, M., Kuhnel, B., Rendon, A., Teumer, A., Willenborg, C., Wright, B., Chen, L., Li, M., Salo, P., Voight, B. F., Burns, P., Laskowski, R. A., Xue, Y., Menzel,

S., Altshuler, D., Bradley, J. R., Bumpstead, S., Burnett, M. S., Devaney, J., Doring, A., Elosua, R., Epstein, S. E., Erber, W., Falchi, M., Garner, S. F., Ghori, M. J., Goodall, A. H., Gwilliam, R., Hakonarson, H. H., Hall, A. S., Hammond, N., Hengstenberg, C., Illig, T., Konig, I. R., Knouff, C. W., McPherson, R., Melander, O., Mooser, V., Nauck, M., Nieminen, M. S., O'Donnell, C. J., Peltonen, L., Potter, S. C., Prokisch, H., Rader, D. J., Rice, C. M., Roberts, R., Salomaa, V., Sambrook, J., Schreiber, S., Schunkert, H., Schwartz, S. M., Serbanovic-Canic, J., Sinisalo, J., Siscovick, D. S., Stark, K., Surakka, I., Stephens, J., Thompson, J. R., Volker, U., Volzke, H., Watkins, N. A., Wells, G. A., Wichmann, H. E., Van Heel, D. A., Tyler-Smith, C., Thein, S. L., Kathiresan, S., Perola, M., Reilly, M. P., Stewart, A. F., Erdmann, J., Samani, N. J., Meisinger, C., Greinacher, A., Deloukas, P., Ouwehand, W. H. and Gieger, C. (2009) A genome-wide meta-analysis identifies 22 loci associated with eight hematological parameters in the HaemGen consortium. *Nat. Genet.* **41**, 1182-1190.

23. Cho, Y. S., Go, M. J., Kim, Y. J., Heo, J. Y., Oh, J. H., Ban, H. J., Yoon, D., Lee, M. H., Kim, D. J., Park, M., Cha, S. H., Kim, J. W., Han, B. G., Min, H., Ahn, Y., Park, M. S., Han, H. R., Jang, H. Y., Cho, E. Y., Lee, J. E., Cho, N. H., Shin, C., Park, T., Park, J. W., Lee, J. K., Cardon, L., Clarke, G., McCarthy, M. I., Lee, J. Y., Oh, B. and Kim, H. L. (2009) A large-scale genome-wide association study of Asian populations uncovers genetic factors influencing eight quantitative traits. *Nat. Genet.* **41**, 527-534.
24. Hindorff, L. A., Sethupathy, P., Junkins, H. A., Ramos, E. M., Mehta, J. P., Collins, F. S. and Manolio, T. A. (2009) Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 9362-9367.
25. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., Maller, J., Sklar, P., de Bakker, P. I., Daly, M. J. and Sham, P. C. (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559-575.
26. Marchini, J., Howie, B., Myers, S., McVean, G. and Donnelly, P. (2007) A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat. Genet.* **39**, 906-913.
27. Barrett, J. C., Fry, B., Maller, J. and Daly, M. J. (2005) Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* **21**, 263-265.
28. Voight, B. F., Kudaravalli, S., Wen, X. and Pritchard, J. K. (2006) A map of recent positive selection in the human genome. *PLoS Biol.* **4**, e72.