Allergy Asthma Immunol Res. 2020 Sep;12(5):806-820
https://doi.org/10.4168/aair.2020.12.5.806
pISSN 2092-7355·eISSN 2092-7363

Allergy, Asthma & Immunology Research AAIR

Original Article

Phenotypic and Molecular Characterization of Risk Loci Associated With Asthma and Lung Function

Mehmet Karaca,[1*] **Nazente Atceken**,[2] **Şefayet Karaca**,[3] **Ersoy Civelek**,[4] **Bülent E. Şekerel**,[5] **Renato Polimanti**[6]

[1]Department of Biology, Faculty of Science and Arts, Aksaray University, Aksaray, Turkey
[2]Graduate School of Natural and Applied Sciences, Aksaray University, Aksaray, Turkey
[3]Department of Nutrition and Dietetics, Faculty of Health Science, Aksaray University, Aksaray, Turkey
[4]Pediatric Allergy and Immunology Clinic, Ankara Child Health and Diseases Hematology Oncology Research Hospital, Ankara, Turkey
[5]Pediatric Allergy and Asthma Unit, Faculty of Medicine, Hacettepe University, Ankara, Turkey
[6]Department of Psychiatry, Yale University School of Medicine and VA CT Healthcare Center, West Haven, CT, United States

Received: Jan 8, 2020
Revised: Feb 25, 2020
Accepted: Feb 27, 2020

Correspondence to
Mehmet Karaca, PhD
Department of Biology, Faculty of Science and Arts, Aksaray University, Bahçesaray District, 68100 Aksaray, Turkey.
Tel: +90-382-2882227
Fax: +90-382-2882299
E-mail: mkaraca@aksaray.edu.tr

ORCID iDs
Mehmet Karaca
https://orcid.org/0000-0002-7688-4223
Nazente Atceken
https://orcid.org/0000-0002-2020-9659
Şefayet Karaca
https://orcid.org/0000-0002-3645-3875
Ersoy Civelek
https://orcid.org/0000-0002-1780-4801
Bülent E. Şekerel
https://orcid.org/0000-0003-4232-3396
Renato Polimanti
https://orcid.org/0000-0003-0745-6046

ABSTRACT

Purpose: Respiratory diseases have a highly multifactorial etiology where different mechanisms contribute to the individual's susceptibility. We conducted a deep characterization of loci associated with asthma and lung function by previous genome-wide association studies (GWAS).
Methods: Sixteen variants were selected from previous GWAS of childhood/adult asthma and pulmonary function tests. We conducted a phenome-wide association study of these loci in 4,083 traits assessed in the UK Biobank (n = 361,194 participants). Data from the Genotype-Tissue Expression (GTEx) project were used to conduct a transcriptomic analysis with respect to tissues relevant for asthma pathogenesis. A pediatric cohort assessed with the International Study of Asthma and Allergies in Children (ISAAC) Phase II tools was used to further explore the association of these variants with 116 traits related to asthma comorbidities.
Results: Our phenome-wide association studies (PheWAS) identified 206 phenotypic associations with respect to the 16 variants identified. In addition to the replication of the phenotypes tested in the discovery GWAS, we observed novel associations related to blood levels of immune cells (eosinophils, neutrophils, monocytes, and lymphocytes) for the asthma-related variants. Conversely, the lung-function variants were associated with phenotypes related to body fat mass. In the ISAAC-assessed cohort, we observed that risk alleles associated with increased fat mass can exacerbate allergic reactions in individuals affected by allergic respiratory diseases. The GTEx-based analysis showed that the variants tested affect the transcriptomic regulation of multiple surrounding genes across several tissues.
Conclusions: This study generated novel data regarding the genetics of respiratory diseases and their comorbidities, providing a deep characterization of loci associated with asthma and lung function.

Keywords: Allergy; asthma; genetics; lung function; transcriptome; loci

Disclosure

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

INTRODUCTION

Respiratory diseases are a major cause of mortality and morbidity worldwide, and children are particularly vulnerable to them.[1] Understanding the molecular processes responsible for respiratory diseases can consistently improve our ability to prevent and treat the disorders, reducing their harmful consequences. Genetic studies have demonstrated that pediatric respiratory diseases and lung function are heritable and several risk variants have been identified by genome-wide association studies (GWAS).[2] Understanding the mechanisms underlying the genetic associations identified remains a challenge. Genetic variants can affect the regulation of multiple genes and be associated with multiple phenotypes via pleiotropic mechanisms including phenotypic causal relationships or shared biological pathways.[3] To date, numerous methods and datasets are available to deepen the biological meaning of genetic associations identified in large GWAS.[4] Information regarding tissue-specific transcriptomic regulation can be used to disentangle which are the genes and the tissues involved in the disease pathogenesis.[5] Phenome-wide association studies (PheWAS) can permit us to detect the full phenotypic spectrum associated with a specific genetic variant and to extricate the possible causal relationships linking the genetic variation to the phenotype of interest.[6,7]

In the present study, we conducted phenomic and transcriptomic analyses with respect to 16 variants (**Table 1**) identified by GWAS of childhood/adult asthma and forced vital capacity.[8-10] Four of the variants investigated were already identified by GWAS of autoimmune diseases including rheumatoid arthritis,[11] ulcerative colitis,[12] eczema,[13] juvenile idiopathic arthritis,[14] and systemic sclerosis.[15] To fully understand the phenotypic spectrum associated with the investigated variants, we conducted a PheWAS including 4,083 phenotypic traits assessed in 361,194 participants from UK Biobank.[16] We leveraged data from the Genotype-Tissue Expression (GTEx) project, which includes expression quantitative trait locus (eQTL) information with respect to 49 tissue sites,[17] to identify changes induced by these variants in tissue-specific transcriptomic regulation. Finally, we tested associations of the investigated variants with phenotypic traits related to pediatric respiratory diseases and their comorbidities derived from a cohort of 919 Turkish children using International Study of Asthma and Allergies in Children (ISAAC) Phase II tools.[18] Through these multiple

Table 1. Variants associated with asthma and lung function

rsID	CHR	BP	Primary phenotype	Additional phenotypes	References
rs1558641	2	102765865	Asthma (childhood onset)	-	8
rs6871536	5	131969874		-	8
rs6967330	7	105658451		-	8
rs928413	9	6213387		Lung function (FEV1/FVC)	8, 13
rs3771166	2	102986222	Asthma (adult onset)	-	9
rs1342326	9	6190076		Eczema	9, 13
rs744910	15	67446785		-	9
rs2284033	22	37534034		Juvenile idiopathic arthritis	9, 14
rs2305480	17	38062196	Asthma (childhood/adult onset)	Rheumatoid arthritis; Ulcerative colitis	8, 9, 11, 12
rs3894194	17	38121993		Systemic sclerosis	8, 9, 15
rs1430193	2	56120853	Lung function (FVC)	-	10
rs6923462	6	7801112		-	10
rs4237643	11	43648368		-	10
rs2863171	11	45250732		-	10
rs1079572	16	78187138		-	10
rs6501431	17	68976415		-	10

CHR, chromosome; BP, base-pair position; FVC, forced vital capacity; FEV1, forced expiratory volume in 1 second.

https://doi.org/10.4168/aair.2020.12.5.806

resources, we provided a deep phenotypic and molecular characterization of risk loci, unraveling the mechanisms by which these genetic variants are associated with asthma and lung function.

MATERIALS AND METHODS

UK biobank

The UK Biobank cohort is an open-access resource available to investigate a wide range of severe and life-threatening illnesses and also normal-range traits.[16] This project has recruited more than 500,000 people assessed through detailed web-based questionnaires on their diet, cognitive function, work history, health status, and other relevant phenotypes. In our analysis, we considered GWAS datasets related to 361,194 unrelated participants of European descent. These data were generated testing 4,083 phenotypes using regression models available in Hail (available at https://github.com/hail-is/hail) including the first 20 ancestry principal components, sex, age, age,^{2} sex×age, and sex×age^2 as covariates. Details regarding QC criteria, GWAS methods, and the GWAS data are available at https://github.com/Nealelab/UK_Biobank_GWAS/tree/master/imputed-v2-gwas. Based on these data, we conducted a phenome-wide investigation, applying a Bonferroni multiple testing correction accounting for the number of variants (n = 16) and phenotypes tested (n = 4,083).

To verify that the associations observed were not false-positive results, we conducted phenome-wide scans with respect to randomly selected variants matched to the variants investigated on the basis of 4 genomic characteristics: i) minor allele frequency, ii) linkage disequilibrium (LD) proxies, iii) distance to the nearest gene, and iv) gene density. This control set of matched variants was selected using the SNPsnap tool[19] according to the following parameters: 1,000 Genomes European reference panel (which is the closest reference panel among those available in SNPsnap); LD distance cutoff of R^2 = 0.05; ± 5% point deviation; ± 50% of gene density relative deviation; ± 50% of relative deviation of the distance to nearest gene; and ± 50% of relative deviation of LD proxies. For each variant investigated (**Table 1**), we extracted up to 100 matched single-nucleotide polymorphism (SNP), excluding human leucocyte antigen region due to its complex LD structure.

GTEx

GTEx version 8 (V8) data were used to analyze the effect of alleles investigated on the expression of the surrounding genes (± 1 Mb of the gene transcription starting site) across human tissues.[17] The association between genetic variants and gene expression was conducted using a linear regression analysis considering the covariates (*i.e.*, top-3 ancestry principal components, genotyping array platform, Probabilistic Estimation of Expression Residuals factors, and sex) applied to gene-cis-SNP pairs using FastQTL[20] and assuming an additive model. A detailed description of the GTEx methods (*i.e.*, preprocessing, expression quantification, and association analysis) used is available at https://gtexportal.org/home/documentationPage#staticTextAnalysisMethods. In our analysis, we initially considered single-tissue eQTL associations surviving genome-wide multiple testing correction (false discovery rate; FDR, q ≤ 0.05). To systematically evaluate the effect across multiple tissues, we considered multi-tissue data calculated using Meta-Tissue.[21] This meta-analytic approach permits us to calculate a posterior probability (m value) that an effect exists in each of the tissues tested. M values > 0.9 were considered to indicate whether the tissue was predicted to show the eQTL association.

https://doi.org/10.4168/aair.2020.12.5.806

ISAAC cohort

The subjects studied (n = 919) were selected from a cross-sectional study conducted using the ISAAC Phase II tools on schoolchildren from 5 city centers (Van, Manisa, Ankara, Antalya, and Trabzon) located in 5 regions of Turkey (East, Aegean, Central Anatolian, Mediterranean, and Black Sea regions, respectively) between September 15, 2005 and May 30, 2006.[18] The procedures used in this study conform to the tenets of the Declaration of Helsinki and received approval from the Ethical Review Board of Ankara Child Health and Diseases Research Hospital (2017-041). Written parental and student consent was obtained separately for each participant. The study was approved by the appropriate ethics committees: the Ankara Child Health and Diseases Research Hospital, the Turkish Ministry of Health, and the central and provincial directors of the Ministry of Education and city governors. Further details regarding the sampling and phenotype assessments, and comorbidity and phenotypic correlation in this cohort are available in previous publications.[7,22,23] Genotyping was performed using the Kompetitive Allele-Specific PCR (KASP) technique.[24] DNA 5–10 ng were used per well, and polymerase chain reaction (PCR) reactions were carried out in the presence of no-template controls (NTCs) at a 10-μL final volume (GeneAmp PCR System 9700, Applied Biosystems). The KASP master mix, assay mix, and cycling conditions were based on manufacturers' protocols (available at http://www.kbioscience.co.uk). KlusterCaller™ software was used to evaluate genotyping data. In the analysis of the ISAAC cohort, Plink 1.07 tool[25] was used to implement logistic and linear regression analyses for the associations between genetic variants (additive model) and phenotypic traits (binary and quantitative, respectively). Quantitative traits were normalized using appropriate Box-Cox power transformations before being entered into the analysis. We adjusted the association analysis according to the following covariates: age, sex, and sampling center. Due to the small sample size, since the ISAAC cohort was used for final external validation of the variants that were surviving Bonferroni multiple testing correction in both UK Biobank and GTEx analyses, we considered a nominal threshold ($P < 0.05$) as a significance criterion.

Data availability

Data supporting the findings of this study are available within this article and its additional files. UK Biobank GWAS summary association data are available at https://github.com/Nealelab/UK_Biobank_GWAS/tree/master/imputed-v2-gwas. GTEx data are available at https://gtexportal.org/home/.

RESULTS

In the UK Biobank PheWAS, we identified 206 genetic associations surviving Bonferroni multiple testing correction ($P < 7.65 \times 10^{-7}$; **Supplementary Table S1**). While asthma risk alleles showed the strongest associations with traits included "medical information" category (**Fig. 1A**), variants associated with lung function showed the strongest association with phenotypes included in the physical characteristics (**Fig. 1B**). Additionally, asthma risk alleles were associated with a larger effect size than those observed for lung-function associated loci. The blood levels of immune system cells showed the strongest significance with the asthma risk alleles: eosinophil percentage (rs928413, $P = 7.61 \times 10^{-264}$), neutrophil count (rs3894194, $P = 4.26 \times 10^{-197}$), leukocyte count (rs3894194, $P = 8.53 \times 10^{-171}$), monocyte percentage (rs3894194, $P = 7.36 \times 10^{-75}$), and lymphocyte percentage (rs3894194, $P = 1.47 \times 10^{-57}$). The variants associated with blood immune cell levels were identified by the asthma GWAS,[8,9] but not by the lung function GWAS.[10] Among the lung function-associated variants,

https://doi.org/10.4168/aair.2020.12.5.806

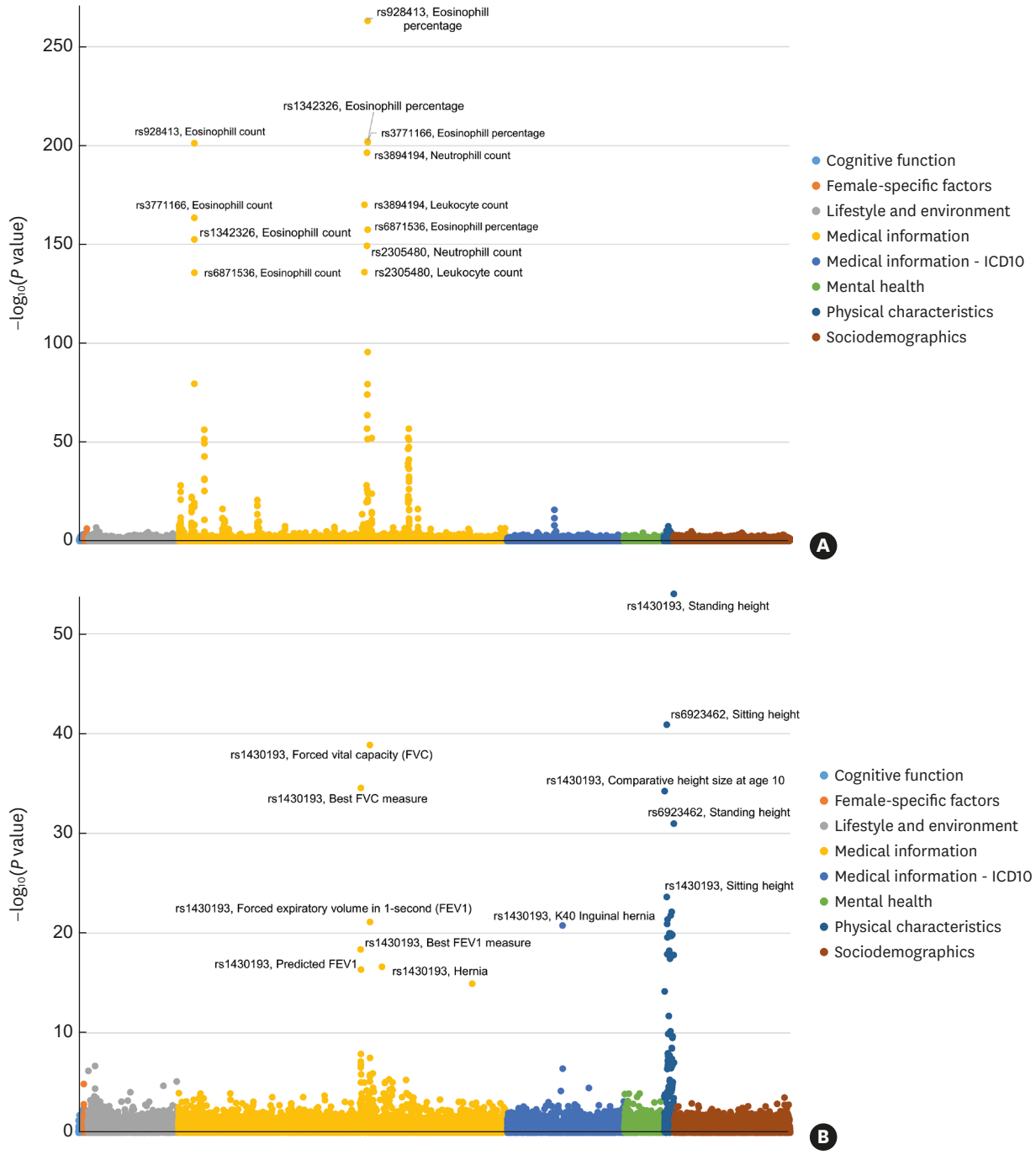

Fig. 1. Manhattan plot of the phenome-wide scan conducted in the UK Biobank with respect to asthma risk alleles (A) and lung function-associated variants (B). As indicated in the legend, the phenotypic categories are color coded.
FVC, forced vital capacity; FEV1, forced expiratory volume in 1 second; ICD, International Statistical Classification of Diseases and Related Health Problems.

rs4237643 showed the largest number of phenotypic associations (n = 33). This variant was identified by the lung function GWAS and, in the UK Biobank cohort, we observed a replication for these phenotypes (forced vital capacity [FVC], $P = 1.28 \times 10^{-8}$; forced expiratory volume in 1 second [FEV1], $P = 2.83 \times 10^{-7}$). However, rs4237643 showed much stronger associations with phenotypes related to fat mass distribution such as arm fat mass ($P = 7.16$

https://doi.org/10.4168/aair.2020.12.5.806

× 10^{-23}), whole-body fat mass ($P = 4.17 \times 10^{-22}$), body mass index ($P = 2.54 \times 10^{-21}$), and leg fat mass ($P = 1.02 \times 10^{-20}$). A similar phenotypic spectrum (*i.e.*, associations with body mass phenotypes, but not with respiratory diseases or blood immune cell levels) was also observed with respect to the other variants identified by lung-function GWAS.

Beyond these associations, we also observed additional findings that may be related to a broader effect of these variants across human health and disease domains (**Table 1**). Multiple variants were associated with basal metabolic rate (rs1430193, $P = 2.39 \times 10^{-8}$; rs4237643, $P = 1.24 \times 10^{-10}$; rs6923462, $P = 3.56 \times 10^{-7}$), number of self-reported non-cancer illnesses (rs1558641, $P = 3.09 \times 10^{-8}$; rs6871536, $P = 9.2 \times 10^{-9}$; rs928413, $P = 3.92 \times 10^{-7}$), and dietary habits (cheese intake: rs4237643, $P = 2 \times 10^{-7}$; salt added to food: rs3894194, $P = 1.56 \times 10^{-7}$). Additional associations were observed for rs1430193 with multiple hernia phenotypes (hernia $P = 1.14 \times 10^{-15}$; inguinal hernia $P = 1.67 \times 10^{-21}$; femoral hernia $P = 3.79 \times 10^{-7}$), for rs744910 with corneal resistance factor ($P = 1.48 \times 10^{-7}$), for rs2305480 with age at menopause ($P = 5.74 \times 10^{-7}$), and for rs2305480 with mouth ulcers ($P = 4.01 \times 10^{-7}$) (**Table 2**).

Our control analysis conducted with respect to randomly selected variants matched to the asthma and lung-function associated variants for multiple genomic features showed that the empirical probability to observe a phenome-wide association surviving multiple testing correction by chance is 0.2%.

Considering single-tissue eQTL, we observed that 14 out of 16 investigated variants are a genome-wide significant eQTL at least with respect to a gene in a tissue. A total of 385 eQTLs reached genome-wide multiple testing correction (FDR, q ≤ 0.05; **Supplementary Table S2**). Rs1342326 and rs6923462 were the only 2 variants that did not show significant effects on gene expression. Among the significant eQTLs, we observed that the asthma risk alleles tend to be eQTLs for multiple genes across several tissues, while variants associated with lung function are eQTLs for a reduced number of genes and tissues (**Fig. 2A**). Across the tissues and the genes tested with respect to each variant, we observed a wide range of normalized effect sizes (NES; **Fig. 2B**). Considering the 2 tissues that expected to be more relevant for asthma pathogenesis and inter-individual variability of lung function, 7 variants showed genome-wide significance with multiple genes in whole blood and lung tissue. The strongest eQTL association was observed for rs2305480 with respect to *GSDMB* gene expression in whole blood (G allele,

Table 2. Broader phenotypic associations with asthma and lung-function risk variants

rsID	Effect allele	Beta	*P* value	Description
rs1430193	T	−0.01	2.39E-08	Basal metabolic rate
		0.00	1.67E-21	Diagnoses - main ICD10: K40 Inguinal hernia
		0.00	3.79E-07	Diagnoses - main ICD10: K41 Femoral hernia
		−0.01	1.14E-15	Hernia
rs1558641	A	−0.01	3.09E-08	Number of self-reported non-cancer illnesses
rs2305480	A	0.02	5.74E-07	Age at menopause (last menstrual period)
		0.00	4.01E-07	Mouth/teeth dental problems: Mouth ulcers
rs3894194	A	0.01	1.56E-07	Salt added to food
rs4237643	G	−0.01	1.24E-10	Basal metabolic rate
		0.01	2.00E-07	Cheese intake
		0.01	6.20E-07	Usual walking pace
rs6871536	C	0.01	9.20E-09	Number of self-reported non-cancer illnesses
rs6923462	C	0.01	3.56E-07	Basal metabolic rate
rs744910	A	0.03	1.48E-07	Corneal resistance factor
rs928413	A	−0.01	3.92E-07	Number of self-reported non-cancer illnesses

ICD, International Statistical Classification of Diseases and Related Health Problems.

https://doi.org/10.4168/aair.2020.12.5.806

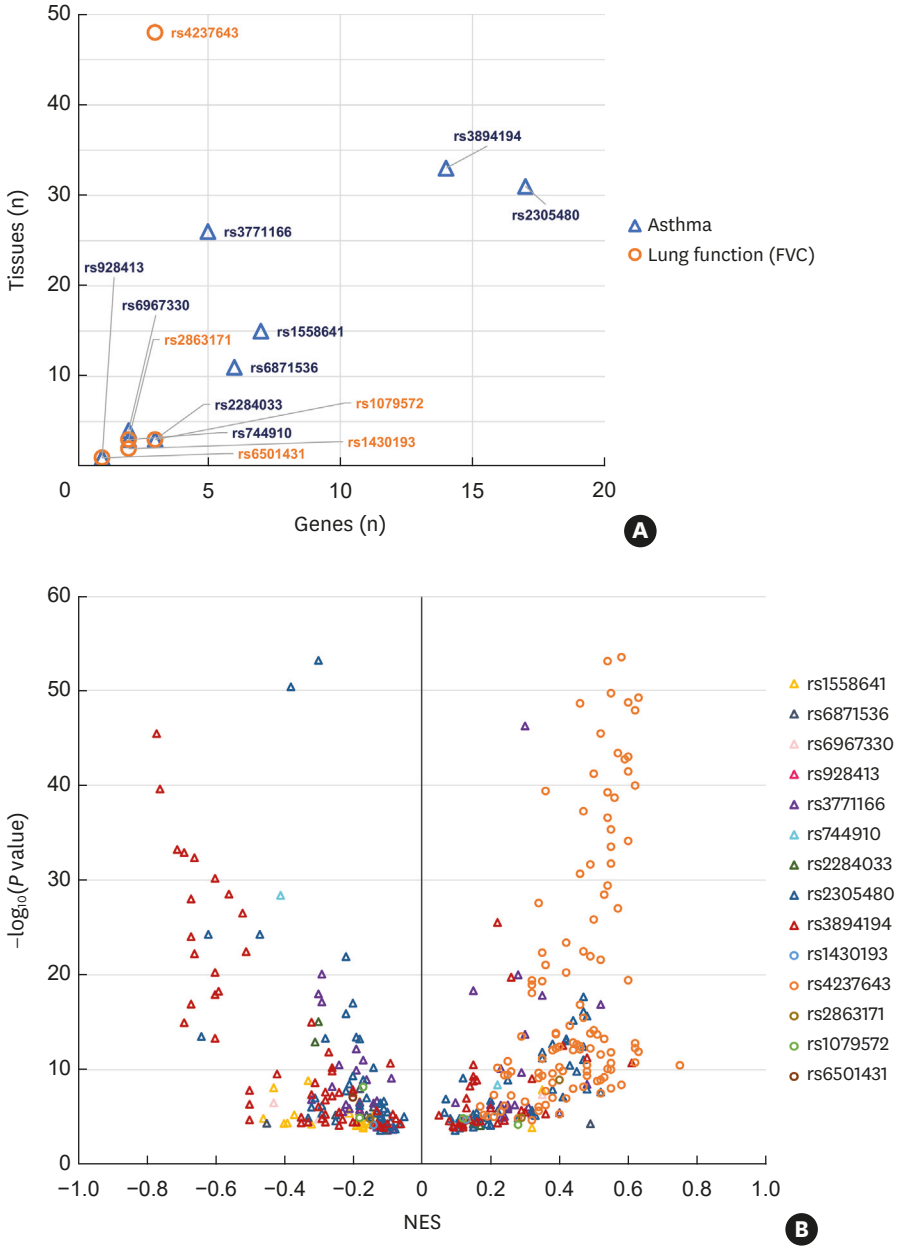

Fig. 2. eQTL analysis of asthma risk alleles (triangles) and lung function-associated variants (circles). (A) The relationship between the number of tissues and genes regulated for each locus investigated. (B) Relationship between normalized effect size and statistical significance for each locus investigated.
eQTL, expression quantitative trait locus; FVC, forced vital capacity; NES, normalized effect sizes.

NES = −0.3, $P = 7.1 \times 10^{-54}$; **Fig. 3A**). Although it showed a weaker association, rs2305480-*GSDMB* eQTL was also genome-wide significant in lung tissue (NES = −0.18, $P = 5.7 \times 10^{-14}$). Considering the multi-tissue analysis, rs2305480-*GSDMB* eQTL (multi-tissue $P = 8.32 \times 10^{-152}$; **Supplementary Fig. S1**) showed a posterior probability greater than 90% in additional tissues including spleen (m = 1), Epstein-Barr virus-transformed lymphocytes (m = 1), small intestine (m = 1), colon transverse (m = 1), and stomach (m = 0.99). Beyond *GSDMB*, rs2305480 affects the transcriptomic regulation of additional genes, including *ORMDL3* (whole blood: NES = −0.38, $P = 4.4 \times 10^{-51}$; lung: NES = −0.22, $P = 1.2 \times 10^{-16}$; multi-tissue eQTL $P = 2.35 \times 10^{-154}$), *GSDMA*

https://doi.org/10.4168/aair.2020.12.5.806

(lung: NES = 0.44, $P = 6.3 \times 10^{-16}$; multi-tissue eQTL $P = 8.44 \times 10^{-217}$), *IKZF3* (whole blood: NES = 0.07, $P = 1.2 \times 10^{-7}$; multi-tissue eQTL $P = 3.83 \times 10^{-13}$), and *PGAP3* (lung: NES = −0.12, $P = 3.2 \times 10^{-7}$; multi-tissue eQTL $P = 9.2 \times 10^{-113}$). A similar eQTL pattern was also observed for rs3894194, which is partially in LD and closely located to rs2305480. On chromosome 11, rs4237643 regulated the transcriptomic of *HSD17B12* in both whole blood (T allele, NES = 0.46, $P = 2.4 \times 10^{-49}$; **Fig. 3B**) and lung (NES = 0.52, $P = 3.6 \times 10^{-46}$). The multi-tissue eQTL analysis showed that rs4237643-*HSD17B12* association has a posterior probability greater than 90% in 48 tissues available in GTEx V8 (multi-tissue eQTL $P < 10^{-300}$; **Supplementary Fig. S2**). Rs4237643 also showed regulatory effects on 2 long non-coding RNAs, RP11-613D13.10 (lung: NES = 0.6, $P = 3.7 \times 10^{-42}$; whole blood: NES = 0.35, $P = 4.8 \times 10^{-20}$; multi-tissue eQTL $P < 10^{-300}$) and RP11-613D13.5 (lung: NES = 0.25, $P = 1.3 \times 10^{-11}$; multi-tissue eQTL $P = 5.84 \times 10^{-155}$). Rs1558641 affects the transcriptomic regulation of *MFSD9* in whole blood (G allele; NES = 0.35, $P = 1.3 \times 10^{-8}$; **Fig. 3C**) and *IL18R1* in lung (NES = −0.18, $P = 5.9 \times 10^{-5}$). Although it reached significance in the multi-tissue eQTL analysis ($P = 7.87 \times 10^{-7}$), whole blood is the only tissue with a posterior probability greater than 90% in the rs1558641-*MFSD9* eQTL (m = 1). Conversely, rs1558641-*IL18R1* association (multi-tissue eQTL $P = 2.35 \times 10^{-44}$) showed a posterior probability greater than 90% in lung (m = 1) together with 22 additional tissues available in GTEx (**Supplementary Fig. S3**). A similar eQTL pattern was observed for rs3771166, which is closely located and in partial LD with rs1558641 (distance: 220 kb, LD $r^2 = 0.13$ in the European population of the 1,000 Genomes Project reference panel).[26] Rs744910 showed genome-wide significant effect on *AAGAB* transcriptomic regulation in whole blood (G allele; NES = 0.1, $P = 4.9 \times 10^{-5}$; **Fig. 3D**). The multi-tissue rs744910-*AAGAB* eQTL analysis showed that only whole blood and esophagus mucosa have a posterior probability greater than 90% (multi-tissue eQTL $P = 7.92 \times 10^{-6}$; whole blood m = 0.98; esophagus mucosa m = 1; **Supplementary Fig. S4**). Similarly, rs2284033 affects *ELFN2* transcriptome profile in lung (G allele; NES = 0.15, $P = 4.6 \times 10^{-5}$; **Fig. 3E**) and the multi-tissue eQTL analysis showed that this association is specific for lung tissue (lung m = 0.90; **Supplementary Fig. S5**).

Considering the 11 variants that showed significant eQTL effects and significant associations in the phenome-wide scan, we tested their association with 116 phenotypic traits (**Supplementary Table S3**) derived from a pediatric cohort (n = 919) assessed using ISAAC Phase II tools. We observed 60 genetic associations with respect to 10 of the variants investigated ($P < 0.05$; **Table 3**). Comparing asthma risk alleles with respect to lung-function associated variants, we observed that risk alleles are associated with a larger number of

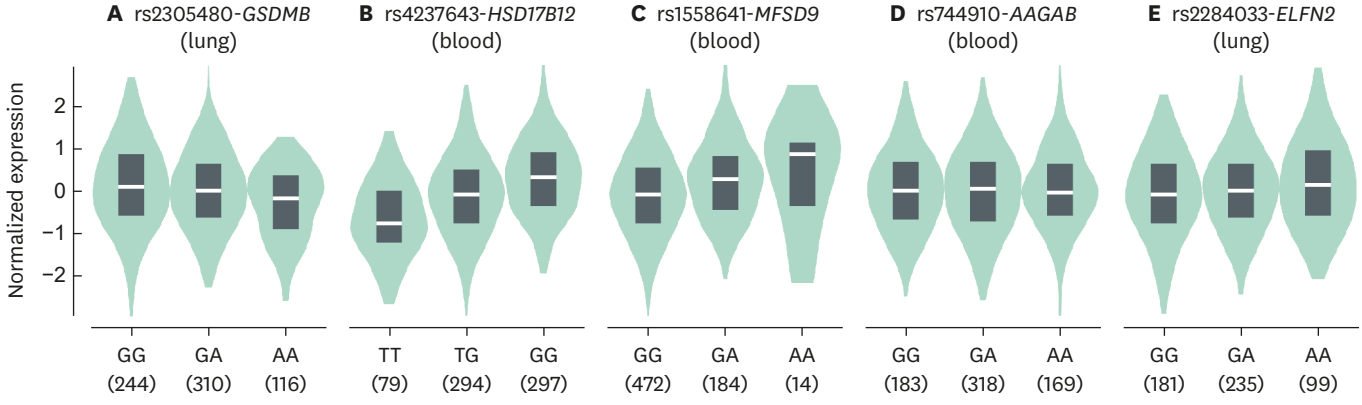

Fig. 3. Most significant eQTLs with respect to whole blood and lung tissue.
eQTL, expression quantitative trait locus.

https://doi.org/10.4168/aair.2020.12.5.806

ISAAC-derived phenotypes thanks (6.5 ± 1.41 vs. 2.67 ± 2.3, respectively). The strongest findings were observed with respect to rs3894194 (an asthma risk allele), which was associated with eight different phenotypes and showed an association with increased risk of

Table 3. Associations of asthma and lung-function risk variants with supporting evidence from phenome-wide and eQTL analyses in the pediatric cohort assessed using ISAAC Phase II tools

SNP	Effect allele	Phenotype	Z score	*P* value
rs1558641	A	Asthmatic bronchitis diagnosis at least twice	3.33	8.68E-04
		At least 32 wheezing episodes	−3.30	9.73E-04
		Allergic bronchitis diagnosis at least once	3.29	1.01E-03
		Ever allergic bronchitis diagnosis	3.23	1.30E-03
		Asthmatic bronchitis diagnosis at least twice	−3.01	2.58E-03
rs2284033	A	Ever asthmatic bronchitis diagnosis	3.01	2.66E-03
		Asthmatic bronchitis diagnosis at least once	2.98	2.88E-03
		Current wheezing and asthma diagnosis at least once	2.95	3.18E-03
		Current Eczema	2.94	3.32E-03
		Olea atopy	2.91	3.58E-03
		Absolute lymphocyte count	−2.85	4.59E-03
		Serum IgE level	−2.84	4.62E-03
rs2305480	A	Serum IgE level	−2.82	4.97E-03
		% neutrophils	−2.77	5.70E-03
		Wheezing severity	−2.70	7.02E-03
		Asthma diagnosis at least twice	2.67	7.61E-03
		% FEV120 compared to healthy individuals	2.63	8.68E-03
		% eosinophils	−2.62	9.10E-03
		Rhinitis severity	2.55	1.09E-02
rs2863171	C	% FEV post compared to healthy individuals	−2.45	1.45E-02
		Asthmatic bronchitis diagnosis at least twice	−2.45	1.45E-02
		Current wheezing, BHR positive, and asthma diagnosis at least once	2.42	1.53E-02
		Horse atopy	2.42	1.55E-02
rs3894194	A	Horse atopy	2.40	1.62E-02
		Asthma diagnosis at least once	2.39	1.69E-02
		Allergic bronchitis diagnosis at least twice	2.38	1.72E-02
		Ever eczema	2.38	1.73E-02
		Ever asthma diagnosis	2.38	1.77E-02
		Diagnosis of asthma, asthmatic bronchitis, allergic bronchitis, or bronchitis at least once	2.37	1.77E-02
		Serum IgE level	−2.36	1.84E-02
		Hematocrit level	−2.36	1.86E-02
		Asthma, asthmatic bronchitis, or allergic bronchitis diagnosis	−2.33	1.98E-02
		% MEF240 compared to healthy individuals	−2.29	2.21E-02
		Flexural dermatitis	−2.29	2.23E-02
		% FEV480 compared to healthy individuals	2.29	2.26E-02
rs4237643	G	Current wheezing or allergic rhino-conjunctivitis symptoms	−2.22	2.68E-02
		Asthma diagnosis at least once	−2.19	2.87E-02
		Cat atopy	2.15	3.14E-02
		Olea atopy	−2.15	3.19E-02
rs6871536	C	Allergic bronchitis diagnosis at least once	2.15	3.20E-02
		Current eczema with flexural dermatitis	2.14	3.20E-02
		% FEV60 compared to healthy individuals	2.15	3.21E-02
		Current wheezing and BHR positive	−2.14	3.28E-02
		Rhino-conjunctivitis	−2.11	3.49E-02
		Absolute neutrophil count	−2.11	3.53E-02
		Allergic bronchitis diagnosis at least once	−2.10	3.61E-02
		> 4% eosinophils	2.09	3.71E-02
rs744910	A	Asthma, asthmatic bronchitis, or allergic bronchitis diagnosis	−2.05	4.02E-02
		% lymphocytes	−2.03	4.26E-02
		Wheezing and eczema	2.02	4.32E-02
		Wheezing and allergic rhinitis	2.02	4.33E-02
		Absolute eosinophil count	−2.02	4.36E-02
		% eosinophils	−2.02	4.39E-02

(continued to the next page)

Table 3. (Continued) Associations of asthma and lung-function risk variants with supporting evidence from phenome-wide and eQTL analyses in the pediatric cohort assessed using ISAAC Phase II tools

SNP	Effect allele	Phenotype	Z score	*P* value
rs928413	A	Diagnosis of asthma, asthmatic bronchitis, allergic bronchitis, or bronchitis at least once	−2.01	4.40E-02
		FEV after bronchodilator drug inhalation subsequently to BTC	−2.01	4.53E-02
		Waking up with wheezing	−1.98	4.75E-02
		Atopic wheezing (Y/N)	−1.98	4.78E-02
		Current eczema	1.97	4.89E-02
		Cockroach atopy	−1.97	4.92E-02
		Waking up with wheezing	1.97	4.94E-02

eQTL, expression quantitative trait loci; ISAAC, International Study of Asthma and Allergies in Children; SNP, single-nucleotide polymorphism; IgE, immunoglobulin E; FEV, forced expiratory volume; BHR, bronchial hyperresponsiveness; MEF, mean expiratory flow; BTC, biliary tract cancer.

being diagnosed with asthma bronchitis at least twice ($P = 8.68 \times 10^{-4}$). Beyond other asthma definitions and comorbidities, rs3894194 was associated with reactions to the skin prick test for cat and horse atopy ($P = 0.031$ and $P = 0.015$, respectively). With respect to the only lung-function associated locus, rs4237643 was associated with four ISAAC-derived phenotypes, including being diagnosed with allergic bronchitis at least twice ($P = 0.017$), mean expiratory flow (FEV) after 480s of hypertonic saline inhalation ($P = 0.023$), absolute neutrophil count ($P = 0.035$), and FEV after bronchodilator drug inhalation subsequently to bronchial challenge test (BCT; $P = 0.045$).

DISCUSSION

Asthma and comorbid respiratory diseases have a highly multifactorial etiology where different pathogenic mechanisms contribute to the individual's susceptibility.[27] In the present study, we explored the phenotypic and molecular spectrum of risk loci associated with asthma and lung function. There is a well-known relationship between respiratory diseases and lung function[28] also including altered lung function in asthmatic patients.[29] However, there is limited knowledge regarding molecular mechanisms linking these phenotypic traits. We leveraged multiple resources to provide a deep characterization of the loci investigated, and verified similarities and differences among them. We leveraged 3 different datasets: i) UK Biobank, a large-scale cohort representative of the characteristics of the general population; ii) GTEx, a dataset providing broad information regarding the role of genetic variation in regulating gene expression in multiple tissues; and iii) a cohort specifically recruited and assessed to investigate asthma and allergies in children.

Our PheWAS conducted in the UK Biobank identified more than 200 phenotypic associations with respect to asthma and lung function-associated loci. In both cases, we replicated the original GWAS phenotypes (asthma and lung function, respectively), but observed much stronger associations with respect to additional traits. Asthma risk loci were strongly associated with blood levels of immune cells and particularly with eosinophil level. This is in line with the known role of immune cells in inflammatory processes at the basis of asthma and allergy pathogenesis.[30] Additionally, eosinophils in asthmatic inflammation have been a recognized risk factor for many years and eosinophilic airway inflammation seems to be associated with risk of severe asthma exacerbation and loss of asthma control with inhaled corticosteroid withdrawal.[31] Our data indicate that risk loci associated with asthma phenotypes are involved in the regulation of immune systems. Based on these findings, we hypothesize that causal mechanisms responsible for the genetic associations reported in the asthma discovery GWAS[8,9] are related to this regulatory effect: individuals predisposed

https://doi.org/10.4168/aair.2020.12.5.806

to have altered immune function have an increased risk for developing asthma. In line with previous studies that observed the effect of eosinophils and other immune cells in both adult and childhood asthma,[32] we observed that the blood levels of immune cells were correlated with risk loci associated with both adult and childhood asthma, suggesting that immune alterations are a pathogenic pathway shared between the different disease-onset phenotypes. Investigating the effect of asthma risk alleles on transcriptomic regulation of the surrounding genes, we were able to gain more information about molecular mechanisms responsible for the association between asthma risk and immune alterations. Our eQTL analyses showed that asthma risk alleles tend to affect multiple genes across several tissues. Among the variants investigated, the asthma-increasing risk allele rs2305480*G is associated with the down-regulation of *GSDMB*, *ORMDL3*, and *PGAP3* and the up-regulation of *GSDMA* and *IKZF3* across multiple tissues. Accordingly, we can hypothesize that the effect of rs2305480 on asthma risk and the blood levels of immune cells is due to mechanisms related to the function of multiple genes, including the *GSDMB* and *GSDMA* regulation of apoptosis in epithelial cells,[33,34] *ORMDL3* regulation of the innate immune system,[35] and *IKZF3* regulation of B lymphocyte proliferation and differentiation.[36,37] Among the other variants tested, the asthma-increasing risk rs1558641*G allele was associated with the up-regulation of *MFSD9* and down-regulation of *IL18R1*. The gene *MFSD9* is involved in transmembrane transporter activities[38] and rs1558641 appears to affect its regulation specifically in whole blood. Conversely, rs1558641 regulates *IL18R1* gene expression across multiple tissues. This gene encodes for the cytokine receptor that binds interleukin 18 (IL18), and it is essential for IL18-mediated signal transduction.[39] Accordingly, we hypothesize that rs1558641 association with asthma and the blood levels of immune cells is due to the regulatory function of the gene *IL18R1*.

Unlike asthma risk loci, variants associated with lung function did not show any association with blood immune cells. Conversely, we observed a consistent association with several traits related to body fat mass together with replication of the lung-function traits. Obesity and pathologic accumulation of fat mass are risk factors for both allergic and non-allergic respiratory diseases.[40,41] Our data support the hypothesis that the causative mechanisms responsible for the genetic associations reported in the lung-function discovery GWAS[10] are related to the effect of body fat mass: individuals with risk variants associated with increased fat mass have an increased predisposition of reduced lung function. Among the lung function-associated variants, rs4237643*T allele is associated with the up-regulation of *HSD17B12* gene which encodes 17-beta-hydroxysteroid dehydrogenase, an enzyme involved in the long-chain fatty acid elongation cycle and the metabolism of steroid hormones.[42] *HSD17B12* locus has been associated with body mass index,[43] type 2 diabetes,[44] and cardiovascular diseases.[13] Accordingly, we hypothesize the effect of lung function-associated loci investigated is due to the mediated pleiotropy,[45] where variants are associated with altered metabolic processes that lead to increased body fat and obesity, which consequently affects lung function. This hypothesis is supported by studies showing the accumulation of adipose tissue causes pulmonary complications directly (mechanical effect of fat accumulation in the chest and abdominal regions) and indirectly (altered immune function due to the numerous cytokines produced by adipocytes).[46-48]

In the pediatric cohort assessed using the ISAAC Phase II tools, we observed that asthma risk alleles are associated with a broader range of asthma- and allergy-related phenotypes than lung function-associated variants. However, we observed that 2 lung function- and fat mass-associated variants (*i.e.*, rs2863171 and rs4237643) are associated with atopy, asthmatic and allergic bronchitis, and the blood levels of immune cells. Although these results will be

https://doi.org/10.4168/aair.2020.12.5.806

needed to be confirmed in a larger cohort, these findings may support the effect of increased fat mass in the positive feedback loop between local inflammation in adipose tissue and altered immune response.[49]

In conclusion, this study generated novel data regarding the genetics of respiratory diseases and their comorbidities, providing a deep characterization of loci associated with asthma and lung function based on multiple sources. Our findings support that asthma and lung function have distinct genetic architecture, but addictive mechanisms may be present, linking these phenotypes.

ACKNOWLEDGMENTS

This study was supported by grants (2017-024 and 2018-057) from the Scientific Research Unit of Aksaray University. The study sponsor had not any involvement regarding the study design, data evaluation, writing of the manuscript and, its submission for publication. The authors are indebted to the field workers, those who collaborated in the participating centers, and all the parents and school staff who assisted in the survey. The authors thank the participants and investigators of the UK Biobank and the Neale lab for generating the genome-wide data used in the present study.

SUPPLEMENTARY MATERIALS

Supplementary Table S1

Significant associations surviving Bonferroni multiple testing correction ($P < 7.65 \times 10^{-7}$) in the UK Biobank PheWAS

Click here to view

Supplementary Table S2

eQTLs surviving genome-wide multiple testing correction

Click here to view

Supplementary Table S3

List of the phenotypes investigated in the pediatric cohort assessed using the ISAAC Phase II tools

Click here to view

Supplementary Fig. S1

Multi-tissue eQTL results of rs2305480 with respect to *GSDMB*.

Click here to view

Supplementary Fig. S2

Multi-tissue eQTL results of rs4237643 with respect to *HSD17B12*.

Click here to view

Supplementary Fig. S3

Multi-tissue eQTL results of rs1558641 with respect to *IL18R1*.

Click here to view

Supplementary Fig. S4

Multi-tissue eQTL results of rs744910 with respect to *AAGAB*.

Click here to view

Supplementary Fig. S5

Multi-tissue eQTL results of rs2284033 with respect to *ELFN2*.

Click here to view

REFERENCES

1. Zar HJ, Ferkol TW. The global burden of respiratory disease-impact on child health. Pediatr Pulmonol 2014;49:430-4.
PUBMED | CROSSREF
2. Ober C. Asthma genetics in the post-GWAS era. Ann Am Thorac Soc 2016;13 Suppl 1:S85-90.
PUBMED | CROSSREF
3. Karaca S, Erge S, Cesuroglu T, Polimanti R. Nutritional habits, lifestyle, and genetic predisposition in cardiovascular and metabolic traits in Turkish population. Nutrition 2016;32:693-701.
PUBMED | CROSSREF
4. Pasaniuc B, Price AL. Dissecting the genetics of complex traits using summary association statistics. Nat Rev Genet 2017;18:117-27.
PUBMED | CROSSREF
5. Greene CS, Krishnan A, Wong AK, Ricciotti E, Zelaya RA, Himmelstein DS, et al. Understanding multicellular function and disease with human tissue-specific networks. Nat Genet 2015;47:569-76.
PUBMED | CROSSREF
6. Polimanti R, Kranzler HR, Gelernter J. Phenome-wide association study for alcohol and nicotine risk alleles in 26394 women. Neuropsychopharmacology 2016;41:2688-96.
PUBMED | CROSSREF
7. Karaca S, Civelek E, Karaca M, Sahiner UM, Ozgul RK, Kocabas CN, et al. Allergy-specific phenome-wide association study for immunogenes in Turkish children. Sci Rep 2016;6:33152.
PUBMED | CROSSREF
8. Bønnelykke K, Sleiman P, Nielsen K, Kreiner-Møller E, Mercader JM, Belgrave D, et al. A genome-wide association study identifies CDHR3 as a susceptibility locus for early childhood asthma with severe exacerbations. Nat Genet 2014;46:51-5.
PUBMED | CROSSREF
9. 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
10. Loth DW, Soler Artigas M, Gharib SA, Wain LV, Franceschini N, Koch B, et al. Genome-wide association analysis identifies six new loci associated with forced vital capacity. Nat Genet 2014;46:669-77.
PUBMED | CROSSREF
11. Eyre S, Bowes J, Diogo D, Lee A, Barton A, Martin P, et al. High-density genetic mapping identifies new susceptibility loci for rheumatoid arthritis. Nat Genet 2012;44:1336-40.
PUBMED | CROSSREF
12. McGovern DP, Gardet A, Törkvist L, Goyette P, Essers J, Taylor KD, et al. Genome-wide association identifies multiple ulcerative colitis susceptibility loci. Nat Genet 2010;42:332-7.
PUBMED | CROSSREF

13. Kichaev G, Bhatia G, Loh PR, Gazal S, Burch K, Freund MK, et al. Leveraging polygenic functional enrichment to improve GWAS power. Am J Hum Genet 2019;104:65-75.
PUBMED | CROSSREF
14. Hinks A, Cobb J, Marion MC, Prahalad S, Sudman M, Bowes J, et al. Dense genotyping of immune-related disease regions identifies 14 new susceptibility loci for juvenile idiopathic arthritis. Nat Genet 2013;45:664-9.
PUBMED | CROSSREF
15. Terao C, Kawaguchi T, Dieude P, Varga J, Kuwana M, Hudson M, et al. Transethnic meta-analysis identifies *GSDMA* and *PRDM1* as susceptibility genes to systemic sclerosis. Ann Rheum Dis 2017;76:1150-8.
PUBMED | CROSSREF
16. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Med 2015;12:e1001779.
PUBMED | CROSSREF
17. GTEx Consortium; Laboratory, Data Analysis &Coordinating Center (LDACC)—Analysis Working Group; Statistical Methods groups—Analysis Working Group; Enhancing GTEx (eGTEx) groups; NIH Common Fund; NIH/NCI Genetic effects on gene expression across human tissues. Nature 2017;550:204-13.
PUBMED | CROSSREF
18. Civelek E, Cakir B, Boz AB, Yuksel H, Orhan F, Uner A, et al. Extent and burden of allergic diseases in elementary schoolchildren: a national multicenter study. J Investig Allergol Clin Immunol 2010;20:280-8.
PUBMED
19. Pers TH, Timshel P, Hirschhorn JN. SNPsnap: a web-based tool for identification and annotation of matched SNPs. Bioinformatics 2015;31:418-20.
PUBMED | CROSSREF
20. Ongen H, Buil A, Brown AA, Dermitzakis ET, Delaneau O. Fast and efficient QTL mapper for thousands of molecular phenotypes. Bioinformatics 2016;32:1479-85.
PUBMED | CROSSREF
21. Sul JH, Han B, Ye C, Choi T, Eskin E. Effectively identifying eQTLs from multiple tissues by combining mixed model and meta-analytic approaches. PLoS Genet 2013;9:e1003491.
PUBMED | CROSSREF
22. Civelek E, Cakir B, Orhan F, Yuksel H, Boz AB, Uner A, et al. Risk factors for current wheezing and its phenotypes among elementary school children. Pediatr Pulmonol 2011;46:166-74.
PUBMED | CROSSREF
23. Civelek E, Yavuz ST, Boz AB, Orhan F, Yuksel H, Uner A, et al. Epidemiology and burden of rhinitis and rhinoconjunctivitis in 9- to 11-year-old children. Am J Rhinol Allergy 2010;24:364-70.
PUBMED | CROSSREF
24. He C, Holme J, Anthony J. SNP genotyping: the KASP assay. Methods Mol Biol 2014;1145:75-86.
PUBMED | CROSSREF
25. 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
26. 1000 Genomes Project Consortium, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, et al. A global reference for human genetic variation. Nature 2015;526:68-74.
PUBMED | CROSSREF
27. Bønnelykke K, Ober C. Leveraging gene-environment interactions and endotypes for asthma gene discovery. J Allergy Clin Immunol 2016;137:667-79.
PUBMED | CROSSREF
28. Lange P, Celli B, Agustí A, Boje Jensen G, Divo M, Faner R, et al. Lung-function trajectories leading to chronic obstructive pulmonary disease. N Engl J Med 2015;373:111-22.
PUBMED | CROSSREF
29. McGeachie MJ, Yates KP, Zhou X, Guo F, Sternberg AL, Van Natta ML, et al. Patterns of growth and decline in lung function in persistent childhood asthma. N Engl J Med 2016;374:1842-52.
PUBMED | CROSSREF
30. Kubo M. Innate and adaptive type 2 immunity in lung allergic inflammation. Immunol Rev 2017;278:162-72.
PUBMED | CROSSREF
31. Price DB, Rigazio A, Campbell JD, Bleecker ER, Corrigan CJ, Thomas M, et al. Blood eosinophil count and prospective annual asthma disease burden: a UK cohort study. Lancet Respir Med 2015;3:849-58.
PUBMED | CROSSREF

32. Holtzman MJ. Asthma as a chronic disease of the innate and adaptive immune systems responding to viruses and allergens. J Clin Invest 2012;122:2741-8.
PUBMED | CROSSREF

33. Qiu S, Liu J, Xing F. 'Hints' in the killer protein gasdermin D: unveiling the secrets of gasdermins driving cell death. Cell Death Differ 2017;24:588-96.
PUBMED | CROSSREF

34. Saeki N, Usui T, Aoyagi K, Kim DH, Sato M, Mabuchi T, et al. Distinctive expression and function of four GSDM family genes (GSDMA-D) in normal and malignant upper gastrointestinal epithelium. Genes Chromosomes Cancer 2009;48:261-71.
PUBMED | CROSSREF

35. Boeck A, Landgraf-Rauf K, Vogelsang V, Siemens D, Prazeres da Costa O, Klucker E, et al. Ca^{2+} and innate immune pathways are activated and differentially expressed in childhood asthma phenotypes. Pediatr Allergy Immunol 2018;29:823-33.
PUBMED | CROSSREF

36. Somasundaram R, Prasad MA, Ungerbäck J, Sigvardsson M. Transcription factor networks in B-cell differentiation link development to acute lymphoid leukemia. Blood 2015;126:144-52.
PUBMED | CROSSREF

37. Tsai DY, Hung KH, Chang CW, Lin KI. Regulatory mechanisms of B cell responses and the implication in B cell-related diseases. J Biomed Sci 2019;26:64.
PUBMED | CROSSREF

38. Perland E, Hellsten SV, Schweizer N, Arapi V, Rezayee F, Bushra M, et al. Structural prediction of two novel human atypical SLC transporters, MFSD4A and MFSD9, and their neuroanatomical distribution in mice. PLoS One 2017;12:e0186325.
PUBMED | CROSSREF

39. Yu Q, Chang HC, Ahyi AN, Kaplan MH. Transcription factor-dependent chromatin remodeling of *Il18r1* during Th1 and Th2 differentiation. J Immunol 2008;181:3346-52.
PUBMED | CROSSREF

40. Pinto LM, Alghamdi M, Benedetti A, Zaihra T, Landry T, Bourbeau J. Derivation and validation of clinical phenotypes for COPD: a systematic review. Respir Res 2015;16:50.
PUBMED | CROSSREF

41. Gonzalez-Garcia M, Caballero A, Jaramillo C, Maldonado D, Torres-Duque CA. Prevalence, risk factors and underdiagnosis of asthma and wheezing in adults 40 years and older: a population-based study. J Asthma 2015;52:823-30.
PUBMED | CROSSREF

42. Kemiläinen H, Adam M, Mäki-Jouppila J, Damdimopoulou P, Damdimopoulos AE, Kere J, et al. The hydroxysteroid (17β) dehydrogenase family gene HSD17B12 is involved in the prostaglandin synthesis pathway, the ovarian function, and regulation of fertility. Endocrinology 2016;157:3719-30.
PUBMED | CROSSREF

43. Turcot V, Lu Y, Highland HM, Schurmann C, Justice AE, Fine RS, et al. Protein-altering variants associated with body mass index implicate pathways that control energy intake and expenditure in obesity. Nat Genet 2018;50:26-41.
PUBMED | CROSSREF

44. Mahajan A, Taliun D, Thurner M, Robertson NR, Torres JM, Rayner NW, et al. Fine-mapping type 2 diabetes loci to single-variant resolution using high-density imputation and islet-specific epigenome maps. Nat Genet 2018;50:1505-13.
PUBMED | CROSSREF

45. Hackinger S, Zeggini E. Statistical methods to detect pleiotropy in human complex traits. Open Biol 2017;7:170125.
PUBMED | CROSSREF

46. Makki K, Froguel P, Wolowczuk I. Adipose tissue in obesity-related inflammation and insulin resistance: cells, cytokines, and chemokines. ISRN Inflamm 2013;2013:139239.
PUBMED | CROSSREF

47. Dixon AE, Peters U. The effect of obesity on lung function. Expert Rev Respir Med 2018;12:755-67.
PUBMED | CROSSREF

48. Mafort TT, Rufino R, Costa CH, Lopes AJ. Obesity: systemic and pulmonary complications, biochemical abnormalities, and impairment of lung function. Multidiscip Respir Med 2016;11:28.
PUBMED | CROSSREF

49. de Heredia FP, Gómez-Martínez S, Marcos A. Obesity, inflammation and the immune system. Proc Nutr Soc 2012;71:332-8.
PUBMED | CROSSREF