Shared and Distinct Genetic Risk Factors for Childhood Onset and Adult Onset Asthma: Genome- and Transcriptome-wide Studies

Milton Pividori, PhD[1,^], Nathan Schoettler, MD, PhD[1,2,^], Dan L. Nicolae, PhD[1,2,3,#], Carole Ober, PhD[2,#], Hae Kyung Im, PhD[1,2,§]

Departments of Medicine[1], Human Genetics[2] and Statistics[3], The University of Chicago, Chicago, IL

[^]These authors contributed equally.

[#]Full Professors

[§]Corresponding Author:

HKI haky@uchicago.edu

Background

Childhood and adult onset asthma differ with respect to severity and co-morbidities. 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 and adult onset asthma, and the genes that may mediate the effects of associated variation.

Methods

We used data from UK Biobank to conduct genome-wide association studies (GWASs) in 37,846 subjects with asthma, including 9,433 childhood onset cases (onset before age 12) and 21,564 adult onset cases (onset between ages 26 and 65), and 318,237 subjects without asthma (controls; older than age 38). We conducted GWASs for childhood onset asthma and adult onset asthma each compared to shared controls, and for age of asthma onset in all 37,846 asthma cases. 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

We detected 61 independent asthma loci: 23 were childhood onset specific, one was adult onset specific, and 37 were shared. Nineteen loci were associated with age of asthma onset. Genes at the childhood onset loci were most highly expressed in skin, blood and small intestine; genes at the adult onset 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.

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. In contrast, the onset of disease in childhood is associated with additional genes with relatively large effect sizes, and SNP-based heritability estimates that are over 3-times larger than for adult onset disease. 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 etiology of adult onset asthma is more lung-centered and environmentally determined, but with immune mediated mechanisms driving disease progression in both children and adults.

Funding

This work was supported by the National Institutes of Health grants R01 MH107666 and P30 DK20595 to HKI, R01 HL129735, R01 HL122712, P01 HL070831, and UG3 OD023282 to CO; NS was supported by T32 HL007605.

Key Words: asthma GWAS, adult onset asthma, childhood onset asthma, transcriptome, UK Biobank

Research in Context

Evidence before this study

Genome-wide association studies in large samples that include both childhood onset and adult onset asthma have 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 may mediate the effects of loci associated with childhoon and/or adult onset asthma.

Added value of this study

Leveraging the resources of 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 vs 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 role in disease initiation, whereas the other associated loci reflect shared mechanisms of disease progression. The childhood onset specific loci highlight skin as a primiary 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.

Introduction

Asthma is the most prevalent chronic respiratory disease worldwide[1]. Its diagnosis 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 is likely many different conditions. In particular, childhood onset asthma and adult onset asthma differ with respect to sex ratios, triggers of exacerbation, associated co-morbidities, severity[2,3], and potentially also for genetic risk factors[4,5]. For example, the most often replicated and most statistically significant genome-wide association study (GWAS) single nucleotide polymorphisms (SNPs) at the 17q12-21 locus are specific to asthma with onset of symptoms in early life[6]. Only one asthma GWAS to date was performed in adult onset cases. The GABRIEL Consortium included 1,947 cases with adult onset asthma (defined as 16 years of age or older) and 3,669 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 compared to the adult onset group, but no loci reached significance in the adult onset cases, likely due to low power. It remains unknown, therefore, whether loci other than 17q12-21 contribute specifically to childhood onset or adult onset asthma. Delineating genetic risk factors affecting age of onset from those that are shared between childhood and adult onset asthma could provide insights into the molecular mechanisms contributing to different clinical manifestations of asthma that is diagnosed at different ages.

To address this question and directly compare genetic risk architectures of adult onset and childhood onset asthma, we leveraged UK Biobank (UKB), a large-scale

101 prospective study collecting demographic, clinical, medical history, and genetic data for
102 nearly 500,000 participants[8]. We performed three asthma GWASs, one in 9,433 adults
103 with a diagnosis of asthma before age 12 years (childhood onset cases) and one in
104 21,564 adults with a diagnosis of asthma between age 26 and 65 years (adult onset
105 cases), each compared to 318,237 adults (>38 years) who did not have a diagnosis of
106 asthma or chronic obstructive lung disease (COPD) (controls). In our study sample, the
107 proportion of self-reported doctor diagnosed asthma was 37,846/318,237 (11·9%),
108 lower than the UK lifetime prevalence of "patient reported clinician diagnosed asthma"
109 of 15·6%[9] and consistent with the known healthy volunteer bias in UK Biobank[10]. In
110 addition, we conducted an age of onset GWAS in 37,846 asthma cases. Our goals were
111 to identify shared and distinct genetic risk loci for childhood and adult onset asthma, and
112 to identify genes that may mediate the effects of associations at age of onset specific
113 loci and those shared between childhood and adult onset asthma.
114

115 **Methods**

116 Sample composition and case definitions

117 Data for 376,358 British white individuals from UKB data release July 2017 were used[8].
118 We extracted disease status (asthma, allergic rhinitis, atopic dermatitis, food allergy,
119 COPD, emphysema, chronic bronchitis), age of onset of asthma, and sex from self-
120 reported questionnaires and hospital records (ICD10 codes) by querying our in-house
121 protected UKB database server[11]. See our 'reproducible research' and other details in
122 the appendix (pp 4-10). Among the 37,846 asthma cases included in our study, all had

123 self-reported doctor diagnosed asthma. Of those, 15,519 also had hospital records with
124 an asthma diagnosis (ICD10 codes).
125 We defined childhood and adult onset asthma using strict age of onset criteria
126 that would minimize the likelihood of misclassification, and considered asthma cases
127 with onset < 12 years of age as childhood onset cases (n=9,433; 3,462/9,433 [36·7%]
128 with ICD10 codes) and with onset > 25 and before 66 years of age as adult onset cases
129 (n=21,564; 9,260/21,564 [42·9%] with ICD10 codes). UKB participants without an
130 asthma diagnosis were included as controls (n=318,237). Individuals with COPD,
131 emphysema or chronic bronchitis were excluded from adult onset cases and controls.
132 For the age of onset GWAS, we included an additional 6,849 subjects with asthma with
133 onset between the ages of 12 and 25 years of age (2,797/6,849 [40·8%] with ICD10
134 codes). Characteristics of the sample are shown in Table 1. After genotype quality
135 control, 10,894,596 variants were available for analyses. See appendix (pp 4-7) for
136 additional details on genotype QC and phenotype definitions.
137
138 Genome-wide association studies
139 We conducted a childhood onset and adult onset GWAS by logistic regression and an
140 age of onset GWAS by linear regression, both using the allele dosages under an
141 additive genetic model, as implemented in Hail (https://github.com/hail-is/hail). All
142 analyses were performed using the Bionimbus Protected Data Cloud[12]. In all three
143 GWASs, we included sex and the first 10 genetic principal components as covariates
144 and used a genome-wide significance threshold of $5x10^{-8}$. FUMA[13], an integrative post-
145 GWAS annotation web-based tool, was used to define independent risk loci and identify

146 enriched tissues. FUMA defines independent loci using linkage disequilibrium (LD)
147 information from the 1000 Genomes project[14]. We specified an r^2 threshold >0·6
148 between genome-wide significant SNPs to represent a single locus. Tissue enrichments
149 were calculated in FUMA using a hypergeometric test to determine overrepresentation
150 of genes mapped to risk loci (by physical distance) among highly expressed genes in
151 each tissue in GTEx relative to all others.
152 Childhood onset and adult onset specific loci were defined as those that were
153 genome-wide significant in either the childhood onset or adult onset GWAS but were not
154 associated with asthma at p<0·05 in the other group, and the 95% confidence intervals
155 (CIs) of the respective ORs did not overlap. All other GWAS loci that were genome-wide
156 signficant in at least one of the two GWASs were considered to be shared.
157
158 SNP-based heritability estimation
159 We used Linkage Disequilibrium Score regression (LDSC) to estimate the SNP-based
160 heritability from the childhood onset, adult onset and age of onset GWAS summary
161 statistics, using SNPs overlapping with HapMap3 variants, as recommended[15], and
162 estimated population prevalence of 8·68% for childhood onset asthma and 9·55% for
163 adult onset asthma[16]. Heritability estimates for binary traits are reported in the liability
164 scale.
165
166 Sensitivity analysis to adult onset misclassification
167 We performed two sets of sensitivity analyses to assess 1) the effects of potential poor
168 recall of age of onset among subjects with adult onset asthma, and 2) the effects of

169 misclassification of COPD as asthma among the adult onset cases, even with exclusion
170 of cases with a reported diagnosis of COPD, emphysma or chronic bronchitis. First, to
171 assure that the adult onset cases did not include a significant proportion of childhood
172 onset asthma in which symptoms remitted in early life but then relapsed in adulthood,
173 we replaced adult onset cases with increasing proportions of randomly selected
174 childhood onset cases, and then tested for association at the two most significant
175 childhood onset specific loci. This procedure was repeated 20 times for each proportion
176 to quantify the sampling variability (appendix pp 7-8). Second, we performed two
177 analyses in which we removed either subjects with ages of asthma onset between 46 to
178 65 years or adult-onset cases and controls with FEV1/FVC <0·70. For each, we
179 compared p-values and ORs to the GWAS including all adult onset cases (appendix pp
180 8-9).
181
182 Predicted transcriptome association test
183 We used the PrediXcan[17] framework to identify genes that may mediate associations
184 between genetic variants and asthma risk. PrediXcan is a software tool that estimates
185 tissue-specific gene expression profiles from an individual's SNP genotype profile using
186 prediction models trained in large references databases of genotypes and tissue-
187 specific gene expression profiles. Using these genotype-imputed expression profiles,
188 PrediXcan can perform gene-based association tests that correlate predicted
189 expression levels with phenotypes (e.g., asthma) to identify candidate causal genes
190 from GWAS data. We used a summary version of PrediXcan, which has high
191 concordance with the individual level version (R^2 >0·99)[18]. For predictions, we

192 downloaded elastic net models trained with reference transcriptome data from the GTEx
193 consortium[19] (http://predictdb.org) for 49 tissues (appendix Table 1).
194 PrediXcan was run separately in the childhood onset and adult onset cases,
195 each with the same controls. Significance was determined using a Bonferroni correction
196 for the 38,608 genes ($p<1{\cdot}29x10^{-6}$) that were expressed in the five tissues determined
197 by FUMA to be enriched (skin, lung tissue, whole blood, small intestine and spleen; see
198 Results). We defined childhood onset specific genes as those whose predicted
199 expression is significantly associated with childhood onset asthma and the variants that
200 predict their expression are within childhood onset specific loci. Adult onset specific
201 genes were similarly defined. Because SNPs at shared loci may predict the expression
202 of genes that are associated only in childhood onset or adult onset cases, we also
203 considered genes to be age of onset specific if they were significantly associated with
204 asthma at $p<1{\cdot}29x10^{-6}$ in one age group and not associated with asthma at $p<0{\cdot}05$ in
205 the other. All other genes were considered shared.
206
207 Role of funding source
208 The funding source did not have any role in the study design; collection, analysis or
209 interpretation of data; in writing the manuscript; in the decision to submit the paper for
210 publication; or determining who has access to the raw data. The corresponding author
211 had full access to all the data in the study and had final responsibility for the decision to
212 submit for publication.
213

214 Results

215 Genome-wide association studies of asthma

216 We first conducted GWASs of childhood onset and adult onset asthma. These studies
217 revealed 61 independent loci associated with asthma, 52 were significant in the
218 childhood onset asthma GWAS and 19 were significant in the adult onset asthma
219 GWAS ($p<5x10^{-8}$) (Figure 1). The GWAS results were robust to inclusion of varying
220 numbers of PCs (10, 14, or 20) (appendix Figure 1) and to limiting the sample to cases
221 with diagnoses based on ICD10 codes (appendix Figure 2). Twenty-eight of the 61 loci,
222 in 27 chromosomal regions, were not previously reported in the GWAS catalog[20],
223 including one study also conducted in UKB subjects but focused on the phenotype
224 asthma+allergies[21]. Among the 28 new loci, 17 were significant in the childhood onset
225 GWAS, one was significant in the adult onset GWAS, and 10 were significant in both
226 (Table 2). Some of these loci contained genes that have been associated with asthma
227 in candidate gene studies (e.g., *FADS2*[22], *MUC5AC*[23,24] and *TBX21*[25]), which provide
228 independent validation of our GWAS findings. The lead SNP or an LD surrogate SNP
229 (r^2>0·40) was reported for 23 of the 28 new loci in the TAGC GWAS[26]. SNPs at 16 of
230 the 23 loci were associated with asthma in TAGC (p<0·05) and had the same direction
231 of effects as in the UKB GWASs (appendix Table 2). To test for potential effects of
232 asthma diagnostic criteria on the GWAS results, we conducted a GWAS comparing
233 individuals with asthma based on self-reported doctor diagnosis alone compared to
234 those with self-reported doctor diagnosis plus ICD10 codes. There were no differences
235 detected between these two groups (appendix Figure 3), indicating that differences
236 between these groups are not influencing the GWAS results discussed above.

237 As in previous GWASs comprised largely of children, the most significant locus in
238 the childhood onset GWAS is at 17q12[27], with the lead SNP in the *GSDMB* gene
239 (Figure 1A). However, the estimated ORs for the lead SNPs at three other loci were
240 similar to or larger than the lead SNP at the 17q locus. The lead SNPs in *IL1RL1* at
241 2q12.1 and in *EMSY* at 11q13.5 had effect sizes on childhood onset asthma similar to
242 the lead SNP in *GSDMB* on 17q (Table 2). Both loci have been prominent in previous
243 asthma GWAS. The lead SNP at the 1q21.3 locus, corresponding to a nonsense
244 mutation (R501X) in the filaggrin (*FLG*) gene at 1q21.3, had the largest OR overall.
245 Variants in *FLG* have been robustly associated with food allergies[28-30] and atopic
246 dermatitis[31-33], but previous associations with asthma have been in the context of other
247 allergic conditions[21,34]. Whether variants in *FLG* are associated with risk for childhood
248 onset asthma independent of its effects on early life allergic disease is yet unknown.
249 To address this possibility, we repeated the childhood onset GWAS after
250 excluding 3,205 childhood onset cases and 5,785 controls who reported having a
251 history of allergic rhinitis, AD or food allergy. As expected in a smaller sample the p-
252 values were overall larger, but the ORs were strikingly similar (appendix Figure 4 and
253 Table 3). Even though the OR at the *FLG* locus on 1q21.3 decreased from 1·97 (95%
254 CI 1·82, 2.13) to 1·61 (95% CI 1·49, 1·74), it remained both highly significant
255 ($p=2{\cdot}45\times10^{-19}$) and the largest OR for childhood onset. These results suggest both a
256 critical role for the allergic diathesis in the development of asthma in childhood and a
257 shared architecture between allergic disease and childhood onset asthma, as previously
258 discussed[34,35].

259 The most significant association in the adult onset GWAS was in the HLA region,
260 with independent associations at the HLA-C/B (6p21·33) and HLA-DR/DQ (6p21·32)
261 loci (Figure 1B). Compared to the childhood onset GWAS, effect sizes were quite small
262 in adult onset cases, with ORs reaching 1·1 at only five loci (2q12.1, 6p21.33, 6p21.32,
263 9p24.1, 10p14) (Table 2).

264 Among the 61 asthma loci, 23 were specific to childhood onset asthma and one
265 was specific to adult onset asthma (Table 2A-B). Regional association plots for the 23
266 loci with childhood or adult onset specific effects are shown in appendix Figure 5.
267 Among the remaining 38 shared loci (Table 2C), mean ORs were larger in the childhood
268 onset cases at all but six loci (permutation test $p<10^{-4}$; appendix pp 8-9), indicating that
269 both more loci contribute to childhood onset asthma and even among shared loci, effect
270 sizes are larger in childhood onset asthma cases (Figure 2). Colocalization analyses
271 using GWAS-PW[36] yielded results that supported our classification of age of onset
272 specific and shared loci (appendix Table 4).

273 Finally, to directly test for loci associated with asthma age of onset, we
274 conducted a third GWAS including all asthma cases in UKB who met our inclusion
275 criteria (n=37,846). In this analysis, 19 loci were associated with age of onset ($p<5x10^{-8}$)
276 (appendix Figure 6 and Table 5). Age of onset loci overlapped with both the childhood
277 and adult onset specific and shared loci, and asthma risk alleles at all but 2 loci were
278 associated with earlier age of onset (11q12 in the *FADS2* gene and 12q13.11 near the
279 *VDR* gene) (Table 2; Figure 2). SNPs at the 1q21.3 locus (*FLG*) had the largest effect
280 on age of onset, with each copy of the asthma risk allele (rs61816761) associated on
281 average with 4·57 (SE 0·43) years earlier onset compared to individuals without the risk

282 allele ($p=8{\cdot}15x10^{-27}$). At the 17q12 locus (rs4795399) each copy of the risk allele was
283 associated on average with 2·29 (SE 0·13) years earlier onset compared to individuals
284 without the risk allele ($p=6{\cdot}76x10^{-65}$). Examples of significant age of onset effects at
285 these and other loci are shown in Figure 3. Overall, both childhood onset specific and
286 shared asthma risk loci were associated with younger ages of onset, and alleles at loci
287 associated with younger ages of onset had larger effects compared to alleles at loci
288 associated with later ages of onset. These results are consistent with a previous GWAS
289 of age of asthma onset in European ancestry subjects that reported five genome-wide
290 significant and three suggestive significant loci, all associated with earlier age of
291 onset[37]. Six of those 8 loci are also associated with age of onset in the UKB GWAS
292 (appendix Table 6).
293
294 SNP-based heritability (h^2)
295 We used LD score regression to estimate the hertiabilities of childhood onset asthma,
296 adult onset asthma, and age of asthma onset. Consistent with the number of associated
297 SNPs and their effect sizes, estimated heritabilities were 0·33 for childhood onset
298 asthma, 0·098 for adult onset asthma, and 0·14 for age of asthma onset. After
299 excluding the significant SNPs (Table 2), estimates were reduced to 0·21, 0·082, and
300 0·087, respectively, indicating that the associated SNPs account for 0·11 of the
301 variance in childhood onset asthma risk, 0·016 of the variance in adult onset risk, and
302 0·049 of the variance in age of asthma onset risk. These results reflect the more
303 significant role for genetic variation in risk for childhood onset compared to adult onset

304 asthma and, conversely, the larger role for environmental variation in risk for adult onset
305 compared to childhood onset asthma.
306
307 Tissue-specific expression of genes at associated loci
308 Using an unbiased approach, we asked whether the tissue-specific expression of genes
309 that map to the 52 childhood onset loci differed from the tissue-specific expression of
310 genes mapped to the 19 adult onset loci. Genes at childhood onset loci (Figure 1A)
311 were most highly expressed in skin, whole blood, and small intestine (lower ileum)
312 compared to all other tissues, whereas genes at adult onset loci (Figure 1B) were most
313 highly expression in lung, whole blood, small intestine (lower ileum), and spleen
314 (enrichment for higher expression, $p<10x10^{-3}$) (appendix Figure 7 and Table 7). These
315 patterns suggest both overlapping and distinct underlying mechanisms associated with
316 asthma that begins in childhood and asthma with onset in adulthood.
317
318 Predicted transcriptome-wide association test
319 To better understand molecular mechanisms and to narrow the list of candidate causal
320 genes at associated loci, we focused on the five tissues that most highly expressed the
321 genes at childhood onset or adult onset loci: skin, lung tissue, whole blood, small
322 intestine, and spleen. We used PrediXcan[17] to identify genes whose expression is
323 predicted by variants associated with asthma in the childhood onset or adult onset
324 GWAS and potentially mediate the effects of associated SNPs on asthma risk.
325 This analysis identified 113 unique, candidate causal genes at 22 of the 61
326 GWAS loci ($p<1{\cdot}4x10^{-6}$) (Figures 4-5) (appendix Figures 8-9 and Table 8). These

327 included 39 genes associated with childhood onset asthma at eight of the childhood
328 onset specific loci and 76 genes associated with childhood and/or adult onset asthma at
329 13 of the shared loci. Variants at the one adult onset specific locus at 2q22·3 did not
330 predict the expression of any genes in the five tissues.

331 The predicted genes most significantly associated with childhood onset asthma
332 were at the 17q12 locus (Z score >10) in skin (*ORMDL3*, *ERBB2*, *PGAP3*, *GSDMA*, 2
333 long noncoding RNAs), lung (*ORMDL3*, *GSDMB*, *GSDMA*, *PGAP3*, *PNMT*), blood
334 (*ORMDL3*, *GSDMB*, *IKZF3*, *MED24*), small intestine (*GSDMA*, *GSDMB*, *PGAP3*), and
335 spleen (*ORMDL3*, *GSDMB*, *ZPBP2*, *MED24*). Some genes were predicted to be more
336 highly expressed in individuals with asthma (e.g., *ORMDL3*, *GSDMB*, *PGAP3*, *ERBB2*),
337 while others were predicted to be less expressed in individuals with asthma (e.g.,
338 *GSDMA*, *MED24*, *IKZF3*) (Figure 4A).This pattern of expression reflects the broad
339 regulatory effects of SNPs and tissue specificifity of gene expression at this locus[27]. The
340 childhood onset asthma locus at 1q21.3 includes genes essential for epidermal
341 differentiation and maintaining essential barrier function. The predicted expression of
342 nine genes at this locus were associated with childhood onset asthma. Higher predicted
343 expression of *CRNN*, *CRCT1* and *THEM5* in skin, of *PSDM4* in lung and blood, and of
344 *LINGO4* in skin, lung, and blood were associated with increased asthma risk. Lower
345 predicted expression of *SPRR2D* in skin, of *S100A12* in lung and blood, of *FLG* in skin,
346 lung and spleen, and of *TDRKH* in skin, lung, blood and spleen were associated with
347 increased asthma risk. *S100A12* has been previously implicated in asthma[38] and *FLG*
348 variants have been associated with atopic dermatitis and food allergies, and asthma in
349 the context of other allergic diseases[21,28-34]. Other childhood onset specific genes

350 previously implicated in asthma but not previously reported in asthma GWASs are
351 *CCL20*[39] at 2q36.3 and *TLR10*[40] at 4p14 in whole blood, and *TLR6*[41,42] at 4p14, *AP5B1*
352 at 11q13.1 and *SERPINB7*[43] at 18q21·33 in skin.
353 The 5q31.1 region had independent loci that were both childhood onset specific
354 and shared in the GWASs. Although the predicted expression of all eight asthma genes
355 at this extended locus were shared, five genes were more significantly associated with
356 childhood onset asthma: higher predicted expression of *RAD50* in skin but lower
357 predicted expression of *SEPT8* in skin and lung, *IL4* in skin, lung and blood, and *AFF4*
358 small intestine were associated with increased risk for asthma (Figure 4B). The
359 remaining three genes had similar associations with childhood and adult onset asthma,
360 with predicted lower expression of *IRF1* in skin and spleen and predicted higher
361 expression of *PDLIM4* in skin and *SLC22A5* in all five tissues associated with increased
362 asthma risk. *IL4, RAD50, SLC22A5* and *PDLIM4* have been highlighted in previous
363 asthma GWAS[26,44], and *KIF3A* was identified in a GWAS of the atopic march[45] and
364 associated with childhood onset asthma in a candidate gene study[46].
365 In contrast to all other loci, predicted expression of 44 genes at two independent
366 shared loci in the HLA region (6p21.32 and 6p21.33; referred to as the HLA region from
367 hereon in) were associated with childhood asthma only (n=1; in skin and lung), adult
368 onset asthma only (n=3; in skin only), or both (n=39; in multiple tissues). (Figure 5). The
369 sheer number of genes in this region with predicted expression associated with asthma,
370 the generally broad tissue expression patterns, and three associated with asthma only
371 in adult onset cases are consistent with this locus being among the two most signficant

372 loci in nearly all asthma GWASs, and the most significant locus in a previous small
373 GWASs of adult onset asthma[7] and in adults with asthma[44,47].
374 Among the remaining 37 shared loci (Figure 2B), SNPs at 12 predicted the
375 expression of 23 unique genes, all of which were associated with both childhood onset
376 and adult onset asthma. These include *IL18R1*, *IL18RAP*, and *ILRL2* at 2q12·1 in
377 multiple tissues, *TSLP* at 5q22.1 in skin, *SMAD3* at 15q22·33 in skin and blood, *LRP1*
378 at 12q13·3 in skin, *IL4R* at 16p12.1 in blood, and *CLEC16A* at 16p13.13 in lung. Loci
379 associated with *IL18R1*, *IL18RAP*, *IL18RAP, ILRL2, TSLP, SMAD3*, *LRP1*, *IL4R* and
380 *CLEC16A* were reported in previous asthma GWAS[21,26,36].
381

382 ## Discussion

383 We report here the first large GWAS of both childhood and adult onset cases. To both
384 maximize our power to detect differences and minimize the likelihood of
385 misclassification, we considered doctor diagnosed asthma before the age of 12 years
386 as childhood onset asthma and doctor diagnosed asthma after the age of 25 years as
387 adult onset asthma. These GWASs revealed 61 independent asthma loci, 23 specific to
388 childhood onset, one specific to adult onset, and 37 shared; with overall larger effect
389 sizes for childhood onset asthma at nearly all loci. Moreover, the predicted expression
390 of 41 of the 113 implicated genes were associated specifically with childhood onset
391 asthma, compared to the predicted expression of three genes associated specifically
392 with adult onset asthma. Our findings of more childhood onset asthma loci and
393 potentially causal genes, and the larger effect sizes of risk alleles in childhood onset
394 cases are particularly striking given that there were nearly 2·5-times more adult onset

395 than childhood onset cases in this study. Thus, despite having substantially less power
396 to detect loci specific to childhood onset asthma, our analyses revealed many more
397 childhood onset asthma loci. Similarly, the asthma risk alleles at 19 loci that were
398 significant in the age of onset GWAS were all associated with younger age of onset.
399 Finally, we showed that the SNP-based heritability of childhood onset asthma is over 3-
400 times larger than the SNP-based heritability of adult onset asthma. These findings are
401 consistent with previous studies showing decreased estimates of asthma heritability
402 with increasing age of onset[48], age of onset SNPs associated with earlier age of
403 onset[37], and an additive, unweighted genetic risk score comprised of 15 SNPs at eight
404 asthma-associated loci associated with earlier age of onset[49]. Our study further shows
405 that genetic risk for adult onset asthma is largely a subset of the genetic risk loci for
406 childhood onset asthma, but with overall smaller effect sizes, consistent with a larger
407 role for environmental risk factors in adult onset asthma.
408 Despite the overlap of adult onset and childhood onset loci, distinct mechanisms
409 contributing to each were suggested by tissue enrichments: childhood onset loci were
410 enriched for genes with highest expression in skin whereas adult onset loci were
411 enriched for genes with highest expression in lung and spleen; both were enriched for
412 genes highly expressed in whole blood and small intestine. The highlighting of skin as a
413 target tissue for childhood onset asthma supports the widely held idea that asthma in
414 childhood is due to impaired barrier function in the skin and other epithelial surfaces.
415 This model proposes that compromised epithelial barriers promote sensitization to food
416 and airway allergens and to wheezing illnesses in early life[35,50]. In fact, childhood onset
417 specific loci identified here have been associated with atopic dermatitis or food allergies,

418 such as *FLG* on 1q21·3 with the atopic march[45], atopic dermatitis[31-33] and food
419 allergies[28-30], *KIF3A* on 5q31.1 and *AP5B1*/*OVOL1* on 11q13.1 with the the atopic
420 march[45] and atopic dermatitis[51], *SERPINB7* on 18q21.33 with food allergies[43], and
421 *CRNN* (cornulin) on 1q21·3 with atopic dermatitis concomitant with asthma and reduced
422 expression in atopic dermatitis-affected skin[52]. Variants at those loci were all associated
423 with earlier age of asthma onset. We further show that these loci are associated with
424 childhood onset asthma, even after exclusion of cases with a history of allergic
425 diseases. In contrast, the enrichment for genes highly expressed in lung and spleen at
426 adult onset loci suggests a more lung-centered, and potentially immune mediated,
427 etiology for asthma with onset later in life. The prominant role of the HLA region in the
428 adult onset asthma GWAS and the fact that predicted expression of three HLA region
429 genes was associated only with adult onset asthma further highlights a central role for
430 immune processes driving asthma pathogenesis in adults. The fact that both childhood
431 onset and adult onset asthma loci were enriched for genes that are most highly
432 expressed in whole blood cells and small intestine further indicate a shared immune
433 etiology, as suggested from a large GWAS that included both children and adults[26].
434 Combining GWAS with a transcriptome-wide association test that uses
435 combinations of associated SNPs to predict gene expression in different tissues
436 revealed significant complexity at the two most highly associated asthma loci. SNPs at
437 the 17q12 locus predicted expression of 18 childhood onset asthma genes and SNPs at
438 the HLA region predicted expression of 42 genes: three were associated with adult
439 onset asthma and most were not HLA genes *per se*. In this regard, it is notable that the
440 *HLA-DRB1*, *HLA-DQB1*, and *HLA-DQA1* genes, which are the most associated HLA

441 loci with autoimmune diseases, are predicted to have reduced expression in both
442 childhood onset and adult onset asthma. Instead, HLA genes with less clear functions
443 have increased predicted expression in asthma (Figure 5). These results strengthen the
444 argument that multiple genes contribute to asthma risk at the HLA and 17q12 loci and
445 probably account for the highly significant GWAS p-values observed at these loci in
446 nearly all studies. It is also likely that these genes have both tissue specific and broad
447 effects in epithelium, lung, and immune tissues.

448 The new loci identified in our study include the first adult onset asthma specific
449 association at 2q22.3. The lead SNP at 2q22.3 is intergenic between *TEX41* and
450 *ACVR2A*. The predicted expression of *ACVR2A* was not associated with asthma in our
451 study, despite it being expressed in lung, blood, small intestine and spleen. *TEX41* was
452 not expressed in any of the five tissues investigated. Interestingly, a GWAS also
453 performed in UKB subjects implicated variants near *TEX41* in heavy vs. never smoking
454 behavior[53]. However, even after removing adult onset cases and controls with reported
455 'ever smoking', the p-value for this SNP remained significant and the OR slightly
456 increased (OR 1·077 [95% CI 1·05, 1.1], $p=2{\cdot}26 \times 10^{-8}$; n=12,132 cases and 176,704
457 controls). Variants in or near this gene, which encodes a lincRNA, have been
458 associated with cardiovascular and immune mediated traits[54], making this a potentially
459 interesting candidate gene for adult onset asthma.

460 Our study had limitations. First, diagnoses of asthma and allergic disease in
461 study subjects were from self reported doctor diagnosis and medical records (ICD10
462 codes). Thus, it is possible that diagnoses, age of onset, or both are misspecified in
463 some subjects. On the one hand, the large sample size and our ability to replicate

464 nearly all previously reported asthma loci (appendix Tables 2 and 9) suggest that our
465 analyses were robust to any inaccuracies in the data. On the other hand, it is possible
466 that subjects with adult onset asthma included cases with poor recall of childhood onset
467 asthma in which symptoms remitted and then relapsed later in life[55] or misclassified
468 cases of COPD among the older age groups. Our sensitivity analysis suggested that if
469 even as few as 5% of the adult onset cases were misclassified we should have
470 observed some signal of association at childhood onset loci, which we did not. The fact
471 that the odds ratios for asthma at shared loci are relatively similar from approximately
472 age 25 to age 65 (Figure 3) and that we do not detect any association signal at the
473 major COPD locus on chromosome 15q25.1 (appendix Figure 10), further suggests that
474 there is negligible misclassification of cases in the older age groups. Second, although
475 we used stringent criteria to classify loci as childhood or adult onset specific, we can't
476 exclude the possibility that in infinitely large sample sizes the effect sizes of some of
477 these loci will have 95% CIs that overlap or the association p-value will become smaller
478 than 0·05. Conversely, some of the shared loci with modest p-values in the adult onset
479 cases may not be true risk loci for asthma with onset at older ages. Third, the gene
480 expression data used to predict candidate target genes included heterogeneous tissues
481 and were collected mostly from adults. As a result, our study may have missed relevant
482 genes whose expression is developmentally regulated or environment specific. Our
483 finding of candidate genes at only 22 of the 61 asthma loci may be due in part to the
484 importance of both in asthma pathogenesis. Moreover, all inference based on gene
485 expression is using imputed expression. It is possible, therefore, that some relevant
486 genes were more difficult to impute and not included in our analysis, although a recent

487 comparative study showed that PrediXcan is a more robust method for prediction of
488 gene expression than other related methods[56]. Fourth, because of the ethnic
489 composition of UKB, this study was limited to individuals of European ancestry only. As
490 a result, we could not evaluate the genetic risk architecture or assess the effects of age
491 of onset specific loci in other populations.

492 In the largest asthma GWAS to date, we show that genetic risk loci for adult
493 onset asthma is largely a subset of the loci associated with childhood onset asthma,
494 with overall smaller effect sizes for onset at later ages. These data suggest that
495 childhood onset specific loci and those associated with age of onset play a role in
496 disease initiation, whereas the other associated loci reflect shared mechanisms of
497 disease progression. The differences in the target tissues that most highly express the
498 genes at associated loci and the predicted expression of genes at age specific and
499 shared loci provides additional genetic and molecular evidence for both shared and
500 distinct pathogenic mechanisms in childhood onset and adult onset asthma. It is
501 therefore possible that the most effective treatments will also differ between these two
502 groups, and that strategies for precision medicine should be further personalized to
503 account for age of asthma onset.
504

505 Author's Contributions

506 All authors were involved in the conception and design of the study and in writing the
507 manuscript. MP and NS conducted all analyses and prepared figures and tables, under
508 the overall supervision of DLN, CO, and HKI.
509

URLs

UK Biobank, [https://www.ukbiobank.ac.uk/]

Hail: https://github.com/hail-is/hail

Acknowledgements

This research has been conducted using UK Biobank Resource under Application Number 19526.

Conflict of Interest

Dr. Im reports personal fees from AbbVie, personal fees from GSK, outside the submitted work. The other authors declare no conflicts of interest.

521 References

522 1. Collaborators GBDCRD. Global, regional, and national deaths, prevalence,
523 disability-adjusted life years, and years lived with disability for chronic obstructive
524 pulmonary disease and asthma, 1990-2015: a systematic analysis for the Global
525 Burden of Disease Study 2015. *Lancet Respir Med* 2017; **5**(9): 691-706.

526 2. Bush A, Menzies-Gow A. Phenotypic differences between pediatric and adult
527 asthma. *Proc Am Thorac Soc* 2009; **6**(8): 712-9.

528 3. Busse W, Banks-Schlegel SP, Larsen GL. Childhood- versus adult-onset
529 asthma. *Am J Respir Crit Care Med* 1995; **151**(5): 1635-9.

530 4. Dijk FN, de Jongste JC, Postma DS, Koppelman GH. Genetics of onset of
531 asthma. *Curr Opin Allergy Clin Immunol* 2013; **13**(2): 193-202.

532 5. Holloway JW, Arshad SH, Holgate ST. Using genetics to predict the natural
533 history of asthma? *J Allergy Clin Immunol* 2010; **126**(2): 200-9; quiz 10-1.

534 6. Bouzigon E, Corda E, Aschard H, et al. Effect of 17q21 variants and smoking
535 exposure in early-onset asthma. *N Engl J Med* 2008; **359**(19): 1985-94.

536 7. Moffatt MF, Gut IG, Demenais F, et al. A large-scale, consortium-based
537 genomewide association study of asthma. *N Engl J Med* 2010; **363**(13): 1211-21.

538 8. Bycroft C, Freeman C, Petkova D, et al. The UK Biobank resource with deep
539 phenotyping and genomic data. *Nature* 2018; **562**(7726): 203-9.

540 9. Mukherjee M, Stoddart A, Gupta RP, et al. The epidemiology, healthcare and
541 societal burden and costs of asthma in the UK and its member nations: analyses of
542 standalone and linked national databases. *BMC Med* 2016; **14**(1): 113.

543 10. Fry A, Littlejohns TJ, Sudlow C, et al. Comparison of Sociodemographic and
544 Health-Related Characteristics of UK Biobank Participants With Those of the General
545 Population. *Am J Epidemiol* 2017; **186**(9): 1026-34.
546 11. Pividori M, Im HK. ukbREST: efficient and streamlined data access for
547 reproducible research in large biobanks. *Bioinformatics* 2018.
548 12. Heath AP, Greenway M, Powell R, et al. Bionimbus: a cloud for managing,
549 analyzing and sharing large genomics datasets. *J Am Med Inform Assoc* 2014; **21**(6):
550 969-75.
551 13. Watanabe K, Taskesen E, van Bochoven A, Posthuma D. Functional mapping
552 and annotation of genetic associations with FUMA. *Nature communications* 2017; **8**(1):
553 1826.
554 14. Consortium TGP, Auton A, Brooks LD, et al. A global reference for human
555 genetic variation. *Nature* 2015; **526**(7571): 68-74.
556 15. Bulik-Sullivan BK, Loh PR, Finucane HK, et al. LD Score regression
557 distinguishes confounding from polygenicity in genome-wide association studies. *Nat*
558 *Genet* 2015; **47**(3): 291-5.
559 16. https://www.asthma.org.uk/get-involved/campaigns/data-visualisations/.
560 17. Gamazon ER, Wheeler HE, Shah KP, et al. A gene-based association method
561 for mapping traits using reference transcriptome data. *Nat Genet* 2015; **47**(9): 1091-8.
562 18. Barbeira AN, Dickinson SP, Bonazzola R, et al. Exploring the phenotypic
563 consequences of tissue specific gene expression variation inferred from GWAS
564 summary statistics. *Nature communications* 2018; **9**(1): 1825.

565 19. Battle A, Brown CD, Engelhardt BE, Montgomery SB. Genetic effects on gene
566 expression across human tissues. *Nature* 2017; **550**(7675): 204-13.
567 20. MacArthur J, Bowler E, Cerezo M, et al. The new NHGRI-EBI Catalog of
568 published genome-wide association studies (GWAS Catalog). *Nucleic Acids Res* 2017;
569 **45**(D1): D896-D901.
570 21. Zhu Z, Lee PH, Chaffin MD, et al. A genome-wide cross-trait analysis from UK
571 Biobank highlights the shared genetic architecture of asthma and allergic diseases. *Nat*
572 *Genet* 2018.
573 22. Standl M, Sausenthaler S, Lattka E, et al. FADS gene cluster modulates the
574 effect of breastfeeding on asthma. Results from the GINIplus and LISAplus studies.
575 *Allergy* 2012; **67**(1): 83-90.
576 23. Rousseau K, Byrne C, Griesinger G, et al. Allelic association and recombination
577 hotspots in the mucin gene (MUC) complex on chromosome 11p15.5. *Ann Hum Genet*
578 2007; **71**(Pt 5): 561-9.
579 24. Johnson L, Shah I, Loh AX, et al. MUC5AC and inflammatory mediators
580 associated with respiratory outcomes in the British 1946 birth cohort. *Respirology* 2013;
581 **18**(6): 1003-10.
582 25. Akahoshi M, Obara K, Hirota T, et al. Functional promoter polymorphism in the
583 TBX21 gene associated with aspirin-induced asthma. *Human genetics* 2005; **117**(1):
584 16-26.
585 26. Demenais F, Margaritte-Jeannin P, Barnes KC, et al. Multiancestry association
586 study identifies new asthma risk loci that colocalize with immune-cell enhancer marks.
587 *Nat Genet* 2018; **50**(1): 42-53.

588 27. Stein MM, Thompson EE, Schoettler N, et al. A decade of research on the
589 17q12-21 asthma locus: Piecing together the puzzle. *J Allergy Clin Immunol* 2018;
590 **142**(3): 749-64 e3.
591 28. Brown SJ, Asai Y, Cordell HJ, et al. Loss-of-function variants in the filaggrin gene
592 are a significant risk factor for peanut allergy. *J Allergy Clin Immunol* 2011; **127**(3): 661-
593 7.
594 29. Marenholz I, Kerscher T, Bauerfeind A, et al. An interaction between filaggrin
595 mutations and early food sensitization improves the prediction of childhood asthma. *J*
596 *Allergy Clin Immunol* 2009; **123**(4): 911-6.
597 30. Venkataraman D, Soto-Ramirez N, Kurukulaaratchy RJ, et al. Filaggrin loss-of-
598 function mutations are associated with food allergy in childhood and adolescence. *J*
599 *Allergy Clin Immunol* 2014; **134**(4): 876-82 e4.
600 31. Bisgaard H, Halkjaer LB, Hinge R, et al. Risk analysis of early childhood eczema.
601 *J Allergy Clin Immunol* 2009; **123**(6): 1355-60 e5.
602 32. Palmer CN, Irvine AD, Terron-Kwiatkowski A, et al. Common loss-of-function
603 variants of the epidermal barrier protein filaggrin are a major predisposing factor for
604 atopic dermatitis. *Nat Genet* 2006; **38**(4): 441-6.
605 33. Weidinger S, Illig T, Baurecht H, et al. Loss-of-function variations within the
606 filaggrin gene predispose for atopic dermatitis with allergic sensitizations. *J Allergy Clin*
607 *Immunol* 2006; **118**(1): 214-9.
608 34. Ferreira MA, Vonk JM, Baurecht H, et al. Shared genetic origin of asthma, hay
609 fever and eczema elucidates allergic disease biology. *Nat Genet* 2017; **49**(12): 1752-7.

610 35. Busse WW. The atopic march: Fact or folklore? *Ann Allergy Asthma Immunol*
611 2018; **120**(2): 116-8.
612 36. Pickrell JK, Berisa T, Liu JZ, Segurel L, Tung JY, Hinds DA. Detection and
613 interpretation of shared genetic influences on 42 human traits. *Nat Genet* 2016; **48**(7):
614 709-17.
615 37. Sarnowski C, Sugier PE, Granell R, et al. Identification of a new locus at 16q12
616 associated with time to asthma onset. *J Allergy Clin Immunol* 2016; **138**(4): 1071-80.
617 38. Yang Z, Yan WX, Cai H, et al. S100A12 provokes mast cell activation: a potential
618 amplification pathway in asthma and innate immunity. *J Allergy Clin Immunol* 2007;
619 **119**(1): 106-14.
620 39. Collison A, Hatchwell L, Verrills N, et al. The E3 ubiquitin ligase midline 1
621 promotes allergen and rhinovirus-induced asthma by inhibiting protein phosphatase 2A
622 activity. *Nat Med* 2013; **19**(2): 232-7.
623 40. Lazarus R, Raby BA, Lange C, et al. TOLL-like receptor 10 genetic variation is
624 associated with asthma in two independent samples. *Am J Respir Crit Care Med* 2004;
625 **170**(6): 594-600.
626 41. Tantisira K, Klimecki WT, Lazarus R, et al. Toll-like receptor 6 gene (TLR6):
627 single-nucleotide polymorphism frequencies and preliminary association with the
628 diagnosis of asthma. *Genes Immun* 2004; **5**(5): 343-6.
629 42. Lau MY, Dharmage SC, Burgess JA, et al. The interaction between farming/rural
630 environment and TLR2, TLR4, TLR6 and CD14 genetic polymorphisms in relation to
631 early- and late-onset asthma. *Sci Rep* 2017; **7**: 43681.

632 43. Marenholz I, Grosche S, Kalb B, et al. Genome-wide association study identifies
633 the SERPINB gene cluster as a susceptibility locus for food allergy. *Nature*
634 *communications* 2017; **8**(1): 1056.
635 44. Li X, Howard TD, Zheng SL, et al. Genome-wide association study of asthma
636 identifies RAD50-IL13 and HLA-DR/DQ regions. *J Allergy Clin Immunol* 2010; **125**(2):
637 328-35 e11.
638 45. Marenholz I, Esparza-Gordillo J, Ruschendorf F, et al. Meta-analysis identifies
639 seven susceptibility loci involved in the atopic march. *Nature communications* 2015; **6**:
640 8804.
641 46. Kovacic MB, Myers JM, Wang N, et al. Identification of KIF3A as a novel
642 candidate gene for childhood asthma using RNA expression and population allelic
643 frequencies differences. *PloS one* 2011; **6**(8): e23714.
644 47. Lasky-Su J, Himes BE, Raby BA, et al. HLA-DQ strikes again: genome-wide
645 association study further confirms HLA-DQ in the diagnosis of asthma among adults.
646 *Clin Exp Allergy* 2012; **42**(12): 1724-33.
647 48. Thomsen SF, Duffy DL, Kyvik KO, Backer V. Genetic influence on the age at
648 onset of asthma: a twin study. *J Allergy Clin Immunol* 2010; **126**(3): 626-30.
649 49. Belsky DW, Sears MR, Hancox RJ, et al. Polygenic risk and the development
650 and course of asthma: an analysis of data from a four-decade longitudinal study. *Lancet*
651 *Respir Med* 2013; **1**(6): 453-61.
652 50. Han H, Roan F, Ziegler SF. The atopic march: current insights into skin barrier
653 dysfunction and epithelial cell-derived cytokines. *Immunol Rev* 2017; **278**(1): 116-30.

654 51. Johansson E, Biagini Myers JM, Martin LJ, et al. KIF3A genetic variation is
655 associated with pediatric asthma in the presence of eczema independent of allergic
656 rhinitis. *J Allergy Clin Immunol* 2017; **140**(2): 595-8 e5.
657 52. Trzeciak M, Sakowicz-Burkiewicz M, Wesserling M, et al. Altered Expression of
658 Genes Encoding Cornulin and Repetin in Atopic Dermatitis. *Int Arch Allergy Immunol*
659 2017; **172**(1): 11-9.
660 53. Wain LV, Shrine N, Miller S, et al. Novel insights into the genetics of smoking
661 behaviour, lung function, and chronic obstructive pulmonary disease (UK BiLEVE): a
662 genetic association study in UK Biobank. *Lancet Respir Med* 2015; **3**(10): 769-81.
663 54. MacArthur DG, Balasubramanian S, Frankish A, et al. A systematic survey of
664 loss-of-function variants in human protein-coding genes. *Science* 2012; **335**(6070): 823-
665 8.
666 55. Sears MR, Greene JM, Willan AR, et al. A longitudinal, population-based, cohort
667 study of childhood asthma followed to adulthood. *N Engl J Med* 2003; **349**(15): 1414-22.
668 56. Fryett JJ, Inshaw J, Morris AP, Cordell HJ. Comparison of methods for
669 transcriptome imputation through application to two common complex diseases. *Eur J*
670 *Hum Genet* 2018.
671

672 Table 1. Characteristics of the asthma cases and controls. NA, not applicable.

	Childhood Onset (N=9,433)	**Adolescent or Young Adult Onset (N=6,849)**	**Adulthood Onset (N=21,564)**	**Controls (N=318,237)**
Age at recruitment (yr)				
Mean (±SD)	55±8	52±8	57±8	57±8
Range	40 - 70	40 - 70	40 - 70	39 - 73
Age at asthma onset (yr)[a]				
Mean (±SD)	6±3	19±4	44±10	NA
Range	0 - 11	12 - 25	26 - 65	NA
Sex (% female)	40·7	57·4	63·6	53·5
Asthma medication use in the past 12 months, N (%)	1,383 (14·7)	1,124 (16·4)	4,081 (18·9)	NA
Current smokers N (%)	802 (8·5)	688 (10·0)	1,379 (6·4)	30,056 (9·4)
Allergic diseases ever				
Allergic rhinitis, N (%)	2,537 (26·9)	2,042 (29·8)	4,660 (21·6)	27,289 (8·6)
Atopic dermatitis, N (%)	1,126 (11·9)	473 (6·9)	927 (4·3)	7,168 (2·3)
Food allergy, N (%)	138 (1·5)	63 (0·9)	172 (0·8)	1,229 (0·4)
Any allergic disease, N (%)	3,205 (34·0)	2,321 (33·9)	5,306 (24·6)	33,808 (10·6)

673

Table 2. Regions with SNPs that were genome-wide significant in the childhood onset or adult onset GWAS. Information for the SNP with smallest p-value at each locus is shown. Summary statistics for the 19 age of asthma onset significant SNPs are shown in appendix Table 5.

Locus[a]	rsID	Position[b]	Nearby Genes[c]	Allele[d]	RAF[e]	Childhood Onset Asthma GWAS			Adult Onset Asthma GWAS			Age of Onset GWAS		
						OR	95% CI	p-value	OR	95% CI	p-value	Beta	SE	p-value
A. Childhood Onset Specific Loci														
1q21.3	rs61816761	1:152285861	*FLG, HRNR, FLG2*	G/A	0·024	1·970	1·823-2·129	1·88E-65	1·016	0·948-1·088	6·56E-01	-4·571	0·426	8·15E-27
1q25.1	rs7518129	1:173163568	*TNFSF4, TNSF18, PRDX6*	A/G	0·310	1·111	1·077-1·146	2·21E-11	0·997	0·977-1·019	8·17E-01	-0·849	0·145	4·89E-09
2q36.3[f]	rs10175070	2:228670575	*CCL20, SLC18A3*	A/G	0·251	1·122	1·086-1·160	4·31E-12	1·000	0·978-1·023	9·92E-01	-0·862	0·154	2·24E-08
3q28	rs12634152	3:188121019	*LPP, FLJ42393, LPP-AS1*	C/T	0·547	1·133	1·100-1·166	1·08E-16	1·007	0·987-1·027	4·98E-01	-0·968	0·136	1·21E-12
4p14	rs5743618	4:38798648	*TLR1, TLR10, TLR6*	A/C	0·774	1·249	1·203-1·295	3·19E-32	1·008	0·985-1·032	5·09E-01	-1·578	0·163	4·53E-22
5q13.2[f]	rs10036789	5:71695918	*PTCD2, ZNF366*	C/G	0·460	1·085	1·054-1·117	4·51E-08	1·016	0·996-1·036	1·12E-01	-0·511	0·137	1·95E-04
5q31.1	rs2051809	5:132056874	*KIF3A, IL4, CCNI2*	C/A	0·247	1·175	1·137-1·214	3·24E-22	1·016	0·994-1·040	1·58E-01	-0·933	0·155	1·59E-09
6p25.3[f]	rs9391997	6:409119	*IRF4, DUSP22, EXOC2*	A/G	0·527	1·102	1·070-1·135	6·89E-11	0·995	0·976-1·015	6·35E-01	-0·597	0·136	1·06E-05
7p15.3[f]	rs34880821	7:22775450	*IL6, TOMM7*	G/A	0·283	1·106	1·071-1·141	4·73E-10	1·000	0·979-1·022	9·88E-01	-0·667	0·149	7·59E-06
8q24.21[f]	rs13277355	8:128777719	*MYC, TMEM75*	G/A	0·274	1·110	1·075-1·146	1·65E-10	1·013	0·991-1·036	2·46E-01	-0·660	0·151	1·20E-05
9q34.3[f]	rs117137535	9:140500443	*ARRDC1, ZMYND19, EHMT1*	G/A	0·026	1·307	1·195-1·429	4·16E-09	0·977	0·913-1·046	5·10E-01	-2·464	0·446	3·42E-08
10p15.1[f]	rs943451	10:6621773	*PRKCQ, PFKFB3, SFMBT2*	C/T	0·318	1·125	1·091-1·161	7·51E-14	1·021	1·000-1·043	5·30E-02	-0·613	0·145	2·54E-05
11q13.1[f]	rs479844	11:65551957	*AP5B1, OVOL1*	A/G	0·555	1·106	1·074-1·139	1·48E-11	1·015	0·995-1·035	1·41E-01	-0·657	0·135	1·22E-06
11q23.3[f]	rs12365699	11:118743286	*CXCR5, DDX6*	A/G	0·833	1·167	1·120-1·216	1·57E-13	1·004	0·978-1·031	7·43E-01	-1·396	0·185	4·42E-14
12q13.2[f]	rs62623446	12:55368291	*TESPA1, MUCL1, NEUROD4*	C/T	0·072	1·188	1·127-1·252	1·40E-10	1·038	0·999-1·077	5·32E-02	-1·079	0·253	1·95E-05
12q24.12[f]	rs10774625	12:111910219	*ATXN2, SH2B3, BRAP*	A/G	0·504	1·087	1·056-1·119	2·07E-08	0·997	0·978-1·017	7·92E-01	-0·713	0·135	1·43E-07
12q24.31[f]	rs1696361	12:121363823	*SPPL3, HNF1A*	C/T	0·359	1·109	1·076-1·142	1·19E-11	1·019	0·998-1·040	6·96E-02	-0·696	0·140	6·84E-07
17q12	rs4795399	17:38061439	*GSDMB, ZPBP2, ORMDL3*	C/T	0·529	1·406	1·365-1·448	1·45E-111	1·005	0·985-1·025	6·28E-01	-2·286	0·134	6·76E-65
17q21.2[f]	rs8066625	17:40390629	*STAT5B, GHDC, STAT5A*	G/A	0·100	1·151	1·098-1·206	4·90E-09	1·006	0·973-1·040	7·33E-01	-0·846	0·227	1·88E-04
17q21.32[f]	rs56308324	17:45819206	*TBX21, TBKBP1, OSBPL7*	A/T	0·132	1·132	1·086-1·180	3·47E-09	1·025	0·996-1·055	8·72E-02	-0·714	0·195	2·54E-04
18q21.33[f]	rs4574025	18:60009814	*TNFRSF11A, KIAA1468, ZCCHC2*	C/T	0·535	1·090	1·058-1·122	9·25E-09	0·986	0·967-1·006	1·69E-01	-0·871	0·136	1·61E-10
18q21.33[f]	rs12964116	18:61442619	*SERPINB7, SERPINB11, SERPINB2*	A/G	0·036	1·336	1·247-1·432	2·56E-16	1·005	0·953-1·059	8·62E-01	-1·916	0·351	4·87E-08
19p13.3[f]	rs4807630	19:1170445	*SBNO2, GPX4, STK11*	C/T	0·308	1·093	1·060-1·128	2·06E-08	1·013	0·992-1·035	2·32E-01	-0·487	0·147	9·25E-04
B. Adult Onset Specific Loci														
2q22.3[f]	rs12617922	2:146156679	*TEX41, ACVR2A*	A/G	0·518	0·983	0·955-1·012	2·36E-01	1·066	1·045-1·087	1·59E-10	0·478	0·136	4·46E-04
C. Shared Asthma Loci														
1q32.1	rs12023876	1:203093201	*ADORA1, PPFIA4*	T/G	0·668	1·102	1·068-1·137	1·11E-09	1·032	1·010-1·053	3·37E-03	-0·567	0·144	8·65E-05
2p25.1	rs13416555	2:8441735	*RNF144A, ID2*	G/C	0·705	1·130	1·094-1·167	1·94E-13	1·040	1·018-1·063	3·50E-04	-0·715	0·150	1·91E-06
2q12.1	rs72823641	2:102936159	*IL1RL1, IL1RL2, IL18R1*	A/T	0·863	1·415	1·349-1·484	3·01E-46	1·109	1·077-1·142	3·72E-12	-1·687	0·208	4·73E-16
2q37.3	rs34290285	2:242698640	*D2HGDH, ING5, GAL3ST2*	A/G	0·745	1·177	1·137-1·219	2·06E-20	1·081	1·056-1·106	2·16E-11	-0·821	0·159	2·56E-07
3p22.3[f]	rs35570272	3:33047662	*GLB1, TRIM71, TMPPE*	G/T	0·396	1·100	1·068-1·133	2·68E-10	1·046	1·025-1·067	9·84E-06	-0·454	0·138	1·03E-03
4q27	rs2069763	4:123377482	*IL2, ADAD1, IL21*	C/A	0·334	1·126	1·092-1·160	1·59E-14	1·032	1·011-1·054	2·44E-03	-0·631	0·142	8·39E-06
5p15.2[f]	rs16903574	5:14610309	*FAM105A, TRIO, OTULIN*	C/G	0·077	1·236	1·174-1·302	9·01E-16	1·052	1·013-1·091	7·75E-03	-1·055	0·250	2·51E-05
5q22.1	rs1837253	5:110401872	*SLC25A46, TSLP*	T/C	0·740	1·211	1·169-1·253	2·33E-27	1·088	1·064-1·113	2·77E-13	-0·733	0·158	3·72E-06
5q31.1	rs17622378	5:131778452	*C5orf56, SLC22A5, IRF1*	G/A	0·573	1·103	1·071-1·136	8·52E-11	1·077	1·055-1·098	3·17E-13	-0·044	0·137	7·46E-01
6p22.1	rs1117490	6:30170510	*TRIM26, TRIM15, TRIM39*	T/C	0·234	1·100	1·063-1·137	2·78E-08	1·059	1·036-1·084	6·83E-07	-0·453	0·158	4·07E-03
6p21.33	rs2428494	6:31322197	*HLA-B, HLA-C, MICA*	T/A	0·477	1·157	1·124-1·191	7·94E-23	1·104	1·082-1·126	4·77E-23	-0·648	0·136	1·78E-06
6p21.32	rs28407950	6:32626348	*HLA-DQA1, HLA-DQB*	T/C	0·756	1·354	1·306-1·405	1·27E-59	1·148	1·121-1·175	7·67E-31	-1·284	0·163	4·03E-15
6q15	rs1321859	6:91011673	*BACH2, MAP3K7*	T/C	0·649	1·102	1·068-1·136	9·32E-10	1·076	1·054-1·098	5·68E-12	-0·268	0·144	6·28E-02
7p21.1[f]	rs4473914	7:20426263	*ITGB6, MACC1, ABCB5*	C/T	0·604	1·092	1·060-1·125	9·02E-09	1·035	1·014-1·056	8·44E-04	-0·338	0·139	1·51E-02
7p15.1[f]	rs917115	7:28172586	*JAZF1, TAX1BP1, CREB5*	T/C	0·208	1·112	1·074-1·152	1·73E-09	1·042	1·017-1·067	7·35E-04	-0·577	0·164	4·17E-04
8q21.13	rs4739738	8:81291645	*TPD52, ZBTB10*	A/G	0·359	1·117	1·084-1·151	3·45E-13	1·057	1·036-1·079	8·63E-08	-0·450	0·140	1·27E-03
9p24.1	rs992969	9:6209697	*RANBP6, IL33*	G/A	0·252	1·248	1·209-1·289	6·76E-42	1·100	1·076-1·125	3·14E-17	-1·026	0·152	1·35E-11
10p14[f]	rs7894791	10:8591369	*GATA3, CELF2*	A/C	0·587	1·103	1·071-1·137	9·43E-11	1·024	1·004-1·045	1·94E-02	-0·494	0·137	3·24E-04
10p14	rs1775554	10:9054340	*GATA3, CELF2*	C/A	0·577	1·117	1·084-1·150	2·85E-13	1·121	1·098-1·143	5·17E-29	-0·002	0·138	9·91E-01
11p15.5[f]	rs12788104	11:1123739	*MUC6, MUC5AC*	A/G	0·688	1·029	0·997-1·061	7·99E-02	1·064	1·041-1·087	1·41E-08	0·147	0·148	3·22E-01

11q12.2 [f]	rs174621	11:61630104	*FADS2, FADS1, FADS3*	A/G	0·772	1·013	0·978-1·049	4·63E-01	1·071	1·046-1·097	1·46E-08	0·508	0·163	1·87E-03
11q13.5 [g]	rs61894547	11:76248630	*EMSY, THAP12, LRRC32*	C/T	0·052	1·463	1·382-1·548	2·17E-39	1·093	1·047-1·141	5·46E-05	-2·227	0·284	4·42E-15
	rs7936312	11:76293726	*EMSY, LRRC32*	G/T	0·477	1·194	1·160-1·230	4·89E-33	1·060	1·040-1·081	4·10E-09	-0·978	0·134	3·49E-13
12q13.11	rs56389811	12:48205358	*HDAC7, SLC48A1, VDR*	T/C	0·761	1·022	0·988-1·058	2·11E-01	1·078	1·054-1·104	2·36E-10	0·485	0·161	2·55E-03
12q13.2	rs705699	12:56384804	*RAB5B, CDK2, SUOX*	G/A	0·425	1·106	1·074-1·139	1·43E-11	1·052	1·031-1·072	6·10E-07	-0·481	0·136	4·20E-04
12q13.3	rs3122929	12:57509102	*STAT6, NAB2, LRP1*	C/T	0·404	1·134	1·101-1·167	6·59E-17	1·061	1·040-1·082	5·73E-09	-0·564	0·138	4·07E-05
12q21.1 [f]	rs11178648	12:71533210	*TSPAN8, PTPRR, LGR5*	T/C	0·592	1·087	1·055-1·120	4·16E-08	1·052	1·031-1·073	6·62E-07	-0·183	0·138	1·86E-01
13q32.3 [f]	rs1887704	13:99974492	*UBAC2, DOCK9, TM9SF2*	C/G	0·681	1·124	1·089-1·161	5·48E-13	1·048	1·026-1·070	1·57E-05	-0·476	0·147	1·15E-03
14q24.1	rs1950897	14:68760141	*RAD51B, ZFYVE26, ZFF36L1*	T/C	0·287	1·096	1·062-1·131	1·38E-08	1·031	1·009-1·053	5·39E-03	-0·520	0·148	4·52E-04
15q22.2	rs11071559	15:61069988	*RORA, ANXA2, VPS13C*	T/C	0·872	1·205	1·151-1·262	2·57E-15	1·069	1·037-1·101	1·22E-05	-0·958	0·210	5·11E-06
15q22.33	rs56062135	15:67455630	*SMAD3, SMAD6, AAGAB*	C/T	0·237	1·194	1·156-1·234	2·97E-26	1·083	1·058-1·108	6·27E-12	-0·681	0·155	1·10E-05
16p13.13	rs35032408	16:11215424	*CLEC16A, CIITA, RMI2*	G/T	0·785	1·152	1·111-1·195	4·59E-14	1·073	1·048-1·100	1·10E-08	-0·568	0·169	7·91E-04
16p12.1	rs3785356	16:27349168	*IL4R, NSMCE, IL21R*	C/T	0·297	1·122	1·087-1·157	5·69E-13	1·036	1·014-1·058	1·24E-03	-0·641	0·147	1·30E-05
16q12.1	rs2066844	16:50745926	*NOD2, SNX20, CYLD*	C/T	0·048	1·194	1·121-1·272	3·52E-08	1·049	1·003-1·097	3·76E-02	-0·743	0·306	1·54E-02
17q21.33	rs28406364	17:47454507	*ZNF652, PHB*	C/T	0·377	1·108	1·076-1·142	1·39E-11	1·036	1·015-1·057	6·55E-04	-0·489	0·140	4·81E-04
19q13.11	rs10414065	19:33721455	*LRP3, CEBPA*	T/C	0·934	1·207	1·133-1·287	6·37E-09	1·108	1·064-1·155	1·01E-06	-1·078	0·291	2·12E-04
21q22.12 [f]	rs11088309	21:36464631	*RUNX1, SETD4*	C/G	0·143	1·039	0·997-1·083	6·82E-02	1·079	1·050-1·109	4·83E-08	0·063	0·189	7·38E-01

[a]Cytogenetic band. [b]SNP position, Genome Reference Consortium Build 37 (hg19). [c]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. [d]Alleles are shown as non-risk/risk alleles, where the risk allele is the allele associated with increased asthma risk. [e]RAF, risk allele frequency in UK Biobank. [f]Not reported in previous GWAS. [g]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.

674 **Figure Legends**

675

676 **Figure 1. GWAS of childhood onset and adult onset asthma.** Miami plot
677 showing results for the childhood onset versus controls GWAS (blue, panel A)
678 and adult onset versus controls GWAS (red, panel B). Each point corresponds to
679 a SNP; the y-axes show the -$\log_{10}$p-values from the childhood onset GWAS
680 (panel A) and adult onset GWAS (panel B). The x-axis shows the position of
681 each SNP along the 22 autosomes. See appendix Figure 6 for the age of asthma
682 onset Manhatten plot.

683

684 **Figure 2.** Forest plot showing the odds ratios (ORs) and 95% confidence
685 intervals from the childhood onset (blue) and adult onset (red) GWASs, and
686 betas and standard errors (gray) from the age of onset GWAS for the 61 asthma
687 associated loci. Left panel: Childhood onset specific loci (top) and adult onset
688 specific locus (bottom); right panel: Shared loci. Loci within each group are
689 sorted by OR in the childhood onset GWAS (largest to smallest).

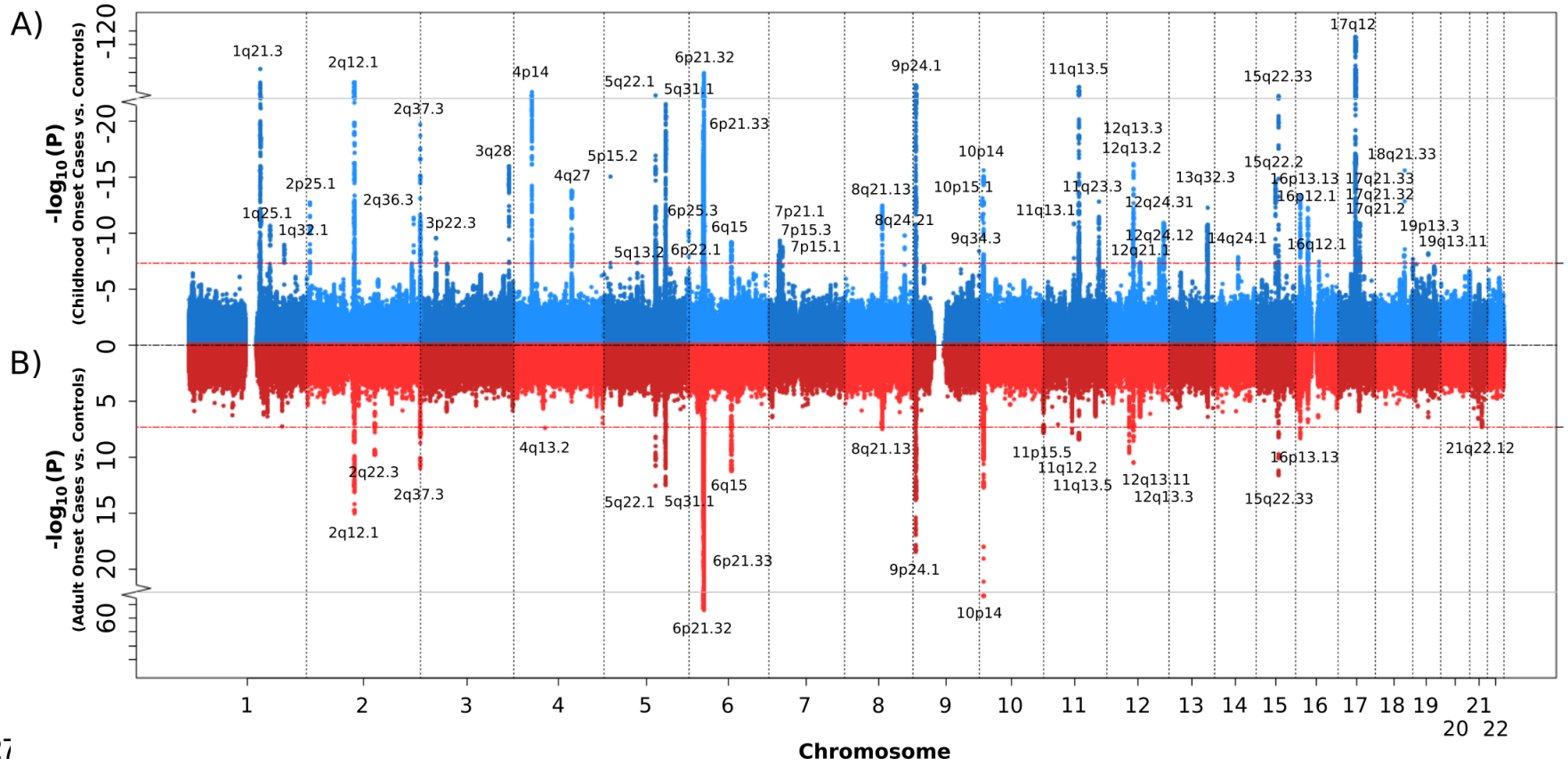
690

691 **Figure 3.** Age of onset effects (ORs) for lead SNPs at three genome-wide
692 signficant childhood onset specific loci (1q21·3, 4p14 and 17q12), and three
693 genome-wide signficant shared loci (2q12·1, 6p21·32, and 11q13·5). The sample
694 sizes for the age of onset bins are [0,5]: n=4,637; [6,10]: n=4,255; [11,15]:
695 n=2,684; [16,20]: n=2,128; [21,25]: n=2,578; [26,30]: n=2,923; [31,35]: n=2,599;

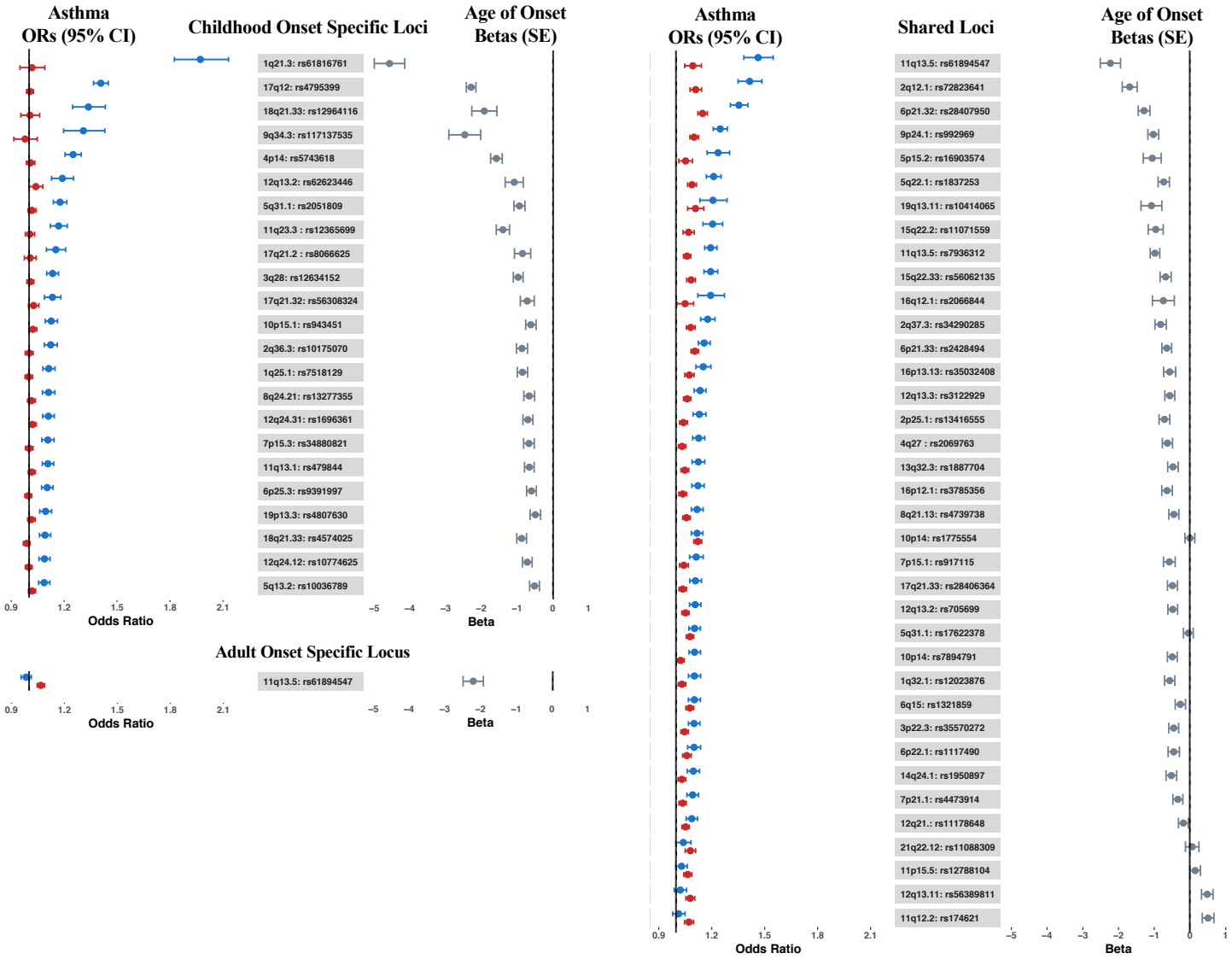
696 [36,40]: n=3,571; [41,45]: n=2,967; [46,50]: n=3,266, [51,55]: n=2,544; [56,65]:
697 n=3,694.
698
699 **Figure 4. Results of PrediXcan studies at non-HLA region loci.** Genes whose
700 predicted expression was significantly associated with asthma in either the
701 childhood onset cases or the adult onset cases are shown for skin (left panels),
702 lung (middle panels) and whole blood (right panels). Values shown for skin
703 combine both sun exposed and not sun exposed skin, showing the most
704 significant statistic of the two. Results using gene expression in spleen and small
705 intestine are shown in appendix Figure 8. The z-scores on the x- and y-axes are
706 from transcriptome-wide tests of association with asthma using SNP sets that
707 predict the expression of that gene. The diagonal dashed lines show the
708 expected when associations are the same in the childhood onset and adult onset
709 cases. The horizontal/vertical dashed lines correspond to z-scores ±4·84
710 ($p=1{\cdot}29 \times 10^{-6}$); the horizontal/vertical dotted lines correspond to z-scores ±1·96
711 (p=0·05). Colored backgrounds correspond to the chromosome location of each
712 gene (see Key). Upper panels: Genes that are associated with childhood onset
713 asthma (no genes were associated with adult onset asthma). Lower panels:
714 Shared genes associated with both childhood onset and adult onset asthma.
715 HLA region genes (6p21·32 and 6p21·33) are shown in Figure 5.
716
717 **Figure 5. Results of PrediXcan studies of HLA region genes.** Genes in the
718 HLA region whose predicted expression was associated with childhood onset or
719 adult onset asthma are shown for skin (left panel), lung tissue (middle panel) and

720 whole blood (right panel). Results using gene expression in spleen and small
721 intestine are shown in appendix Figure 9. The genes on darker shaded
722 backgrounds correspond to shared genes; the four genes on lighter color
723 backgrounds are age-specific. See Figure 4 for additional details.
724

Figure 1

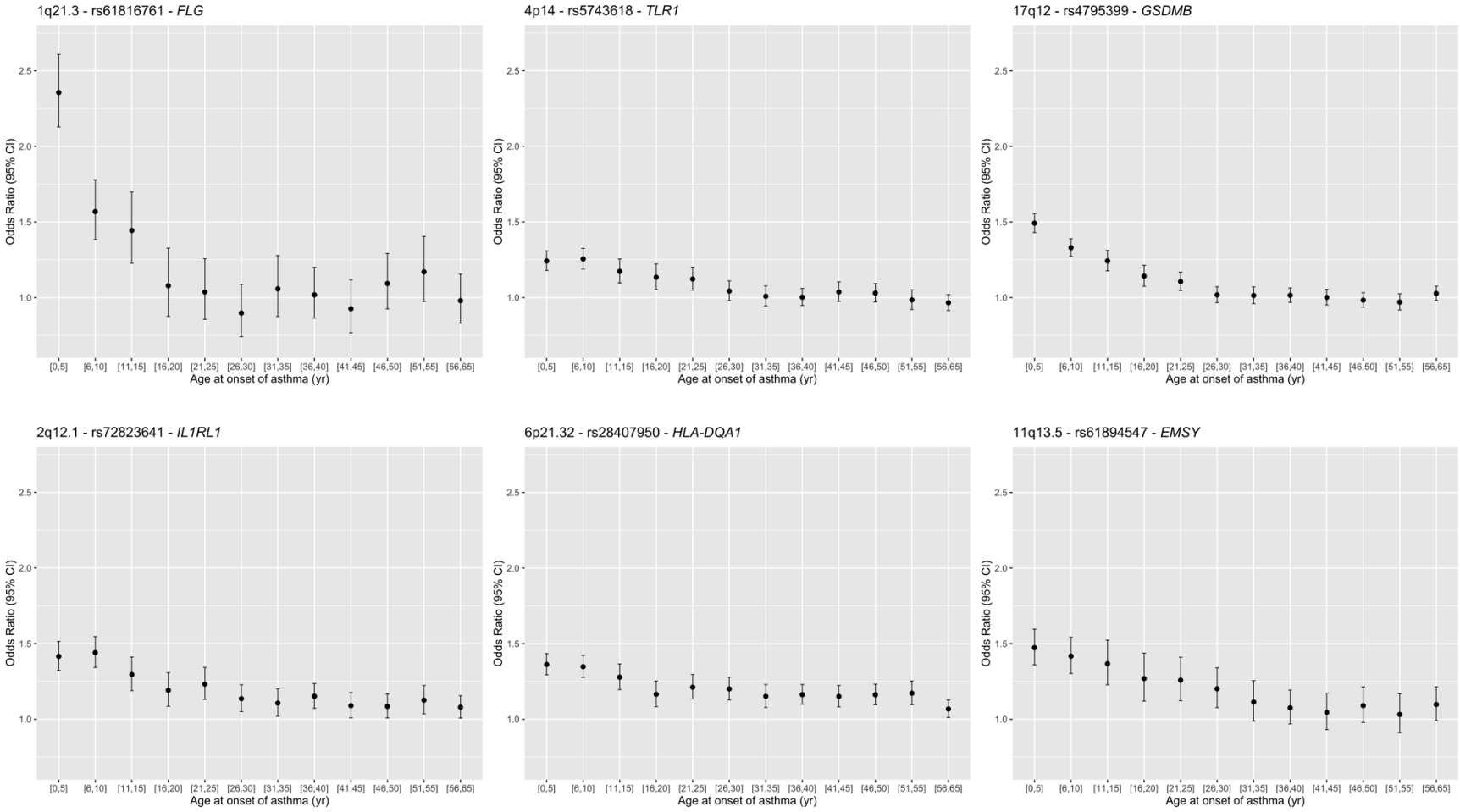

A)
B)
-log10(P)
(Childhood Onset Cases vs. Controls)
(Adult Onset Cases vs. Controls)
-120
-20
-15
-10
-5
0
5
10
15
20
60
1q21.3
2q12.1
2q37.3
4p14
5q22.1
6p21.32
5q31.1
6p21.33
9p24.1
11q13.5
15q22.33
17q12
3q28
5p15.2
4q27
2p25.1
2q36.3
1q25.1
1q32.1
3p22.3
6p25.3
6q15
5q13.2
6p22.1
7p21.1
7p15.3
7p15.1
8q21.13
8q24.21
10p14
10p15.1
9q34.3
11q23.3
11q13.1
12q13.3
12q13.2
12q24.31
12q24.12
12q21.1
13q32.3
14q24.1
15q22.2
16p13.13
16p12.1
16q12.1
17q21.33
17q21.32
17q21.2
18q21.33
19p13.3
19q13.11
4q13.2
2q22.3
2q37.3
2q12.1
5q22.1
5q31.1
6q15
6p21.33
6p21.32
8q21.13
9p24.1
10p14
11p15.5
11q12.2
11q13.5
12q13.11
12q13.3
16p13.13
15q22.33
21q22.12
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
Chromosome

730 **Figure 2**



731

732 **Figure 3**

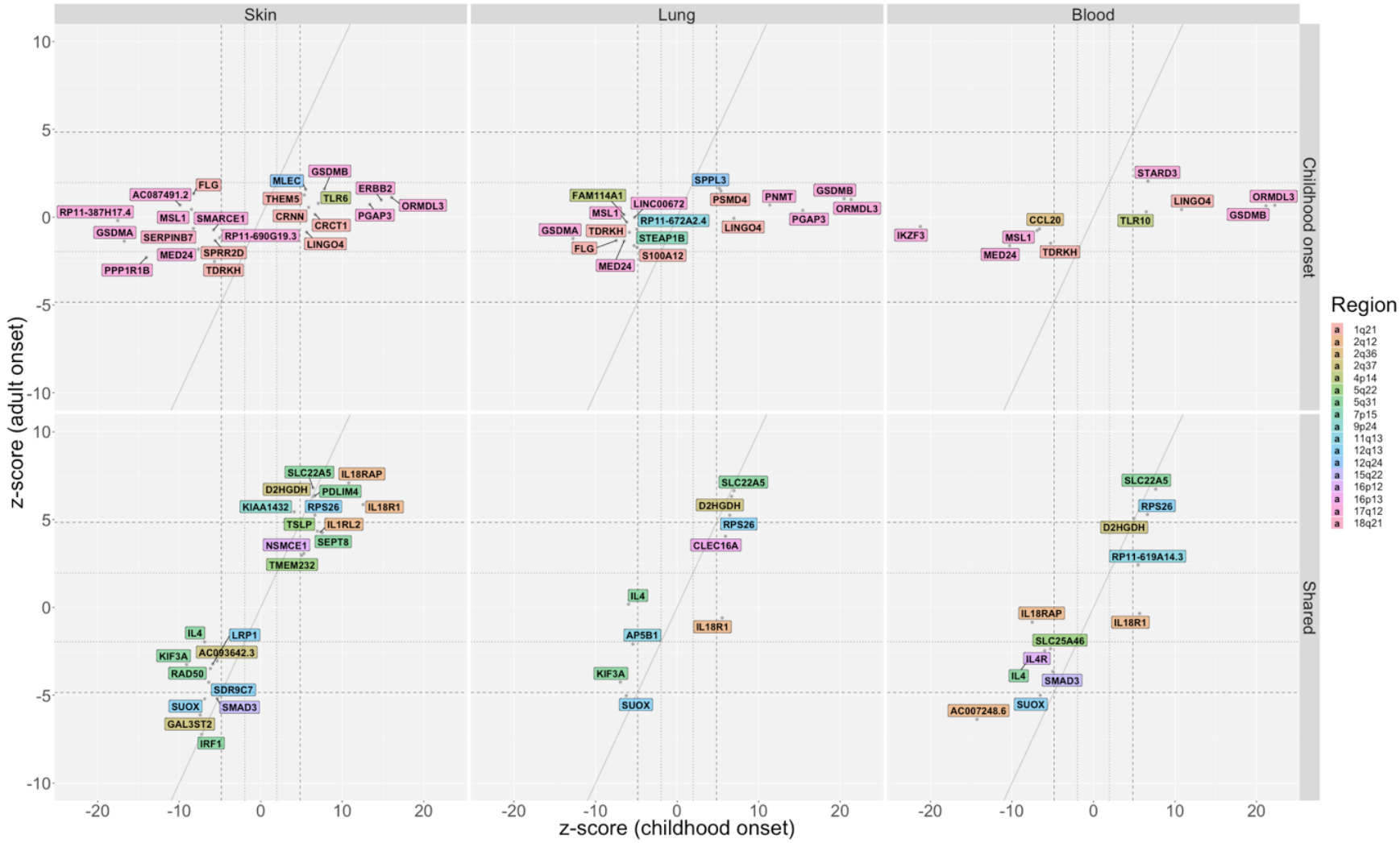


733

734

735

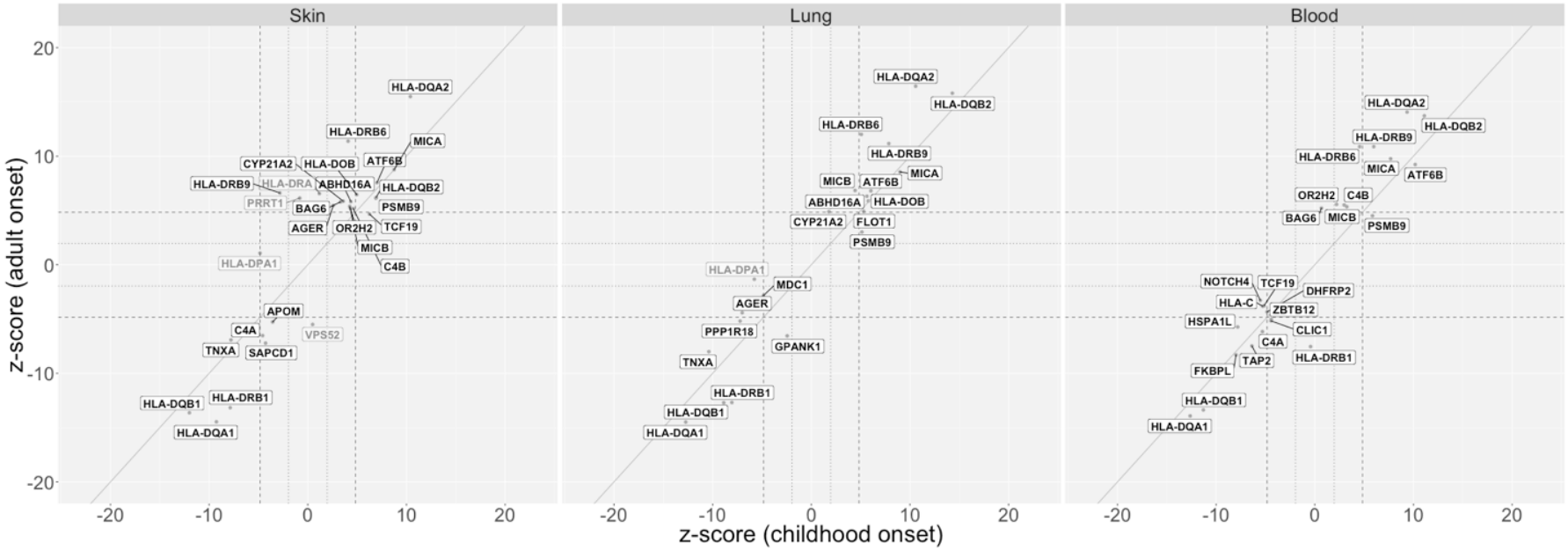
736 **Figure 4**



737

738

Figure 5


Skin
Lung
Blood
z-score (adult onset)
z-score (childhood onset)
20
10
0
-10
-20
HLA-DQA2
HLA-DRB6
MICA
CYP21A2
HLA-DOB
ATF6B
HLA-DRB9
HLA-DRA
ABHD16A
HLA-DQB2
PRRT1
BAG6
PSMB9
AGER
OR2H2
TCF19
MICB
C4B
HLA-DPA1
APOM
C4A
VPS52
TNXA
SAPCD1
HLA-DQB1
HLA-DRB1
HLA-DQA1
HLA-DQA2
HLA-DQB2
HLA-DRB6
HLA-DRB9
MICA
MICB
ATF6B
ABHD16A
HLA-DOB
CYP21A2
FLOT1
PSMB9
HLA-DPA1
MDC1
AGER
PPP1R18
GPANK1
TNXA
HLA-DRB1
HLA-DQB1
HLA-DQA1
HLA-DQA2
HLA-DQB2
HLA-DRB9
HLA-DRB6
MICA
ATF6B
OR2H2
C4B
BAG6
MICB
PSMB9
NOTCH4
TCF19
DHFRP2
HLA-C
ZBTB12
HSPA1L
CLIC1
C4A
TAP2
HLA-DRB1
FKBPL
HLA-DQB1
HLA-DQA1