

Genome-wide prediction of childhood asthma and related phenotypes in a longitudinal birth cohort

Ben D. Spycher, PhD,^{a,b} John Henderson, MD, FRCP,^a Raquel Granell, PhD,^a David M. Evans, PhD,^a George Davey Smith, MD, DSc,^a Nicholas J. Timpson, PhD,^a and Jonathan A. C. Sterne, PhD^a *Bristol, United Kingdom, and Bern, Switzerland*

Background: Childhood wheezing and asthma vary greatly in clinical presentation and time course. The extent to which phenotypic variation reflects heterogeneity in disease pathways is unclear.

Objective: We sought to assess the extent to which single nucleotide polymorphisms (SNPs) associated with childhood asthma in a genome-wide association study are predictive of asthma-related phenotypes.

Methods: In 8365 children from a population-based birth cohort, the Avon Longitudinal Study of Parents and Children, allelic scores were derived based on between 10 and 215,443 SNPs ranked according to the inverse of the *P* value for their association with physician-diagnosed asthma in an independent genome-wide association study (6176 cases and 7111 control subjects). We assessed the predictive value of allelic scores for asthma-related outcomes at age 7 to 9 years (physician's diagnosis, longitudinal wheezing phenotypes, and measurements of pulmonary function, bronchial responsiveness, and atopy).

Results: Scores based on the 46 highest-ranked SNPs were associated with the symptom-based phenotypes early onset persistent wheeze ($P < 10^{-11}$; area under the receiver operating characteristic curve [AUC], 0.59) and intermediate-onset wheeze ($P < 10^{-3}$; AUC, 0.58). Among lower-ranked SNPs (ranks, 21,545–46,416), there was evidence for associations with diagnosed asthma ($P < 10^{-4}$; AUC, 0.54) and atopy ($P < 10^{-5}$; AUC, 0.55). We found little evidence of associations with transient early wheezing, reduced pulmonary function, or nonasthma phenotypes.

Conclusion: The genetic origins of asthma are diverse, and some pathways are specific to wheezing syndromes, whereas others

are shared with atopy and bronchial hyperresponsiveness. Our study also provides evidence of etiologic differences among wheezing syndromes. (*J Allergy Clin Immunol* 2012;130:503–9.)

Key words: Respiratory sounds, hypersensitivity, respiratory function tests, genome-wide association study, forecasting

Childhood asthma and wheezing vary greatly in clinical presentation and time course¹ and are thought to have heterogeneous causes. Children with wheeze can be classified into more homogeneous phenotypes,^{1–7} but it is unclear whether these reflect distinct disease pathways. The etiologic relationship of asthma with intermediate phenotypes, such as atopy, bronchial hyperresponsiveness (BHR), and reduced pulmonary function, is unclear, although twin and family studies have found evidence of shared genetic components.^{8–15}

Genome-wide association studies (GWASs) of asthma have provided evidence for both common and distinct pathways.^{16–19} Among loci robustly associated with asthma susceptibility,¹⁶ one (*ORMDL3/GSDML*) was primarily associated with childhood asthma but not with atopy,^{16,17} whereas associations with eosinophil count and allergic disease have been reported for others (*IL1RL1* and *IL33*).^{17,20} The hypothesis that loci associated with lung function^{18,19} also confer susceptibility to asthma was not confirmed, suggesting distinct pathways for these phenotypes.¹⁶

Stringent genome-wide significance thresholds are used to address the problem of false discovery associated with multiple testing in GWASs. As a result, it is likely that many true associations do not reach these thresholds (false-negative results) and are therefore neither reported nor considered for further replication.²¹ Genome-wide prediction, examining the combined effect of a large number of variants simultaneously rather than separately, retrieves some of the predictive information lost by dismissing false-negative results.^{22–25} Including genome-wide information rather than only the variants reaching stringent significance thresholds can improve prediction of disease status for a number of complex diseases.^{24,25} Furthermore, this approach can be used to discover shared genetic liability between different complex phenotypes.²⁵

The work presented here aims to construct genetic prediction scores by using genome-wide data from the childhood cohorts of the GABRIEL (A Multidisciplinary Study to Identify the Genetic and Environmental Causes of Asthma in the European Community) consortium¹⁶ and use these to predict childhood asthma and a range of asthma-related and non-asthma-related phenotypes in independent data from a large population-based birth cohort.

METHODS

More detail is provided in the [Methods](#) section of this article's Online Repository at www.jacionline.org.

From ^athe School of Social and Community Medicine, University of Bristol, and ^bthe Institute of Social and Preventive Medicine, University of Bern.

The UK Medical Research Council and the Wellcome Trust (grant reference 092731) and the University of Bristol provide core support for the Avon Longitudinal Study of Parents and Children. B.D.S. is the recipient of a European Respiratory Society/Marie Curie Joint Research Fellowship (no. MC 1614-2010). The research leading to these results has received funding from the European Respiratory Society and the European Community's Seventh Framework Programme FP7/2007-2013–Marie Curie Actions under grant agreement RESPIRE, PCOFUND-GA-2008-229571. R.G. was supported by the UK Medical Research Council (grant no. 0401540). The GABRIEL project was supported by a contract from the European Commission (018996) and grants from the French Ministry of Research, the Wellcome Trust (WT084703MA), and Asthma UK. Disclosure of potential conflict of interest: J. Henderson, G. D. Smith, and J. A. C. Sterne receive research support from the Medical Research Council and the Wellcome Trust. R. Granell receives research support from the Medical Research Council. The rest of the authors declare that they have no relevant conflicts of interest.

Received for publication January 25, 2012; revised May 24, 2012; accepted for publication June 5, 2012.

Corresponding author: Jonathan A. C. Sterne, PhD, School of Social and Community Medicine, University of Bristol, Canynge Hall, 39 Whatley Road, Bristol BS8 2PS, United Kingdom. E-mail: jonathan.sterne@bristol.ac.uk.

0091-6749/\$36.00

© 2012 American Academy of Allergy, Asthma & Immunology
<http://dx.doi.org/10.1016/j.jaci.2012.06.002>

Abbreviations used

ALSPAC:	Avon Longitudinal Study of Parents and Children
AUC:	Area under receiver operating characteristic curve
BHR:	Bronchial hyperresponsiveness
FVC:	Forced vital capacity
GABRIEL:	A Multidisciplinary Study to Identify the Genetic and Environmental Causes of Asthma in the European Community
GWAS:	Genome-wide association study
IQ:	Intelligence quotient
IQOR:	Interquartile odds ratio
OR:	Odds ratio
SNP:	Single nucleotide polymorphism
SPT:	Skin prick test
WISC:	Wechsler Intelligence Scale for Children

Study population and data collection

The Avon Longitudinal Study of Parents and Children (ALSPAC) is a longitudinal population-based birth cohort study that recruited 14,541 pregnant women in Avon, United Kingdom with expected dates of delivery from April 1991 to December 1992. An additional 542 eligible pregnancies were recruited when children were aged 7 years. These 15,083 pregnancies resulted in 14,610 live births. The study has been described elsewhere.^{26,27} Each year from 0 to 8 years of age, study mothers were sent child health questionnaires, including detailed questions on respiratory symptoms. We included only unrelated children of European ancestry. Ethical approval for the study was obtained from the ALSPAC Law and Ethics Committee and the Local Research Ethics Committees.

Genotyping and quality control

Children of the GABRIEL consortium, including ALSPAC, were genotyped for more than 500,000 single nucleotide polymorphisms (SNPs). GWAS results for asthma from all childhood cohorts of the GABRIEL consortium were downloaded from <http://www.cng.fr/gabriel/>.¹⁶ We included SNPs satisfying the following quality control criteria in 1 or more cohorts: minor allele frequency of 5% or greater, consistency with Hardy-Weinberg equilibrium ($P > 10^{-4}$) in control subjects, and a call rate of 97% or greater.

Construction of genetic scores

Log odds ratios (ORs) for childhood asthma from GABRIEL cohorts excluding ALSPAC (18 cohorts with 6,176 cases and 7,111 control subjects referred to as the discovery set) were combined by using random-effects meta-analysis, and SNPs were ranked in inverse order of resulting P values. In ALSPAC 440,524 genotyped SNPs satisfied quality control criteria and overlapped with SNPs included in the discovery set. Prediction scores were calculated by using the k highest-ranking SNPs (with $k = 10, 22, 46, 100, 215, 464, 1,000, 2,154, 4,642, 10,000, 21,544, 46,416, 100,000$, and $215,443$; multiplying successively by $10^{0.33}$ to achieve an even spacing on the log scale). Two types of scores were calculated as follows: log odds scores using pooled log-ORs as weights for the allele count and allele count scores using the signs of the log-ORs (+1 if log-OR > 0 and -1 if log-OR < 0). The prediction scores were constructed using the PLINK program (version 1.07).²⁸

Outcomes in mid-childhood

ALSPAC children were classified according to whether they had ever received a physician's diagnosis of asthma or reported use of asthma medication in the past 12 months (based on the specific question "Has asthma medication been given to your child in the last 12 months?") at age 7 to 8 years. Results of skin prick tests (SPTs) at ages 7 to 8 years were used to define atopy (mean wheal diameter ≥ 2 mm to house dust mite, cat, or grass pollen) and to classify asthma as atopic or nonatopic. Phenotypes of wheezing in the first 7 years of life were defined by using latent class analysis of parental

reports of wheezing at 7 time points from 6 months to 7 years. We identified 6 phenotypes characterized by different longitudinal trajectories of wheezing. These were classified as follows: persistent, late-onset, intermediate-onset, prolonged early, transient early, and never/infrequent wheezing. The nature of these phenotypes; their associations with asthma, atopy, and lung function in mid-childhood; and their reproducibility in independent data have been described elsewhere.^{3,29} At ages 8 to 9 years, lung function was measured by spirometry (Vitalograph 2120, Maids Moreton, United Kingdom), and bronchial responsiveness to methacholine was measured using the method of Yan et al.³⁰ To evaluate the specificity of findings, we also defined the following nonasthma phenotypes: highest quintiles of systolic blood pressure at age 7 to 8 years, intelligence quotient (IQ) at age 8 to 9 years, and height at age 7 to 8 years (Table I).

Statistical analysis

We used logistic regression to examine the association of each outcome with each genetic prediction score. We quantified the predictive ability of scores by using area under the receiver operating characteristic curve (AUC), a measure of discrimination.³¹ We also quantified the magnitude of association by using interquartile odds ratios (IQORs) comparing the highest and lowest quintiles of the prediction score. Analyses were repeated using scores based on independent sets of SNPs with rankings between consecutive values of k . The R statistical software version 2.11.0 (R Foundation for Statistical Computing) was used for all analyses.

RESULTS

Genetic data were available for 8365 unrelated white European ALSPAC children. Of these, 5748 (68.7%) replied to the questionnaire at age 7 years, and 6128 (73.3%) and 5621 (67.2%) participated in study clinics at ages 7 to 8 and 8 to 9 years, respectively. Wheezing phenotypes based on repeated assessment of current wheeze throughout the first 7 years of life³ were available in 7176 (85.8%) children. Of these, 506 (7.1%) had persistent, 338 (4.7%) had late-onset, 168 (2.3%) had intermediate-onset, 538 (7.5%) had prolonged early, 726 (10.1%) had transient early, and the remaining 4900 (68.3%) had never or rare wheezing. At age 7 years, 1132 (19.9%) of 5677 children had ever been given a physician's diagnosis of asthma. Physician-diagnosed asthma was associated with most defined outcomes: associations were particularly strong (OR, >15) for use of asthma medication at age 7 years and the persistent and intermediate-onset wheeze phenotypes (Table I). Among the nonasthma outcomes, high IQ ($P = .085$) and height ($P = .008$) were negatively associated with diagnosed asthma (Table I).

Meta-analysis of GABRIEL childhood cohorts excluding ALSPAC

Combining GWAS results for asthma from the GABRIEL childhood cohorts excluding ALSPAC yielded similar results to the original study combining all childhood cohorts (Fig 1 in this article and see supplementary Fig 2a in Moffatt et al¹⁶). The top 18 SNPs, and therefore all SNPs included in the 10-SNP scores, were in the *ORMDL3/GSDMB* locus on chromosome 17q21, a well-established asthma susceptibility locus primarily associated with childhood-onset disease (Fig 1 and see Table E1 in this article's Online Repository at www.jacionline.org).¹⁸ Additional loci contributing to the 46-SNP scores included 6 SNPs on chromosome 2 implicating the *IL1RL1* and *IL18R1* genes and 2 SNPs flanking the *IL33* gene on chromosome 9. Other contributions to the 100-SNP score included loci on chromosomes 1, 5, 6, 8, 12, 15, and 22 (Fig 1 and see Table E1).

TABLE I. Defined outcomes in mid-childhood and association with physician-diagnosed asthma at age 7 years

Outcome	Definition	Age at assessment (y)	Physician-diagnosed asthma at 7 y		P value*
			No (n = 4545)	Yes (n = 1132)	
Atopic asthma	Physician-diagnosed asthma and positive SPT response	7-8	0/4545 (0.0)	332/861 (38.6)	—
Nonatopic asthma	Physician-diagnosed asthma and negative SPT response	7-8	0/4545 (0.0)	529/861 (61.4)	—
Asthma medication	Use of asthma medication in past 12 mo	7-8	66/4520 (1.5)	689/1120 (61.5)	<.001
Persistent wheeze	Wheeze phenotype “persistent”†	0-7	47/4466 (1.1)	322/1112 (29.0)	<.001
Late-onset wheeze	Wheeze phenotype “late onset”†	0-7	109/4466 (2.4)	174/1112 (15.6)	<.001
Intermediate-onset wheeze	Wheeze phenotype “intermediate onset”†	0-7	14/4466 (0.3)	119/1112 (10.7)	<.001
Prolonged early wheeze	Wheeze phenotype “prolonged early”†	0-7	272/4466 (6.1)	142/1112 (12.8)	<.001
Transient early wheeze	Wheeze phenotype “transient early”†	0-7	451/4466 (10.1)	104/1112 (9.4)	.502
Positive SPT response	At least 1 positive SPT response‡	7-8	539/3484 (15.5)	332/861 (38.6)	<.001
BHR	Third tertile of bronchial challenge dose-response slope§	8-9	560/2410 (23.2)	241/524 (46.0)	<.001
Low FEV ₁	FEV ₁ z score <20th percentile of normal distribution	8-9	593/3456 (17.2)	199/873 (22.8)	<.001
Low FVC	FVC z score <20th percentile of normal distribution	8-9	625/3510 (17.8)	152/882 (17.2)	.730
High systolic BP	Fifth quintile of systolic blood pressure	7-8	705/3946 (17.9)	196/984 (19.9)	.140
High IQ	Fifth quintile of IQ total score	8-9	736/3680 (20.0)	161/921 (17.5)	.085
High body height	Fifth quintile of body height	7-8	727/3995 (18.2)	146/997 (14.6)	.008

Prevalence among children with and without diagnosed asthma are reported as n/N (percentage).

BP, Blood pressure.

*Fisher exact test.

†Phenotypes of wheezing in the first 7 years of life were defined by using latent class analysis, as reported previously.³

‡Mean wheal diameter of 2 mm or greater to house dust mite, cat, or grass pollen.

§Tertiles were formed among children with a positive dose-response slope. Children with a dose-response slope of 0 or less were then grouped together with children in the lowest tertile.

||Based on parental report.

Prediction of asthma phenotypes

Fig 2 shows the predictive performance of log odds scores for 4 groups of outcomes. For the first group (asthma, atopic and non-atopic asthma, and asthma medication), there was evidence of association with genetic prediction scores ($P < 10^{-4}$), but discrimination was weak (AUC < 0.6; Fig 2, A1-C1). The AUC for physician-diagnosed asthma with the 10-SNP score was 0.55: the corresponding interquartile odds ratio (IQOR) was 1.75 (95% CI, 1.44-2.12), and the P value for association was 1.4×10^{-7} . Similar values of AUCs and IQORs were observed for nonatopic asthma and use of asthma medication. Prediction of atopic asthma, which was weaker than for the other asthma outcomes for scores with smaller numbers of SNPs, improved with increasing numbers of SNPs: the AUC reached 0.57 for 46,416 SNPs. By contrast, the AUC for nonatopic asthma and use of asthma medication decreased with increasing numbers of SNPs.

For the second outcome group, the early childhood wheezing phenotypes, there was evidence of association ($P < 10^{-3}$) for all phenotypes except transient early wheeze ($P > .25$ for all scores). Discrimination and strength of association varied between phenotypes (Fig 2, A2-C2). The strongest associations and best discrimination were observed for persistent and intermediate-onset wheezing. Prediction of persistent wheezing based on 46 SNPs resulted in an AUC of 0.59, an IQOR of 2.54 (95% CI, 1.92-3.37), and a P value of 4.9×10^{-12} . The AUC for these 2 phenotypes decreased to 0.54 or less for scores including 10^4 SNPs or more. The strongest evidence for association was with the scores including 46 SNPs (IQOR, 1.51; 95% CI, 1.10-2.06; $P = 5.5 \times 10^{-4}$) for late-onset wheeze and for the score including 46,416 SNPs (IQOR, 1.49; 95% CI,

1.15-1.93; $P = 4.0 \times 10^{-5}$) for prolonged early wheeze, but the AUC did not exceed 0.56.

Among physiologic measurements (the third outcome group), BHR was associated ($P < 10^{-3}$) with scores including 2154 and 4642 SNPs and a positive SPT response was associated with scores including 4642 or more SNPs, whereas little evidence of association with any score was found for FEV₁ ($P > .029$) and forced vital capacity (FVC; $P > .078$; Fig 2, A3-C3). Maximum AUC values for BHR were reached with a score including 2154 SNPs (AUC, 0.54) and for a positive SPT response with a score including 10^5 SNPs (0.56), whereas AUC was consistently 0.52 or less for FEV₁ and FVC.

There was little evidence of association between scores and the fourth group (nonasthma outcomes; Fig 2, A4-C4). Results were closely similar when using allele count scores instead of log odds scores (see Fig E1 in this article's Online Repository at www.jacionline.org).

When using scores based on independent sets of SNPs with ranks between successive cutoffs, scores including SNPs ranked 1 to 10, 11 to 22, and 23 to 46 were predictive of physician-diagnosed asthma ($P \leq 10^{-6}$; AUC, 0.54-0.55), asthma medication ($P \leq 10^{-4}$; AUC, 0.54-0.55), and atopic ($P < .02$; AUC, 0.53-0.56) and nonatopic ($P < .002$; AUC, 0.54-0.55) asthma (see Fig E2 in this article's Online Repository at www.jacionline.org). There was evidence of association ($P < .007$) between scores including SNPs ranked 10,001 to 21,544 and 21,545 to 46,416 and physician-diagnosed asthma, atopic asthma, and medication use, with the AUC approaching 0.56 for atopic asthma. Among early wheezing phenotypes, persistent wheezing was associated with scores including SNPs ranked 1 to 10, 11 to

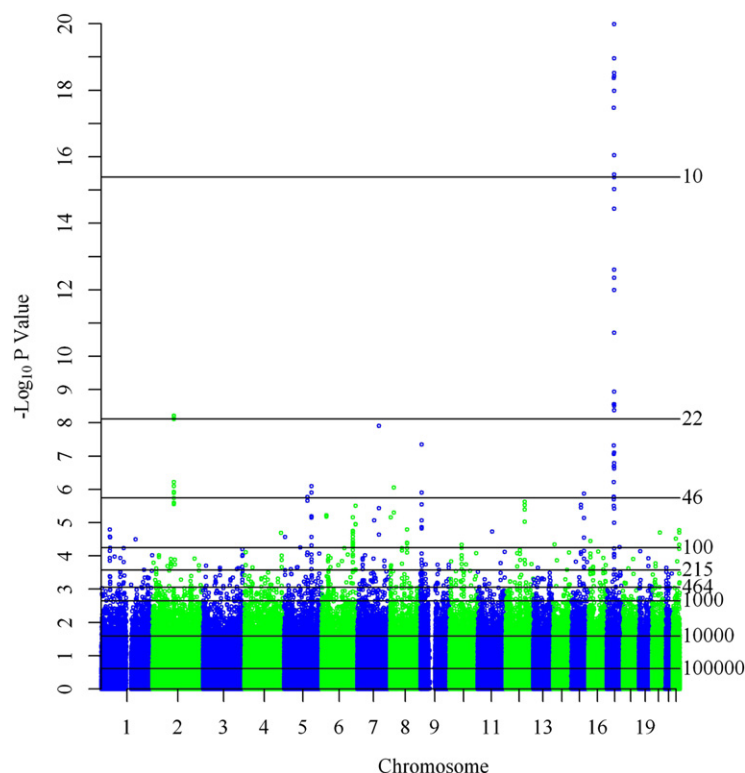


FIG 1. Manhattan plot of pooled results for physician-diagnosed asthma in GABRIEL childhood cohorts excluding ALSPAC. Horizontal lines represent genetic prediction scores with a varying number (shown to the right of each line) of included SNPs. All SNPs with *P* values above a line are included in the respective score.

22, and 23 to 46 ($P < 10^{-7}$; AUC, 0.57-0.58), but there was little evidence of associations with scores based on lower-ranked SNPs. A similar pattern was observed for intermediate-onset and late-onset wheezing, although associations with the first 3 sets of SNPs were weaker. The score including ranks 10,001 to 21,544 was associated with prolonged early wheeze ($P = 3.4 \times 10^{-4}$; AUC, 0.55). Among physiologic measurements, notable associations were only found for scores based on lesser ranks, with SNPs ranked 21,545 to 46,416 being most predictive of SPT positivity ($P = 4.0 \times 10^{-7}$; AUC, 0.56) and those ranked 1,001 to 2,154 being most predictive of BHR ($P = 2.8 \times 10^{-4}$; AUC, 0.54). There was little evidence of associations with low FEV₁ and FVC values and the nonasthma outcomes.

DISCUSSION

Main findings

We used SNPs associated with asthma in the GABRIEL GWAS¹⁶ to construct genome-wide prediction scores and evaluated their performance in an independent population-based birth cohort study. Prediction scores were most strongly associated with the persistent and intermediate-onset wheezing phenotypes, which were defined previously by using latent class analysis.³ We found associations with asthma among the lower-ranked SNPs with *P* values for association that did not reach genome-wide thresholds in the GABRIEL GWAS. Although the highest-ranked SNPs were not associated with atopy, there was evidence of associations of atopy with asthma-associated SNPs that did not reach genome-wide significance. The specificity of

the prediction scores for asthma-related phenotypes was confirmed by the absence of associations with the nonasthma phenotypes, blood pressure, height, and IQ.

Strengths and limitations

Our study exploited the combined signal of SNPs across the genome that is usually ignored in GWASs. By constructing prediction scores based on the logORs (or directions of association) from the GWASs and examining their associations with asthma phenotypes in independent data, we avoided making large numbers of multiple comparisons, hence the need for stringent *P* value thresholds. The study sample was from a well-phenotyped, large, population-based birth cohort. A limitation of the study is that identified associations with asthma-linked SNPs with low *P* value rankings did not permit identification of the specific SNPs responsible for these associations because only the combined contribution to prediction scores of numerous SNPs, many of which are not associated with asthma, was considered. The allele scores used in our study can, by construction, only account for additive genetic effects and not for dominant allelic effects or gene-by-gene interaction. We did not have medical records of diagnoses but relied on parental reports of physician-diagnosed asthma. This could have resulted in misclassification of asthma, which might have attenuated the strength of associations with disease diagnosis compared with symptom-based phenotypes. Similarly, use of asthma medication was based on parental report. Although we asked specifically about asthma medication, it is possible that parents erroneously reported other treatments as being

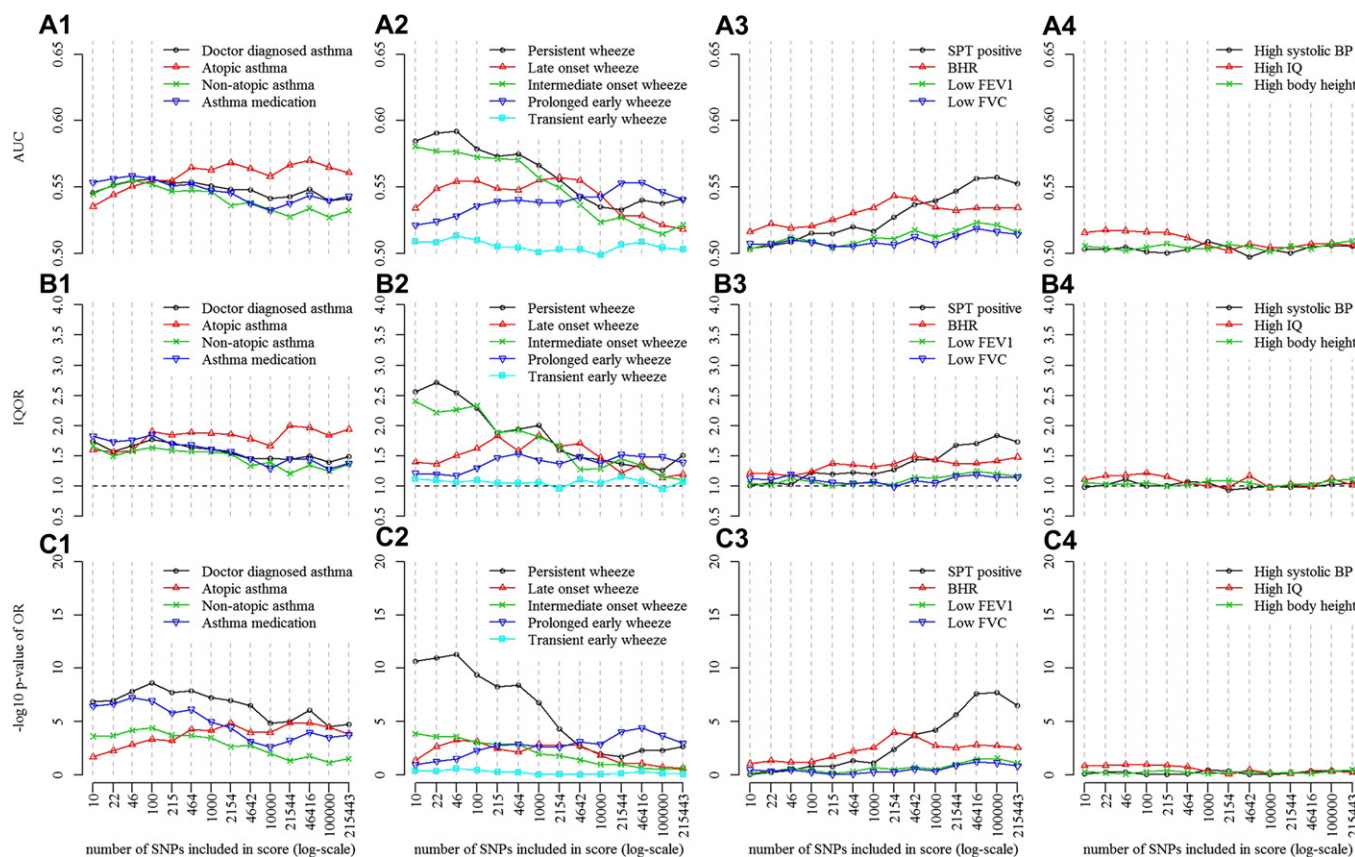


FIG 2. Comparison of predictive value of allelic log odds scores for the defined outcomes in mid-childhood. AUC (A1-A4), IQOR (B1-B4), and *P* values for association (C1-C4) are shown. Outcomes are based on a physician's diagnosis (A1-C1), phenotypes of wheeze in the first 7 years of life (A2-C2), physiologic measurements (A3-C3), and nonasthma outcomes (A4-C4). SNPs are included in inverse order of their *P* values in an independent GWAS for physician-diagnosed asthma.

for asthma. The magnitude of any such misclassification was small because only 6.8% of children reported asthma medication without a reported diagnosis of asthma or current symptoms.

Relationship to other genome-wide prediction studies

To our knowledge, the present study is the first to use genome-wide prediction to demonstrate the existence of common genetic components of different asthma-related phenotypes. In a recent asthma GWAS, Ferreira et al³² also used genome-wide allelic scores based on results from the GABRIEL study to predict physician-diagnosed asthma in independent data. However, they did not investigate associations between these scores and other related phenotypes. Another study used a related approach to predict severe asthma exacerbations in children.³³ The authors reported that the discriminative ability of the prediction score improved when they included more variants (up to 160). In 2 other studies with multiple genetic markers to predict asthma-related phenotypes, the variants were either preselected based on whether they belonged to a particular candidate gene³⁴ or whether they reached genome-wide significance.¹⁶

Previous studies have demonstrated the utility of genome-wide prediction in other complex diseases.^{24,25} Applying similar methodology, Evans et al²⁴ found that for some diseases, including bipolar disorder, coronary heart disease, and type II diabetes, the

discriminative ability of genome-wide prediction scores increased as more SNPs satisfying less stringent *P* value thresholds were included. Although this increase was small and unlikely to be of diagnostic utility, it implies the existence of genetic variants contributing to disease risk that remain unidentified, even in large GWASs. A genome-wide prediction score for schizophrenia was recently shown to be predictive of bipolar disorder, suggesting that these diseases share a common polygenic component.²⁵

Interpretation

In interpreting our findings, it is important to note that the predictive ability of genome-wide scores is expected to decrease as the number of included variants increases, unless the additional variants include loci truly associated with disease. This is because, among the lower-ranked SNPs, increasing proportions are included because, by chance alone, they show evidence of association in the discovery set. These variants dilute signals originating from truly associated loci. If discrimination improves with the number of SNPs included in the prediction scores, as observed for atopy and atopic asthma, or if it is maintained, as observed for diagnosed asthma and asthma medication use, this suggests the presence of variants truly associated with these phenotypes among the lower-ranked SNPs. The validity of this interpretation is supported by the fact that little or no discriminative ability was found for nonasthma phenotypes. It is likely that contamination by

noise arising from sampling variation in the discovery dataset partly explains the relatively poor discrimination found for all phenotypes (all AUCs were less than 0.60). However, even in the absence of sampling variation, the AUC of genome-wide prediction scores is limited by other factors, such as the prevalence and heritability of the phenotype, as discussed extensively elsewhere.³⁵

Early wheezing phenotypes and asthma

We found that the top 46 ranked SNPs associated with childhood asthma in the GABRIEL GWAS predicted 2 early wheezing phenotypes, persistent wheeze and intermediate-onset wheeze, better than they predicted physician-diagnosed asthma (the discovery phenotype). We found little evidence that lower-ranked SNPs predicted these wheezing phenotypes, although this cannot be excluded. The first 46 SNPs had *P* values of less than 2×10^{-6} and included previously reported loci,¹⁶ such as the *ORMDL3/GSDMB* locus on chromosome 17q21, the *IL1RL1* and *IL18R1* genes on chromosome 2, and *IL33* on chromosome 9 but also some others (see Table E1). Given the strong signal observed for the first 22 SNPs, 18 of which are in the 17q21 region, this locus appears to be the main single contributor to the discrimination reached for these wheezing phenotypes. Our results agree with previous findings suggesting that the 17q21 locus is primarily associated with early-onset severe asthma³⁶⁻³⁸ but not with atopy^{16,36} or lung function.³⁶ Our study suggests that a symptom-based phenotype defined by wheezing beginning in the first 3 years of life and persisting up to 6 years more specifically reflects the disease susceptibility conferred by this locus than physician-diagnosed asthma.

Our study provides evidence of still-undiscovered asthma susceptibility loci and adds to the evidence that asthma is a polygenic disease: we found that genome-wide prediction scores based on independent sets of SNPs that had marker-specific *P* values for association with childhood asthma of greater than .025 (ranks >10,000, Fig 1) were predictive of the asthma phenotypes, physician-diagnosed asthma, atopic asthma, and medication use. This is consistent with findings by Ferreira et al³² in their recent GWAS-diagnosed asthma and with findings for other complex diseases^{24,25} and suggests the presence of numerous loci of small effect. Identifying these loci will be difficult, even in well-powered GWASs. Apart from the SNPs identified by the GABRIEL study, most SNPs identified as associated with asthma in recent GWASs^{32,39-43} had ranks of greater than 10,000 in our study (see Table E2 in this article's Online Repository at www.jacionline.org). The 11q13.5 locus (*LRR32*), which was identified as associated with atopic status among asthmatic patients by Ferreira et al,³² had a *P* value of .032 and ranked 12,607 in the present study and might have contributed to the discrimination of atopic asthma observed for the set of SNPs ranked 10,001 to 21,544 (see Fig E2).

Atopy, BHR, and asthma

Our study provides evidence that childhood asthma shares common polygenic components with atopy and BHR, suggesting that there are shared genetic pathways between these phenotypes and asthma. This has been previously shown in twin studies that use differences in cross-trait correlations between monozygotic and dizygotic twins to infer correlation between the heritable components (genetic correlation) underlying 2 traits.

Several such studies have reported genetic correlation between asthma or wheeze and different measurements of allergy, including self-reported allergy,¹⁴ positive SPT responses,^{10,13} and total serum IgE levels,¹⁰ although estimates vary widely between studies (0.15-0.71). Similar genetic correlation was found between asthma or wheeze and bronchial responsiveness (0.11-0.85),^{10,13} whereas estimates for genetic correlation between measurements of allergy and bronchial responsiveness (0.15-0.59)^{10,11,13,15} and between FEV₁ and FEV₁/FVC ratio with other asthma-related traits (−0.52 to −0.11)^{10,15} tended to be smaller in magnitude. The evidence provided by our study is more direct because it is based on allelic information from individual SNPs. Other attempts to find a genetic overlap between asthma phenotypes based on individual SNPs have been less successful, probably because they relied on SNPs reaching stringent thresholds in GWASs. The GABRIEL study, for instance, found little overlap between loci associated with asthma and those associated with total serum IgE levels, concluding that increased IgE levels have a minor role in the development of asthma and might be secondary to asthma.¹⁶ Our study shows that such an overlap is detectable among variants that do not meet genome-wide significance thresholds in GWASs for asthma.

Conclusions

Our study shows that the genetic origins of asthma phenotypes are diverse. Some pathways appear to be specific to wheezing syndromes, but there is also evidence for shared genetic components with atopy and BHR. Discovering the genes underlying this shared heritability will require larger sample sizes, more specific phenotypic definitions, or both. Among wheezing syndromes, further evidence for etiologic differences comes from the heterogeneity in patterns of genetic association with different symptom-based phenotypes. Understanding these differences could be key to devising interventions that alter the natural history of disease in targeted subgroups.

We are extremely grateful to all the subjects who took part in this study, the midwives for their help in recruiting them, and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists, and nurses. We thank the Sanger Centre, Centre National de Génotypage, and 23andMe for generating the ALSPAC GWAS data and John Kemp for assistance with cleaning up these data.

Key messages

- By using genetic prediction scores based on up to 215,443 SNPs, this study found that not only high-ranked (reaching stringent significance thresholds in an independent GWAS for childhood asthma) but also lower-ranked SNPs contributed to predicting asthma-related phenotypes in children.
- Scores based on high-ranked SNPs were most predictive for symptom-based phenotypes.
- Scores based on lower-ranked SNPs predicted atopy and BHR, providing evidence of shared genetic influences with asthma, whereas scores did not predict lung function.

REFERENCES

- Martinez FD, Wright AL, Taussig LM, Holberg CJ, Halonen M, Morgan WJ. Asthma and wheezing in the first six years of life. *N Engl J Med* 1995;332:133-8.
- Silverman M, Grigg J, McKean M. Virus-induced wheeze in young children—a separate disease?. In: Johnston S, Papadopoulos N, editors. *Respiratory infections in allergy and asthma*. New York: Marcel Dekker; 2002. p. 427-71.
- Henderson J, Granell R, Heron J, Sherriff A, Simpson A, Woodcock A, et al. Associations of wheezing phenotypes in the first 6 years of life with atopy, lung function and airway responsiveness in mid-childhood. *Thorax* 2008;63:974-80.
- Spycher BD, Silverman M, Brooke AM, Minder CE, Kuehni CE. Distinguishing phenotypes of childhood wheeze and cough using latent class analysis. *Eur Respir J* 2008;31:974-81.
- Henderson J, Granell R, Sterne J. The search for new asthma phenotypes. *Arch Dis Child* 2009;94:333-6.
- Spycher BD, Silverman M, Kuehni CE. Phenotypes of childhood asthma: are they real? *Clin Exp Allergy* 2010;40:1130-41.
- Schultz A, Brand PL. Episodic viral wheeze and multiple trigger wheeze in pre-school children: a useful distinction for clinicians? *Paediatr Respir Rev* 2011;12:160-4.
- Duffy DL, Mitchell CA, Martin NG. Genetic and environmental risk factors for asthma: a cotwin-control study. *Am J Respir Crit Care Med* 1998;157:840-5.
- Palmer LJ, Burton PR, Faux JA, James AL, Musk AW, Cookson WO. Independent inheritance of serum immunoglobulin E concentrations and airway responsiveness. *Am J Respir Crit Care Med* 2000;161:1836-43.
- Ferreira MA, O'Gorman L, Le Souef P, Burton PR, Toelle BG, Robertson CF, et al. Variance components analyses of multiple asthma traits in a large sample of Australian families ascertained through a twin proband. *Allergy* 2006;61:245-53.
- Thomsen SF, Ferreira MA, Kyvik KO, Fenger M, Backer V. A quantitative genetic analysis of intermediate asthma phenotypes. *Allergy* 2009;64:427-30.
- Thomsen SF, Kyvik KO, Backer V. Etiological relationships in atopy: a review of twin studies. *Twin Res Hum Genet* 2008;11:112-20.
- Thomsen SF, Ulrik CS, Kyvik KO, Ferreira MA, Backer V. Multivariate genetic analysis of atopy phenotypes in a selected sample of twins. *Clin Exp Allergy* 2006;36:1382-90.
- Willemsen G, van Beijsterveldt TC, van Baal CG, Postma D, Boomsma DI. Heritability of self-reported asthma and allergy: a study in adult Dutch twins, siblings and parents. *Twin Res Hum Genet* 2008;11:132-42.
- Wu T, Boezen HM, Postma DS, Los H, Postmus PE, Snieder H, et al. Genetic and environmental influences on objective intermediate asthma phenotypes in Dutch twins. *Eur Respir J* 2010;36:261-8.
- 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.
- Ober C, Yao TC. The genetics of asthma and allergic disease: a 21st century perspective. *Immunol Rev* 2011;242:10-30.
- Repapi E, Sayers I, Wain LV, Burton PR, Johnson T, Obeidat M, et al. Genome-wide association study identifies five loci associated with lung function. *Nat Genet* 2010;42:36-44.
- Hancock DB, Eijgelsheim M, Wilk JB, Gharib SA, Loehr LR, Marcianti KD, et al. Meta-analyses of genome-wide association studies identify multiple loci associated with pulmonary function. *Nat Genet* 2010;42:45-52.
- Gudbjartsson DF, Bjornsdottir US, Halapi E, Helgadóttir A, Sulem P, Jonsdóttir GM, et al. Sequence variants affecting eosinophil numbers associate with asthma and myocardial infarction. *Nat Genet* 2009;41:342-7.
- Williams SM, Haines JL. Correcting away the hidden heritability. *Ann Hum Genet* 2011;75:348-50.
- de los Campos G, Gianola D, Allison DB. Predicting genetic predisposition in humans: the promise of whole-genome markers. *Nat Rev Genet* 2010;11:880-6.
- Yang J, Benyamin B, McEvoy BP, Gordon S, Henders AK, Nyholt DR, et al. Common SNPs explain a large proportion of the heritability for human height. *Nat Genet* 2010;42:565-9.
- Evans DM, Visscher PM, Wray NR. Harnessing the information contained within genome-wide association studies to improve individual prediction of complex disease risk. *Hum Mol Genet* 2009;18:3525-31.
- Purcell SM, Wray NR, Stone JL, Visscher PM, O'Donovan MC, Sullivan PF, et al. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* 2009;460:748-52.
- Golding J, Pembrey M, Jones R, Team AS. ALSPAC—the Avon Longitudinal Study of Parents and Children. I. Study methodology. *Paediatr Perinat Epidemiol* 2001;15:74-87.
- Jones RW, Ring S, Tyfield L, Hamvas R, Simmons H, Pembrey M, et al. A new human genetic resource: a DNA bank established as part of the Avon longitudinal study of pregnancy and childhood (ALSPAC). *Eur J Hum Genet* 2000;8:653-60.
- 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.
- Savenije OE, Granell R, Caudri D, Koppelman GH, Smit HA, Wijga A, et al. Comparison of childhood wheezing phenotypes in 2 birth cohorts: ALSPAC and PIAMA. *J Allergy Clin Immunol* 2011;127:1505-12, e14.
- Yan K, Salome C, Woolcock AJ. Rapid method for measurement of bronchial responsiveness. *Thorax* 1983;38:760-5.
- Steyerberg EW. *Clinical prediction models—a practical approach to development, validation, and updating*. New York: Springer; 2009.
- Ferreira MA, Matheson MC, Duffy DL, Marks GB, Hui J, Le Souef P, et al. Identification of IL6R and chromosome 11q13.5 as risk loci for asthma. *Lancet* 2011;378:1006-14.
- Xu M, Tantisira KG, Wu A, Litonjua AA, Chu JH, Himes BE, et al. Genome wide association study to predict severe asthma exacerbations in children using random forests classifiers. *BMC Med Genet* 2011;12:90.
- Bureau A, Dupuis J, Falls K, Lunetta KL, Hayward B, Keith TP, et al. Identifying SNPs predictive of phenotype using random forests. *Genet Epidemiol* 2005;28:171-82.
- Wray NR, Yang J, Goddard ME, Visscher PM. The genetic interpretation of area under the ROC curve in genomic profiling. *PLoS Genet* 2010;6:e1000864.
- Bouzigon E, Corda E, Aschard H, Dizier MH, Boland A, Bousquet J, et al. Effect of 17q21 variants and smoking exposure in early-onset asthma. *N Engl J Med* 2008;359:1985-94.
- Halapi E, Gudbjartsson DF, Jonsdóttir GM, Bjornsdottir US, Thorleifsson G, Helgadóttir H, et al. A sequence variant on 17q21 is associated with age at onset and severity of asthma. *Eur J Hum Genet* 2010;18:902-8.
- Binia A, Khorasani N, Bhavsar PK, Adcock I, Brightling CE, Chung KF, et al. Chromosome 17q21 SNP and severe asthma. *J Hum Genet* 2011;56:97-8.
- Himes BE, Hunninghake GM, Baurley JW, Rafaels NM, Sleiman P, Strachan DP, et al. Genome-wide association analysis identifies PDE4D as an asthma-susceptibility gene. *Am J Hum Genet* 2009;84:581-93.
- Sleiman PM, Flory J, Imielinski M, Bradfield JP, Annaiah K, Willis-Owen SA, et al. Variants of DENND1B associated with asthma in children. *N Engl J Med* 2010;362:36-44.
- Torgerson DG, Ampleford EJ, Chiu GY, Gauderman WJ, Gignoux CR, Graves PE, et al. Meta-analysis of genome-wide association studies of asthma in ethnically diverse North American populations. *Nat Genet* 2011;43:887-92.
- Hirota T, Takahashi A, Kubo M, Tsunoda T, Tomita K, Doi S, et al. Genome-wide association study identifies three new susceptibility loci for adult asthma in the Japanese population. *Nat Genet* 2011;43:893-6.
- Noguchi E, Sakamoto H, Hirota T, Ochiai K, Imoto Y, Sakashita M, et al. Genome-wide association study identifies HLA-DP as a susceptibility gene for pediatric asthma in Asian populations. *PLoS Genet* 2011;7:e1002170.

METHODS

Genotyping

Genotyping of ALSPAC children was carried out at 2 different centers (Wellcome Trust Sanger Centre, Cambridge, United Kingdom, and Laboratory Corporation of America, Burlington, NC) using the Illumina Human550 quad array (Illumina, San Diego, Calif).

Quality control

Subjects were excluded on the basis of sex mismatches, minimal or excessive heterozygosity, disproportionate levels of individual missingness ($>3\%$), and cryptic relatedness measured as a proportion of identity by descent (>0.1). The remaining subjects were assessed for evidence of population stratification by means of multidimensional scaling analysis and compared with the HapMap II (release 22) European descent (CEU), Han Chinese, Japanese, and Yoruba reference populations; all subjects with non-European ancestry were removed. SNPs with a minor allele frequency of less than 1%, a call rate of less than 95%, or evidence for violations of Hardy-Weinberg equilibrium ($P < 5 \times 10^{-7}$) were removed.

Construction of genetic scores

GWAS results for asthma from the GABRIEL consortium study were downloaded from <http://www.cng.fr/gabriel/>. Details on the methods used in that study are published elsewhere.^{E1} We included all GABRIEL childhood cohorts except ALSPAC: B58C (United Kingdom: 213 cases and 200 control subjects), BAMSE (Sweden: 239 cases and 246 control subjects), Busselton (Australia: 188 cases and 390 control subjects), CAPPS and SAGE (Canada: 88 cases and 182 control subjects), ECRHS (multiple countries: 279 cases and 620 control subjects), EGEA (France: 482 cases and 598 control subjects), GABRIEL Advanced Surveys (Germany: 841 cases and 851 control subjects), GAIN (multiple countries: 1475 cases and 1668 control subjects), Karelia Allergy Study (Finland: 33 cases and 36 control subjects), KSMU (Russia: 112 cases and 116 control subjects), MAGICs and MAS (Germany: 801 cases and 572 control subjects), MRCA-UKC (United Kingdom: 177 cases and 399 control subjects), PIAMA (The Netherlands, 172 cases and 187 control subjects), SAPALDIA (Switzerland: 237 cases and 356 control subjects), SLSJ Familial Connection (Canada: 373 cases and 390 control subjects), TOMSK (Russia: 197 cases and 91 control subjects), and UFA (Russia: 269 cases and 209 control subjects). In total, this dataset (hereafter referred to as the “discovery set”) included 6176 cases and 7111 control subjects. We included SNPs fulfilling the following quality control criteria in at least 1 cohort: minor allele frequency of 5% or greater, consistency with Hardy-Weinberg equilibrium by 1 *df* tests in control subjects ($P > 10^{-4}$), and a call rate of 97% or greater in 1 or more cohorts ($P > 10^{-4}$, variable “select5” = 1 in results file; see description of results file at <http://www.cng.fr/gabriel/>). Pooled effect estimates and SEs were computed for all included SNPs using a random effects model and the “metan” command in Stata software (Stata 10; StataCorp, College Station, Tex).

On the basis of these results, we computed allelic scores for children in ALSPAC. SNPs satisfying the above quality control criteria in ALSPAC were then ranked in inverse order of *P* values of pooled logORs obtained from the discovery dataset. We calculated allelic scores as the weighted sum of allele dosages over the *k* highest-ranking SNPs, with *k* taking values of 10, 22, 46, 100, 215, 464, 1,000, 2,154, 4,642, 10,000, 21,544, 46,416, 100,000, and 215,443. These numbers are obtained by successively multiplying by $10^{0.33}$ and were chosen to achieve an even spacing on the log scale. Two types of scores were calculated: log odds scores using pooled logORs as weights for the allele count and allele count scores using the signs of the logORs (+1 if logOR > 0 and -1 if logOR < 0) as weights. In the case of missing genotypes, allele counts were imputed with their expected value based on the sample allele frequency. We also computed scores based on independent

sets of SNPs with rankings between consecutive values of *k* (ie, SNPs ranked 1-10, 11-22, 23-46, ..., 100,001-215,443). The prediction scores were constructed using the PLINK program (version 1.07).^{E2}

Outcome measurements in mid-childhood

Spirometry. Lung function was measured by spirometry (Vitalograph 2120), according to American Thoracic Society criteria.^{E3} Flow-volume curves were reviewed by 1 respiratory physician (JH) to ensure adherence to standards, resulting in the rejection of 338 (4.6%) and the correction of 883 (11.5%) measurements in which the automated program had selected an inappropriate curve. Variables (FEV₁ and FVC) were converted to sex-, age- and height-adjusted *z* scores.^{E4} For both FEV₁ and FVC values, children were classified according to whether they were in the lowest quintiles of the *z* scores.

Bronchial challenge tests. Children whose FEV₁ was greater than 70% of predicted value for height participated in methacholine challenge tests. Airway responsiveness to methacholine was measured by using the method of Yan et al^{E5} and expressed for each subject as the dose-response slope of FEV₁ (percentage decrease from baseline) per millimole of methacholine.

The following conditions meant that the child was unsuitable for testing:

- chest infection in the last 3 weeks treated by a general practitioner with antibiotics;
- children with FEV₁ of less than 70% of predicted value;
- acute asthma in the past 3 weeks treated with prednisolone; or
- upper respiratory tract infection or cold in the last 3 weeks.

Children with a recent cold or asthma episode were invited back for a future visit. BHR was defined as being in the highest tertile of the dose-response slope.

SPTs. The atopic status of the children was determined at 7 to 8 years of age based on SPT responses to a panel of up to 12 common allergens, including house dust mite (*Dermatophagoides pteronyssinus*), mixed grasses, and cat (ALK-Abelló, Hørsholm, Denmark). Sensitization to one of these 3 allergens has been shown to identify 95% of all sensitized children in this population.^{E6} A positive response was defined as a mean wheal diameter of greater than 2 mm with an absent response to negative control solution, and atopy was defined as a positive response to 1 or more of house dust mite, cat, or grass pollen.

Blood pressure. Blood pressure was measured with a Dinamap 9301 Vital Signs Monitor (GE Healthcare, Fairfield, Conn). Children were classified according to whether they were in the highest quintile of systolic blood pressure.

IQ. The Wechsler Intelligence Scale for Children (WISC), 3rd edition, UK Edition^{E7} was used to assess cognitive function. A short form of the measure was used in which alternate items (always starting with item number 1 in the standard form) were used for all subtests, with the exception of the coding subtest, which was administered in its full form. The 10 WISC subtests comprise 5 verbal subtests (information, similarities, arithmetic, vocabulary, and comprehension) and 5 performance subtests (picture completion, coding, picture arrangement, block design, and object assembly). Raw scores were calculated according to the items used in the alternate item form of the WISC. This was achieved by summing the individual items within each subtest and multiplying by 2 for picture completion, information, arithmetic, vocabulary, comprehension, and picture arrangement; multiplying by 5/3 for similarities; and multiplying by 3/2 for object assembly and block design, thus making the raw scores comparable with those that would have been obtained had the full test been administered. By using the look-up tables provided in the WISC manual, age-scaled scores were obtained from the raw scores, and the total IQ score was calculated as the sum of all scores (verbal and performance). Children were classified according to whether they were in the highest quintile of total IQ score.

REFERENCES

- E1. 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.
- E2. 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.
- E3. Standardization of spirometry, 1994 update. American Thoracic Society. *Am J Respir Crit Care Med* 1995;152:1107-36.
- E4. Chinn S, Rona RJ. Height and age adjustment for cross sectional studies of lung function in children aged 6-11 years. *Thorax* 1992;47:707-14.
- E5. Yan K, Salome C, Woolcock AJ. Rapid method for measurement of bronchial responsiveness. *Thorax* 1983;38:760-5.
- E6. Roberts G, Peckitt C, Northstone K, Strachan D, Lack G, Henderson J, et al. Relationship between aeroallergen and food allergen sensitization in childhood. *Clin Exp Allergy* 2005;35:933-40.
- E7. Wechsler D, Golombok S, Rust J. WISC-III^{UK} Wechsler Intelligence Scale for Children—Third Edition UK Manual. Sidcup (United Kingdom): The Psychological Corporation; 1992.
- E8. Himes BE, Hunninghake GM, Baurley JW, Rafaels NM, Sleiman P, Strachan DP, et al. Genome-wide association analysis identifies PDE4D as an asthma-susceptibility gene. *Am J Hum Genet* 2009;84:581-93.
- E9. Sleiman PM, Flory J, Imielinski M, Bradfield JP, Annaiah K, Willis-Owen SA, et al. Variants of DENND1B associated with asthma in children. *N Engl J Med* 2010;362:36-44.
- E10. Torgerson DG, Ampleford EJ, Chiu GY, Gauderman WJ, Gignoux CR, Graves PE, et al. Meta-analysis of genome-wide association studies of asthma in ethnically diverse North American populations. *Nat Genet* 2011;43:887-92.
- E11. Hirota T, Takahashi A, Kubo M, Tsunoda T, Tomita K, Doi S, et al. Genome-wide association study identifies three new susceptibility loci for adult asthma in the Japanese population. *Nat Genet* 2011;43:893-6.
- E12. Noguchi E, Sakamoto H, Hirota T, Ochiai K, Imoto Y, Sakashita M, et al. Genome-wide association study identifies HLA-DP as a susceptibility gene for pediatric asthma in Asian populations. *PLoS Genet* 2011;7:e1002170.
- E13. Ferreira MA, Matheson MC, Duffy DL, Marks GB, Hui J, Le Souef P, et al. Identification of IL6R and chromosome 11q13.5 as risk loci for asthma. *Lancet* 2011;378:1006-14.

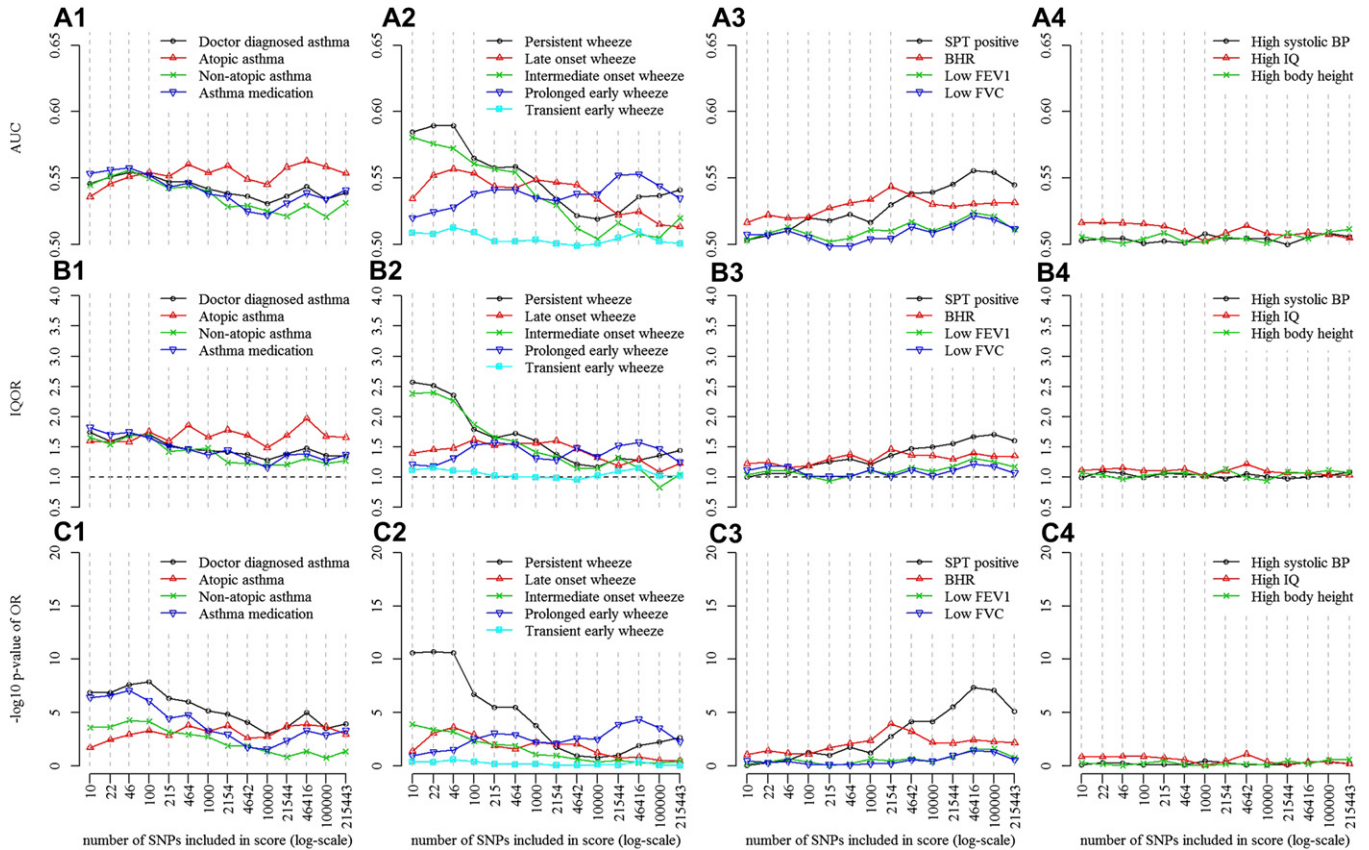


FIG E1. Analysis with allele count scores: comparison of predictive value for outcomes in mid-childhood. AUC (A1-A4), IQOR (B1-B4), and P values for association (C1-C4) are shown. Outcomes are based on a physician's diagnosis (A1-C1), phenotypes of wheeze in the first 7 years of life (A2-C2), physiologic measurements (A3-C3), and nonasthma outcomes (A4-C4). SNPs are included in inverse order of their P values in an independent GWAS for physician-diagnosed asthma.

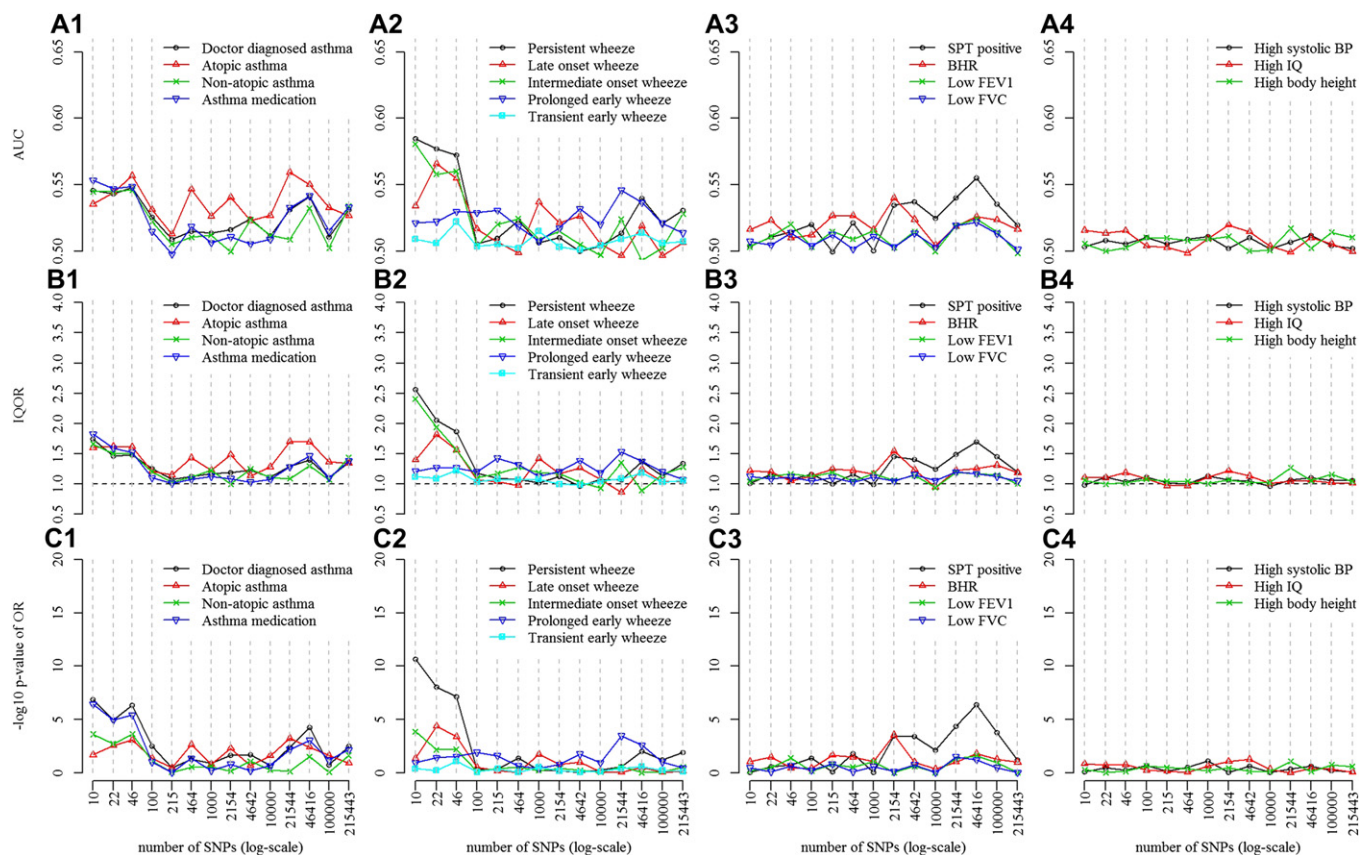


FIG E2. Analysis with scores based on independent sets of SNPs: comparison of predictive value for outcomes in mid-childhood. AUC (A1-A4), IQOR (B1-B4), and P values for association (C1-C4) are shown. Outcomes are based on a physician's diagnosis (A1-C1), phenotypes of wheeze in the first 7 years of life (A2-C2), physiologic measurements, and nonasthma outcomes (A4-C4). SNPs are included in inverse order of their P values in an independent GWAS for physician-diagnosed asthma. Scores include SNPs ranked between 2 sequential numbers on the y -axis; for example, the score corresponding to the number 1000 includes SNPs ranked from 465 to 1000.

TABLE E1. First 100 SNPs included for prediction from pooled GWAS results of GABRIEL childhood cohorts excluding ALSPAC

Order	Chromosome	Gene	rs Number	Position	Ref. allele	Alt. allele	Freq. alt. allele	logOR	P value
1	17	<i>GSDMB</i>	rs2305480	35315722	A	G	0.53	0.27	1.0E-20
2	17	<i>ZPBP2</i>	rs2872507	35294289	A	G	0.52	0.26	1.1E-19
3	17	<i>ORMDL3</i>	rs4795405	35341943	C	T	0.46	-0.26	3.0E-19
4	17	<i>GSDMB</i>	rs8067378	35304874	A	G	0.52	-0.25	3.7E-19
5	17	<i>ZNFN1A3</i>	rs9303277	35229995	C	T	0.52	-0.25	4.2E-19
6	17	<i>ZPBP2</i>	rs11557467	35282160	G	T	0.52	-0.25	1.0E-18
7	17	<i>ZNFN1A3</i>	rs907092	35175785	A	G	0.52	0.25	3.3E-18
8	17	<i>GSDMB</i>	rs7216389	35323475	C	T	0.48	0.26	8.9E-17
9	17	<i>ORMDL3</i>	rs8079416	35346239	C	T	0.55	-0.23	3.4E-16
10	17	<i>GSDM1</i>	rs3894194	35375519	A	G	0.55	-0.23	4.1E-16
11	17	<i>GSDMB</i>	rs2290400	35319766	C	T	0.49	0.25	9.2E-16
12	17	<i>GSDM1</i>	rs3902025	35372780	G	T	0.56	0.23	3.6E-15
13	17	<i>GSDM1</i>	rs1007654	35364880	A	G	0.62	0.22	2.4E-13
14	17	<i>GSDM1</i>	rs7219080	35368042	A	C	0.62	0.22	4.4E-13
15	17	<i>GSDM1</i>	rs3859192	35382174	C	T	0.47	0.20	1.0E-12
16	17	<i>THRAP4</i>	rs11078936	35451440	C	T	0.63	0.18	1.1E-09
17	17	<i>PSMD3</i>	rs8075668	35391149	C	T	0.46	-0.17	3.0E-09
18	17	<i>THRAP4</i>	rs9916158	35435755	G	T	0.37	-0.17	4.1E-09
19	2	<i>IL1RL1</i>	rs10206753	102334794	C	T	0.62	0.17	6.1E-09
20	2	<i>IL18R1</i>	rs3771166	102352654	A	G	0.62	0.17	6.2E-09
21	2	<i>IL1RL1</i>	rs10204137	102334644	A	G	0.38	-0.17	7.2E-09
22	2	<i>IL1RL1</i>	rs10192157	102334788	C	T	0.38	-0.17	7.7E-09
23	7		rs6967330	105445687	A	G	0.83	-0.22	1.2E-08
24	9	<i>IL33</i>	rs1342326	6180076	A	C	0.17	0.25	4.4E-08
25	17		rs868150	35466885	A	G	0.59	-0.15	7.9E-08
26	17	<i>STARD3</i>	rs1877031	35067606	A	G	0.34	0.16	8.5E-08
27	17	<i>ORMDL3</i>	rs3744246	35337876	C	T	0.22	-0.19	1.6E-07
28	17	<i>STARD3</i>	rs2271308	35071008	C	T	0.27	0.16	2.0E-07
29	17	<i>PSMD3</i>	rs8070454	35414280	C	T	0.39	0.15	2.0E-07
30	17		rs9635726	35273667	C	T	0.20	0.18	2.4E-07
31	17		rs4252627	35122241	C	T	0.67	-0.16	5.9E-07
32	2	<i>IL18R1</i>	rs1420094	102382119	C	T	0.48	-0.14	6.0E-07
33	2	<i>IL1RL1</i>	rs13431828	102321085	C	T	0.14	-0.21	7.9E-07
34	5	<i>IL13</i>	rs1295686	132023742	C	T	0.19	0.17	7.9E-07
35	8		rs10503648	19062324	G	T	0.82	-0.19	8.7E-07
36	2	<i>IL18R1</i>	rs2287033	102377669	C	T	0.52	0.14	1.2E-06
37	9	<i>IL33</i>	rs2066362	6209176	G	T	0.16	0.25	1.2E-06
38	5	<i>IL13</i>	rs20541	132023863	A	G	0.82	-0.17	1.3E-06
39	2	<i>IL18R1</i>	rs4851004	102375969	C	T	0.48	-0.14	1.3E-06
40	15		rs1317060	79193470	C	T	0.33	-0.15	1.3E-06
41	17	<i>PERLD1</i>	rs2941504	35084426	A	G	0.69	-0.17	1.6E-06
42	17	<i>PERLD1</i>	rs2941503	35082271	A	G	0.69	-0.17	1.7E-06
43	5		rs2303718	112954654	A	G	0.95	0.52	1.7E-06
44	17	<i>ERBB2</i>	rs1810132	35119531	C	T	0.68	-0.16	1.7E-06
45	17		rs12453682	35023531	C	T	0.69	-0.15	1.7E-06
46	2	<i>IL1RL2</i>	rs11674302	102253560	C	T	0.87	0.21	1.8E-06
47	17	<i>PERLD1</i>	rs1565922	35084561	A	G	0.69	-0.17	2.0E-06
48	12		rs10492323	92616130	C	T	0.94	0.29	2.3E-06
49	2	<i>IL18RAP</i>	rs3755266	102409144	A	G	0.52	0.14	2.5E-06
50	2	<i>IL18RAP</i>	rs2310300	102415506	A	G	0.48	-0.14	2.8E-06
51	9		rs406322	6136121	C	T	0.26	0.16	2.8E-06
52	15	<i>SMAD3</i>	rs744910	65233839	A	G	0.49	0.13	2.9E-06
53	12		rs2364871	92626260	C	T	0.06	-0.28	2.9E-06
54	17		rs3764351	35077865	A	G	0.35	0.15	3.1E-06
55	6		rs16894934	163725732	A	G	0.16	0.18	3.1E-06
56	15	<i>SMAD3</i>	rs12708492	65254595	C	T	0.48	0.13	3.5E-06
57	7		rs10488047	105452573	C	T	0.11	0.22	3.6E-06
58	17		rs4795402	35338911	A	C	0.74	0.16	3.7E-06
59	12		rs17732565	92600582	A	G	0.06	-0.31	4.0E-06
60	5	<i>RAD50</i>	rs6871536	131997773	C	T	0.81	-0.15	6.3E-06
61	6		rs9322206	150004310	C	T	0.65	-0.13	6.9E-06
62	5	<i>RAD50</i>	rs2240032	132005026	C	T	0.19	0.15	7.0E-06

(Continued)

TABLE E1. (Continued)

Order	Chromosome	Gene	rs Number	Position	Ref. allele	Alt. allele	Freq. alt. allele	logOR	P value
63	15	IL33	rs1533055	79197161	A	G	0.35	-0.14	7.2E-06
64	9		rs928413	6203387	A	G	0.26	0.22	8.5E-06
65	12		rs17021577	92711110	A	C	0.94	0.32	9.4E-06
66	6	IL33	rs3763197	163759002	C	T	0.85	-0.17	1.1E-05
67	6		rs1764020	163724957	C	T	0.68	-0.13	1.1E-05
68	9		rs992969	6199697	A	G	0.74	-0.22	1.4E-05
69	1		rs4660531	41612409	G	T	0.34	-0.13	1.6E-05
70	8		rs952559	81451189	A	G	0.65	-0.14	1.6E-05
71	22		rs738932	45671152	A	G	0.74	0.15	1.7E-05
72	6		rs6903998	150194186	C	T	0.64	-0.13	1.8E-05
73	6		rs3798761	150046472	A	G	0.35	0.13	1.8E-05
74	6		rs9393175	150040416	A	G	0.35	0.13	1.8E-05
75	11		rs471931	70187425	C	T	0.54	0.12	1.9E-05
76	20	SLC22A5	rs10153982	38331069	A	G	0.31	-0.14	2.0E-05
77	22		rs140535	45650411	C	T	0.27	-0.16	2.0E-05
78	6		rs12529698	150013780	A	G	0.65	-0.13	2.1E-05
79	6		rs2275046	150198694	A	G	0.36	0.12	2.1E-05
80	1		rs4660192	41614937	C	T	0.53	-0.12	2.6E-05
81	6		rs9767555	150009851	C	T	0.64	-0.12	2.6E-05
82	1		rs12145783	42918044	A	G	0.12	0.19	2.8E-05
83	22		rs16988828	29333561	A	G	0.10	-0.20	3.0E-05
84	1		rs864537	165678008	A	G	0.42	-0.12	3.2E-05
85	6		rs1934534	150074725	C	T	0.65	-0.12	3.3E-05
86	16	SLC22A5	rs12934193	11011226	C	T	0.82	0.17	3.5E-05
87	6		rs9689723	150018764	A	C	0.65	-0.12	3.8E-05
88	8		rs7822328	81409214	A	G	0.63	-0.13	4.0E-05
89	8		rs6473225	81443562	A	G	0.38	0.13	4.2E-05
90	6		rs10872646	150109368	A	G	0.35	0.12	4.2E-05
91	14		rs7145790	25823876	G	T	0.10	-0.19	4.5E-05
92	10		rs10763419	58234737	C	T	0.46	0.12	4.6E-05
93	6		rs9689084	150002477	A	G	0.65	-0.12	4.6E-05
94	22		rs5769204	45664150	A	G	0.70	0.14	4.6E-05
95	6		rs9766886	149999877	C	T	0.65	-0.12	4.8E-05
96	5	SLC22A5	rs2073643	131751187	C	T	0.44	0.12	4.8E-05
97	6		rs1108889	150208262	A	G	0.34	0.12	5.2E-05
98	17		rs8064360	62324021	A	G	0.85	-0.16	5.4E-05
99	5		rs4704524	78144211	C	T	0.90	0.20	5.5E-05
100	1		rs12121067	42907042	C	T	0.12	0.18	5.6E-05

Alt., Alternative; Freq., frequency; Ref., reference.

TABLE E2. Pooled GWAS results of GABRIEL childhood cohorts excluding ALSPAC for specific SNPs reported in previous GWASs for asthma

Study	rs Number	Chromosome	Position	Gene	Ref. allele	Alt. allele	Freq. alt. allele	OR	95% CI	P value	Rank*
Himes et al ^{E8}	rs1588265	5	59405551	<i>PDE4D</i>	A	G	0.30	1.06	1.00-1.13	4.6E-02	17,975
Moffatt et al ^{E1}	rs3771166	2	102352654	<i>IL18R1</i>	A	G	0.62	1.19	1.12-1.26	6.2E-09	20
Moffatt et al ^{E1}	rs2073643	5	131751187	<i>SLC22A5</i>	C	T	0.44	1.13	1.06-1.19	4.8E-05	96
Moffatt et al ^{E1}	rs1295686	5	132023742	<i>IL13</i>	C	T	0.19	1.18	1.11-1.27	7.9E-07	34
Moffatt et al ^{E1}	rs9273349	6	32733847	<i>HLA-DQ</i>	C	T	1.00	0.87	0.81-0.93	9.9E-05	†
Moffatt et al ^{E1}	rs1342326	9	6180076	<i>IL33</i>	A	C	0.17	1.28	1.17-1.40	4.4E-08	24
Moffatt et al ^{E1}	rs11071559	15	58857280	<i>RORA</i>	C	T	0.13	0.89	0.82-0.96	4.5E-03	1,862
Moffatt et al ^{E1}	rs744910	15	65233839	<i>SMAD3</i>	A	G	0.49	1.14	1.08-1.21	2.9E-06	52
Moffatt et al ^{E1}	rs2305480	17	35315722	<i>GSDMB</i>	A	G	0.53	1.31	1.24-1.39	1.0E-20	1
Himes et al ^{E8}	rs3894194	17	35375519	<i>GSDMA</i>	A	G	0.55	0.79	0.75-0.84	4.1E-16	10
Moffatt et al ^{E1}	rs2284033	22	35863980	<i>IL2RB</i>	A	G	1.00	1.10	1.04-1.16	1.5E-03	†
Sleiman et al ^{E9}	rs1775456	1	195999678	<i>DENND1B</i>	A	G	0.21	1.01	0.94-1.08	7.7E-01	†
Torgerson et al ^{E10} Hirota et al ^{E11}	rs1837253	5	110429771	<i>TSLP</i>	C	T	0.26	0.87	0.79-0.96	5.2E-03	2,130
Hirota et al ^{E11}	rs7686660	4	144222609	<i>USP38-GAB1</i>	G	T	0.75	1.06	1.00-1.13	7.1E-02	27,993
Hirota et al ^{E11}	rs404860	6	32292323	<i>NOTCH4</i>	C	T	0.82	1.10	0.97-1.25	1.3E-01	52,884
Hirota et al ^{E11}	rs10508372	10	9012024		A	G	0.94	1.12	0.97-1.29	1.1E-01	43,459
Hirota et al ^{E11}	rs1701704	12	54698754		G	T	0.66	0.96	0.89-1.05	3.8E-01	157,171
Noguchi et al ^{E12}	rs987870	6	33150858	<i>HLA-DPA1</i>	A	G	0.14	1.09	1.01-1.18	3.5E-02	13,454
Ferreira et al ^{E13}	rs4129267	1	152692888	<i>ILR6</i>	C	T	0.41	1.06	0.98-1.14	1.2E-01	49,504
Ferreira et al ^{E13}	rs7130588	11	75948331	<i>LRR32</i>	A	G	0.37	1.10	1.01-1.20	3.2E-02	12,607

*Rank according to inverse of *P* value. This represents the order in which the allele was included in scores.

†SNP not genotyped and therefore not included in score.