Shared and distinct genetic risk factors for childhood-onset and adult-onset asthma: genome-wide and transcriptome-wide studies

Milton Pividori, Nathan Schoettler*, Dan L Nicolae, Carole Ober, Hae Kyung Im*

Summary

Background Childhood-onset and adult-onset asthma differ with respect to severity and comorbidities. Whether they also differ with respect to genetic risk factors has not been previously investigated in large samples. The goals of this study were to identify shared and distinct genetic risk loci for childhood-onset and adult-onset asthma, and to identify the genes that might mediate the effects of associated variation.

Methods We did genome-wide and transcriptome-wide studies, using data from the UK Biobank, in individuals with asthma, including adults with childhood-onset asthma (onset before 12 years of age), adults with adult-onset asthma (onset between 26 and 65 years of age), and adults without asthma (controls; aged older than 38 years). We did genome-wide association studies (GWAS) for childhood-onset asthma and adult-onset asthma each compared with shared controls, and for age of asthma onset in all asthma cases, with a genome-wide significance threshold of $p<5\times10^{-8}$. Enrichment studies determined the tissues in which genes at GWAS loci were most highly expressed, and PrediXcan, a transcriptome-wide gene-based test, was used to identify candidate risk genes.

Findings Of 376 358 British white individuals from the UK Biobank, we included 37 846 with self-reports of doctor-diagnosed asthma: 9433 adults with childhood-onset asthma; 21 564 adults with adult-onset asthma; and an additional 6849 young adults with asthma with onset between 12 and 25 years of age. For the first and second GWAS analyses, 318 237 individuals older than 38 years without asthma were used as controls. We detected 61 independent asthma loci: 23 were childhood-onset specific, one was adult-onset specific, and 37 were shared. 19 loci were associated with age of asthma onset. The most significant asthma-associated locus was at 17q12 (odds ratio 1·406, 95% CI 1·365–1·448; $p=1\cdot45\times10^{-111}$) in the childhood-onset GWAS. Genes at the childhood onset-specific loci were most highly expressed in skin, blood, and small intestine; genes at the adult onset-specific loci were most highly expressed in lung, blood, small intestine, and spleen. PrediXcan identified 113 unique candidate genes at 22 of the 61 GWAS loci. Single-nucleotide polymorphism-based heritability estimates were more than three times larger for childhood-onset asthma (0·327) than for adult-onset disease (0·098). The onset of disease in childhood was associated with additional genes with relatively large effect sizes, with the largest odds ratio observed at the *FLG* locus at 1q21.3 (1·970, 95% CI 1·823–2·129).

Interpretation Genetic risk factors for adult-onset asthma are largely a subset of the genetic risk for childhood-onset asthma but with overall smaller effects, suggesting a greater role for non-genetic risk factors in adult-onset asthma. Combined with gene expression and tissue enrichment patterns, we suggest that the establishment of disease in children is driven more by dysregulated allergy and epithelial barrier function genes, whereas the cause of adult-onset asthma is more lung-centred and environmentally determined, but with immune-mediated mechanisms driving disease progression in both children and adults.

Funding US National Institutes of Health.

Lancet Respir Med 2019

Published Online
April 26, 2019
http://dx.doi.org/10.1016/S2213-2600(19)30055-4

See Online/Comment
http://dx.doi.org/10.1016/S2213-2600(19)30047-5

*Contributed equally

Department of Medicine (M Pividori PhD, N Schoettler MD, Prof D L Nicolae PhD, H K Im PhD), **Department of Human Genetics** (N Schoettler, Prof D L Nicolae, Prof C Ober PhD), **and Department of Statistics** (Prof D L Nicolae), **The University of Chicago, Chicago, IL, USA**

Correspondence to:
Dr Hae Kyung Im, Department of Medicine, The University of Chicago, Chicago, IL 60637, USA
haky@uchicago.edu

Introduction

Asthma is the most prevalent chronic respiratory disease worldwide.[1] The diagnosis of asthma is based on the presence of reversible airflow obstruction and clinical symptoms that include wheeze, cough, and shortness of breath. Despite these shared features, asthma probably comprises many different conditions. In particular, childhood-onset asthma and adult-onset asthma differ with respect to sex ratios, triggers of exacerbation, associated comorbidities, severity,[2,3] and potentially genetic risk factors.[4,5] For example, a genome-wide association study (GWAS)[6] showed that the most often replicated and most statistically significant single-nucleotide polymorphisms (SNPs) at the 17q12–21 locus are specific to asthma with onset of symptoms in early life. Only one asthma GWAS to date has been conducted on adult-onset asthma, and this was as a subanalysis of a larger sample that also included childhood-onset asthma and asthma with unknown age of onset. The GABRIEL Consortium included 1947 people with adult-onset

Research in context

Evidence before this study

We searched PubMed for genome-wide association studies of childhood-onset asthma, adult-onset asthma, and age of asthma onset from database inception to July 1, 2018, using the terms "asthma", "GWAS", and "age of onset", and manually searched those papers for age of onset-specific studies of asthma. No language restriction was applied. We included studies with more than 1000 cases and 1000 controls and considered only loci that met genome-wide significance ($p<5\times10^{-8}$). These previous studies reported five genome-wide significant loci for childhood-onset asthma, none for adult-onset asthma, and five for age to asthma onset. Genome-wide association studies in large samples that include both childhood-onset and adult-onset asthma identified many loci associated with asthma risk. However, little was known about the shared or distinct effects of those or other loci on age of asthma onset, or about the genes that might mediate the effects of loci associated with childhood-onset or adult-onset asthma.

Added value of this study

Leveraging the resources of the UK Biobank, we identified loci with both age of onset-specific effects and shared effects. We further showed a significantly greater contribution of genetic variation to childhood-onset asthma, implying a greater role for environmental risk factors in adult-onset asthma, and different biological pathways and tissue enrichments for genes at loci associated with childhood-onset versus adult-onset asthma.

Implications of all the available evidence

Our results suggest that childhood onset-specific loci and those associated with age of onset play a part in disease initiation, whereas the other associated loci reflect shared mechanisms of disease progression. The childhood onset-specific loci highlight skin as a primary target tissue for early-onset disease and support the idea that asthma in childhood is due to impaired barrier function in the skin and other epithelial surfaces.

asthma (defined as 16 years of age or older) and 3669 adult controls.[7] Overall, genome-wide significant variants in the combined sample of 10 365 cases and 16 110 controls revealed larger odds ratios (ORs) in the childhood-onset group than in the adult-onset group,[7] but no loci reached significance in the adult-onset cases, probably because of low power. Therefore, whether loci other than 17q12–21 contribute specifically to childhood-onset or adult-onset asthma remains unknown. Delineation of genetic risk factors that affect age of onset from those that are shared between childhood-onset and adult-onset asthma could provide insights into the molecular mechanisms contributing to different clinical manifestations of asthma that is diagnosed at different ages.

See Online for appendix

To address this question and directly compare genetic risk architectures of adult-onset and childhood-onset asthma, we used data from the UK Biobank, a large-scale prospective study collecting demographic, clinical, medical history, and genetic data for nearly 500 000 participants.[8] Our goals were to identify shared and distinct genetic risk loci for childhood-onset and adult-onset asthma, and to identify genes that might mediate the effects of associations at age of onset-specific loci and those shared between childhood-onset and adult-onset asthma.

Methods

Study design and participants

We did GWAS and transcriptome-wide studies across three categories: childhood-onset cases (compared with controls without asthma), adult-onset cases (compared with controls without asthma), and asthma cases with onset in childhood, adolescence and young adulthood, or adulthood (to assess age of onset).

For more on **Hail** see https://github.com/hail-is/hail

We used data for British white individuals from UK Biobank data release July 19, 2017.[8] We extracted disease status (asthma, allergic rhinitis, atopic dermatitis, food allergy, chronic obstructive pulmonary disease (COPD), emphysema, and chronic bronchitis), age of onset of asthma, and sex from self-reported questionnaires and hospital records (International Classification of Diseases 10th revision [ICD-10] codes) by querying our in-house protected UK Biobank database server.[9] For our main case analysis, we included individuals who self-reported that they had doctor-diagnosed asthma. Further details of our research approach are provided in the appendix (pp 4–7).

We defined childhood-onset and adult-onset asthma using strict age of onset criteria that would minimise the likelihood of misclassification; we considered asthma cases with onset at younger than 12 years as childhood-onset cases and those with onset between 26 years and 65 years of age as adult-onset cases. UK Biobank participants aged older than 38 years without an asthma diagnosis (ie, no self-reports and no ICD-10 codes) were included as controls. Individuals with COPD, emphysema, or chronic bronchitis (self-reports or ICD-10 codes) were excluded from adult-onset cases and controls. Additionally, for the age-of-onset GWAS, we included UK Biobank participants with asthma (according to the inclusion and exclusion criteria above) with onset between the ages of 12 and 25 years. Additional details on genotype quality control and phenotype definitions are provided in the appendix (pp 4–7).

Genome-wide association studies

We did a childhood-onset and an adult-onset GWAS by logistic regression and an age-of-onset GWAS by linear regression, all using the allele dosages under an additive genetic model, as implemented in Hail. We used the

Bionimbus Protected Data Cloud for all analyses.[10] In all three GWAS, we included sex and the first ten genetic principal components as covariates (to adjust for ancestry) and used a genome-wide significance threshold of $p<5\times10^{-8}$. We used FUMA,[11] an integrative post-GWAS annotation web-based tool, to define independent risk loci and identify enriched tissues. FUMA defines independent loci using linkage disequilibrium information from the 1000 Genomes Project.[12] We specified an r^2 threshold of more than 0·6 between genome-wide significant SNPs to represent a single locus. We calculated tissue enrichments in FUMA using a hypergeometric test to determine over-representation of genes mapped to risk loci (by physical distance) among highly expressed genes in each tissue in the Genotype-Tissue Expression project relative to all others.

Childhood-onset and adult-onset specific loci were defined as those that were genome-wide significant in either the childhood-onset or adult-onset GWAS but were not associated with asthma at p<0·05 in the other group and the 95% CIs of the respective ORs did not overlap. Shared loci were defined as those that were genome-wide significant in at least one of the childhood-onset or adult-onset GWAS and had p<0·05 in both GWAS and overlapping 95% CIs. For regions where the lead SNP differs between GWAS, the SNP with the lowest p value across all three GWAS is referenced. Colocalisation of variants associated with childhood-onset and adult-onset asthma were also tested using pairwise GWAS.[13] For this analysis, adult-onset GWAS results were input as phenotype 1 and childhood-onset GWAS results were input as phenotype 2.

SNP-based heritability estimation

We used linkage disequilibrium score regression to estimate the SNP-based heritability from the childhood-onset, adult-onset, and age-of-onset GWAS summary statistics. For this analysis, we used SNPs identified as risk loci in this study, and selected those that overlapped with known HapMap3 variants (a database of SNP population frequency),[14] with an estimate of population prevalence of 8·68% for childhood-onset asthma and 9·55% for adult-onset asthma.[15] Heritability estimates for binary traits are reported in the liability scale.

Sensitivity analysis

We did two sets of post-hoc sensitivity analyses to assess the effects of potential poor recall of age of onset among individuals with adult-onset asthma, and the effects of misclassification of COPD as asthma among the adult-onset cases, even with exclusion of cases with a reported diagnosis of COPD, emphysema, or chronic bronchitis. First, to assure that the adult-onset cases did not include a significant proportion of childhood-onset asthma in which symptoms remitted in early life but then relapsed in adulthood, we replaced adult-onset cases with increasing proportions of randomly selected childhood-onset cases, and then tested for association at the two most significant childhood onset-specific loci. This procedure was repeated 20 times for each proportion to quantify the sampling variability (appendix pp 7–8). Second, we did two analyses in which we removed either individuals with ages of asthma onset between 46 and 65 years or adult-onset cases and controls with FEV_1/FVC <0·70. For each, we compared p values and ORs with the GWAS including all adult-onset cases (appendix pp 8–9).

We also repeated the GWAS analyses while varying the numbers of principal components (10, 14, or 20; appendix p 12) and limiting the sample to cases with diagnoses based on ICD-10 codes (appendix pp 13–14). To test for potential effects of asthma diagnostic criteria on the GWAS results, we did a GWAS comparing individuals with asthma based on self-reported doctor diagnosis alone compared with those with self-reported doctor diagnosis plus ICD-10 codes.

For more on the **Genotype-Tissue Expression project** see https://gtexportal.org/home/

Predicted transcriptome association test

We used the PrediXcan[16] framework to identify genes that might mediate associations between genetic variants and asthma risk. PrediXcan is a software tool that estimates tissue-specific gene expression profiles from an individual's SNP genotype profile by use of prediction models trained in large reference databases of genotypes and tissue-specific gene expression profiles. With these genotype-imputed expression profiles, PrediXcan can perform gene-based association tests that correlate predicted expression levels with phenotypes (eg, asthma) to identify candidate causal genes from GWAS data. We used a summary version of PrediXcan, which has high concordance with the individual-level version ($r^2>0·99$).[17] For predictions, we downloaded elastic net models trained with reference transcriptome data from the

	Childhood onset (n=9433)	Adolescent or young adult onset (n=6849)	Adult onset (n=21 564)	Controls (n=318 237)
Age at recruitment, years	55 (8); 40–70	52 (8); 40–70	57 (8); 40–70	57 (8); 39–73
Age at asthma onset, years	6 (3); 0–11	19 (4); 12–25	44 (10); 26–65	NA
Sex				
Female	3840 (40·7%)	3934 (57·4%)	13 708 (63·6%)	170 144 (53·5%)
Male	5593 (59·3%)	2915 (42·6%)	7856 (36·4%)	148 093 (46·5%)
Asthma medication use in the past 12 months	1383 (14·7%)	1124 (16·4%)	4081 (18·9%)	NA
Current smokers	802 (8·5%)	688 (10·0%)	1379 (6·4%)	30 056 (9·4%)
Allergic diseases ever				
Allergic rhinitis*	2537 (26·9%)	2042 (29·8%)	4660 (21·6%)	27 289 (8·6%)
Atopic dermatitis*	1126 (11·9%)	473 (6·9%)	927 (4·3%)	7168 (2·3%)
Food allergy†	138 (1·5%)	63 (0·9%)	172 (0·8%)	1229 (0·4%)
Any allergic disease‡	3205 (34·0%)	2321 (33·9%)	5306 (24·6%)	33 808 (10·6%)

Data are mean (SD); range or n (%). NA=not applicable. *Based on self-reports and International Classification of Diseases 10th revision codes. †Based on self-reports. ‡Based on self-reports, International Classification of Diseases 10th revision codes, or both.

Table 1: **Baseline characteristics**

Genotype-Tissue Expression consortium[18] for 49 tissues (appendix pp 9, 47).

PrediXcan was run separately in the childhood-onset and adult-onset cases, each with the same controls. For the transcriptome-wide analysis, we analysed genes that were expressed in the tissues determined by FUMA to be enriched, using a Bonferroni correction to determine significance ($p<1{\cdot}29\times10^{-6}$). We defined childhood onset-specific genes as those with predicted expression significantly associated with childhood-onset asthma and the variants that predicted their expression using the PrediXcan models were within childhood onset-specific loci in our GWAS. Adult onset-specific genes were similarly defined. Because SNPs at shared loci might predict the expression of genes that are associated only in childhood-onset or adult-onset cases, we also considered genes to be age of onset-specific if they were significantly associated with asthma at $p<1{\cdot}29\times10^{-6}$ in one age group and not associated with asthma at $p<0{\cdot}05$ in the other. All other genes were considered to be shared.

Role of the funding source

The funder of the study had no role in study design; collection, analysis, or interpretation of data; in writing of the manuscript; in the decision to submit the paper for publication; or in determining who had access to the raw data. The corresponding author had full access to all of the data in the study and had final responsibility for the decision to submit for publication.

Results

Of 376 358 British white individuals from the UK Biobank, we included 37 846 with self-reports of doctor-diagnosed asthma: 9433 adults with childhood-onset asthma (3462 [36·7%] of whom also had hospital records with an asthma diagnosis [ICD-10 codes]); 21 564 adults with adult-onset asthma (9260 [42·9%] with ICD-10 codes); and an additional 6849 young adults with asthma with onset between the ages of 12 and 25 years of age (2797 [40·8%] with ICD-10 codes). For the first and second GWAS analyses, 318 237 individuals older than 38 years without asthma were used as controls. 20 275 individuals from UK Biobank did not have asthma but were excluded as controls because they had been diagnosed with COPD, emphysema, or chronic bronchitis. Characteristics of the total sample (n=37 846) are shown in table 1. After genotype quality control, 10 894 596 variants were available for analyses.

The GWAS of childhood-onset and adult-onset asthma revealed 61 independent loci associated with asthma. 56 were significant in the childhood-onset asthma GWAS, and 19 were significant in the adult-onset asthma GWAS ($p<5\times10^{-8}$; figure 1). 28 of the 61 loci, in 27 chromosomal regions, were not previously reported in the GWAS catalogue,[19] including one study also done in UK Biobank participants but focused on the phenotype

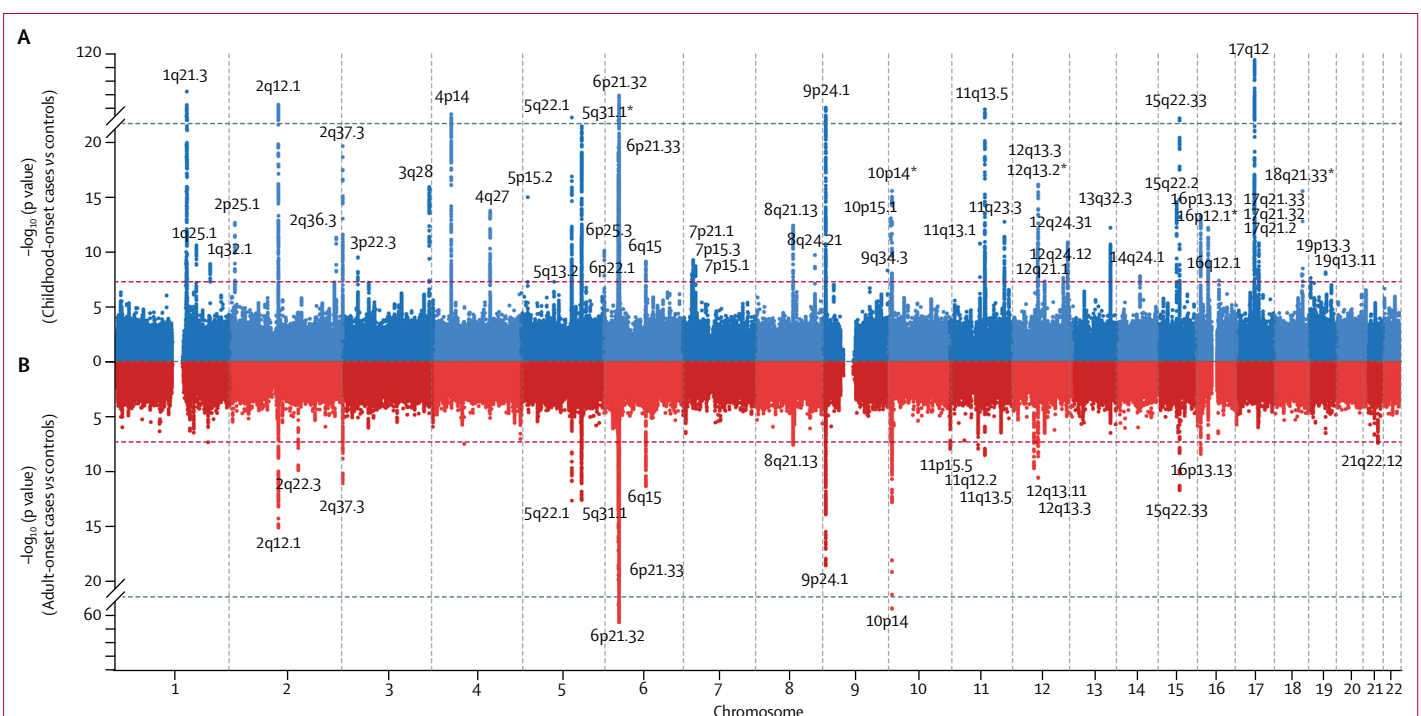

Figure 1: **Miami plot of childhood-onset and adult-onset asthma GWAS**
Miami plot showing results for the childhood-onset versus controls GWAS (A) and adult-onset versus controls GWAS (B). Each point corresponds to a SNP; the y-axes show the $-\log_{10}$ p values from the childhood-onset GWAS and adult-onset GWAS. The x-axis shows the position of each SNP along the 22 autosomes. The red dashed line indicates the genome-wide significance threshold (5×10^{-8}). See appendix p 41 for the age of asthma onset Manhattan plot. GWAS=genome-wide association study. SNP=single-nucleotide polymorphism. *Chromosomal bands with two independent loci.

asthma plus allergies.[20] Among the 28 new loci, 17 were significant in the childhood-onset GWAS, one was significant in the adult-onset GWAS, and ten were shared (table 2). Some of these loci contained genes that have been associated with asthma in candidate gene studies (eg, *FADS2*,[21] *MUC5AC*,[22,23] and *TBX21*[24]). The lead SNP or a linkage disequilibrium surrogate SNP (r^2>0·40) was reported for 23 of the 28 new loci in the Transnational Asthma Genetics Consortium GWAS.[25] SNPs at 16 of the 23 loci were associated with asthma in the Transnational Asthma Genetics Consortium GWAS (p<0·05) and had the same direction of effects as in the UK Biobank GWAS (appendix p 58).

As in previous GWAS comprising largely children,the most significant locus in the childhood-onset GWAS was at 17q12 (OR 1·406, 95% CI 1·365–1·448; p=$1·45\times10^{-111}$), with the lead SNP in *GSDMB* (figure 1A, table 2).[26] However, the estimated ORs for the lead SNPs at three other loci were similar to or larger than the OR for the lead SNP at the 17q12 locus. The lead SNPs in *IL1RL1* at 2q12.1 and in *EMSY* at 11q13.5, both of which were shared risk loci between adult-onset and childhood-onset asthma, had effect sizes on childhood-onset asthma similar to the lead SNP in *GSDMB* on 17q12 (table 2). Both loci have been prominent in previous asthma GWAS.[7,25] The lead SNP at the 1q21.3 locus, corresponding to a nonsense mutation (Arg501X) in the filaggrin (*FLG*) gene at 1q21.3, had the largest OR overall and was specific to childhood-onset asthma (OR 1·970, 95% CI 1·823–2·129; p=$1·88\times10^{-65}$; table 2). Variants in *FLG* have been robustly associated with food allergies[27–29] and atopic dermatitis,[30–32] but previous associations with asthma have been in the context of other allergic conditions.[20,33] Whether variants in *FLG* are associated with risk for childhood-onset asthma independent of its effects on early life allergic disease is yet unknown. To address this possibility, we repeated the childhood-onset GWAS after excluding 3205 individuals with childhood-onset asthma and 5785 controls who reported having a history of allergic rhinitis, atopic dermatitis, or food allergy. As expected in a smaller sample, the p values were larger overall, but the ORs were strikingly similar (appendix pp 16, 59–60). Even though the OR at the *FLG* locus on 1q21.3 decreased from 1·970 (95% CI 1·823–2·129) to 1·612 (1·492–1·743), it remained both highly significant (p=$2·45\times10^{-19}$) and the largest OR for childhood onset.

The most significant association in the adult-onset GWAS was in the HLA region, with independent associations at the HLA-C/B (6p21.33) and HLA-DR/DQ (6p21.32) loci, both of which were shared risk loci between adult-onset and childhood-onset asthma (figure 1B). Compared with the childhood-onset GWAS, effect sizes were quite small in adult-onset cases, with ORs reaching 1·1 at only six loci (2q12.1, 6p21.33, 6p21.32, 9p24.1, 10p14, and 19q13.11; table 2).

Among the 61 asthma loci, 23 were specific to childhood-onset asthma and one was specific to adult-onset asthma (table 2). Regional association plots for the 24 loci with childhood onset-specific or adult onset-specific effects are shown in the appendix (pp 17–40). Among the remaining 37 shared loci (table 2), mean ORs were larger in the childhood-onset cases at all but six loci (permutation test p<10^{-4}; appendix pp 8–9), indicating both that more loci contribute to childhood-onset asthma, and that even among shared loci the effect sizes are larger in childhood-onset asthma cases (figure 2). Colocalisation analyses using pairwise GWAS[13] yielded results that supported our classification of age of onset-specific and shared loci (appendix pp 61–62).

To directly test for loci associated with asthma age of onset, we did a third GWAS of all individuals with asthma in the UK Biobank who met our inclusion criteria (n=37 846), including those with onset during adolescence or young adulthood (n=6849). In this analysis, 19 loci were associated with age of onset (p<5×10^{-8}; appendix pp 41, 63). Age-of-onset loci overlapped with the childhood onset-specific, adult onset-specific, and shared loci, and asthma risk alleles at all but two loci were associated with earlier age of onset in the age-of-onset GWAS (11q12 in the *FADS2* gene and 12q13.11 near the *VDR* gene; table 2; figure 2). SNPs at the 1q21.3 locus (*FLG*) had the largest effect on age of onset, with an average 4·57 years (SE 0·43) earlier onset in individuals with a copy of the risk allele (rs61816761) than in individuals with the non-risk allele (p=$8·15\times10^{-27}$). At the 17q12 locus (rs4795399), each copy of the risk allele was associated on average with 2·29 years (SE 0·13) earlier onset than individuals without the risk allele (p=$6·76\times10^{-65}$), so individuals who were homozygous for the risk allele had on average 4·6 years earlier onset asthma than those who were homozygous for the non-risk allele. Examples of significant age-of-onset effects at these and other loci are shown in figure 3. Overall, both childhood onset-specific and shared asthma risk loci were associated with younger ages of onset than were adult onset-specifc loci, and alleles at loci associated with younger ages of onset had larger effects than did alleles at loci associated with later ages of onset.

We used linkage disequilibrium score regression to estimate the heritabilities of childhood-onset asthma, adult-onset asthma, and age of asthma onset. Consistent with the number of associated SNPs and their effect sizes, estimated heritabilities were 0·327 for childhood-onset asthma, 0·098 for adult-onset asthma, and 0·136 for age of asthma onset. After excluding the significant SNPs (table 2), estimates were reduced to 0·213 for childhood-onset asthma, 0·082 for adult-onset asthma, and 0·087 for age of asthma onset, indicating that the associated SNPs account for 0·114 of the variance in childhood-onset asthma risk, 0·016 of the variance in adult-onset risk, and 0·049 of the variance in age of asthma onset risk.

We assessed whether the tissue-specific expression of genes that mapped to the 56 childhood-onset loci differed from the tissue-specific expression of genes that mapped to the 19 adult-onset loci. Genes at childhood-onset loci

	rsID	Position*	Nearby genes†	Allele‡	RAF§	Childhood-onset asthma GWAS			Adult-onset asthma GWAS			Age-of-onset GWAS		
						OR	95% CI	p value	OR	95% CI	p value	β value	SE	p value
Childhood-onset asthma-specific loci														
1q21.3	rs61816761	1:152285861	*FLG, HRNR, FLG2*	G/A	0·024	1·970	1·823–2·129	$1·88×0^{-65}$	1·016	0·948–1·088	$6·56×10^{-01}$	−4·571	0·426	$8·15×10^{-27}$
1q25.1	rs7518129	1:173163568	*TNFSF4, TNSF18, PRDX6*	A/G	0·310	1·111	1·077–1·146	$2·21×10^{-11}$	0·997	0·977–1·019	$8·17×10^{-01}$	−0·849	0·145	$4·89×10^{-09}$
2q36.3¶	rs10175070	2:228670575	*CCL20, SLC18A3*	A/G	0·251	1·122	1·086–1·160	$4·31×10^{-12}$	1·000	0·978–1·023	$9·92×10^{-01}$	−0·862	0·154	$2·24×10^{-08}$
3q28	rs12634152	3:188121019	*LPP, FLJ42393, LPP-AS1*	C/T	0·547	1·133	1·100–1·166	$1·08×10^{-16}$	1·007	0·987–1·027	$4·98×10^{-01}$	−0·968	0·136	$1·21×10^{-12}$
4p14	rs5743618	4:38798648	*TLR1, TLR10, TLR6*	A/C	0·774	1·249	1·203–1·295	$3·19×10^{-32}$	1·008	0·985–1·032	$5·09×10^{-01}$	−1·578	0·163	$4·53×10^{-22}$
5q13.2¶	rs10036789	5:71695918	*PTCD2, ZNF366*	C/G	0·460	1·085	1·054–1·117	$4·51×10^{-08}$	1·016	0·996–1·036	$1·12×10^{-01}$	−0·511	0·137	$1·95×10^{-04}$
5q31.1	rs2051809	5:132056874	*KIF3A, IL4, CCNI2*	C/A	0·247	1·175	1·137–1·214	$3·24×10^{-22}$	1·016	0·994–1·040	$1·58×10^{-01}$	−0·933	0·155	$1·59×10^{-09}$
6p25.3¶	rs9391997	6:409119	*IRF4, DUSP22, EXOC2*	A/G	0·527	1·102	1·070–1·135	$6·89×10^{-11}$	0·995	0·976–1·015	$6·35×10^{-01}$	−0·597	0·136	$1·06×10^{-05}$
7p15.3¶	rs34880821	7:22775450	*IL6, TOMM7*	G/A	0·283	1·106	1·071–1·141	$4·73×10^{-10}$	1·000	0·979–1·022	$9·88×10^{-01}$	−0·667	0·149	$7·59×10^{-06}$
8q24.21¶	rs13277355	8:128777719	*MYC, TMEM75*	G/A	0·274	1·110	1·075–1·146	$1·65×10^{-10}$	1·013	0·991–1·036	$2·46×10^{-01}$	−0·660	0·151	$1·20×10^{-05}$
9q34.3¶	rs117137535	9:140500443	*ARRDC1, ZMYND19, EHMT1*	G/A	0·026	1·307	1·195–1·429	$4·16×10^{-09}$	0·977	0·913–1·046	$5·10×10^{-01}$	−2·464	0·446	$3·42×10^{-08}$
10p15.1¶	rs943451	10:6621773	*PRKCQ, PFKFB3, SFMBT2*	C/T	0·318	1·125	1·091–1·161	$7·51×10^{-14}$	1·021	1·000–1·043	$5·30×10^{-02}$	−0·613	0·145	$2·54×10^{-05}$
11q13.1¶	rs479844	11:65551957	*AP5B1, OVOL1*	A/G	0·555	1·106	1·074–1·139	$1·48×10^{-11}$	1·015	0·995–1·035	$1·41×10^{-01}$	−0·657	0·135	$1·22×10^{-06}$
11q23.3¶	rs12365699	11:118743286	*CXCR5, DDX6*	A/G	0·833	1·167	1·120–1·216	$1·57×10^{-13}$	1·004	0·978–1·031	$7·43×10^{-01}$	−1·396	0·185	$4·42×10^{-14}$
12q13.2¶	rs62623446	12:55368291	*TESPA1, MUCL1, NEUROD4*	C/T	0·072	1·188	1·127–1·252	$1·40×10^{-10}$	1·038	0·999–1·077	$5·32×10^{-02}$	−1·079	0·253	$1·95×10^{-05}$
12q24.12¶	rs10774625	12:111910219	*ATXN2, SH2B3, BRAP*	A/G	0·504	1·087	1·056–1·119	$2·07×10^{-08}$	0·997	0·978–1·017	$7·92×10^{-01}$	−0·713	0·135	$1·43×10^{-07}$
12q24.31¶	rs1696361	12:121363823	*SPPL3, HNF1A*	C/T	0·359	1·109	1·076–1·142	$1·19×10^{-11}$	1·019	0·998–1·040	$6·96×10^{-02}$	−0·696	0·140	$6·84×10^{-07}$
17q12	rs4795399	17:38061439	*GSDMB, ZPBP2, ORMDL3*	C/T	0·529	1·406	1·365–1·448	$1·45×10^{-111}$	1·005	0·985–1·025	$6·28×10^{-01}$	−2·286	0·134	$6·76×10^{-65}$
17q21.2¶	rs8066625	17:40390629	*STAT5B, GHDC, STAT5A*	G/A	0·100	1·151	1·098–1·206	$4·90×10^{-09}$	1·006	0·973–1·040	$7·33×10^{-01}$	−0·846	0·227	$1·88×10^{-04}$
17q21.32¶	rs56308324	17:45819206	*TBX21, TBKBP1, OSBPL7*	A/T	0·132	1·132	1·086–1·180	$3·47×10^{-09}$	1·025	0·996–1·055	$8·72×10^{-02}$	−0·714	0·195	$2·54×10^{-04}$
18q21.33¶	rs4574025	18:60009814	*TNFRSF11A, KIAA1468, ZCCHC2*	C/T	0·535	1·090	1·058–1·122	$9·25×10^{-09}$	0·986	0·967–1·006	$1·69×10^{-01}$	−0·871	0·136	$1·61×10^{-10}$
18q21.33¶	rs12964116	18:61442619	*SERPINB7, SERPINB11, SERPINB2*	A/G	0·036	1·336	1·247–1·432	$2·56×10^{-16}$	1·005	0·953–1·059	$8·62×10^{-01}$	−1·916	0·351	$4·87×10^{-08}$
19p13.3¶	rs4807630	19:1170445	*SBNO2, GPX4, STK11*	C/T	0·308	1·093	1·060–1·128	$2·06×10^{-08}$	1·013	0·992–1·035	$2·32×10^{-01}$	−0·487	0·147	$9·25×10^{-04}$
Adult-onset asthma-specific loci														
2q22.3¶	rs12617922	2:146156679	*TEX41, ACVR2A*	A/G	0·518	0·983	0·955–1·012	$2·36×10^{-01}$	1·066	1·045–1·087	$1·59×10^{-10}$	0·478	0·136	$4·46×10^{-04}$
Shared asthma loci														
1q32.1	rs12023876	1:203093201	*ADORA1, PPFIA4*	T/G	0·668	1·102	1·068–1·137	$1·11×10^{-09}$	1·032	1·010–1·053	$3·37×10^{-03}$	−0·567	0·144	$8·65×10^{-05}$
2p25.1	rs13416555	2:8441735	*RNF144A, ID2*	G/C	0·705	1·130	1·094–1·167	$1·94×10^{-13}$	1·040	1·018–1·063	$3·50×10^{-04}$	−0·715	0·150	$1·91×10^{-06}$
2q12.1	rs72823641	2:102936159	*IL1RL1, IL1RL2, IL18R1*	A/T	0·863	1·415	1·349–1·484	$3·01×10^{-46}$	1·109	1·077–1·142	$3·72×10^{-12}$	−1·687	0·208	$4·73×10^{-16}$
2q37.3	rs34290285	2:242698640	*D2HGDH, ING5, GAL3ST2*	A/G	0·745	1·177	1·137–1·219	$2·06×10^{-20}$	1·081	1·056–1·106	$2·16×10^{-11}$	−0·821	0·159	$2·56×10^{-07}$
3p22.3¶	rs35570272	3:33047662	*GLB1, TRIM71, TMPPE*	G/T	0·396	1·100	1·068–1·133	$2·68×10^{-10}$	1·046	1·025–1·067	$9·84×10^{-06}$	−0·454	0·138	$1·03×10^{-03}$
4q27	rs2069763	4:123377482	*IL2, ADAD1, IL21*	C/A	0·334	1·126	1·092–1·160	$1·59×10^{-14}$	1·032	1·011–1·054	$2·44×10^{-03}$	−0·631	0·142	$8·39×10^{-06}$
5p15.2¶	rs16903574	5:14610309	*FAM105A, TRIO, OTULIN*	C/G	0·077	1·236	1·174–1·302	$9·01×10^{-16}$	1·052	1·013–1·091	$7·75×10^{-03}$	−1·055	0·250	$2·51×10^{-05}$
5q22.1	rs1837253	5:110401872	*SLC25A46, TSLP*	T/C	0·740	1·211	1·169–1·253	$2·33×10^{-27}$	1·088	1·064–1·113	$2·77×10^{-13}$	−0·733	0·158	$3·72×10^{-06}$

(Table 2 continues on next page)

	rsID	Position*	Nearby genes†	Allele‡	RAF§	Childhood-onset asthma GWAS			Adult-onset asthma GWAS			Age-of-onset GWAS		
						OR	95% CI	p value	OR	95% CI	p value	β value	SE	p value
(Continued from previous page)														
5q31.1	rs17622378	5:131778452	*C5orf56, SLC22A5, IRF1*	G/A	0·573	1·103	1·071–1·136	$8·52×10^{-11}$	1·077	1·055–1·098	$3·17×10^{-13}$	−0·044	0·137	$7·46×10^{-01}$
6p22.1	rs1117490	6:30170510	*TRIM26, TRIM15, TRIM39*	T/C	0·234	1·100	1·063–1·137	$2·78×10^{-08}$	1·059	1·036–1·084	$6·83×10^{-07}$	−0·453	0·158	$4·07×10^{-03}$
6p21.33	rs2428494	6:31322197	*HLA-B, HLA-C, MICA*	T/A	0·477	1·157	1·124–1·191	$7·94×10^{-23}$	1·104	1·082–1·126	$4·77×10^{-23}$	−0·648	0·136	$1·78×10^{-06}$
6p21.32	rs28407950	6:32626348	*HLA-DQA1, HLA-DQB*	T/C	0·756	1·354	1·306–1·405	$1·27×10^{-59}$	1·148	1·121–1·175	$7·67×10^{-31}$	−1·284	0·163	$4·03×10^{-15}$
6q15	rs1321859	6:91011673	*BACH2, MAP3K7*	T/C	0·649	1·102	1·068–1·136	$9·32×10^{-10}$	1·076	1·054–1·098	$5·68×10^{-12}$	−0·268	0·144	$6·28×10^{-02}$
7p21.1¶	rs4473914	7:20426263	*ITGB6, MACC1, ABCB5*	C/T	0·604	1·092	1·060–1·125	$9·02×10^{-09}$	1·035	1·014–1·056	$8·44×10^{-04}$	−0·338	0·139	$1·51×10^{-02}$
7p15.1¶	rs917115	7:28172586	*JAZF1, TAX1BP1, CREB5*	T/C	0·208	1·112	1·074–1·152	$1·73×10^{-09}$	1·042	1·017–1·067	$7·35×10^{-04}$	−0·577	0·164	$4·17×10^{-04}$
8q21.13	rs4739738	8:81291645	*TPD52, ZBTB10*	A/G	0·359	1·117	1·084–1·151	$3·45×10^{-13}$	1·057	1·036–1·079	$8·63×10^{-08}$	−0·450	0·140	$1·27×10^{-03}$
9p24.1	rs992969	9:6209697	*RANBP6, IL33*	G/A	0·252	1·248	1·209–1·289	$6·76×10^{-42}$	1·100	1·076–1·125	$3·14×10^{-17}$	−1·026	0·152	$1·35×10^{-11}$
10p14¶	rs7894791	10:8591369	*GATA3, CELF2*	A/C	0·587	1·103	1·071–1·137	$9·43×10^{-11}$	1·024	1·004–1·045	$1·94×10^{-02}$	−0·494	0·137	$3·24×10^{-04}$
10p14	rs1775554	10:9054340	*GATA3, CELF2*	C/A	0·577	1·117	1·084–1·150	$2·85×10^{-13}$	1·121	1·098–1·143	$5·17×10^{-29}$	−0·002	0·138	$9·91×10^{-01}$
11p15.5¶	rs12788104	11:1123739	*MUC6, MUC5AC*	A/G	0·688	1·029	0·997–1·061	$7·99×10^{-02}$	1·064	1·041–1·087	$1·41×10^{-08}$	0·147	0·148	$3·22×10^{-01}$
11q12.2¶	rs174621	11:61630104	*FADS2, FADS1, FADS3*	A/G	0·772	1·013	0·978–1·049	$4·63×10^{-01}$	1·071	1·046–1·097	$1·46×10^{-08}$	0·508	0·163	$1·87×10^{-03}$
11q13.5‖	rs61894547	11:76248630	*EMSY, THAP12, LRRC32*	C/T	0·052	1·463	1·382–1·548	$2·17×10^{-39}$	1·093	1·047–1·141	$5·46×10^{-05}$	−2·227	0·284	$4·42×10^{-15}$
11q13.5‖	rs7936312	11:76293726	*EMSY, LRRC32*	G/T	0·477	1·194	1·160–1·230	$4·89×10^{-33}$	1·060	1·040–1·081	$4·10×10^{-09}$	−0·978	0·134	$3·49×10^{-13}$
12q13.11	rs56389811	12:48205358	*HDAC7, SLC48A1, VDR*	T/C	0·761	1·022	0·988–1·058	$2·11×10^{-01}$	1·078	1·054–1·104	$2·36×10^{-10}$	0·485	0·161	$2·55×10^{-03}$
12q13.2	rs705699	12:56384804	*RAB5B, CDK2, SUOX*	G/A	0·425	1·106	1·074–1·139	$1·43×10^{-11}$	1·052	1·031–1·072	$6·10×10^{-07}$	−0·481	0·136	$4·20×10^{-04}$
12q13.3	rs3122929	12:57509102	*STAT6, NAB2, LRP1*	C/T	0·404	1·134	1·101–1·167	$6·59×10^{-17}$	1·061	1·040–1·082	$5·73×10^{-09}$	−0·564	0·138	$4·07×10^{-05}$
12q21.1¶	rs11178648	12:71533210	*TSPAN8, PTPRR, LGR5*	T/C	0·592	1·087	1·055–1·120	$4·16×10^{-08}$	1·052	1·031–1·073	$6·62×10^{-07}$	−0·183	0·138	$1·86×10^{-01}$
13q32.3¶	rs1887704	13:99974492	*UBAC2, DOCK9, TM9SF2*	C/G	0·681	1·124	1·089–1·161	$5·48×10^{-13}$	1·048	1·026–1·070	$1·57×10^{-05}$	−0·476	0·147	$1·15×10^{-03}$
14q24.1	rs1950897	14:68760141	*RAD51B, ZFYVE26, ZFF36L1*	T/C	0·287	1·096	1·062–1·131	$1·38×10^{-08}$	1·031	1·009–1·053	$5·39×10^{-03}$	−0·520	0·148	$4·52×10^{-04}$
15q22.2	rs11071559	15:61069988	*RORA, ANXA2, VPS13C*	T/C	0·872	1·205	1·151–1·262	$2·57×10^{-15}$	1·069	1·037–1·101	$1·22×10^{-05}$	−0·958	0·210	$5·11×10^{-06}$
15q22.33	rs56062135	15:67455630	*SMAD3, SMAD6, AAGAB*	C/T	0·237	1·194	1·156–1·234	$2·97×10^{-26}$	1·083	1·058–1·108	$6·27×10^{-12}$	−0·681	0·155	$1·10×10^{-05}$
16p13.13	rs35032408	16:11215424	*CLEC16A, CIITA, RMI2*	G/T	0·785	1·152	1·111–1·195	$4·59×10^{-14}$	1·073	1·048–1·100	$1·10×10^{-08}$	−0·568	0·169	$7·91×10^{-04}$
16p12.1	rs3785356	16:27349168	*IL4R, NSMCE, IL21R*	C/T	0·297	1·122	1·087–1·157	$5·69×10^{-13}$	1·036	1·014–1·058	$1·24×10^{-03}$	−0·641	0·147	$1·30×10^{-05}$
16q12.1	rs2066844	16:50745926	*NOD2, SNX20, CYLD*	C/T	0·048	1·194	1·121–1·272	$3·52×10^{-08}$	1·049	1·003–1·097	$3·76×10^{-02}$	−0·743	0·306	$1·54×10^{-02}$
17q21.33	rs28406364	17:47454507	*ZNF652, PHB*	C/T	0·377	1·108	1·076–1·142	$1·39×10^{-11}$	1·036	1·015–1·057	$6·55×10^{-04}$	−0·489	0·140	$4·81×10^{-04}$
19q13.11	rs10414065	19:33721455	*LRP3, CEBPA*	T/C	0·934	1·207	1·133–1·287	$6·37×10^{-09}$	1·108	1·064–1·155	$1·01×10^{-06}$	−1·078	0·291	$2·12×10^{-04}$
21q22.12¶	rs11088309	21:36464631	*RUNX1, SETD4*	C/G	0·143	1·039	0·997–1·083	$6·82×10^{-02}$	1·079	1·050–1·109	$4·83×10^{-08}$	0·063	0·189	$7·38×10^{-01}$

Information for the SNP with smallest p value at each locus is shown. Summary statistics for the 19 significant SNPs in the age of asthma onset GWAS are shown in appendix p 63. GWAS=genome-wide association study. OR=odds ratio. RAF=risk allele frequency. SNP=single-nucleotide polymorphism. *SNP position according to Genome Reference Consortium Build 37 (hg19). †The gene in which the SNP is located is indicated first, followed by the previous gene and the next gene; for intergenic SNPs, only the previous and next genes are shown. ‡Alleles are shown as non-risk/risk alleles, where the risk allele is the allele associated with increased asthma risk. §RAF in the UK Biobank. ¶Not reported in previous GWAS. ‖The SNP with the smallest p value differs in the childhood-onset GWAS and the adult-onset GWAS at the same locus; both SNPs are shown.

Table 2: **Regions with SNPs that were genome-wide significant in the childhood-onset or adult-onset GWAS**

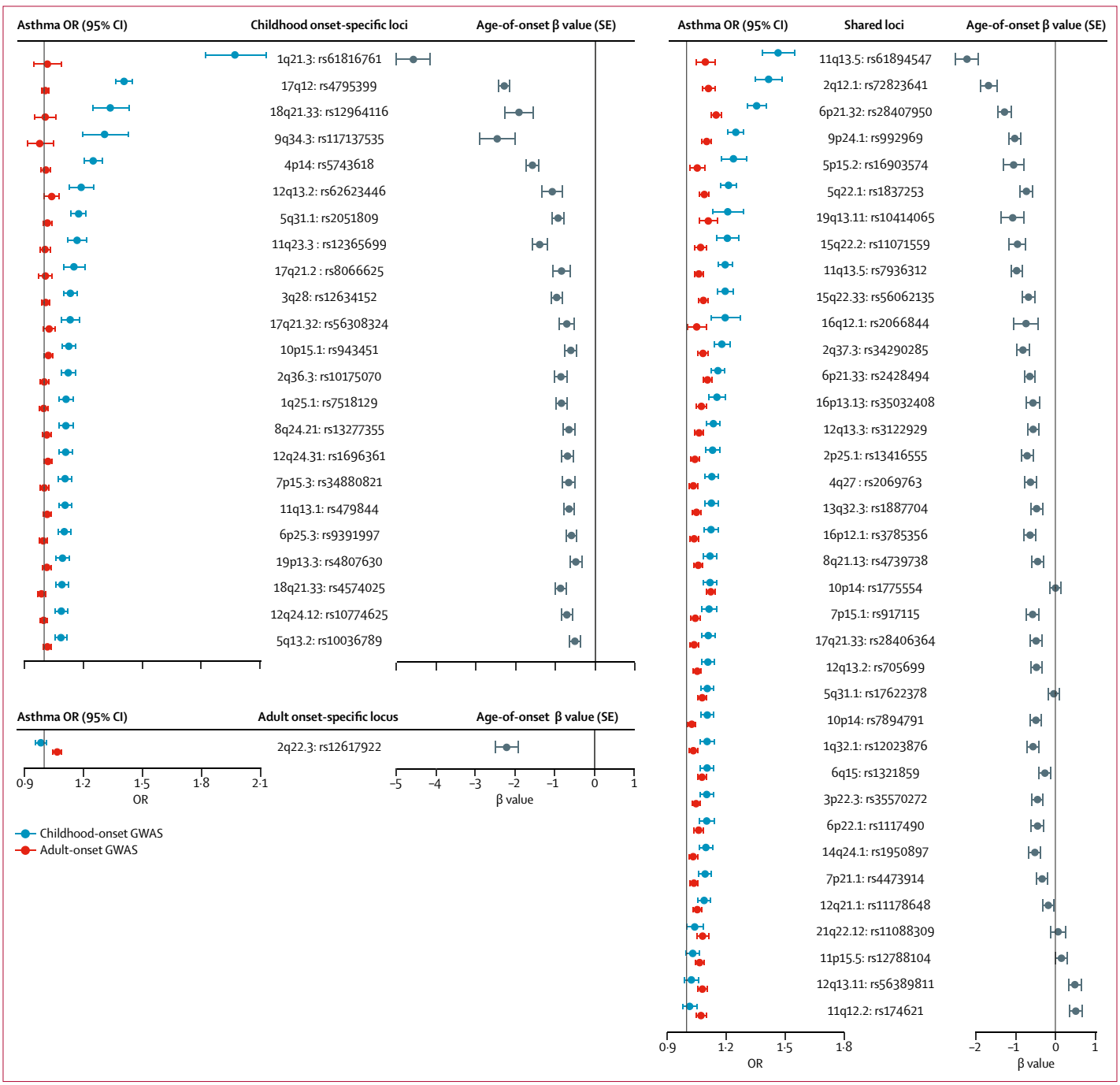

***Figure 2*: Forest plot of childhood-onset asthma, adult-onset asthma, and age-of-onset GWAS**
Error bars show 95% CIs. GWAS=genome-wide association study. OR=odds ratio.

(figure 1A) were most highly expressed in skin, whole blood, and small intestine (lower ileum) compared with all other tissues, whereas genes at adult-onset loci (figure 1B) were most highly expressed in lung, whole blood, small intestine (lower ileum), and spleen (enrichment for higher expression, $p<10\times10^{-3}$; appendix pp 42, 65).

To better understand molecular mechanisms and to narrow the list of candidate causal genes at associated loci, we focused on the five tissues that most highly expressed the genes at loci associated with childhood-onset asthma or those associated with adult-onset asthma: skin, lung, whole blood, small intestine, and spleen. We used PrediXcan[16] to identify genes whose

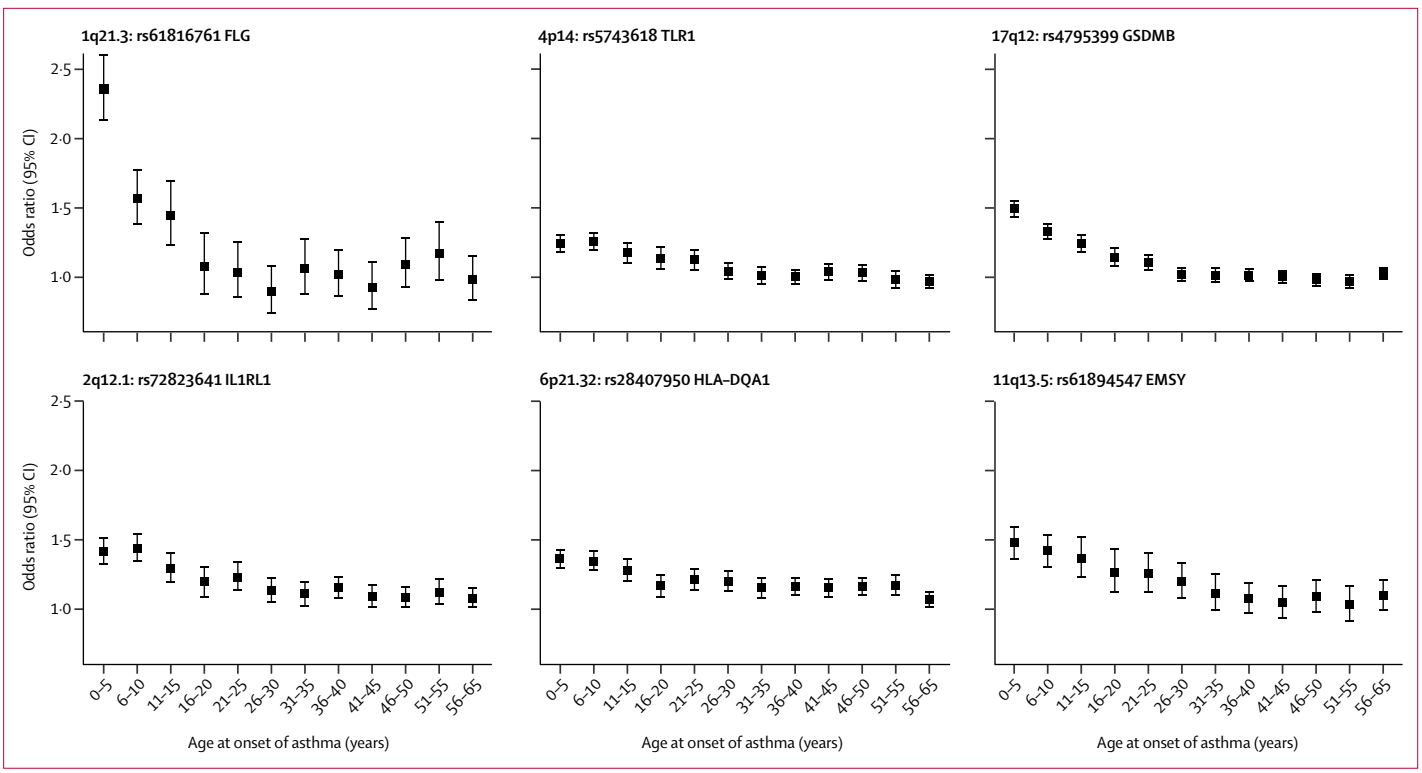

Figure 3: **Examples of age-of-onset effect sizes**
Age-of-onset effects for lead single-nucleotide polymorphisms at three genome-wide signficant childhood onset-specific loci (1q21.3, 4p14 and 17q12), and three genome-wide signficant shared loci (2q12.1, 6p21.32, and 11q13.5). The sample sizes for the age-of-onset bins are n=4637 for 0–5 years, n=4255 for 6–10 years, n=2684 for 11–15 years, n=2128 for 16–20 years, n=2578 for 21–25 years, n=2923 for 26–30 years, n=2599 for 31–35 years, n=3571 for 36–40 years, n=2967 for 41–45 years, n=3266 for 46–50 years, n=2544 for 51–55 years, and n=3694 for 56–65 years.

expression is predicted by variants associated with asthma in the childhood-onset or adult-onset GWAS and potentially mediate the effects of associated SNPs on asthma risk.

Of 38 608 genes in the five tissues of interest, the transcriptome-wide association analysis identified 113 unique, candidate causal genes at 22 of the 61 GWAS loci ($p<1{\cdot}4\times10^{-6}$; figures 4,5; appendix pp 43–44, 66). These included 39 genes associated with childhood-onset asthma at eight of the childhood onset-specific loci and 74 genes associated with childhood-onset or adult-onset asthma at 13 of the shared loci. Variants at the one adult onset-specific locus at 2q22.3 did not predict the expression of any genes in the five tissues.

The predicted genes most significantly associated with childhood-onset asthma were at the 17q12 locus (Z score >10 or <–10) in skin (*ORMDL3*, *ERBB2*, *PGAP3*, and *GSDMA*, as well as two long non-coding RNAs), lung (*ORMDL3*, *GSDMB*, *GSDMA*, *PGAP3*, and *PNMT*), blood (*ORMDL3*, *GSDMB*, *IKZF3*, and *MED24*), small intestine (*GSDMA*, *GSDMB*, and *PGAP3*), and spleen (*ORMDL3*, *GSDMB*, *ZPBP2*, and *MED24*; figure 4; appendix p 43). Some genes were predicted to be highly expressed in individuals with asthma (eg, *ORMDL3*, *GSDMB*, *PGAP3*, and *ERBB2*), whereas others were predicted to have reduced expression in individuals with asthma (eg, *GSDMA*, *MED24*, and *IKZF3*; figure 4). This pattern of expression reflects the broad regulatory effects of SNPs and tissue specificifity of gene expression at this locus.[26] The childhood-onset asthma locus at 1q21.3 includes genes that are essential for epidermal differentiation and maintainance of essential barrier function. The predicted expression of nine genes at the 1q21.3 locus was associated with childhood-onset asthma: higher predicted expression of *CRNN*, *CRCT1*, and *THEM5* in skin, of *PSDM4* in lung and of *LINGO4* in skin, lung, and blood was associated with increased asthma risk, whereas lower predicted expression of *SPRR2D* in skin, of *S100A12* in lung, of *FLG* in skin, lung, and spleen, and of *TDRKH* in skin, lung, blood, and spleen was associated with increased asthma risk. *S100A12* has been previously implicated in asthma[34] and *FLG* variants have been associated with atopic dermatitis and food allergies, and asthma in the context of other allergic diseases.[20,27–33] Other childhood onset-specific genes previously implicated in asthma but not previously reported in asthma GWAS are *CCL20*[35] at 2q36.3 and *TLR10*[36] at 4p14 in whole blood, and *TLR6*[37,38] at 4p14, *AP5B1* at 11q13.1, and *SERPINB7*[39] at 18q21.33 in skin.

The 5q31.1 region had independent loci that were both childhood-onset specific and shared in the GWAS. Although the predicted expression of all eight asthma

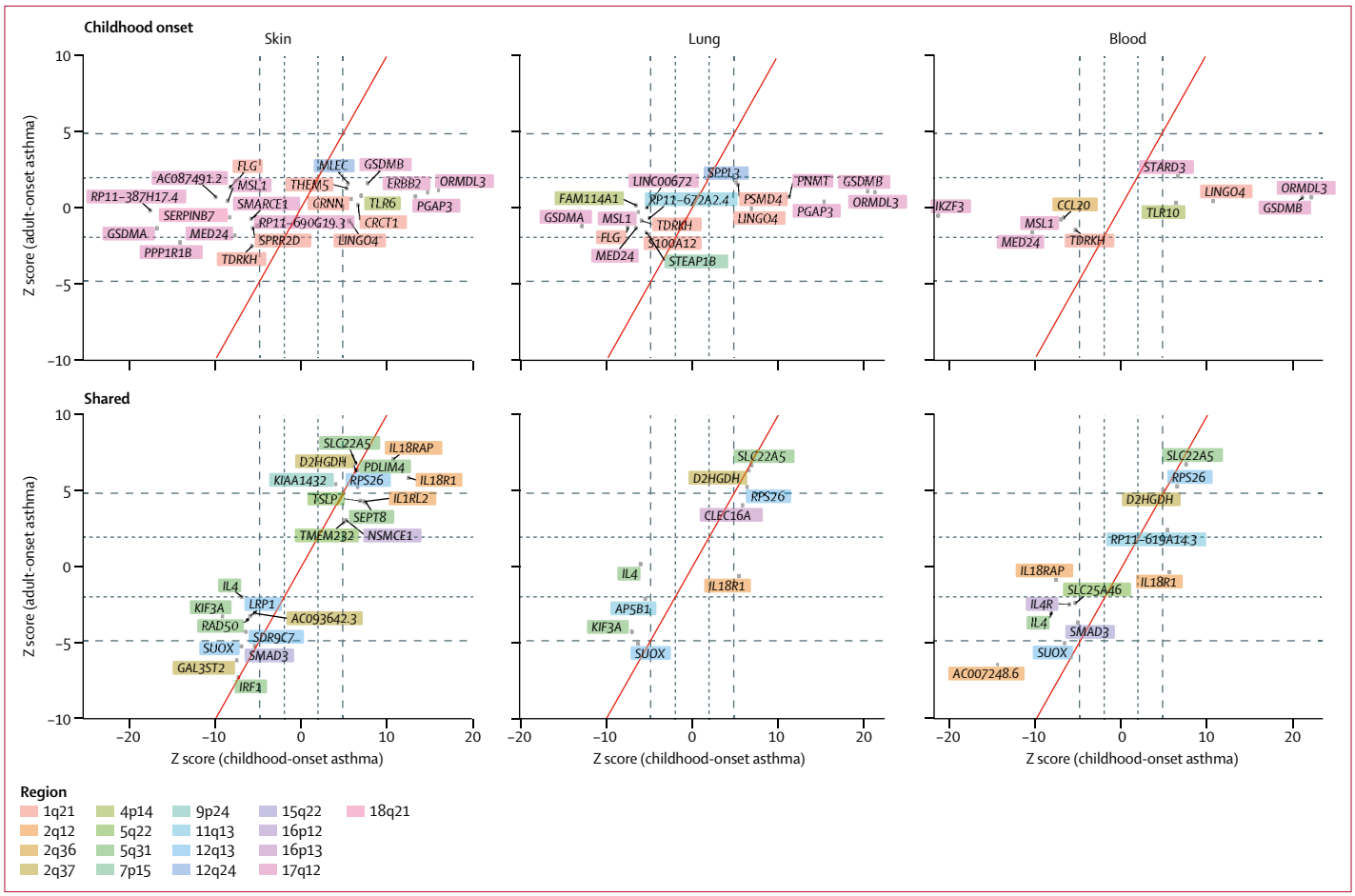

Figure 4: **Results of PrediXcan studies at non-HLA region loci**

Genes whose predicted expression was significantly associated with asthma in either the childhood-onset asthma cases or the adult-onset cases are shown. Values for skin combine both sun-exposed and non-sun-exposed skin, showing the most significant statistic of the two. Results using gene expression in spleen and small intestine are shown in appendix p 43. The Z scores on the x-axes and y-axes are from transcriptome-wide tests of association with asthma using single-nucleotide polymorphism sets that predict the expression of that gene. The diagonal lines show the expected values when associations are the same in the childhood-onset and adult-onset cases. The horizontal and vertical dashed lines correspond to Z scores ±4·84 ($p=1·29\times10^{-6}$); the horizontal and vertical dotted lines correspond to Z scores ±1·96 (p=0·05).

genes at this extended locus were shared, five genes were more significantly associated with childhood-onset asthma: higher predicted expression of *RAD50* in skin but lower predicted expression of *SEPT8* in skin and lung, *IL4* in skin, lung, and blood, and *AFF4* small intestine were associated with increased risk for asthma (figure 4). The remaining three genes had similar associations with childhood-onset and adult-onset asthma, with predicted lower expression of *IRF1* in skin and spleen and predicted higher expression of *PDLIM4* in skin and *SLC22A5* in all five tissues associated with increased asthma risk. *IL4, RAD50, SLC22A5*, and *PDLIM4* have been highlighted in previous asthma GWAS,[25,40] and *KIF3A* was identified in a GWAS of the progression of allergic disease in childhood (ie, the atopic march)[41] and associated with childhood-onset asthma in a candidate gene study.[42]

Predicted expression of 44 genes at two independent shared loci in the HLA region (ie, 6p21.32 and 6p21.33) were associated with childhood-onset asthma only (n=1; in skin and lung), adult-onset asthma only (n=3; in skin only), or both (n=39; in multiple tissues; figure 5).

Among the remaining 37 shared loci (figure 2B), SNPs at 12 predicted the expression of 23 unique genes, all of which were associated with both childhood-onset and adult-onset asthma. These include *IL18R1, IL18RAP*, and *ILRL2* at 2q12·1 in multiple tissues, *TSLP* at 5q22.1 in skin, *SMAD3* at 15q22.33 in skin and blood, *LRP1* at 12q13.3 in skin, *IL4R* at 16p12.1 in blood, and *CLEC16A* at 16p13.13 in lung. Loci associated with *IL18R1, IL18RAP, IL18RAP, ILRL2, TSLP, SMAD3, LRP1, IL4R*, and *CLEC16A* were reported in previous asthma GWAS.[13,20,25]

In sensitivity analyses, the GWAS results were robust to inclusion of varying numbers of principal components (10, 14, or 20; appendix p 12) and to limiting the sample to cases with diagnoses based on ICD-10 codes (appendix p 13).

In the GWAS comparing individuals with asthma based on self-reported doctor diagnosis alone compared

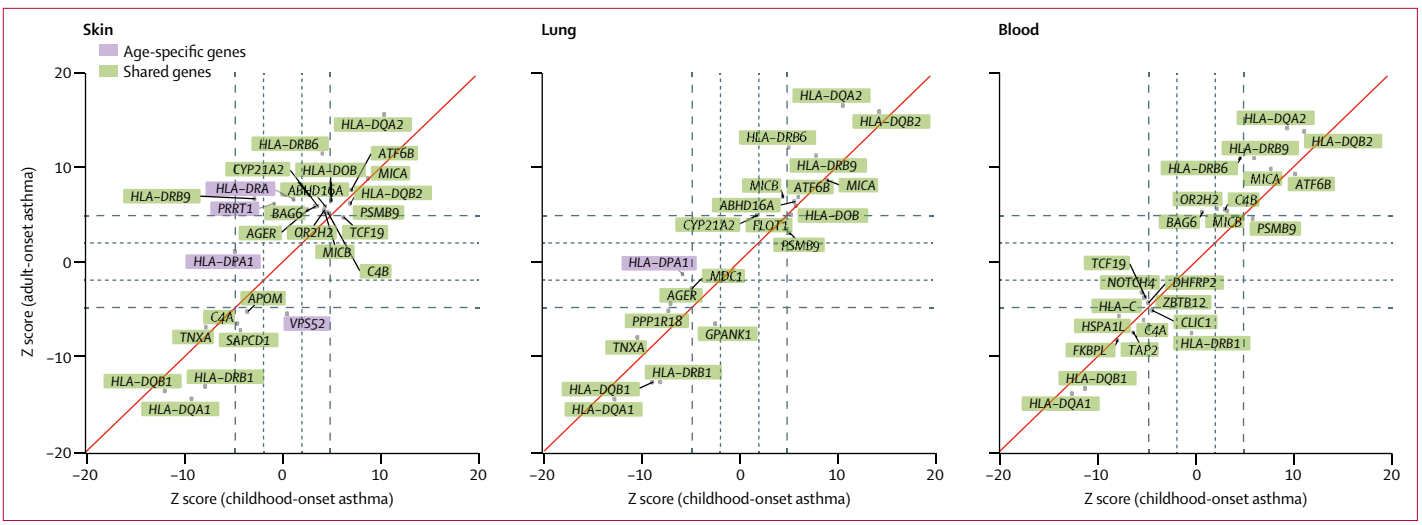

Figure 5: **Results of PrediXcan studies of HLA region genes**
Genes in the HLA region whose predicted expression was associated with childhood-onset or adult-onset asthma are shown. Results using gene expression in spleen and small intestine are shown in appendix p 44.

with those with self-reported doctor diagnosis plus ICD-10 codes, no differences in risk loci were detected between these two groups (appendix p 15), indicating that differences in identifying asthma cases between these groups did not affect the GWAS results. We also did a COPD GWAS using all cases and controls, and there was no overlap between regions identified in the COPD GWAS and any of the asthma groups, including the locus on chromosome 15q25.1 (appendix p 45).

Discussion

We report here the first large GWAS of both childhood-onset and adult-onset asthma cases. These GWAS revealed 61 independent asthma loci, 23 specific to childhood onset, one specific to adult onset, and 37 shared, with overall larger effect sizes for childhood-onset asthma at nearly all loci. Moreover, the predicted expression of 42 of the 113 implicated genes were associated specifically with childhood-onset asthma, compared with the predicted expression of three genes associated specifically with adult-onset asthma. Several newly discovered risk loci in this study contain genes (eg, *FADS2*,[21] *MUC5AC*,[22,23] and *TBX21*[24]) that have been associated with asthma in candidate gene studies, which supports the external validity of our GWAS findings.

We found that childhood-specific or shared loci were associated with development of asthma at younger ages and with alleles with larger effects by comparison with adult-specific loci. These results are consistent with a previous GWAS of age of asthma onset in individuals with European ancestry that reported five genome-wide significant and three suggestive significant loci, all of which were associated with earlier age of onset;[43] six of those eight loci are also associated with age of onset in the UK Biobank GWAS (appendix p 64). Our findings are particularly striking given that there were nearly 2·5 times more adult-onset than childhood-onset cases in this study.

Thus, despite having substantially less power to detect loci specific to childhood-onset asthma, our analyses revealed many more childhood-onset asthma loci. Similarly, the asthma risk alleles at 19 loci that were significant in the age-of-onset GWAS were all associated with younger age of onset. Finally, we showed that the SNP-based heritability of childhood-onset asthma is more than three times larger than that of adult-onset asthma, suggesting a more important role for genetic variation in risk for childhood-onset than in adult-onset asthma and, conversely, the larger role for environmental variation in risk for adult-onset than for childhood-onset asthma. Regardless, the loci with genome-wide significant SNPs in these GWAS accounted for only 32·7% of the heritability of childhood-onset asthma and 9·8% of the heritability of adult-onset asthma, suggesting that many more asthma loci remain to be discovered. These findings are consistent with previous studies showing decreased estimates of asthma heritability with increasing age of onset,[44] SNPs associated with earlier age of onset,[43] and an additive, unweighted genetic risk score comprising 15 SNPs at eight asthma-associated loci associated with earlier age of onset.[45] Our study further shows that genetic risk for adult-onset asthma is largely a subset of the genetic risk loci for childhood-onset asthma, but with overall smaller effect sizes, consistent with a larger role for environmental risk factors in adult-onset asthma.

Despite the overlap of adult-onset and childhood-onset loci, distinct mechanisms contributing to each were suggested by tissue enrichments: childhood-onset loci were enriched for genes with highest expression in skin, whereas adult-onset loci were enriched for genes with highest expression in lung and spleen; both were

enriched for genes highly expressed in whole blood and small intestine. These findings suggest both overlapping and distinct underlying mechanisms associated with asthma that begins in childhood and asthma with onset in adulthood. The highlighting of skin as a target tissue for childhood-onset asthma supports the widely held idea that asthma in childhood is due to impaired barrier function in the skin and other epithelial surfaces. This model proposes that compromised epithelial barriers promote sensitisation to food and airway allergens and to wheezing illnesses in early life.[46,47] In fact, childhood onset-specific loci identified in this study have been associated with atopic dermatitis or food allergies, such as *FLG* on 1q21.3 with the atopic march,[41] atopic dermatitis,[30–32] and food allergies,[27–29] *KIF3A* on 5q31.1 and *AP5B1*/*OVOL1* on 11q13.1 with the atopic march[41] and atopic dermatitis,[48] *SERPINB7* on 18q21.33 with food allergies,[39] and *CRNN* (which encodes cornulin) on 1q21.3 with atopic dermatitis concomitant with asthma and reduced expression in atopic dermatitis-affected skin.[49] Variants at those loci were all associated with earlier age of asthma onset. We further showed that these loci are associated with childhood-onset asthma, even after exclusion of patients with a history of allergic diseases in prespecified analyses, suggesting both a crucial role for the allergic diathesis in the development of asthma in childhood and a shared architecture between allergic disease and childhood-onset asthma.[33,46]

By contrast, the enrichment for genes highly expressed in lung and spleen at adult-onset loci suggests a more lung-centred, and potentially immune-mediated, cause for asthma with onset later in life. The prominant role of the HLA region in the adult-onset asthma GWAS further highlights a central role for immune processes driving asthma pathogenesis in adults. Our findings are consistent with previous research suggesting that the HLA region (6p21.32 and 6p21.33) is among the most signficant loci in nearly all asthma GWAS, and was the most significant locus in a previous small GWAS of adult-onset asthma[7] and in adults with asthma.[40,50] The fact that both childhood-onset and adult-onset asthma loci were enriched for genes that are most highly expressed in whole blood cells and small intestine further indicate a shared immune cause, as suggested from a large GWAS that included both children and adults.[25]

Combining GWAS with a transcriptome-wide association test that uses combinations of associated SNPs to predict gene expression in different tissues revealed significant complexity at the two most highly associated asthma loci. SNPs at the 17q12 locus predicted expression of 18 childhood-onset asthma genes and SNPs at the HLA region predicted expression of 42 genes: three were associated with adult-onset asthma and most were not HLA genes per se. These results strengthen the argument that multiple genes contribute to asthma risk at the HLA and 17q12 loci and probably account for the highly significant GWAS p values observed at these loci in nearly all studies. It is also likely that these genes have both tissue-specific and broad effects in epithelium, lung, and immune tissues.

The new loci identified in our study include the first adult-onset asthma-specific association at 2q22.3. The lead SNP at 2q22.3 is intergenic between *TEX41* and *ACVR2A*. The predicted expression of *ACVR2A* was not associated with asthma in our study, despite it being expressed in lung, blood, small intestine, and spleen. *TEX41* was not expressed in any of the five tissues investigated. Interestingly, a GWAS also done in UK Biobank participants implicated variants near *TEX41* in heavy versus never smoking behaviour.[51] However, even after removing adult-onset cases and controls with reported ever smoking, the p value for this SNP remained significant and the OR slightly increased (OR 1·077 [95% CI 1·05–1·1], $p=2·26\times10^{-8}$; n=12 132 cases and 176 704 controls).[51] Variants in or near this gene, which encodes a long intergenic non-coding RNA, have been associated with cardiovascular and immune-mediated traits,[52] making this a potentially interesting candidate gene for adult-onset asthma.

Our study had limitations. First, diagnoses of asthma and allergic disease in study participants were from self reported doctor diagnosis and medical records (ICD-10 codes). Thus, it is possible that diagnoses, age of onset, or both are misspecified in some participants. On the one hand, the large sample size and our ability to replicate nearly all previously reported asthma loci (appendix pp 58, 67) suggest that our analyses were robust to any inaccuracies in the data. On the other hand, it is possible that individuals with adult-onset asthma included people with poor recall of childhood-onset asthma in which symptoms remitted and then relapsed later in life[53] or misclassified cases of COPD among the older age groups. Our sensitivity analysis suggested that if even as few as 5% of the adult-onset cases were misclassified, we should have observed some signal of association at childhood-onset loci, which we did not. The fact that the ORs for asthma at shared loci are relatively similar from approximately 25 to 65 years of age (figure 3), and that we do not detect any association signal at the major COPD locus on chromosome 15q25.1 (appendix p 45), further suggests that there is negligible misclassification of cases in the older age groups. Second, although we used stringent criteria to classify loci as childhood-onset or adult-onset specific, we cannot exclude the possibility that in infinitely large sample sizes the effect sizes of some of these loci will have 95% CIs that overlap or the association p value will become smaller than 0·05. Conversely, some of the shared loci with modest p values in the adult-onset cases might not be true risk loci for asthma with onset at older ages. Third, the gene expression data used to predict candidate target genes included heterogeneous tissues and were collected mostly from adults. As a result, our study might have missed relevant genes with expression

that is developmentally regulated or environment specific. Our finding of candidate genes at only 22 of the 61 asthma loci mght be due in part to the importance of both in asthma pathogenesis. Moreover, all inference based on gene expression is using imputed expression. It is possible, therefore, that some relevant genes were more difficult to impute and not included in our analysis, although a comparative study[54] showed that PrediXcan is a more robust method for prediction of gene expression than other related methods. Fourth, because of the ethnic composition of the UK Biobank, this study was limited to individuals of European ancestry only. As a result, we could not evaluate the genetic risk architecture or assess the effects of age of onset-specific loci in other populations. Finally, in our study sample, the proportion of self-reported, doctor-diagnosed asthma was 37 846 of 318 237 (11·9%) cases, lower than the UK lifetime prevalence of patient-reported, clinician-diagnosed asthma of 15·6%[55] and consistent with the known healthy volunteer bias in UK Biobank.[56]

In the largest asthma GWAS so far, we show that genetic risk loci for adult-onset asthma are largely a subset of the loci associated with childhood-onset asthma, with overall smaller effect sizes for onset at later ages. These data suggest that childhood onset-specific loci and those associated with age of onset play a part in disease initiation, whereas the other associated loci reflect shared mechanisms of disease progression. The differences in the target tissues that most highly express the genes at associated loci and the predicted expression of genes at age-specific and shared loci provides additional genetic and molecular evidence for both shared and distinct pathogenic mechanisms in childhood-onset and adult-onset asthma. It is therefore possible that the most effective treatments will also differ between these two groups, and that strategies for precision medicine should be further personalised to account for age of asthma onset.

Contributors

All authors were involved in the conception and design of the study and in writing the manuscript. MP and NS did all analyses and prepared figures and tables, under the overall supervision of DLN, CO, and HKI.

Declaration of interests

HKI reports personal fees from AbbVie and GlaxoSmithKline. The other authors declare no competing interests.

Acknowledgments

This research has been conducted using UK Biobank Resource under application number 19526 and 44300. The study was funded by US National Institutes of Health grants R01 MH107666 and P30 DK20595 to HKI and R01 HL129735, R01 HL122712, P01 HL070831, and UG3 OD023282 to CO; NS was supported by T32 HL007605.

References

1 GBD 2015 Chronic Respiratory Disease Collaborators. Global, regional, and national deaths, prevalence, disability-adjusted life years, and years lived with disability for chronic obstructive pulmonary disease and asthma, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet Respir Med* 2017; **5:** 691–706.

2 Bush A, Menzies-Gow A. Phenotypic differences between pediatric and adult asthma. *Proc Am Thorac Soc* 2009; **6:** 712–19.

3 Busse W, Banks-Schlegel SP, Larsen GL. Childhood- versus adult-onset asthma. *Am J Respir Crit Care Med* 1995; **151:** 1635–39.

4 Dijk FN, de Jongste JC, Postma DS, Koppelman GH. Genetics of onset of asthma. *Curr Opin Allergy Clin Immunol* 2013; **13:** 193–202.

5 Holloway JW, Arshad SH, Holgate ST. Using genetics to predict the natural history of asthma? *J Allergy Clin Immunol* 2010; **126:** 200–09.

6 Bouzigon E, Corda E, Aschard H, et al. Effect of 17q21 variants and smoking exposure in early-onset asthma. *N Engl J Med* 2008; **359:** 1985–94.

7 Moffatt MF, Gut IG, Demenais F, et al. A large-scale, consortium-based genomewide association study of asthma. *N Engl J Med* 2010; **363:** 1211–21.

8 Bycroft C, Freeman C, Petkova D, et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature* 2018; **562:** 203–09.

9 Pividori M, Im HK. ukbREST: efficient and streamlined data access for reproducible research in large biobanks. *Bioinformatics* 2018; published online Nov 5. DOI:10.1093/bioinformatics/bty925.

10 Heath AP, Greenway M, Powell R, et al. Bionimbus: a cloud for managing, analyzing and sharing large genomics datasets. *J Am Med Inform Assoc* 2014; **21:** 969–75.

11 Watanabe K, Taskesen E, van Bochoven A, Posthuma D. Functional mapping and annotation of genetic associations with FUMA. *Nat Commun* 2017; **8:** 1826.

12 1000 Genomes Project Consortium, Auton A, Brooks LD, et al. A global reference for human genetic variation. *Nature* 2015; **526:** 68–74.

13 Pickrell JK, Berisa T, Liu JZ, Segurel L, Tung JY, Hinds DA. Detection and interpretation of shared genetic influences on 42 human traits. *Nat Genet* 2016; **48:** 709–17.

14 Bulik-Sullivan BK, Loh PR, Finucane HK, et al. LD score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet* 2015; **47:** 291–95.

15 Asthma UK. Asthma data visualisations. https://www.asthma.org.uk/get-involved/campaigns/data-visualisations/ (accessed Nov 18, 2018).

16 Gamazon ER, Wheeler HE, Shah KP, et al. A gene-based association method for mapping traits using reference transcriptome data. *Nat Genet* 2015; **47:** 1091–98.

17 Barbeira AN, Dickinson SP, Bonazzola R, et al. Exploring the phenotypic consequences of tissue specific gene expression variation inferred from GWAS summary statistics. *Nat Commun* 2018; **9:** 1825.

18 Battle A, Brown CD, Engelhardt BE, Montgomery SB. Genetic effects on gene expression across human tissues. *Nature* 2017; **550:** 204–13.

19 MacArthur J, Bowler E, Cerezo M, et al. The new NHGRI–EBI Catalog of published genome-wide association studies (GWAS Catalog). *Nucleic Acids Res* 2017; **45:** D896–901.

20 Zhu Z, Lee PH, Chaffin MD, et al. A genome-wide cross-trait analysis from UK Biobank highlights the shared genetic architecture of asthma and allergic diseases. *Nat Genet* 2018; **50:** 857–64.

21 Standl M, Sausenthaler S, Lattka E, et al. *FADS* gene cluster modulates the effect of breastfeeding on asthma. Results from the GINIplus and LISAplus studies. *Allergy* 2012; **67:** 83–90.

22 Rousseau K, Byrne C, Griesinger G, et al. Allelic association and recombination hotspots in the mucin gene (*MUC*) complex on chromosome 11p15.5. *Ann Hum Genet* 2007; **71:** 561–69.

23 Johnson L, Shah I, Loh AX, et al. MUC5AC and inflammatory mediators associated with respiratory outcomes in the British 1946 birth cohort. *Respirology* 2013; **18:** 1003–10.

24 Akahoshi M, Obara K, Hirota T, et al. Functional promoter polymorphism in the *TBX21* gene associated with aspirin-induced asthma. *Hum Genet* 2005; **117:** 16–26.

25 Demenais F, Margaritte-Jeannin P, Barnes KC, et al. Multiancestry association study identifies new asthma risk loci that colocalize with immune-cell enhancer marks. *Nat Genet* 2018; **50:** 42–53.

26 Stein MM, Thompson EE, Schoettler N, et al. A decade of research on the 17q12-21 asthma locus: piecing together the puzzle. *J Allergy Clin Immunol* 2018; **142:** 749–64.e3.

27 Brown SJ, Asai Y, Cordell HJ, et al. Loss-of-function variants in the filaggrin gene are a significant risk factor for peanut allergy. *J Allergy Clin Immunol* 2011; **127:** 661–67.

28 Marenholz I, Kerscher T, Bauerfeind A, et al. An interaction between filaggrin mutations and early food sensitization improves the prediction of childhood asthma. *J Allergy Clin Immunol* 2009; **123:** 911–16.

29 Venkataraman D, Soto-Ramirez N, Kurukulaaratchy RJ, et al. Filaggrin loss-of-function mutations are associated with food allergy in childhood and adolescence. *J Allergy Clin Immunol* 2014; **134:** 876–82.e4.

30 Bisgaard H, Halkjaer LB, Hinge R, et al. Risk analysis of early childhood eczema. *J Allergy Clin Immunol* 2009; **123:** 1355–60.e5.

31 Palmer CN, Irvine AD, Terron-Kwiatkowski A, et al. Common loss-of-function variants of the epidermal barrier protein filaggrin are a major predisposing factor for atopic dermatitis. *Nat Genet* 2006; **38:** 441–46.

32 Weidinger S, Illig T, Baurecht H, et al. Loss-of-function variations within the filaggrin gene predispose for atopic dermatitis with allergic sensitizations. *J Allergy Clin Immunol* 2006; **118:** 214–19.

33 Ferreira MA, Vonk JM, Baurecht H, et al. Shared genetic origin of asthma, hay fever and eczema elucidates allergic disease biology. *Nat Genet* 2017; **49:** 1752–57.

34 Yang Z, Yan WX, Cai H, et al. S100A12 provokes mast cell activation: a potential amplification pathway in asthma and innate immunity. *J Allergy Clin Immunol* 2007; **119:** 106–14.

35 Collison A, Hatchwell L, Verrills N, et al. The E3 ubiquitin ligase midline 1 promotes allergen and rhinovirus-induced asthma by inhibiting protein phosphatase 2A activity. *Nat Med* 2013; **19:** 232–37.

36 Lazarus R, Raby BA, Lange C, et al. TOLL-like receptor 10 genetic variation is associated with asthma in two independent samples. *Am J Respir Crit Care Med* 2004; **170:** 594–600.

37 Tantisira K, Klimecki WT, Lazarus R, et al. Toll-like receptor 6 gene (*TLR6*): single-nucleotide polymorphism frequencies and preliminary association with the diagnosis of asthma. *Genes Immun* 2004; **5:** 343–46.

38 Lau MY, Dharmage SC, Burgess JA, et al. The interaction between farming/rural environment and *TLR2, TLR4, TLR6* and *CD14* genetic polymorphisms in relation to early- and late-onset asthma. *Sci Rep* 2017; **7:** 43681.

39 Marenholz I, Grosche S, Kalb B, et al. Genome-wide association study identifies the *SERPINB* gene cluster as a susceptibility locus for food allergy. *Nat Commun* 2017; **8:** 1056.

40 Li X, Howard TD, Zheng SL, et al. Genome-wide association study of asthma identifies RAD50-IL13 and HLA-DR/DQ regions. *J Allergy Clin Immunol* 2010; **125:** 328–35.e11.

41 Marenholz I, Esparza-Gordillo J, Ruschendorf F, et al. Meta-analysis identifies seven susceptibility loci involved in the atopic march. *Nat Commun* 2015; **6:** 8804.

42 Kovacic MB, Myers JM, Wang N, et al. Identification of *KIF3A* as a novel candidate gene for childhood asthma using RNA expression and population allelic frequencies differences. *PLoS One* 2011; **6:** e23714.

43 Sarnowski C, Sugier PE, Granell R, et al. Identification of a new locus at 16q12 associated with time to asthma onset. *J Allergy Clin Immunol* 2016; **138:** 1071–80.

44 Thomsen SF, Duffy DL, Kyvik KO, Backer V. Genetic influence on the age at onset of asthma: a twin study. *J Allergy Clin Immunol* 2010; **126:** 626–30.

45 Belsky DW, Sears MR, Hancox RJ, et al. Polygenic risk and the development and course of asthma: an analysis of data from a four-decade longitudinal study. *Lancet Respir Med* 2013; **1:** 453–61.

46 Busse WW. The atopic march: fact or folklore? *Ann Allergy Asthma Immunol* 2018; **120:** 116–18.

47 Han H, Roan F, Ziegler SF. The atopic march: current insights into skin barrier dysfunction and epithelial cell-derived cytokines. *Immunol Rev* 2017; **278:** 116–30.

48 Johansson E, Biagini Myers JM, Martin LJ, et al. *KIF3A* genetic variation is associated with pediatric asthma in the presence of eczema independent of allergic rhinitis. *J Allergy Clin Immunol* 2017; **140:** 595–98 e5.

49 Trzeciak M, Sakowicz-Burkiewicz M, Wesserling M, et al. Altered expression of genes encoding cornulin and repetin in atopic dermatitis. *Int Arch Allergy Immunol* 2017; **172:** 11–19.

50 Lasky-Su J, Himes BE, Raby BA, et al. HLA-DQ strikes again: genome-wide association study further confirms HLA-DQ in the diagnosis of asthma among adults. *Clin Exp Allergy* 2012; **42:** 1724–33.

51 Wain LV, Shrine N, Miller S, et al. Novel insights into the genetics of smoking behaviour, lung function, and chronic obstructive pulmonary disease (UK BiLEVE): a genetic association study in UK Biobank. *Lancet Respir Med* 2015; **3:** 769–81.

52 MacArthur DG, Balasubramanian S, Frankish A, et al. A systematic survey of loss-of-function variants in human protein-coding genes. *Science* 2012; **335:** 823–28.

53 Sears MR, Greene JM, Willan AR, et al. A longitudinal, population-based, cohort study of childhood asthma followed to adulthood. *N Engl J Med* 2003; **349:** 1414–22.

54 Fryett JJ, Inshaw J, Morris AP, Cordell HJ. Comparison of methods for transcriptome imputation through application to two common complex diseases. *Eur J Hum Genet* 2018; **26:** 1658–67.

55 Mukherjee M, Stoddart A, Gupta RP, et al. The epidemiology, healthcare and societal burden and costs of asthma in the UK and its member nations: analyses of standalone and linked national databases. *BMC Med* 2016; **14:** 113.

56 Fry A, Littlejohns TJ, Sudlow C, et al. Comparison of sociodemographic and health-related characteristics of UK Biobank participants with those of the general population. *Am J Epidemiol* 2017; **186:** 1026–34.