

1 **A meta-analysis of genome-wide association studies of childhood wheezing**  
2 **phenotypes identifies ANXA1 as a susceptibility locus for persistent wheezing**

3 Granell R<sup>1\*</sup>, Curtin JA<sup>2</sup>, Haider S<sup>3</sup>, Kitaba N<sup>4</sup>, Mathie S<sup>3</sup>, Gregory L<sup>3</sup>, Yates LL<sup>3</sup>, Tutino M<sup>2</sup>, Hankinson  
4 J<sup>2</sup>, Perretti M<sup>5</sup>, Vonk JM<sup>6,7</sup>, Arshad SH<sup>8,9,10</sup>, Cullinan P<sup>3</sup>, Fontanella S<sup>3</sup>, Roberts G<sup>4,8,9</sup>, Koppelman  
5 GH<sup>7,11</sup>, Simpson A<sup>2</sup>, Turner S<sup>12</sup>, Murray CS<sup>2</sup>, Lloyd CM<sup>3§</sup>, Holloway JW<sup>4,8,§</sup>, Custovic A<sup>3, §</sup>

6 **on behalf of UNICORN<sup>13</sup> and Breathing Together<sup>14</sup> Investigators**

7 <sup>§</sup>Equal contribution, joint senior authors

8 1 MRC Integrative Epidemiology Unit, Department of Population Health Sciences, Bristol Medical School,  
9 University of Bristol, Bristol, United Kingdom;

10 2 Division of Infection, Immunity and Respiratory Medicine, School of Biological Sciences, The University of  
11 Manchester, Manchester Academic Health Science Centre, and Manchester University NHS Foundation  
12 Trust;

13 3 National Heart and Lung Institute, Imperial College London, London, United Kingdom;

14 4 Human Development and Health, Faculty of Medicine, University of Southampton, Southampton, United  
15 Kingdom;

16 5 William Harvey Research Institute, Barts and The London School of Medicine Queen Mary University of  
17 London, London, United Kingdom;

18 6 University of Groningen, University Medical Center Groningen, Department of Epidemiology, Groningen,  
19 the Netherlands.

20 7 University of Groningen, University Medical Center Groningen, Groningen Research Institute for Asthma  
21 and COPD (GRIAC), Groningen, the Netherlands.

22 8 NIHR Southampton Biomedical Research Centre, University Hospitals Southampton NHS Foundation  
23 Trust, Southampton, United Kingdom;

24 9 David Hide Asthma and Allergy Research Centre, Isle of Wight, United Kingdom;

25 10 Clinical and Experimental Sciences, Faculty of Medicine, University of Southampton, Southampton,  
26 United Kingdom;

27 11 University of Groningen, University Medical Center Groningen, Beatrix Children's Hospital, Department  
28 of Pediatric Pulmonology and Pediatric Allergology, Groningen, the Netherlands.

29 12 Child Health, University of Aberdeen, Aberdeen, United Kingdom

30 \*Correspondence and requests for reprints should be addressed to Dr Raquel Granell, Population

31 Health Sciences, Bristol Medical School, University of Bristol, BRISTOL BS8 2BN, UK

32 E-mail [Raquel.Granell@bristol.ac.uk](mailto:Raquel.Granell@bristol.ac.uk)

33 **This article has an online appendix.**

34 This article is dedicated to the memory of our wonderful colleague and friend Professor John

35 Henderson (1958-2019), whose contribution to the understanding of the heterogeneity of

36 childhood asthma cannot be overstated. Rainbow-chasers and UNICORN riders forever.

37 <sup>13</sup>**STELAR/UNICORN investigators**

38 Professor Graham Devereux

39 Dr Dimitrios Charalampopoulos

40 <sup>41</sup>**Breathing Together investigators**

41 Prof Andrew Bush

42 Prof Sejal Saglani

43 Prof Benjamin Marsland

44 Prof Jonathan Grigg

45 Prof Jurgen Schwarze

46 Prof Mike Shields

47 Prof Peter Ghazal

48 Prof Ultan Power

49 Dr Ceyda Oksel

50 **Word count: 3,500**

51 **Keywords:** GWAS, asthma, wheeze phenotypes, Annexin 1

52 **Data availability:** The informed consent obtained from all included participants does not allow the

53 data to be made freely available through any third party maintained public repository. However,

54 data used for this submission can be made available on request to the corresponding cohort  
55 Executive. The ALSPAC website provides information on how to request and access its data (  
56 <http://www.bristol.ac.uk/alspac/researchers/access/>). For queries regarding access of data from  
57 MAAS, IoW, SEATON or Ashford please contact Philip Couch [philip.couch@manchester.ac.uk](mailto:philip.couch@manchester.ac.uk)).

58 All summary data used to plot the figures in our manuscript has been deposited in Dryad.

59 **Code availability:** Code used for this submission has been deposited in Dryad.

60 **Author Contributions:** R.G., J.C., S.H. and A.C. conceived and planned the study. R.G., A.C., J.W.H.,  
61 L.L.Y. and C.M.L. wrote the manuscript. R.G. and J.C. analysed the data. J.C. and N.K. provided  
62 eQTL analysis for *ANXA1*, J.W.H. supervised eQTL analyses. J.H. and M.T. performed the PChi-C  
63 experiments and analysis. J.C. and M.T. interpreted the genetic architecture results for *ANXA1*.  
64 G.H.K and J.M.V. performed a replication study in PIAMA. C.M.L, S.A.M. and L.G.G. conceived and  
65 designed murine experiments. S.A.M and L.G.G. conducted experiments in mouse models. M.P.  
66 provided annexin<sup>-/-</sup> mice and advice on annexin. All authors provided critical feedback and helped  
67 shape the research, analysis, and manuscript. This publication is the work of the authors and  
68 Raquel Granell and Adnan Custovic will serve as guarantors for the contents of this paper.

69 **Competing Interests statement:** Graham Roberts: MRC grant to my institution President of the  
70 British Society of Allergy and Clinical Immunology. Gerard Koppelman: Dutch Lung Foundation,  
71 Ubbo Emmius Foundation (Money to insitution) Dutch Lung Foundation, Vertex, TEVA the  
72 Netherlands, GSK, ZON-MW (VICI grant), European Union (Money to institution) Astra Zeneca,  
73 Pure IMS, GSK (Money to institution) Sanofi, Boehringer Ingelheim (Money to institution). Angela  
74 Simpson: Medical research council Research grant JP Moulton Charitable Foundation Research  
75 grant Asthma UK Research grant. Clare Murray: Asthma Uk National Institute for Health Research  
76 Moulton Charitable Foundation North West Lung Centre Charity GSK (Lecture fees) Novartis

77 (Lecture fees). Clare Lloyd: Wellcome Trust 107059/Z/15/Z. John Holloway: Medial Research  
78 Council grant MR/S025340/1 (to institution) American Academy of Allergy Asthma and  
79 Immunology (AAAI) (Support for speaker travel to AAAAI annual congress). Adnan Custovic: MRC  
80 (research grants) EPSRC (research grant) Wellcome Trust (research grant) Worg Pharmaceuticals  
81 (Personal payment <US\$5000) GSK Honorarium for lecture, personal, <US\$ 5000 AstraZeneca  
82 Honorarium for lecture, personal, <US\$ 5000 Sanofi Honorarium for lecture, personal, <US\$ 5000  
83 Stallergens-Greer Honorarium for lecture, personal, <US\$ 5000 WAO (Board of officers, unpaid).  
84 **Funding:** Supported by the UK Medical Research Council (MRC) Programme Grant MR/S025340/1,  
85 the Wellcome Trust (WT) Strategic Award (108,818/15/Z) and a WT Senior Fellowship to CML  
86 (107059/Z/15/Z). The MRC and Wellcome (Grant ref: 217065/Z/19/Z) and the University of Bristol  
87 provide core support for ALSPAC. ALSPAC GWAS data was generated by Sample Logistics and  
88 Genotyping Facilities at Wellcome Sanger Institute and LabCorp (Laboratory Corporation of  
89 America) using support from 23andMe. A comprehensive list of grants funding is available on the  
90 ALSPAC website ([http://www.bristol.ac.uk/alspac/external/documents/grant-](http://www.bristol.ac.uk/alspac/external/documents/grant-acknowledgements.pdf)  
91 [acknowledgements.pdf](http://www.bristol.ac.uk/alspac/external/documents/grant-acknowledgements.pdf)).  
92 PIAMA was funded by the Netherlands Lung Foundation (grant 3.4.01.26, 3.2.06.022, 3.4.09.081  
93 and 3.2.10.085CO), the ZON-MW Netherlands Organization for Health Research and Development  
94 (grant 912-03-031), the Stichting Astmabestrijding and the Ministry of the Environment. Genome-  
95 wide genotyping was funded by the European Commission as part of GABRIEL (grant number  
96 018996) and a grant from BBMRI-NL (CP 29). GHK is supported by a ZON-MW VICI grant.  
97 This research was funded in part by the Wellcome Trust [217065/Z/19/Z]. For the purpose of Open  
98 Access, the author has applied a CC BY public copyright license to any Author Accepted Manuscript  
99 version arising from this submission.

100 **ABSTRACT**

101 **Background:** Many genes associated with asthma explain only a fraction of its heritability. Most  
102 genome-wide association studies (GWASs) used a broad definition of “doctor-diagnosed asthma”,  
103 thereby diluting genetic signals by not considering asthma heterogeneity. The objective of our  
104 study was to identify genetic associates of childhood wheezing phenotypes.

105 **Methods:** We conducted a novel multivariate GWAS meta-analysis of wheezing phenotypes jointly  
106 derived using unbiased analysis of data collected from birth to 18 years in 9,568 individuals from  
107 five UK birth-cohorts.

108 **Results:** 44 independent SNPs were associated with early-onset persistent, 25 with preschool  
109 remitting, 33 with mid-childhood remitting and 32 with late-onset wheeze. We identified a novel  
110 locus on chr9q21.13 (close to annexin 1 (*ANXA1*),  $p < 6.7 \times 10^{-9}$ ), associated exclusively with early-  
111 onset persistent wheeze. We identified rs75260654 as the most likely causative single nucleotide  
112 polymorphism (SNP) using Promoter Capture Hi-C loops, and then showed that the risk allele (T)  
113 confers a reduction in *ANXA1* expression. Finally, in a murine model of house dust mite (HDM)-  
114 induced allergic airway disease, we demonstrated that *anxa1* protein expression increased and  
115 *anxa1* mRNA was significantly induced in lung tissue following HDM exposure. Using *anxa1*<sup>-/-</sup>  
116 deficient mice, we showed that loss of *anxa1* results in heightened airway hyperreactivity and Th2  
117 inflammation upon allergen challenge.

118 **Conclusions:** We discovered a novel locus uniquely associated with early-onset persistent wheeze,  
119 identified the most likely causative variant, and showed that *ANXA1* may play a role in regulating  
120 the pulmonary immune response to allergens. Targeting this pathway in persistent disease may  
121 represent an exciting therapeutic prospect.

122 **Abstract word count:** 250

## 123 INTRODUCTION

124 Asthma is a complex disorder caused by a variety of mechanisms which result in multiple clinical  
125 phenotypes(1). It has a strong genetic component, and twin studies estimate its heritability to be ~  
126 60-70%(2). “Asthma genes” have been identified through a range of approaches, from candidate  
127 gene association studies(3) and family-based genome-wide linkage analyses(4) to genome-wide  
128 association studies (GWASs)(5-7). The first asthma GWAS (2007) identified multiple markers on  
129 chromosome 17q21 associated with childhood-onset asthma(5). A comprehensive review  
130 summarising the results of 42 GWASs of asthma and asthma-related traits has been published  
131 recently(8). The most widely replicated locus is 17q12-21, followed by 6p21 (*HLA* region), 2q12  
132 (*IL1RL1/IL18R1*), 5q22 (*TSLP*) and 9p24 (*IL33*)(9). Overall, the evidence suggests that multiple genes  
133 are underlying the association peaks(9).

134 However, despite undeniable successes, genetic studies of asthma have produced relatively  
135 heterogeneous results, and only a small proportion of the heritability is accounted for(10). One  
136 part of the explanation for the paucity of precise replication are numerous gene-environment  
137 interactions(11). Another important consideration is asthma heterogeneity, in that asthma  
138 diagnosis comprises several conditions with distinct pathophysiology(12, 13), each potentially  
139 underpinned by different genetic associations(14). However, in order to maximise sample size,  
140 most GWASs used a definition of “doctor-diagnosed asthma”(15). Such aggregated outcome  
141 definitions are imprecise(16) and phenotypically and mechanistically heterogeneous(17), and this  
142 heterogeneity may dilute important genetic signals(14).

143 One way of disaggregating asthma diagnosis is to use data-driven methods to derive subtypes in a  
144 hypothesis-neutral way(18). For example, we jointly modelled data on wheezing from birth to  
145 adolescence in five UK population-based birth cohorts and identified five distinct phenotypes(19).

146 However, although latent modelling approaches have been instrumental in elucidating the  
147 heterogenous nature of childhood asthma diagnosis([13](#)), there has been little research into the  
148 genetic associations of phenotypes derived using data-driven methods. This is the first study to  
149 investigate the genetic architecture of wheezing phenotypes from infancy to adolescence, to  
150 identify genes specific to each phenotype and better understand the genetic heterogeneity  
151 between the disease class profiles.

## 152 **MATERIALS & METHODS**

### 153 **Study design, setting, participants and data sources/measurement**

154 The Study Team for Early Life Asthma Research (STELAR) consortium(20) brings together five UK  
155 population-based birth cohorts: Avon Longitudinal Study of Parents and Children (ALSPAC)(21),  
156 Ashford(22) and Isle of Wight (IOW)(23) cohorts, Manchester Asthma and Allergy Study  
157 (MAAS)(24) and the Aberdeen Study of Eczema and Asthma to Observe the Effects of Nutrition  
158 (SEATON)(25). The cohorts are described in detail in the **Online Supplement** (OLS). All studies  
159 were approved by research ethics committees. Informed consent was obtained from parents, and  
160 study subjects gave their assent/consent when applicable.

161 Validated questionnaires were completed on multiple occasions from infancy to adolescence(19).

162 A list of variables, per cohort, is shown in **Table E1**, and the cohort-specific time points and sample  
163 sizes in **Table E2**. Data were harmonised and imported into Asthma eLab web-based knowledge  
164 management platform to facilitate joint analyses(20).

### 165 **Definition of primary outcome (wheeze phenotypes from infancy to adolescence)**

166 In the pooled analysis among 15,941 subjects with at least two observations on current wheeze,  
167 we used latent class analysis (LCA) to derive wheeze phenotypes from birth to age 18 years(19). A  
168 detailed description of the analysis is presented in(19) and in the OLS. A five-class solution was  
169 selected as the optimal model(19), and the classes (wheeze phenotypes) were labeled as: (1)  
170 *never/infrequent wheeze* (52.4%); (2) *early-onset pre-school remitting wheeze* (18.6%); (3)  
171 *early-onset middle-childhood remitting wheeze* (9.8%); (4) *early-onset persistent wheeze* (10.4%);  
172 and (5) *late-onset wheeze* (8.8%). These latent classes were used in the subsequent GWAS.



173 **Genotyping, imputation and GWAS Meta-Analysis**

174 Genotyping, quality control, imputation and exclusions are described in the OLS. Analyses were  
175 performed independently in ALSPAC, MAAS and the combined IOW-SEATON-ASHFORD (genotyped  
176 on the same platform, at the same time, and imputed together). We used SNPTEST v2.5.2(26) with  
177 a multinomial logistic regression model (-method newml), using the never/infrequent wheeze as  
178 the reference. A meta-analysis of the three GWASs was performed using METAL(27) with a total of  
179 8,057,852 single nucleotide polymorphism (SNPs). LD pruning, pre-selection and gene annotation  
180 is described in the OLS.

181 **Post-GWAS studies**

182 Our GWAS identified a novel locus in chr9q21 nearby *Annexin A1* (*ANXA1*), exclusively associated  
183 with early-onset persistent wheeze (see results section). We therefore proceeded with studies to  
184 identify causal variants and explore the biological mechanisms underlying this locus (detail in OLS).  
185 To this end, we firstly identified the most likely causative SNP using Promoter Capture Hi-C loops.  
186 We then ascertained genotype effect on gene expression and assessed the potential biological  
187 function of *ANXA1* in asthma. Finally, we used a murine model of house dust mite (HDM)-induced  
188 allergic airway disease to investigate whether *ANXA1* was important in regulating immune  
189 responses to a clinically relevant aeroallergen and used knock-out mice to derive further *in vivo*  
190 functional data to support our GWAS-finding.

## 191 RESULTS

### 192 Participants and descriptive data

193 We included a total of 9,568 subjects with European ancestry: ALSPAC, n=6833; MAAS, n=887;  
194 SEATON, n=548; ASHFORD, n=348; and IOW, n=952. Demographic characteristics of the study  
195 population(s) are shown in **Table E3**. Comparison of included vs. excluded participants across  
196 cohorts (per cohort and time point) is shown in **Table E4**.

### 197 GWAS Meta-Analysis

198 We conducted three GWASs (ALSPAC, MAAS, IOW-SEATON-Ashford) in parallel and results were  
199 meta-analyzed. QQ-plots are shown in **Figure E1**. A circular Manhattan plot showing an overview  
200 of the GWAS results by wheeze phenotype is shown in **Figure 1**. A total of 589 SNPs were  
201 associated with at least one phenotype with  $p < 10^{-5}$ . After clumping, we identified 134  
202 independent SNPs uniquely associated with different phenotypes ( $p < 10^{-5}$ ): of these, 44 were  
203 associated with early-onset persistent, 25 with early-onset preschool remitting, 33 with early-  
204 onset mid-childhood remitting and 32 with late-onset wheeze (**Table E5**). For some SNPs there  
205 were nominal associations with more than one phenotype. For example, chr17q21 was identified  
206 as a top locus for early-onset persistent wheeze ( $p = 5.42 \times 10^{-9}$ ), but some of the SNPs in this region  
207 were also associated with the early-onset mid-childhood remitting phenotype ( $p < 0.0001$ ).

208 To help identify functional elements located near the GWAS-associated variants (potential causal  
209 variants), we used locus zoom plots (LZP) for the 134 independent SNPs ( $p < 10^{-5}$ ). Following close  
210 inspection of all plots, we short-listed 85 independent SNPs for which the LZPs potentially  
211 indicated more than one causal variant (**Figures E2-E5**) and followed them up for further  
212 annotation. The results of GWAS meta-analysis for these 85 SNPs with main associations across  
213 the four wheeze phenotypes are presented in **Table 1**. Previously associated traits for each

214 region/gene associated with the different wheeze phenotypes are shown in **Table E6**. Briefly, one  
215 region (6q27) among the top hits for early-onset preschool remitting wheeze was previously  
216 associated with asthma, but in the context of obesity with a nominal association with asthma and  
217 BMI(28). Another region/gene (3q26.31/*NAALADL2*) identified as top hit for early-onset preschool  
218 remitting wheeze, was reported as an associate of severe asthma exacerbations, but only at  
219 nominal level (29). No regions/genes identified as top hits for early-onset mid-childhood remitting  
220 wheeze were found to have previous associations with asthma. Several genes/loci identified as top  
221 hits for late-onset wheeze were previously associated with asthma: *ACOXL* chr2q13 (later onset  
222 asthma and obesity(30)), *PRKAA2* chr1p32.2 (lymphocyte count and asthma susceptibility(31)),  
223 *CD200* 3q13.2 (adult onset non-allergic asthma(32)), *GIMAP* family 7q36.1 (autoimmune diabetes,  
224 asthma, allergy(33)), 9p22.3 (asthma in <16 year-old (34)) and *16p12.1* (asthma and rhino-  
225 conjunctivitis at 10-15 years(35)).

226 We identified two GWAS-significant loci for early-onset persistent wheeze: 17q21,  $p < 5.5 \times 10^{-9}$ , and  
227 a novel locus on 9q21.13 (*ANXA1*),  $p < 6.7 \times 10^{-9}$ . The *ANXA1* locus was the only GWAS-significant  
228 locus that had not previously been associated with asthma or atopic traits, with one previous  
229 study showing an association with FEV<sub>1</sub>/FVC and bronchodilator response in smokers(36). *ANXA1*  
230 is strongly expressed in bronchial mast cells and has anti-inflammatory properties(37), and may be  
231 involved in epithelial airway repair(38) (**Table E6**). We therefore followed up top SNPs from this  
232 locus.

### 233 ***ANXA1* locus and persistent wheeze**

234 Two SNPs (rs75260654, the lead SNP, and rs116849664 located downstream of *ANXA1*) were  
235 associated with early-onset persistent wheeze at genome-wide significance (GWS), with an

236 additional SNP rs78320984 almost reaching GWS (**Table E7**). These SNPs are in LD with each other  
237 (**Figure E6**), but not with any other SNPs.

#### 238 *Promoter Capture identifies rs75260654 as the most likely causative variant*

239 To identify the most likely causative variant, we investigated the overlap of the SNPs with  
240 Promoter Capture Hi-C interactions involving the *ANXA1* promoter in CD4+ cells in MAAS cohort  
241 subjects. Of the three SNPs, only rs75260654 overlapped a region interacting with the *ANXA1*  
242 promoter (**Figure 2**). Moreover, rs75260654 overlapped a *POLR2A* ChIP-seq peaks and an ATAC-  
243 seq peak and active enhancer in the type II pneumocyte derived A549 cell line. This shows that  
244 rs75260654 is located in a region directly interacting with the *ANXA1* promoter and is  
245 transcriptionally active in relevant cell types.

246 Allele Frequencies of rs75260654 (MAF=0.02) across wheeze phenotypes are shown in **Table E8**.  
247 Two individuals (one in MAAS and one in ALSPAC) were homozygote for the minor allele (T), and  
248 both were in the early-onset persistent wheeze class. One subject reported current wheeze and  
249 asthma through childhood, with hospitalizations for lower respiratory tract infection in the 1<sup>st</sup> year  
250 of life confirmed in healthcare records. The second individual reported current wheezing at 1.5,  
251 2.5 and 8-9 years and doctor-diagnosed asthma and the use of asthma medication at 8-9 years.

#### 252 *rs75260654: Effect on Genomic Features*

253 VEP prediction shows the SNP rs75260654 (C changed to T) to be located downstream of three  
254 protein-coding transcripts of *AXNA1* and overlapping the known regulatory region ID  
255 ENSR00000882742 on Chromosome 9: 73,173,001-73,173,200. This region is active in the GI tract,  
256 M2 macrophages, neural progenitor cells, and trophoblasts, but is repressed in T lymphocytes  
257 including CD4+ CD25+, Treg, and CD8+ cells.

258 *rs75260654: Effect on gene expression*

259 The effect of rs75260654 on the expression of nearby genes was investigated by browsing the  
260 eQTL GTEX data available in Ensembl. Compared to C, the T allele was found to reduce the  
261 expression of *ANXA1* in naive B-cells (effect size=-2.36795, p=0.01) and to increase expression in  
262 Lymphoblasoid Cell Lines (LCL) (effect size=0.848856, pe=0.001) (**Figure 3**). This SNP affects  
263 expression of the neighboring gene *ALDH1A1* (aldehyde dehydrogenase-1 family member A1)  
264 (effect size=-2.40446, p=0.0039 in macrophages infected with Salmonella). In the eQTL catalogue,  
265 rs75260654 is identified as an eQTL of *ANXA1* in various immune cells (at nominal significance)  
266 including T cells, monocytes, fibroblasts, whole blood, Th2 memory cells, naive B cells. rs75260654  
267 is also an eQTL of *ANXA1* in monocytes that were stimulated with R848 (agonist of TLRs 7 and 8)  
268 and a human seasonal influenza A virus (39) (at nominal significance) (**Table E9**). In the lung  
269 rs116849664 and rs78320984 (both in LD with rs75260654) were eQTLs of *ANXA1* (**Table E10**) as  
270 well as LINC01474 at nominal significance levels.

271 Additional supporting evidence regarding the significance of the T-allele on the expression of  
272 these genes was provided using eQTLGene Consortium meta-analysis of 24 Cohorts and 24331  
273 samples(40). This method reproduced the previous modest results showing a cis-eQTL effect of  
274 rs75260654 on both the *ANXA1* ( $p=6.02\times 10^{-23}$ ) and *ALDH1A1* ( $p=1.11\times 10^{-19}$ ) at FDR=0. No  
275 significant trans-eQTLs were observed.

276 *Potential biological function of ANXA1 in asthma*

277 Protein-protein network analysis demonstrated that *ANXA1* interacts directly with genes enriched  
278 for asthma (including *IL4* and *IL13*) and inflammatory regulation (*NR3C1*, Glucocorticoid receptor)  
279 showing its significance in dysregulation of the immune response (see **Figure E7** and **Table E11**).

280 **Functional studies of *anxa1* in a murine model**

281 *Pulmonary expression of *anxa1* is modulated by aeroallergen exposure*

282 We first analysed expression of *anxa1* using a model of HDM-induced allergic airway disease  
283 (**Figure 4A**)([41](#)). Consistently, immunohistochemistry analysis revealed *anxa1* protein expression  
284 increased following HDM challenge (**Figure 4B,C**). *Anxa1* mRNA was significantly induced in lung  
285 tissue following HDM exposure (**Figure 4D**). This increase suggests that the pro-resolving *anxa1*  
286 may play a role in regulating the pulmonary immune response to allergen.

287 *Anxa1 suppresses allergen-induced airway hyperresponsiveness (AHR) and type-2 inflammation*

288 To confirm a functional role for *anxa1* in allergic airway disease, we exposed *anxa1*<sup>-/-</sup> mice to  
289 intranasal HDM. Wildtype mice given HDM over 3 weeks developed significant AHR compared to  
290 PBS control mice. Mice deficient in *anxa1* had significantly worse allergen-induced lung function  
291 (greater airway resistance) compared to WT treated mice (**Figure 4E**). *Anxa1*<sup>-/-</sup> mice exhibited  
292 significantly increased airway eosinophilia and elevated numbers of Th2 lymphocytes (**Figure**  
293 **4F,G**). Lung tissue cytokine levels reflected the exacerbated airway Th2 inflammation, with  
294 elevation in IL-4, and significant induction of IL-5 and IL-13 (**Figure 4H,J**). Thus, *anxa1* deficiency  
295 results in an alteration of the pulmonary immune response, with uncontrolled eosinophilia and an  
296 exacerbation of type-2 inflammation and AHR in response to allergen.

## 297 **DISCUSSION**

298 Herein, we present a comprehensive description of the genetic architecture of childhood wheezing  
299 disorders. Using a novel approach applied to a unique dataset from five UK birth cohorts, we  
300 identified subsets of SNPs differentially associated across four wheezing phenotypes: early-onset  
301 persistent (44 SNPs, 19 loci), early-onset preschool remitting (25 SNPs, 10 loci), early-onset mid-  
302 childhood remitting (33 SNPs, 9 loci) and late-onset (32 SNPs, 20 loci). We found little evidence of  
303 genetic associations spanning across different phenotypes. This suggests that genetic architecture  
304 of different wheeze phenotypes comprises a limited number of variants likely underpinning  
305 mechanisms which are shared across phenotypes, but that each phenotype is also characterized  
306 by unique phenotype-specific genetic associations. Importantly, we identified a novel locus in  
307 chr9q21 nearby *ANXA1* exclusively associated with early-onset persistent wheeze ( $p < 6.7 \times 10^{-9}$ ). To  
308 identify the most likely causative variant, we investigated the overlap of the associated SNPs with  
309 Promoter Capture Hi-C interactions to demonstrate that SNP rs75260654 overlapped a region  
310 interacting with the *ANXA1* promoter. Using eQTL data, we identified that the risk allele (T) of  
311 rs75260654 associated with early-onset persistent wheeze is also associated with *ANXA1*  
312 expression. Further investigation of the biological function of *ANXA1* revealed that it interacts with  
313 genes enriched for asthma (including *IL4* and *IL13*) and inflammatory regulation (*NR3C1*,  
314 Glucocorticoid receptor). In functional mouse experiments, *anxa1* protein expression increased  
315 and *anxa1* mRNA was significantly induced in lung tissue following HDM exposure, suggesting that  
316 the pro-resolving *anxa1* may play a role in regulating the pulmonary immune response to allergen.  
317 Concurrently, by utilizing *anxa1*<sup>-/-</sup> deficient mice we demonstrated that loss of *anxa1* results in  
318 heightened AHR and Th2 inflammation upon allergen challenge, providing important *in vivo*  
319 functional data to support our GWAS-finding.

320 *ANXA1* is a 37-kDa glycoprotein with potent anti-inflammatory and pro-resolving properties that  
321 are mediated by interaction with a specific G-protein coupled receptor FPR2(42). This axis  
322 represents an important resolution pathway in chronic inflammatory settings such as those of  
323 rheumatoid arthritis(43) and ulcerative colitis(44). *ANXA1* belongs to the annexin family of Ca<sup>2+</sup>-  
324 dependent phospholipid-binding proteins, and through inhibition of phospholipase A2, it reduces  
325 eicosanoid production, which also contributes to its anti-inflammatory activities. Modulation of  
326 M2 macrophage phenotype is also promoted by *ANXA1* to attenuate tissue inflammation(45).  
327 Corticosteroids (a mainstay of asthma treatment) increase the synthesis of *ANXA1*(46). Plasma  
328 *ANXA1* levels are significantly lower in asthmatic patients with frequent exacerbations compared  
329 to those with stable disease, suggesting a link between this mediator and disease state(47).  
330 Furthermore, children with wheeze have reduced airway levels of *ANXA1*(48).  
331 Previous functional studies using *anxa1*<sup>-/-</sup> deficient mice challenged with ovalbumin showed  
332 *anxa1*-deficient mice to have elevated AHR compared to WT mice(49). Ng *et al.* demonstrated that  
333 untreated *anxa1*-deficient mice have spontaneous AHR that predisposes them to exacerbated  
334 response to allergen(49). In the current study, we demonstrated in the murine lung the induction  
335 of *Anxa1* in response to HDM exposure. In addition, genetic deletion of *anxa1* potentiated the  
336 development of AHR and enhanced eosinophilia and markers of Th2 inflammation in mice treated  
337 with HDM, which is consistent with and extends previous reports. Of interest, in mice, *anxa1*  
338 expression was recently found to be characteristic of a novel cell-type called the Hillock cell, which  
339 may be involved in squamous barrier function and immunomodulation(50). These data identify  
340 the *ANXA1*/FPR2 signaling axis as an important regulator of allergic disease, that could be  
341 manipulated for therapeutic benefit.



342 Our study has several limitations. By GWAS standards, our study is comparatively small and may  
343 be considered to be underpowered. The sample size may be an issue when using an aggregated  
344 definition (such as “doctor-diagnosed asthma”) but is less likely to be an issue when primary  
345 outcome is determined by deep phenotyping. This is indirectly confirmed in our analyses. Our  
346 primary outcome was derived through careful phenotyping over a period of more than two  
347 decades in five independent birth cohorts, and although comparatively smaller than some asthma  
348 GWASs, our study proved to be powered enough to detect previously identified key associations  
349 (e.g. chr17q21 locus). Precise phenotyping has the potential to identify new risk loci. For example,  
350 a comparatively small GWAS (1,173 cases and 2,522 controls) which used a specific subtype of  
351 early-onset childhood asthma with recurrent severe exacerbations as an outcome, identified a  
352 functional variant in a novel susceptibility gene *CDHR3* (SNP rs6967330) as an associate of this  
353 disease subtype, but not of doctor-diagnosed asthma(51). This important discovery was made with  
354 a considerably smaller sample size but using a more precise asthma subtype. In contrast, the  
355 largest asthma GWAS to date had a ~40-fold higher sample size(7), but reported no significant  
356 association between *CDHR3* and aggregated asthma diagnosis. Therefore, with careful  
357 phenotyping, smaller sample sizes may be adequately powered to identify larger effect sizes than  
358 those in large GWASs with broader outcome definitions(52).

359 The importance of the precise outcome definition was highlighted in our previous studies in  
360 ALSPAC which explored genetic associates of wheeze phenotypes derived by LCA(53, 54). Our  
361 current findings are consistent with our earlier report suggesting that 17q21 SNPs are associated  
362 with early-onset persistent, but not with early transient or late-onset wheeze (53). Further analysis  
363 using genetic prediction scores based on 10-200,000 SNPs ranked according to their associations  
364 with physician-diagnosed asthma found that the 46 highest ranked SNPs predicted persistent  
365 wheeze more strongly than doctor-diagnosed asthma(54). Finally, a candidate gene study

366 combining data from ALSPAC and PIAMA found different associations of IL33-IL1RL1 pathway  
367 polymorphisms with different phenotypes(55).

368 We are cognisant that there may be a perception of the lack of replication of our GWAS findings.

369 We would argue that direct replication is almost certainly not possible in other cohorts, as

370 phenotypes for replication studies should be homogenous(56). However, there is a considerable

371 heterogeneity in LCA-derived wheeze phenotypes between studies, and although phenotypes in

372 different studies are usually designated with the same names, they differ between studies in

373 temporal trajectories, distributions within a population, and associated risk factors(57). This

374 heterogeneity is in part consequent on the number and the non-uniformity of the timepoints

375 used, and is likely one of the factors responsible for the lack of consistent associations of

376 discovered phenotypes with risk factors reported in previous studies(58). This will also adversely

377 impact the ability to identify phenotype-specific genetic associates. For example, we have

378 previously shown that less distinct wheeze phenotypes in PIAMA were identified compared to

379 those derived in ALSPAC(59). Thus, phenotypes that are homogeneous to those in our study

380 almost certainly cannot readily be derived in available populations. This is exemplified in our

381 attempted replication of *ANXA1* findings in PIAMA cohort (see OLS, **Table E12**). In this analysis, the

382 number of individuals assigned to persistent wheezing in PIAMA was small (40), associates of this

383 phenotype differed to those in STELAR cohorts, and the SNPs' imputation scores were low (<0.60),

384 which meant the conditions for replication were not met.

385 Our study population is of European descent, and we cannot generalize the results to different

386 ethnicities or environments. It is important to highlight the under-representation of ethnically

387 diverse populations in most GWASs(9). To mitigate against this, large consortia have been formed,

388 which combine the results of multiple ethnically diverse GWASs to increase the overall power to

389 identify asthma-susceptibility loci. Examples include the GABRIEL(6), EVE(60) and TAGC(7)  
390 consortia, and the value of diverse, multi-ethnic participants in large-scale genomic studies has  
391 recently been shown(61). However, such consortia do not have the depth of longitudinal data to  
392 allow the type of analyses which we carried out to derive a multivariable primary outcome.  
393 One strength of our approach is that we used data from five birth cohorts with detailed and  
394 lifelong phenotyping, which were harmonised in a common knowledge management platform(20),  
395 allowing joint analyses. We performed three parallel GWASs that produced estimates with  
396 remarkably consistent directions of effects.

397 In conclusion, using unique data from five UK birth cohorts, we identified subsets of SNPs  
398 differentially associated across four wheezing phenotypes from infancy to adolescence. We found  
399 little evidence of genetic associations spanning across different phenotypes. We discovered a  
400 novel locus in chr9q21 uniquely associated with early-onset persistent wheeze ( $p < 6.7 \times 10^{-9}$ ),  
401 identified SNP rs75260654 as the most likely causative variant, and demonstrated that the risk  
402 allele (T) confers a reduction in *ANXA1* expression. In mouse experiments, *ANXA1* expression  
403 increased in lung tissue following allergen exposure, suggesting that the pro-resolving *ANXA1* may  
404 play a role in regulating the pulmonary immune response to allergen. Using *ANXA1*-deficient mice,  
405 we demonstrated that loss of *ANXA1* results in heightened AHR and Th2 inflammation upon  
406 allergen challenge, providing important *in vivo* functional data to support our GWAS finding.  
407 Targeting these pathways to promote the clearance of chronic inflammation in persistent disease  
408 may represent an exciting therapeutic prospect.

## 409 REFERENCES

- 410 1. Pavord ID, Beasley R, Agusti A, Anderson GP, Bel E, Brusselle G, Cullinan P, Custovic A, Ducharme FM,  
411 Fahy JV, Frey U, Gibson P, Heaney LG, Holt PG, Humbert M, Lloyd CM, Marks G, Martinez FD, Sly  
412 PD, von Mutius E, Wenzel S, Zar HJ, Bush A. After asthma: redefining airways diseases. *Lancet* 2018;  
413 391: 350-400.
- 414 2. Duffy DL, Martin NG, Battistutta D, Hopper JL, Mathews JD. Genetics of asthma and hay fever in  
415 Australian twins. *Am Rev Respir Dis* 1990; 142: 1351-1358.
- 416 3. Simpson A, Custovic A, Tepper R, Graves P, Stern DA, Jones M, Hankinson J, Curtin JA, Wu J, Blekic M,  
417 Bukvic BK, Aberle N, Marinho S, Belgrave D, Morgan WJ, Martinez FD. Genetic variation in vascular  
418 endothelial growth factor- $\alpha$  and lung function. *Am J Respir Crit Care Med* 2012; 185: 1197-1204.
- 419 4. Daniels SE, Bhattacharya S, James A, Leaves NI, Young A, Hill MR, Faux JA, Ryan GF, le Souef PN, Lathrop  
420 GM, Musk AW, Cookson WO. A genome-wide search for quantitative trait loci underlying asthma.  
421 *Nature* 1996; 383: 247-250.
- 422 5. Moffatt MF, Kabesch M, Liang L, Dixon AL, Strachan D, Heath S, Depner M, von Berg A, Bufe A, Rietschel  
423 E, Heinzmann A, Simma B, Frischer T, Willis-Owen SA, Wong KC, Illig T, Vogelberg C, Weiland SK,  
424 von Mutius E, Abecasis GR, Farrall M, Gut IG, Lathrop GM, Cookson WO. Genetic variants regulating  
425 ORMDL3 expression contribute to the risk of childhood asthma. *Nature* 2007; 448: 470-473.
- 426 6. Moffatt MF, Gut IG, Demenais F, Strachan DP, Bouzigon E, Heath S, von Mutius E, Farrall M, Lathrop M,  
427 Cookson W, Consortium G. A large-scale, consortium-based genomewide association study of  
428 asthma. *N Engl J Med* 2010; 363: 1211-1221.
- 429 7. Demenais F, Margaritte-Jeannin P, Barnes KC, Cookson WOC, Altmuller J, Ang W, Barr RG, Beaty TH,  
430 Becker AB, Beilby J, Bisgaard H, Bjornsdottir US, Bleecker E, Bonnelykke K, Boomsma DI, Bouzigon E,  
431 Brightling CE, Brossard M, Brusselle GG, Burchard E, Burkart KM, Bush A, Chan-Yeung M, Chung KF,  
432 Couto Alves A, Curtin JA, Custovic A, Daley D, de Jongste JC, Del-Rio-Navarro BE, Donohue KM,  
433 Duijts L, Eng C, Eriksson JG, Farrall M, Fedorova Y, Feenstra B, Ferreira MA, Australian Asthma  
434 Genetics Consortium c, Freidin MB, Gajdos Z, Gauderman J, Gehring U, Geller F, Genuneit J, Gharib  
435 SA, Gilliland F, Granell R, Graves PE, Gudbjartsson DF, Haahtela T, Heckbert SR, Heederik D, Heinrich  
436 J, Heliovaara M, Henderson J, Himes BE, Hirose H, Hirschhorn JN, Hofman A, Holt P, Hottenga J,  
437 Hudson TJ, Hui J, Imboden M, Ivanov V, Jaddoe VWV, James A, Janson C, Jarvelin MR, Jarvis D, Jones  
438 G, Jonsdottir I, Jousilahti P, Kabesch M, Kahonen M, Kantor DB, Karunas AS, Khusnutdinova E,  
439 Koppelman GH, Kozyrskyj AL, Kreiner E, Kubo M, Kumar R, Kumar A, Kuokkanen M, Lahousse L,  
440 Laitinen T, Laprise C, Lathrop M, Lau S, Lee YA, Lehtimaki T, Letort S, Levin AM, Li G, Liang L, Loehr  
441 LR, London SJ, Loth DW, Manichaikul A, Marenholz I, Martinez FJ, Matheson MC, Mathias RA,  
442 Matsumoto K, Mbarek H, McArdle WL, Melbye M, Melen E, Meyers D, Michel S, Mohamdi H, Musk  
443 AW, Myers RA, Nieuwenhuis MAE, Noguchi E, O'Connor GT, Ogorodova LM, Palmer CD, Palotie A,  
444 Park JE, Pennell CE, Pershagen G, Polonikov A, Postma DS, Probst-Hensch N, Puzyrev VP, Raby BA,  
445 Raitakari OT, Ramasamy A, Rich SS, Robertson CF, Romieu I, Salam MT, Salomaa V, Schlunssen V,  
446 Scott R, Selivanova PA, Sigsgaard T, Simpson A, Siroux V, Smith LJ, Solodilova M, Standl M,  
447 Stefansson K, Strachan DP, Stricker BH, Takahashi A, Thompson PJ, Thorleifsson G, Thorsteinsdottir  
448 U, Tiesler CMT, Torgerson DG, Tsunoda T, Uitterlinden AG, van der Valk RJP, Vaysse A, Vedantam S,  
449 von Berg A, von Mutius E, Vonk JM, Waage J, Wareham NJ, Weiss ST, White WB, Wickman M,  
450 Widen E, Willemsen G, Williams LK, Wouters IM, Yang JJ, Zhao JH, Moffatt MF, Ober C, Nicolae DL.  
451 Multiancestry association study identifies new asthma risk loci that colocalize with immune-cell  
452 enhancer marks. *Nat Genet* 2018; 50: 42-53.
- 453 8. El-Husseini ZW, Gosens R, Dekker F, Koppelman GH. The genetics of asthma and the promise of  
454 genomics-guided drug target discovery. *Lancet Respir Med* 2020; 8: 1045-1056.
- 455 9. Kim KW, Ober C. Lessons Learned From GWAS of Asthma. *Allergy Asthma Immun* 2019; 11: 170-187.
- 456 10. Ober C, Yao TC. The genetics of asthma and allergic disease: a 21st century perspective. *Immunol Rev*  
457 2011; 242: 10-30.

- 458 11. Custovic A, Marinho S, Simpson A. Gene-environment interactions in the development of asthma and  
459 atopy. *Expert Rev Respir Med* 2012; 6: 301-308.
- 460 12. Custovic A. "Asthma" or "Asthma Spectrum Disorder"? *J Allergy Clin Immunol Pract* 2020; 8: 2628-2629.
- 461 13. Haider S, Granell R, Curtin J, Fontanella S, Cucco A, Turner S, Simpson A, Roberts G, Murray CS,  
462 Holloway JW, Devereux G, Cullinan P, Arshad SH, Custovic A. Modeling Wheezing Spells Identifies  
463 Phenotypes with Different Outcomes and Genetic Associates. *Am J Respir Crit Care Med* 2022; 205:  
464 883-893.
- 465 14. Custovic A, Henderson J, Simpson A. Does understanding endotypes translate to better asthma  
466 management options for all? *J Allergy Clin Immunol* 2019; 144: 25-33.
- 467 15. Aaron SD, Vandemheen KL, FitzGerald JM, Ainslie M, Gupta S, Lemiere C, Field SK, Mclvor RA,  
468 Hernandez P, Mayers I, Mulpuru S, Alvarez GG, Pakhale S, Mallick R, Boulet LP, Canadian  
469 Respiratory Research N. Reevaluation of Diagnosis in Adults With Physician-Diagnosed Asthma.  
470 *JAMA* 2017; 317: 269-279.
- 471 16. Looijmans-van den Akker I, van Luijn K, Verheij T. Overdiagnosis of asthma in children in primary care: a  
472 retrospective analysis. *Br J Gen Pract* 2016; 66: e152-157.
- 473 17. Robinson PFM, Fontanella S, Ananth S, Martin Alonso A, Cook J, Kaya-de Vries D, Polo Silveira L, Gregory  
474 L, Lloyd C, Fleming L, Bush A, Custovic A, Saglani S. Recurrent Severe Preschool Wheeze: From  
475 Prespecified Diagnostic Labels to Underlying Endotypes. *Am J Respir Crit Care Med* 2021; 204: 523-  
476 535.
- 477 18. Howard R, Rattray M, Prosperi M, Custovic A. Distinguishing asthma phenotypes using machine learning  
478 approaches. *Current allergy and asthma reports* 2015; 15: 1-10.
- 479 19. Oksef C, Granell R, Haider S, Fontanella S, Simpson A, Turner S, Devereux G, Arshad SH, Murray CS,  
480 Roberts G, Holloway JW, Cullinan P, Henderson J, Custovic A, Stelar investigators bTi. Distinguishing  
481 Wheezing Phenotypes from Infancy to Adolescence. A Pooled Analysis of Five Birth Cohorts. *Ann*  
482 *Am Thorac Soc* 2019; 16: 868-876.
- 483 20. Custovic A, Ainsworth J, Arshad H, Bishop C, Buchan I, Cullinan P, Devereux G, Henderson J, Holloway J,  
484 Roberts G, Turner S, Woodcock A, Simpson A. The Study Team for Early Life Asthma Research  
485 (STELAR) consortium 'Asthma e-lab': team science bringing data, methods and investigators  
486 together. *Thorax* 2015; 70: 799-801.
- 487 21. Golding J, Pembrey M, Jones R. ALSPAC-the Avon longitudinal study of parents and children. I. Study  
488 methodology. *Paediatric and perinatal epidemiology* 2001; 15: 74-87.
- 489 22. Cullinan P, MacNeill SJ, Harris JM, Moffat S, White C, Mills P, Newman Taylor AJ. Early allergen  
490 exposure, skin prick responses, and atopic wheeze at age 5 in English children: a cohort study.  
491 *Thorax* 2004; 59: 855-861.
- 492 23. Arshad SH, Holloway JW, Karmaus W, Zhang H, Ewart S, Mansfield L, Matthews S, Hodgekiss C, Roberts  
493 G, Kurukulaaratchy R. Cohort Profile: The Isle Of Wight Whole Population Birth Cohort (IOWBC). *Int*  
494 *J Epidemiol* 2018; 47: 1043-1044i.
- 495 24. Custovic A, Simpson BM, Murray CS, Lowe L, Woodcock A, Asthma NACM, Allergy Study G. The National  
496 Asthma Campaign Manchester Asthma and Allergy Study. *Pediatr Allergy Immunol* 2002; 13: 32-37.
- 497 25. Martindale S, McNeill G, Devereux G, Campbell D, Russell G, Seaton A. Antioxidant intake in pregnancy  
498 in relation to wheeze and eczema in the first two years of life. *American journal of respiratory and*  
499 *critical care medicine* 2005; 171: 121-128.
- 500 26. Marchini J, Howie B. Genotype imputation for genome-wide association studies. *Nat Rev Genet* 2010;  
501 11: 499-511.
- 502 27. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans.  
503 *Bioinformatics* 2010; 26: 2190-2191.
- 504 28. Melen E, Himes BE, Brehm JM, Boutaoui N, Klanderma BJ, Sylvia JS, Lasky-Su J. Analyses of shared  
505 genetic factors between asthma and obesity in children. *J Allergy Clin Immunol* 2010; 126: 631-637  
506 e631-638.
- 507 29. Herrera-Luis E, Espuela-Ortiz A, Lorenzo-Diaz F, Keys KL, Mak ACY, Eng C, Huntsman S, Villar J,  
508 Rodriguez-Santana JR, Burchard EG, Pino-Yanes M. Genome-wide association study reveals a novel

- 509 locus for asthma with severe exacerbations in diverse populations. *Pediatr Allergy Immunol* 2021;  
510 32: 106-115.
- 511 30. Zhu Z, Guo Y, Shi H, Liu CL, Panganiban RA, Chung W, O'Connor LJ, Himes BE, Gazal S, Hasegawa K,  
512 Camargo CA, Jr., Qi L, Moffatt MF, Hu FB, Lu Q, Cookson WOC, Liang L. Shared genetic and  
513 experimental links between obesity-related traits and asthma subtypes in UK Biobank. *J Allergy Clin*  
514 *Immunol* 2020; 145: 537-549.
- 515 31. Cusanovich DA, Billstrand C, Zhou X, Chavarria C, De Leon S, Michelini K, Pai AA, Ober C, Gilad Y. The  
516 combination of a genome-wide association study of lymphocyte count and analysis of gene  
517 expression data reveals novel asthma candidate genes. *Hum Mol Genet* 2012; 21: 2111-2123.
- 518 32. Siroux V, Gonzalez JR, Bouzigon E, Curjuric I, Boudier A, Imboden M, Anto JM, Gut I, Jarvis D, Lathrop M,  
519 Omenaas ER, Pin I, Wjst M, Demenais F, Probst-Hensch N, Kogevinas M, Kauffmann F. Genetic  
520 heterogeneity of asthma phenotypes identified by a clustering approach. *Eur Respir J* 2014; 43: 439-  
521 452.
- 522 33. Heinonen MT, Laine AP, Soderhall C, Gruziova O, Rautio S, Melen E, Pershagen G, Lahdesmaki HJ, Knip  
523 M, Ilonen J, Henttinen TA, Kere J, Laheesmaa R, Finnish Pediatric Diabetes R. GIMAP GTPase family  
524 genes: potential modifiers in autoimmune diabetes, asthma, and allergy. *J Immunol* 2015; 194:  
525 5885-5894.
- 526 34. Denham S, Koppelman GH, Blakey J, Wjst M, Ferreira MA, Hall IP, Sayers I. Meta-analysis of genome-  
527 wide linkage studies of asthma and related traits. *Respir Res* 2008; 9: 38.
- 528 35. Sottile G, Ferrante G, Torregrossa M, Cibella F, Cilluffo G, Fasola S, Alessandro R, Seidita G, Viegi G, La  
529 Grutta S. An association analysis to identify genetic variants linked to asthma and rhino-  
530 conjunctivitis in a cohort of Sicilian children. *Ital J Pediatr* 2019; 45: 16.
- 531 36. Lutz SM, Cho MH, Young K, Hersh CP, Castaldi PJ, McDonald ML, Regan E, Mattheisen M, DeMeo DL,  
532 Parker M, Foreman M, Make BJ, Jensen RL, Casaburi R, Lomas DA, Bhatt SP, Bakke P, Gulsvik A,  
533 Crapo JD, Beaty TH, Laird NM, Lange C, Hokanson JE, Silverman EK, Investigators E, Investigators  
534 CO. A genome-wide association study identifies risk loci for spirometric measures among smokers  
535 of European and African ancestry. *BMC Genet* 2015; 16: 138.
- 536 37. Vieira Braga FA, Kar G, Berg M, Carpaij OA, Polanski K, Simon LM, Brouwer S, Gomes T, Hesse L, Jiang J,  
537 Fasouli ES, Efremova M, Vento-Tormo R, Talavera-Lopez C, Jonker MR, Affleck K, Palit S, Strzelecka  
538 PM, Firth HV, Mahubani KT, Cvejic A, Meyer KB, Saeb-Parsy K, Luinge M, Brandsma CA, Timens W,  
539 Angelidis I, Strunz M, Koppelman GH, van Oosterhout AJ, Schiller HB, Theis FJ, van den Berge M,  
540 Nawijn MC, Teichmann SA. A cellular census of human lungs identifies novel cell states in health  
541 and in asthma. *Nat Med* 2019; 25: 1153-1163.
- 542 38. Leoni G, Neumann PA, Kamaly N, Quiros M, Nishio H, Jones HR, Sumagin R, Hilgarth RS, Alam A,  
543 Fredman G, Argyris I, Rijcken E, Kusters D, Reutelingsperger C, Perretti M, Parkos CA, Farokhzad OC,  
544 Neish AS, Nusrat A. Annexin A1-containing extracellular vesicles and polymeric nanoparticles  
545 promote epithelial wound repair. *J Clin Invest* 2015; 125: 1215-1227.
- 546 39. Quach H, Rotival M, Pothlichet J, Loh YE, Dannemann M, Zidane N, Laval G, Patin E, Harmant C, Lopez  
547 M, Deschamps M, Naffakh N, Duffy D, Coen A, Leroux-Roels G, Clement F, Boland A, Deleuze JF,  
548 Kelso J, Albert ML, Quintana-Murci L. Genetic Adaptation and Neandertal Admixture Shaped the  
549 Immune System of Human Populations. *Cell* 2016; 167: 643-656 e617.
- 550 40. Vösa U, Claringbould A, Westra H-J, Bonder MJ, Deelen P, Zeng B, Kirsten H, Saha A, Kreuzhuber R,  
551 Kasela S, Pervjakova N, Alvaes I, Fave M-J, Agbessi M, Christiansen M, Jansen R, Seppälä I, Tong L,  
552 Teumer A, Schramm K, Hemani G, Verlouw J, Yaghootkar H, Sönmez R, Brown A, Kukushkina V,  
553 Kalnapienkis A, Rüeger S, Porcu E, Kronberg-Guzman J, Kettunen J, Powell J, Lee B, Zhang F,  
554 Arindrarto W, Beutner F, Brugge H, Dmitreva J, Elansary M, Fairfax BP, Georges M, Heijmans BT,  
555 Kähönen M, Kim Y, Knight JC, Kovacs P, Krohn K, Li S, Loeffler M, Marigorta UM, Mei H, Momozawa  
556 Y, Müller-Nurasyid M, Nauck M, Nivard M, Penninx B, Pritchard J, Raitakari O, Rotzchke O,  
557 Slagboom EP, Stehouwer CDA, Stumvoll M, Sullivan P, Hoen PACT, Thiery J, Tönjes A, van Dongen J,  
558 van Iterson M, Veldink J, Völker U, Wijmenga C, Swertz M, Andiappan A, Montgomery GW, Ripatti  
559 S, Perola M, Kutalik Z, Dermitzakis E, Bergmann S, Frayling T, van Meurs J, Prokisch H, Ahsan H,  
560 Pierce B, Lehtimäki T, Boomsma D, Psaty BM, Gharib SA, Awadalla P, Milani L, Ouwehand W,

- 561 Downes K, Stegle O, Battle A, Yang J, Visscher PM, Scholz M, Gibson G, Esko T, Franke L. Unraveling  
562 the polygenic architecture of complex traits using blood eQTL metaanalysis. 2018: 447367.
- 563 41. Gregory LG, Causton B, Murdoch JR, Mathie SA, O'Donnell V, Thomas CP, Priest FM, Quint DJ, Lloyd CM.  
564 Inhaled house dust mite induces pulmonary T helper 2 cytokine production. *Clin Exp Allergy* 2009;  
565 39: 1597-1610.
- 566 42. Perretti M, Chiang N, La M, Fierro IM, Marullo S, Getting SJ, Solito E, Serhan CN. Endogenous lipid- and  
567 peptide-derived anti-inflammatory pathways generated with glucocorticoid and aspirin treatment  
568 activate the lipoxin A4 receptor. *Nat Med* 2002; 8: 1296-1302.
- 569 43. D'Acquisto F, Paschalidis N, Raza K, Buckley CD, Flower RJ, Perretti M. Glucocorticoid treatment inhibits  
570 annexin-1 expression in rheumatoid arthritis CD4+ T cells. *Rheumatology (Oxford)* 2008; 47: 636-  
571 639.
- 572 44. Vong L, Ferraz JG, Dufton N, Panaccione R, Beck PL, Sherman PM, Perretti M, Wallace JL. Up-regulation  
573 of Annexin-A1 and lipoxin A(4) in individuals with ulcerative colitis may promote mucosal  
574 homeostasis. *PLoS One* 2012; 7: e39244.
- 575 45. McArthur S, Juban G, Gobetti T, Desgeorges T, Theret M, Gondin J, Toller-Kawahisa JE,  
576 Reutelingsperger CP, Chazaud B, Perretti M, Mounier R. Annexin A1 drives macrophage skewing to  
577 accelerate muscle regeneration through AMPK activation. *J Clin Invest* 2020; 130: 1156-1167.
- 578 46. Rhen T, Cidlowski JA. Antiinflammatory action of glucocorticoids--new mechanisms for old drugs. *N Engl*  
579 *J Med* 2005; 353: 1711-1723.
- 580 47. Lee SH, Lee PH, Kim BG, Seo HJ, Baek AR, Park JS, Lee JH, Park SW, Kim DJ, Park CS, Jang AS. Annexin A1  
581 in plasma from patients with bronchial asthma: its association with lung function. *BMC Pulm Med*  
582 2018; 18: 1.
- 583 48. Eke Gungor H, Tahan F, Gokahmetoglu S, Saraymen B. Decreased levels of lipoxin A4 and annexin A1 in  
584 wheezy infants. *Int Arch Allergy Immunol* 2014; 163: 193-197.
- 585 49. Ng FS, Wong KY, Guan SP, Mustafa FB, Kajiji TS, Bist P, Biswas SK, Wong WS, Lim LH. Annexin-1-deficient  
586 mice exhibit spontaneous airway hyperresponsiveness and exacerbated allergen-specific antibody  
587 responses in a mouse model of asthma. *Clin Exp Allergy* 2011; 41: 1793-1803.
- 588 50. Montoro DT, Haber AL, Biton M, Vinarsky V, Lin B, Birket SE, Yuan F, Chen S, Leung HM, Villoria J, Rogel  
589 N, Burgin G, Tsankov AM, Waghray A, Slyper M, Waldman J, Nguyen L, Dionne D, Rozenblatt-Rosen  
590 O, Tata PR, Mou H, Shivaraju M, Bihler H, Mense M, Tearney GJ, Rowe SM, Engelhardt JF, Regev A,  
591 Rajagopal J. A revised airway epithelial hierarchy includes CFTR-expressing ionocytes. *Nature* 2018;  
592 560: 319-324.
- 593 51. Bonnelykke K, Sleiman P, Nielsen K, Kreiner-Moller E, Mercader JM, Belgrave D, den Dekker HT, Husby  
594 A, Sevelsted A, Faura-Tellez G, Mortensen LJ, Paternoster L, Flaaten R, Molgaard A, Smart DE,  
595 Thomsen PF, Rasmussen MA, Bonas-Guarch S, Holst C, Nohr EA, Yadav R, March ME, Blicher T,  
596 Lackie PM, Jaddoe VW, Simpson A, Holloway JW, Duijts L, Custovic A, Davies DE, Torrents D, Gupta  
597 R, Hollegaard MV, Hougaard DM, Hakonarson H, Bisgaard H. A genome-wide association study  
598 identifies CDHR3 as a susceptibility locus for early childhood asthma with severe exacerbations. *Nat*  
599 *Genet* 2014; 46: 51-55.
- 600 52. Schoettler N, Rodriguez E, Weidinger S, Ober C. Advances in asthma and allergic disease genetics: Is  
601 bigger always better? *J Allergy Clin Immunol* 2019; 144: 1495-1506.
- 602 53. Granell R, Henderson AJ, Timpson N, St Pourcain B, Kemp JP, Ring SM, Ho K, Montgomery SB,  
603 Dermizakis ET, Evans DM, Sterne JA. Examination of the relationship between variation at 17q21  
604 and childhood wheeze phenotypes. *J Allergy Clin Immunol* 2013; 131: 685-694.
- 605 54. Spycher B, Henderson J, Granell R, Evans DM, Smith GD, Timpson NJ, Sterne JAC. Genome-wide  
606 prediction of childhood asthma and related phenotypes in a longitudinal birth cohort. *Eur Respir J*  
607 2012; 40.
- 608 55. Savenije OE, Mahachie John JM, Granell R, Kerkhof M, Dijk FN, de Jongste JC, Smit HA, Brunekreef B,  
609 Postma DS, Van Steen K, Henderson J, Koppelman GH. Association of IL33-IL-1 receptor-like 1  
610 (IL1RL1) pathway polymorphisms with wheezing phenotypes and asthma in childhood. *J Allergy Clin*  
611 *Immunol* 2014; 134: 170-177.

- 612 56. Crawford F CG, Chappell FM, et al. A systematic review and individual patient data meta-analysis of  
613 prognostic factors for foot ulceration in people with diabetes: the international research  
614 collaboration for the prediction of diabetic foot ulcerations (PODUS). Southampton (UK): NIHR  
615 Journals Library; 2015 Jul. (Health Technology Assessment, No. 19.57.) Chapter 5, Validation of the  
616 model. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK305598/>.
- 617 57. Oksel C, Haider S, Fontanella S, Frainay C, Custovic A. Classification of Pediatric Asthma: From  
618 Phenotype Discovery to Clinical Practice. *Front Pediatr* 2018; 6: 258.
- 619 58. Oksel C, Granell R, Mahmoud O, Custovic A, Henderson AJ, Stelar, Breathing Together i. Causes of  
620 variability in latent phenotypes of childhood wheeze. *J Allergy Clin Immunol* 2019; 143: 1783-1790  
621 e1711.
- 622 59. Savenije OE, Granell R, Caudri D, Koppelman GH, Smit HA, Wijga A, de Jongste JC, Brunekreef B, Sterne  
623 JA, Postma DS, Henderson J, Kerkhof M. Comparison of childhood wheezing phenotypes in 2 birth  
624 cohorts: ALSPAC and PIAMA. *J Allergy Clin Immunol* 2011; 127: 1505-1512 e1514.
- 625 60. Torgerson DG, Ampleford EJ, Chiu GY, Gauderman WJ, Gignoux CR, Graves PE, Himes BE, Levin AM,  
626 Mathias RA, Hancock DB, Baurley JW, Eng C, Stern DA, Celedon JC, Rafaels N, Capurso D, Conti DV,  
627 Roth LA, Soto-Quiros M, Togias A, Li X, Myers RA, Romieu I, Van Den Berg DJ, Hu D, Hansel NN,  
628 Hernandez RD, Israel E, Salam MT, Galanter J, Avila PC, Avila L, Rodriguez-Santana JR, Chapela R,  
629 Rodriguez-Cintron W, Diette GB, Adkinson NF, Abel RA, Ross KD, Shi M, Faruque MU, Dunston GM,  
630 Watson HR, Mantese VJ, Ezurum SC, Liang L, Ruczinski I, Ford JG, Huntsman S, Chung KF, Vora H, Li  
631 X, Calhoun WJ, Castro M, Sienra-Monge JJ, del Rio-Navarro B, Deichmann KA, Heinzmann A, Wenzel  
632 SE, Busse WW, Gern JE, Lemanske RF, Jr., Beaty TH, Bleecker ER, Raby BA, Meyers DA, London SJ,  
633 Mexico City Childhood Asthma S, Gilliland FD, Children's Health S, study H, Burchard EG, Genetics of  
634 Asthma in Latino Americans Study SoG-E, Admixture in Latino A, Study of African Americans AG,  
635 Environments, Martinez FD, Childhood Asthma R, Education N, Weiss ST, Childhood Asthma  
636 Management P, Williams LK, Study of Asthma P, Pharmacogenomic Interactions by R-E, Barnes KC,  
637 Genetic Research on Asthma in African Diaspora S, Ober C, Nicolae DL. Meta-analysis of genome-  
638 wide association studies of asthma in ethnically diverse North American populations. *Nat Genet*  
639 2011; 43: 887-892.
- 640 61. Wojcik GL, Graff M, Nishimura KK, Tao R, Haessler J, Gignoux CR, Highland HM, Patel YM, Sorokin EP,  
641 Avery CL, Belbin GM, Bien SA, Cheng I, Cullina S, Hodonsky CJ, Hu Y, Huckins LM, Jeff J, Justice AE,  
642 Kocarnik JM, Lim U, Lin BM, Lu Y, Nelson SC, Park SL, Poisner H, Preuss MH, Richard MA, Schurmann  
643 C, Setiawan VW, Sockell A, Vahi K, Verbanck M, Vishnu A, Walker RW, Young KL, Zubair N, Acuna-  
644 Alonso V, Ambite JL, Barnes KC, Boerwinkle E, Bottinger EP, Bustamante CD, Caberto C, Canizales-  
645 Quinteros S, Conomos MP, Deelman E, Do R, Doheny K, Fernandez-Rhodes L, Fornage M, Hailu B,  
646 Heiss G, Henn BM, Hindorff LA, Jackson RD, Laurie CA, Laurie CC, Li Y, Lin DY, Moreno-Estrada A,  
647 Nadkarni G, Norman PJ, Pooler LC, Reiner AP, Romm J, Sabatti C, Sandoval K, Sheng X, Stahl EA,  
648 Stram DO, Thornton TA, Wassel CL, Wilkens LR, Winkler CA, Yoneyama S, Buyske S, Haiman CA,  
649 Kooperberg C, Le Marchand L, Loos RJF, Matise TC, North KE, Peters U, Kenny EE, Carlson CS.  
650 Genetic analyses of diverse populations improves discovery for complex traits. *Nature* 2019; 570:  
651 514-518.
- 652





654 **Table 1.** GWAS Meta-analysis: Short-listed 85 top independent SNPs across the four wheezing phenotypes

Early-onset Persistent Wheezing									
Locus	Independent SNPs	Nearby Genes (SNPnexus)	Effect allele (freq)/ other allele	Beta	SE	P-value	Effect Direction (3 GWAS)	min_pval _other*	Previous relevant associations†
1q43	rs4620530	<i>CHRM3</i>	g(0.56)/t	0.25	0.05	2.45E-06	+++	0.79	FEV <sub>1</sub> , FEV <sub>1</sub> /FVC, asthma- high priority drug target
2p25.1	rs13398488	<i>RNF144A</i>	g(0.29)/a	0.25	0.05	2.18E-06	+-	0.13	asthma, allergy, childhood onset asthma, allergic rhinitis
2q12.2	rs6543291	<i>FHL2</i>	c(0.4)/t	0.23	0.05	6.97E-06	+++	0.10	bronchial hyper-responsiveness, airway inflammation; novel gene associated with asthma severity in human
3q21.3	rs77655717	<i>EFCC1, RAB43, RAB7A</i>	c(0.05)/t	0.47	0.10	6.40E-06	+++	0.39	<i>RAB43</i> : response to bronchodilator, FEV/FEC ratio; <i>RAB7A</i> : eosinophil count
4p16.3	rs7680608 <sup>eQTL</sup>	<i>RNF212, IDUA, DGKQ, SLC26A1</i>	g(0.93)/c	-0.42	0.09	1.31E-06	---	0.15	4p16: asthma
	rs77822621 <sup>eQTL</sup>		c(0.96)/t	-0.50	0.11	7.16E-06	---	0.01	
4q31.21	rs115228498	<i>INPP4B</i>	c(0.02)/t	0.79	0.17	2.70E-06	+++	0.02	atopic asthma
5p15.31	rs116494115	<i>ADCY2</i>	g(0.01)/a	0.75	0.17	6.49E-06	+++	0.09	asthma × air pollution, childhood asthma
7q22.3	rs76871421	<i>CDHR3</i>	c(0.12)/t	0.37	0.07	5.71E-07	+++	0.22	childhood asthma
9q21.13	rs75260654	<i>ANXA1, TMC1, LOC101927258, ALDH1A1</i>	c(0.98)/t	-0.90	0.16	<b>6.66E-09</b>	---	0.05	<b>ANXA1</b> : FEV <sub>1</sub> /FVC, response to bronchodilators in smokers, with anti-inflammatory properties, strongly expressed in bronchial mast cells and potentially involved in epithelial airway
	rs116849664		c(0.98)/t	-0.89	0.16	<b>1.99E-08</b>	---	0.06	

									repair
10q24.2	rs7088157	<i>LOXL4, R3HCC1L</i>	g(0.5)/a	-0.23	0.05	7.34E-06	---	0.26	<i>R3HCC1L</i> : eosinophil count, atopic eczema, psoriasis, BMI
11p15.4	rs112474574	<i>TRIM5, TRIM6, TRIM22</i>	c(0.96)/t	-0.55	0.12	2.29E-06	---	0.14	severe asthma and insulin resistance
11q23.3	rs116861530 <sup>eQTL</sup>	<i>SIK3</i>	g(0.94)/a	-0.42	0.09	9.07E-06	---	0.01	triglycerides, glucose metabolism, eosinophil count
14q22.1	rs1105683	<i>KTN1</i>	c(0.07)/t	0.41	0.09	9.15E-06	+++	0.24	severe asthma
5q13.3	rs2202714 <sup>eQTL</sup>	<i>FAM227B</i>	g(0.36)/a	0.23	0.05	8.71E-06	+++	0.01	rs35251997 and FEV <sub>1</sub> ; FEV <sub>1</sub> /FVC
15q25.2	rs117540214 <sup>eQTL</sup>	<i>ADAMTSL3</i>	g(0.06)/a	0.42	0.10	9.82E-06	+++	3.91E-03	FEV <sub>1</sub> /FVC
17q12	rs17676191	<i>IKZF3</i>	g(0.10)/a	0.36	0.08	2.18E-06	+++	3.06E-03	early-onset asthma, persistent wheezing (chr17q12-q21)
	rs79026872		c(0.03)/t	0.64	0.13	2.08E-06	+++	2.56E-03	
	rs4795400		c(0.53)/t	0.30	0.05	<b>5.42E-09</b>	+++	1.96E-04	
	rs1031460		<i>GSDMB</i>	g(0.50)/t	0.27	0.05	8.71E-08	+++	
	rs56199421		c(0.45)/t	-0.23	0.05	4.50E-06	---	9.61E-04	
17q21	rs4795406	<i>LRRC3C</i>	g(0.55)/c	-0.24	0.05	9.91E-07	---	1.51E-03	
	rs72832972		c(0.92)/t	-0.38	0.08	8.91E-06	---	0.01	
	rs4794821	<i>GSDMA</i>	c(0.47)/t	0.27	0.05	9.43E-08	+++	1.07E-03	
	rs59843584		c(0.78)/a	-0.31	0.06	6.38E-08	---	6.63E-03	
	rs4804311		g(0.08)/a	0.42	0.09	9.65E-07	++	0.05	
19p13.2	rs2013694	<i>MARCH2, HNRNPM, MYO1F</i>	c(0.89)/t	-0.38	0.08	8.29E-07	++	0.39	triglycerides, HDL-cholesterol, metabolic syndrome; <i>MYO1F</i> :
	rs73501545		g(0.16)/a	0.31	0.07	8.39E-06	+++	0.29	FEV <sub>1</sub> and FVC
	rs111644945		g(0.9)/a	-0.41	0.08	4.01E-07	---	0.02	
22q11.1	rs5994170	<i>CECR5</i>	g(0.4)/a	0.23	0.05	4.95E-06	+++	0.58	triglycerides, eosinophil count
	rs34902370		c(0.75)/t	-0.25	0.06	6.80E-06	---	0.41	and body height

## Early-onset Pre-school Remitting Wheezing

Locus	SNP	Nearby Genes (SNPnexus)	coded(freq)/other allele	Beta	SE	P-value	Direction	min_pval _other	Previous relevant associations
1q32.3	rs12730098 <sup>eQTL</sup>	PPP2R5A	c(0.79)/t	-0.22	0.05	8.44E-06	---	0.53	waist circumference & obesity
2p24.2	rs2880066	FAM49A or	t(0.09)/a	0.32	0.07	4.34E-06	+++	0.20	airway repair in non-atopic asthma
	rs10180268	CYRIA	c(0.06)/t	0.43	0.09	6.56E-07	+++	0.19	
3q26.31	rs3861377	NLGN1	g(0.89)/a	-0.28	0.06	7.75E-06	---	0.28	smoking
	rs10513743	NAALADL2	c(0.84)/t	-0.25	0.06	4.97E-06	+-	0.06	Exacerbations requiring hospitalisation in asthma- suggestive pvalue
5q13.3	rs10075253	SV2C	c(0.85)/t	-0.27	0.06	1.20E-06	---	0.17	BMI
6q27	rs2453395	PDE10A	g(0.33)/a	0.19	0.04	9.51E-06	+++	0.01	Birthweight; asthma and BMI
7q21.11	rs4730561	MAGI2	g(0.36)/a	-0.20	0.04	6.78E-06	---	0.13	allergic diseases & atopy, smoking, BMI, airway wall thickness
	rs73144976		g(0.97)/a	-0.47	0.11	9.41E-06	---	0.26	
	rs67259321		c(0.06)/t	0.43	0.08	1.65E-07	++	0.76	
9p13.3	rs10758259 <sup>eQTL</sup>	C9orf24	g(0.17)/a	-0.27	0.06	4.64E-06	---	0.01	airway repair
11q22.3	rs72994149	GUCY1A2	c(0.84)/t	-0.24	0.05	8.33E-06	+-	0.06	systolic blood pressure
13q21.1	rs2872948	PRR20A/B/C/D/E	t(0.96)/a	-0.54	0.10	5.93E-08	---	0.27	systolic blood pressure
	rs73527654		g(0.08)/a	0.34	0.07	2.85E-06	+++	0.41	
15q21.1	rs116966886	SEMA6D	g(0.99)/a	-0.82	0.18	7.55E-06	+-	0.57	smoking
	rs117565527		g(0.99)/a	-0.87	0.17	2.38E-07	+-	0.43	

## Early-onset Mid-childhood Remitting Wheezing

Locus	SNP	Nearby Genes (SNPnexus)	coded(freq)/other allele	Beta	SE	P-value	direction	min_pval_other	Previous relevant associations
1q23.2	rs35725789	<i>CADM3</i> ,	c(0.95)/a	-0.56	0.12	5.42E-06	+-	0.01	neutrophil count, CRP
	rs146141555	<i>FCER1A</i> ,	c(0.98)/t	-0.89	0.17	2.04E-07	+-	0.08	
	rs146575092	<i>MPTX1</i> , <i>OR10J1</i>	g(0.98)/a	-0.85	0.17	8.73E-07	+-	0.07	
2p22.3	rs7595553	<i>MRPL50P1</i>	g(0.16)/c	-0.46	0.10	3.26E-06	---	0.12	PM 2.5 exposure level and global DNA methylation level
3p25.3	rs34315999 <sup>eQTL</sup>	<i>RAD18</i>	c(0.03)/t	0.69	0.14	1.11E-06	+++	0.14	atopy/SPT
3q29	rs146961758	<i>MRPL50P1</i> , <i>LSG1</i> , <i>TMEM44-AS1</i> , <i>TMEM44</i> , <i>ATP13A3</i>	t(0.05)/a	0.57	0.12	6.01E-06	++	0.11	3q29: BMI <i>TMEM44-AS1</i> , <i>TMEM44</i> , <i>ATP13A3</i> : diastolic blood pressure; <i>LSG1</i> : BMI, eosinophil count
4q24	rs138794367	<i>SLC9B1</i>	c(0.99)/t	-1.02	0.22	5.47E-06	---	0.13	eosinophil count, allergic rhinitis
5q14.1	rs115719402	<i>AP3B1</i>	g(0.96)/a	-0.60	0.13	7.20E-06	---	0.06	Vital capacity, BMI
13q31.1	rs9602218	<i>RNU6-67P</i> , <i>SLITRK1</i>	c(0.06)/a	0.58	0.12	1.74E-06	+++	0.05	RNU6-67P/ rs976078: food allergy
	rs61960366		g(0.97)/a	-0.79	0.15	7.09E-08	+-	0.12	
	rs74589927	<i>VENTXP2</i> , <i>UBE2D3P4</i> , <i>MTND4P1</i>	g(0.02)/a	0.73	0.16	3.78E-06	++	0.02	
	rs2210726		c(0.91)/t	-0.47	0.10	1.33E-06	---	0.02	
	rs4390476		c(0.08)/a	0.46	0.10	8.81E-06	+++	0.12	
14q24.2	rs117443464	<i>ZFYVE1</i>	g(0.95)/a	-0.57	0.12	4.68E-06	+++	0.19	LDL cholesterol and systolic blood pressure
20p12.3-p12.2	rs6077514	<i>PLCB4</i>	c(0.88)/t	-0.39	0.09	4.03E-06	---	0.43	neutrophil count

Late-onset Wheezing									
Locus	SNP	Nearby Genes (SNPnexus)	coded(freq)/other allele	Beta	SE	P-value	direction	min_pval_other	Previous relevant associations
1p36.13	rs9439669	KLHDC7A	t(0.82)/a	-0.34	0.07	5.15E-06	---	0.31	1p36.13: metabolic syndrome
1p32.2	rs2051039	PPAP2B, PRKAA2	c(0.08)/t	0.47	0.10	6.06E-06	+++	0.08	PRKAA2: lymphocyte count and asthma susceptibility
1p31.1	rs72673642	HMGB1P18	g(0.77)/a	-0.31	0.07	6.25E-06	---	0.01	smoking, BMI
2q13	rs140983998	ACOXL, BUB1	c(0.98)/t	-0.88	0.19	4.71E-06	---	0.40	ACOXL: later onset asthma and obesity
2q14.3	rs148008098	AMMECR1L	c(0.96)/t	-0.69	0.15	3.41E-06	---	0.01	body height, blood protein; growth, bone, and heart alterations
3p24.2	rs4072729	RARB	c(0.03)/t	0.61	0.13	4.20E-06	+++	0.23	FEV1/FVC, adult lung function
3q13.2	rs145629570	KIAA2018, NAA50, SIDT1, CD200	c(0.02)/t	0.92	0.18	6.83E-07	+++	0.10	SIDT1: FEV1/FVC; CD200: adult-onset non-allergic asthma
3q23	rs113643470	TFDP2, XRN1	c(0.98)/t	-0.91	0.19	1.68E-06	---	0.03	XRN1: eosinophil count; 3q23: allergic disease and atopic sensitisation
4p11	rs17472015	SLAIN2, SLC10A4, FRYL	c(0.01)/t	1.00	0.23	9.49E-06	+++	0.46	FRYL: body height, age at menopause
7q36.1	rs117660982	KRBA1, ZNF467	g(0.97)/a	-0.74	0.16	7.63E-06	+-	0.18	systolic blood pressure
	rs118027705	GIMAP family, AOC1	c(0.97)/t	-0.77	0.17	6.48E-06	+-	0.01	AOC1: CV disease, smoking; GIMAP family: autoimmune diabetes, asthma, and allergy

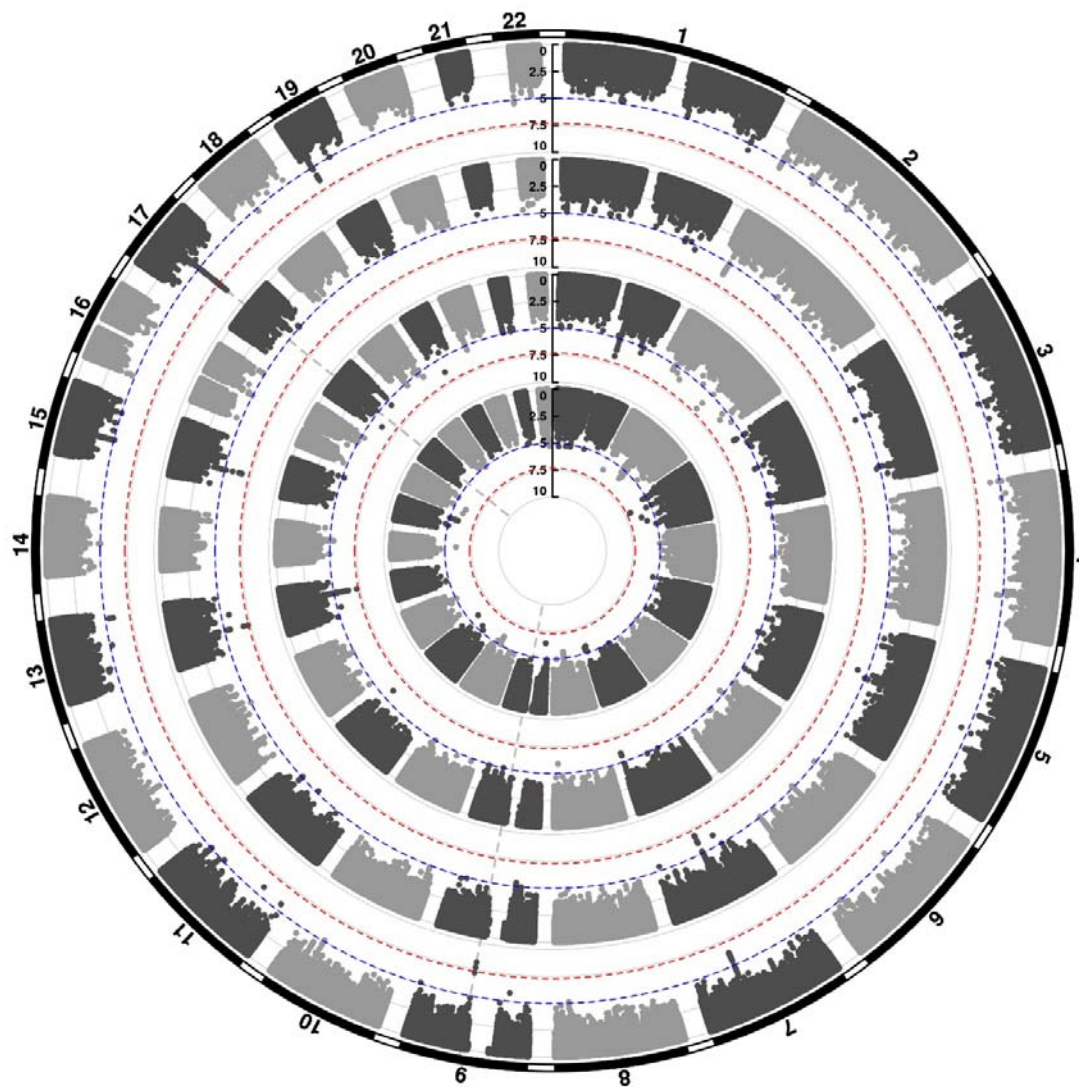
	rs139489493	<i>LOC105375566</i>	c(0.98)/t	-0.95	0.20	2.28E-06	---	0.03	
7q36.3	rs144271668	<i>PTPRN2</i>	c(0.01)/a	0.88	0.19	2.91E-06	+++	0.28	eczema
8q21.3	rs990182	<i>LOC105375631</i>	t(0.42)/a	0.28	0.06	2.57E-06	+++	0.46	8q21.3: type 1 diabetes
9p22.3	rs79110962	<i>NFIB, ZDHHC21</i>	c(0.08)/t	0.51	0.10	3.98E-07	+++	0.05	9p22.3: asthma (mean age<16 years)
10q23.31	rs7896106	<i>SLC16A12, IFIT family, PANK1</i>	g(0.35)/t	0.30	0.06	1.35E-06	+++	0.05	<i>SLC16A12</i> : Body height; <i>PANK1</i> : insulin
11q23.3	rs141958628	<i>CBL, CCDC84, MCAM</i>	c(0.98)/t	-0.98	0.20	1.33E-06	-+-	0.27	<i>CCDC84</i> : asthma, allergy
15q15.3-q21.1	rs139134265	<i>SPG11, CTDSPL2</i>	g(0.02)/c	0.87	0.20	9.11E-06	+++	0.13	<i>CTDSPL2</i> : alcohol drinking
15q25.2	rs143862030	<i>ADAMTSL3, GOLGA6L4, UBE2Q2P8</i>	c(0.04)/t	0.64	0.13	1.65E-06	+++	0.08	<i>ADAMTSL3</i> : FEV1/FVC; lean mass
16p13.3	rs113390367	<i>SSTR5-AS1, CACNA1H</i>	g(0.86)/a	-0.40	0.08	1.04E-06	---	0.16	<i>CACNA1H</i> : eosinophil count
16p12.1	rs4788025	<i>GSG1L</i>	g(0.46)/a	-0.30	0.06	7.99E-07	---	0.19	16p12.1: current asthma and rhino-conjunctivitis at 10-15 years
22q13.32	rs133498	<i>FAM19A5 or TAF5</i>	g(0.94)/a	-0.48	0.11	5.35E-06	---	0.84	Obesity and Metabolic Dysfunction

655 \*minimum p-value across associations with the other three wheezing phenotypes, using the never/infrequent wheeze as the baseline  
656 phenotype

657 eQTL: identified in expression analyses of whole blood and/or lung tissues using Genotype-Tissue Expression database (<https://gtexportal.org>)  
658 using the European reference panel.

659 † List of references or sources (GeneCards, GWAS Catalog, PhenoScanner) available in **Table E6**.

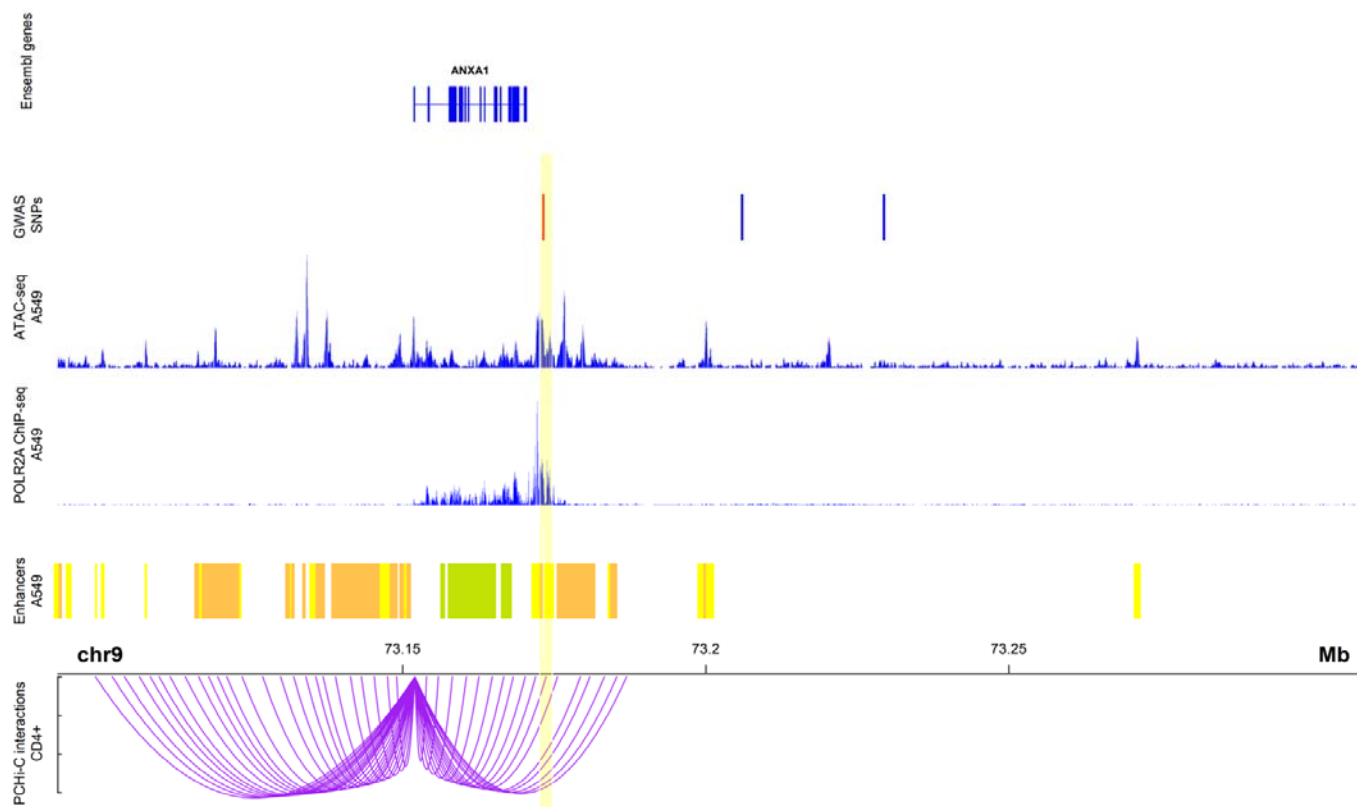
660 **Figure 1.** Circular Manhattan plot showing an overview of the GWAS results by wheeze phenotype (from  
661 outside to inside: early-onset persistent, early-onset pre-school remitting, early-onset mid-childhood  
662 remitting and late-onset wheeze). The red line indicates the genome-wide significance threshold ( $P < 5 \times$   
663  $10^{-8}$ ), while the blue line indicates the threshold for genetic variants that showed a suggestive significant  
664 association ( $P < 10^{-5}$ ).



665  
666  
667  
668



669 **Figure 2.** Chromatin interactions between rs75260654 and the *ANXA1* promoter in CD4+ cells in MAAS

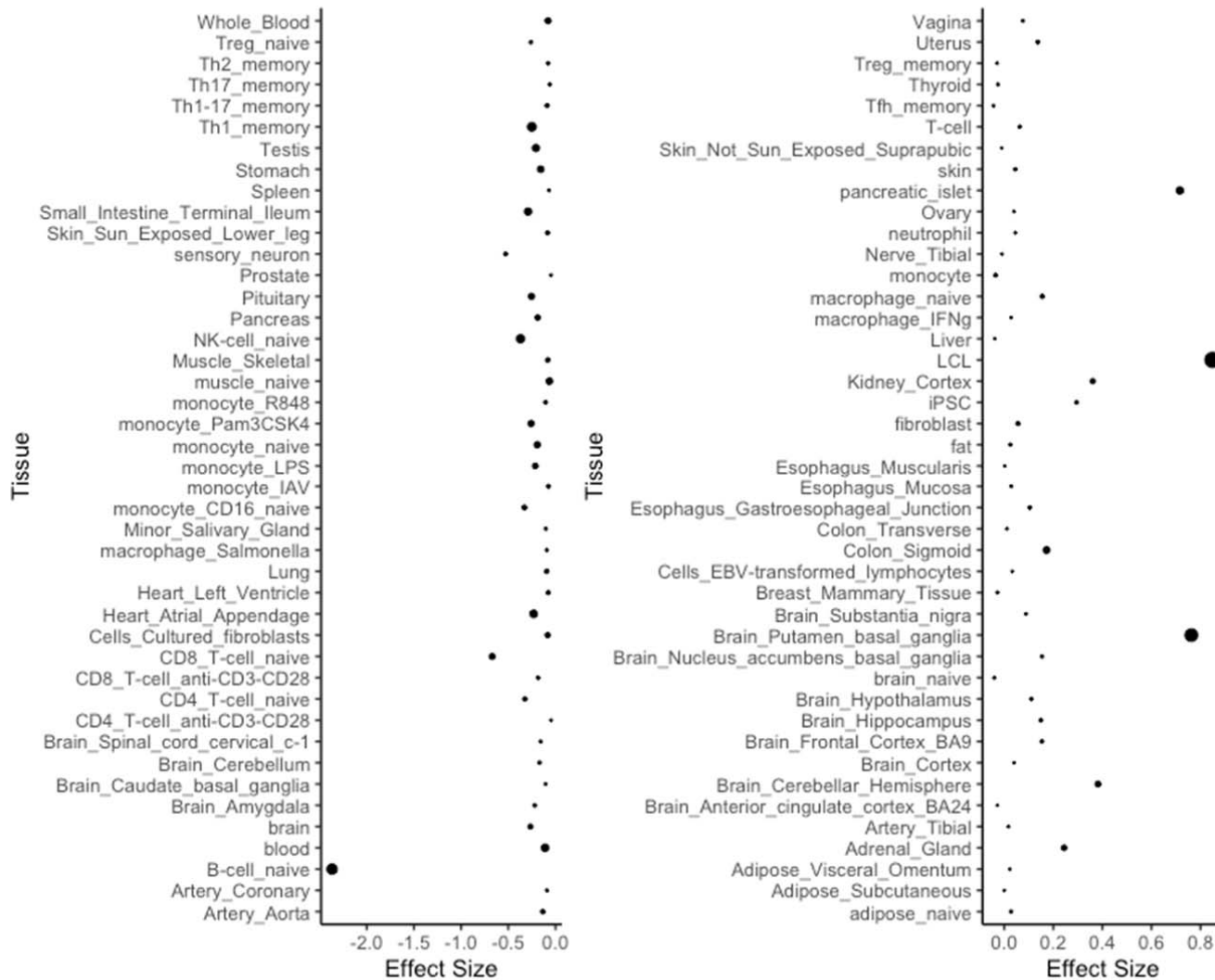


670

671 *rs75260654* physically interacts with *ANXA1* promoter in CD4+ T-cells and overlaps a region of active (POLR2AphosphoS2 CHIP-seq) open (ATAC-  
672 seq) chromatin in A549 cell line (lung epithelial carcinoma). The region is also predicted to be an active enhancer (ChromHMM 18-state model) in  
673 the A549 cell type. Only ChromHMM enhancer chromatin are displayed. Yellow shaded area indicates the PChi-C fragment overlapping  
674 *rs75260654* (red bar) and interacting with the *ANXA1* promoter.

675 **Figure 3.** eQTL ANXA1 and rs75260654 across different tissue types. Point size is proportional

Tissue specific effect of rs75260654 on ANXA1 expression

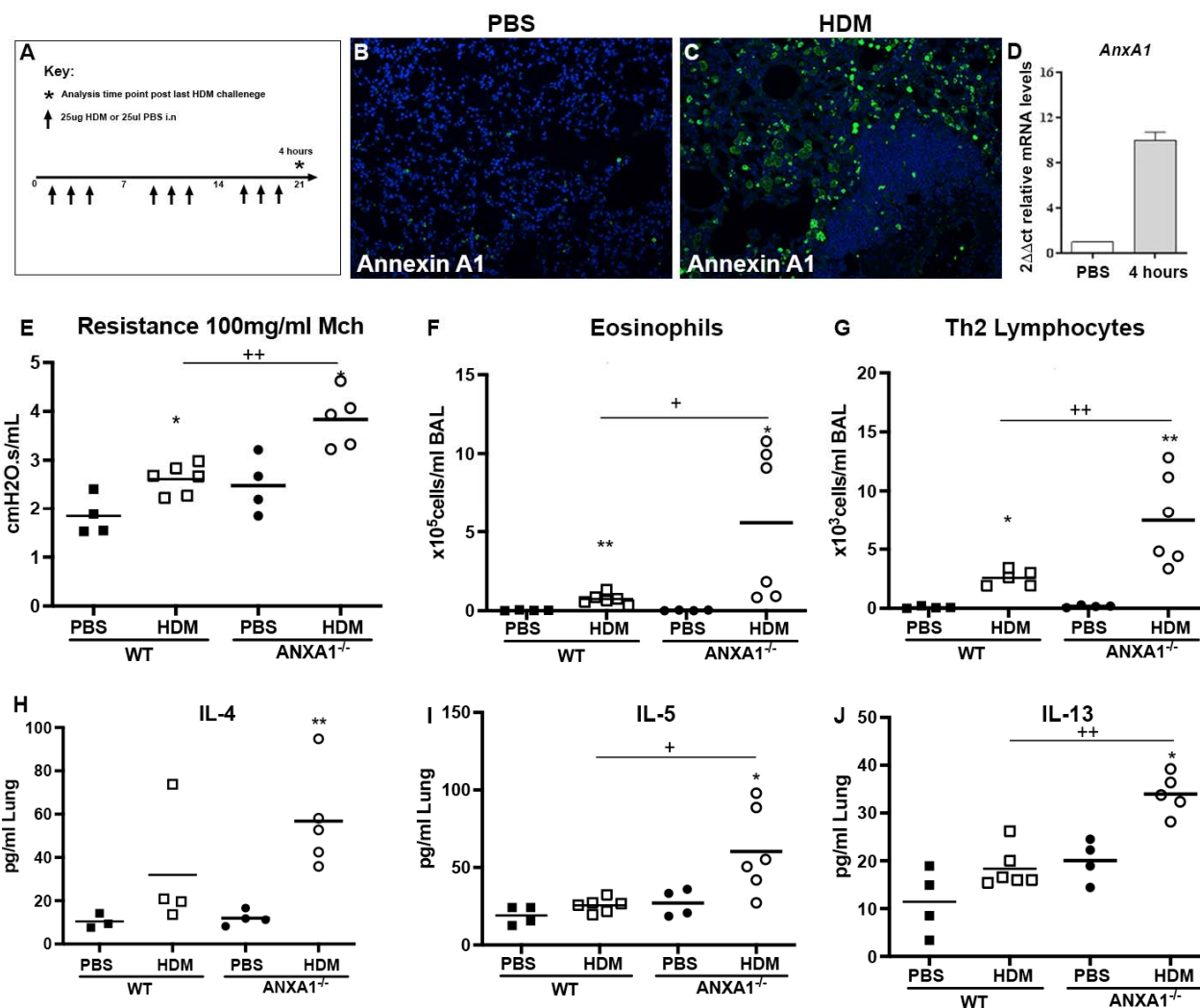


676 to  $-\log_{10}$  p-value.

677

678

679 **Figure 4.** Annexin A1 is induced following HDM challenge and mice deficient in *ANXA1* have  
 680 exacerbated airway hyperreactivity. (A) Schematic of house dust mite allergen dosing  
 681 protocol, N=4-6 per group, data representative of two animal experiments. (B, C)  
 682 Immunofluorescent staining of paraffin embedded lung tissue sections incubated with anti-  
 683 Annexin A1, counterstained with DAPI (N=4 per group). (D) mRNA expression of annexin A1 in  
 684 lung tissue following HDM exposure, expression normalized to housekeeping gene *hprt*. Mice  
 685 receiving HDM were analysed for changes in airway hyper-reactivity following methacholine  
 686 (MCh) challenge in tracheotomized restrained mice. (E) Airway resistance at top MCh dose  
 687 100mg/ml (F) Eosinophils quantified in BAL, (F) T1/ST2+ lymphocytes quantified in the BAL.  
 688 (H) IL-4 (I) IL-5 and (J) IL-13 quantified in lung tissue by ELISA. \*  $p < 0.05$  and \*\*  $p < 0.01$  relative  
 689 to PBS control group by Mann Whitney test. +  $p < 0.05$  and ++  $p < 0.01$  comparing HDM



690 AnnexinA1 KO mice relative to HDM WT group by Mann Whitney test.

691

1 **A meta-analysis of genome-wide association studies of childhood wheezing**  
2 **phenotypes identifies *ANXA1* as a susceptibility locus for persistent wheezing**

3 Granell R, Curtin JA, Haider S, Kitaba N, Mathie S, Gregory L, Yates LL, Tutino M, Hankinson J, Perretti  
4 M, Vonk JM, Arshad SH, Cullinan P, Fontanella S, Roberts G, Koppelman GH, Simpson A, Turner S,  
5 Murray CS, Lloyd CM, Holloway JW, Custovic A

6 **on behalf of UNICORN and Breathing Together Investigators**

7

8 **SUPPLEMENTARY INFORMATION**

## 9 ONLINE METHODS

### 10 Description of cohorts

11 The Study Team for Early Life Asthma Research (STELAR) consortium<sup>1</sup> brings together five UK  
12 population-based birth cohorts as described below. All studies were approved by research ethics  
13 committees. Informed consent was obtained from parents, and study subjects gave their  
14 assent/consent when applicable. Data were harmonised and imported into Asthma eLab web-based  
15 knowledge management platform to facilitate joint analyses ([www.asthmaelab.org](http://www.asthmaelab.org))<sup>1</sup>.

#### 16 The Avon Longitudinal Study of Parents and Children (ALSPAC)

17 ALSPAC is a birth cohort study established in 1991 in Avon, UK<sup>2,3</sup>. Pregnant women with expected  
18 dates of delivery 1<sup>st</sup> April 1991 to 31<sup>st</sup> December 1992 were invited to take part in the study. The initial  
19 number of pregnancies enrolled is 14,541. Of these initial pregnancies, there was a total of 14,676  
20 fetuses, resulting in 14,062 live births and 13,988 children who were alive at 1 year of age.

21 When the oldest children were approximately 7 years of age, an attempt was made to bolster the  
22 study with eligible cases who had failed to join originally. As a result, when considering variables  
23 collected from the age of seven onwards (and potentially abstracted from obstetric notes) there are  
24 data available for more than the 14,541 pregnancies mentioned above. The number of new  
25 pregnancies not in the initial sample (known as Phase I enrolment) that are currently represented on  
26 the built files and reflecting enrolment status at the age of 24 is 913 (456, 262 and 195 recruited  
27 during Phases II, III and IV respectively), resulting in an additional 913 children being enrolled. The  
28 phases of enrolment are described in more detail in the cohort profile paper and its update. The total  
29 sample size for analyses using any data collected after the age of seven is therefore 15,454  
30 pregnancies, resulting in 15,589 fetuses. Of these 14,901 were alive at 1 year of age.

31 Ethical approval: Ethical approval for the study was obtained from the ALSPAC Ethics and Law  
32 Committee and the Local Research Ethics Committees. Informed consent for the use of data collected  
33 via questionnaires and clinics was obtained from participants following the recommendations of the  
34 ALSPAC Ethics and Law Committee at the time.

35 Data dictionary: The study website contains details of available data through a fully searchable data  
36 dictionary: <http://www.bristol.ac.uk/alspac/researchers/our-data/>

37 We are extremely grateful to all the families who took part in this study, the midwives for their help in  
38 recruiting them, and the whole ALSPAC team, which includes interviewers, computer and laboratory  
39 technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses.

#### 40 The Manchester Asthma and Allergy Study (MAAS)

41 MAAS is an unselected birth cohort study established in 1995 in Manchester, UK<sup>4</sup>. It consists of a  
42 mixed urban-rural population within 50 square miles of South Manchester and Cheshire, United  
43 Kingdom located within the maternity catchment area of Wythenshawe and Stepping Hill Hospitals.  
44 All pregnant women were screened for eligibility at antenatal visits (8-10<sup>th</sup> week of pregnancy). Of the  
45 1499 couples who met the inclusion criteria ( $\leq 10$  weeks of pregnancy, maternal age  $\geq 18$  years, and  
46 questionnaire and skin prick data test available for both parents), 288 declined to take part in the  
47 study and 27 were lost to follow-up between recruitment and the birth of a child. A total of 1184  
48 children were born into the study between February 1996 and April 1998. They were followed  
49 prospectively for 19 years to date and attended follow-up clinics for assessments, which included lung  
50 function measurements, skin prick testing, biological samples (serum, plasma and urine), and  
51 questionnaire data collection. The study was approved by the North West – Greater Manchester East  
52 Research Ethics Committee.

## 53 The Study of Eczema and Asthma to Observe the influence of Nutrition (SEATON)

54 SEATON is an unselected birth cohort study established in 1997 in Aberdeen, UK, which was designed  
55 to explore the relationship between antenatal dietary exposures and asthma outcomes in childhood<sup>5</sup>.  
56 2000 healthy pregnant women attending an antenatal clinic, at median 12 weeks gestation, were  
57 recruited. An interviewer administered a questionnaire to the women and atopic status was  
58 ascertained by skin prick test (SPT). The cohort included 1924 children born between April 1998 and  
59 December 1999. Participants were recruited prenatally and followed up by self-completion  
60 questionnaire to 15 years of age using postal questionnaires to record the presence of asthma and  
61 allergic diseases. Lung function measurements and SPT to common allergens was performed at 5, 10  
62 and 15 years. The study was approved by the North of Scotland Research Ethics Committee.

## 63 ASHFORD

64 The Ashford study is an unselected birth cohort study established in 1991 in Ashford, UK<sup>6</sup>. It included  
65 642 children born between 1992 and 1993. Participants were recruited prenatally and followed to age  
66 14 years. Detailed standardised questionnaires were administered at each follow-up to collect  
67 information on the natural history of asthma and other allergic diseases. Lung function measurements  
68 and SPT was carried out at 5, 8 and 14 years of age. In 2015, the study children aged 20 were sent a  
69 self-completion questionnaire, which was returned by 60% of the participants.

## 70 The Isle of Wight (IOW) cohort

71 IOW is an unselected birth cohort study established in 1989 on the Isle of Wight, UK<sup>7-9</sup>. After the  
72 exclusion of adoptions, perinatal deaths, and refusal for follow-up, written informed consent was  
73 obtained from parents to enrol 1,456 newborns born between 1<sup>st</sup> January 1989 and 28<sup>th</sup> February  
74 1990. Follow-up assessments were conducted to 26 years of age to prospectively study the  
75 development of asthma and allergic diseases. At each follow-up, validated questionnaires were  
76 completed by the parents. Additionally, the Skin Prick Test (SPT) was performed on 980, 1036 and 853  
77 participants at 4, 10 and 18 years of age to check allergic reactions to common allergens. At 10, 18,  
78 and 26 years, spirometry and methacholine challenge tests were performed to diagnose lung  
79 problems. Ethics approvals were obtained from the Isle of Wight Local Research Ethics Committee  
80 (now named the National Research Ethics Service, NRES Committee South Central – Southampton B)  
81 at recruitment and for the subsequent follow-ups.

## 82 **Definition of variables**

83 A list of all variables used in the current study, per cohort, is shown in **Table E1**.

## 84 Demographic, exposures and outcomes

85 Postal questionnaires were used in ALSPAC and SEATON, while interviewer-administered  
86 questionnaires were employed in other cohorts.

87 Parental history of asthma, eczema and hay fever was defined based on the responses given to the  
88 question “have you (and/or your partner) ever had asthma/eczema/hay fever”. Maternal and paternal  
89 smoking were defined based on the response given to the question “do you (or does your partner)  
90 smoke”, administered during pregnancy. Low birth weight was defined as birth weight less than 2500  
91 g based on NHS birth records.

92 Asthma in MAAS was defined as a case if positive for two of the following criteria: doctor diagnosis of  
93 asthma in the past 12 months, current wheeze in the last 12 months, doctor prescription for asthma.  
94 Asthma in ALSPAC was defined as a mothers’ report of doctor ever diagnosis of asthma.

95 Current wheeze in MAAS was defined as a questionnaire report to the question “have you wheezed in  
96 the last 12 months” upon attendance at a follow up clinic. Current wheeze in ALSPAC was defined as a  
97 mothers’ report to the question “has your child had any wheezing or whistling in the last 12 months?”.

98 Asthma medication in ALSPAC was defined as a mothers' report to the question "has your child taken  
99 any asthma medication in the last 12 months?". Lower respiratory hospital admissions: Data on  
100 hospital admissions in MAAS were obtained by manually inspecting the General Practice (GP) records  
101 for each individual.

102 Early-life risk factors were divided into four groups according to timing of exposure; maternal and  
103 child characteristics (gender, maternal smoking during pregnancy and maternal history of asthma),  
104 perinatal (low birth weight adjusted for gestational age), environmental (pet ownership, smoke  
105 exposure after birth) and allergic sensitization (defined based on positive skin prick test to cat, house  
106 dust mite or grass) variables.

#### 107 Primary outcome: Joint wheeze phenotypes

108 We used latent class analysis (LCA) to identify longitudinal trajectories of wheeze<sup>10</sup> based on pooled  
109 analysis among 15,941 children with at least two observations on wheezing at five time periods that  
110 were approximately shared across all cohorts: infancy (½-1 year); early childhood (2-3 years); pre-  
111 school/early school age (4-5 years); middle childhood (8-10 years); and adolescence (14-18 years).  
112 Cohort-specific definitions other variables derived from the questionnaires are provided in **Table E2**.

113 To control for cohort-specific variation, Cohort ID was included in the LCA model as an additional  
114 predictor by transforming the 5-category variable into a set of four dummy variables and including  
115 them as covariates. The largest cohort, ALSPAC, was treated as the non-coded category to which all  
116 other cohorts were compared. The expectation maximization algorithm was used to estimate relevant  
117 parameters, with 100,000 iterations and 500 replications.

118 To assess model fit, we used (1) the Bayesian information criterion (BIC), (2) the Akaike information  
119 criterion (AIC), (3) Lo-Mendell-Rubin likelihood ratio test (LMR), (4) Bootstrapped likelihood ratio and,  
120 (4) quality of classification certainty (model entropy). The BIC is an index used in Bayesian statistics to  
121 choose among a set of competing models; the model with the lowest BIC is preferred. Using the  
122 lowest BIC as a selection criterion, the best fitting model was chosen as the five-class solution with a  
123 nominal covariate (BIC:31340). Analyses were carried out using Mplus 8, R  
124 (<http://www.r-project.org/>) and Stata 14 (StataCorp, College Station, Tex).

125 Based on the statistical fit, a five-class solution was selected as the optimal model<sup>10</sup>, and the classes  
126 (wheeze phenotypes) were labeled as: (1) *Never/Infrequent wheeze* (52.4%); (2) *Early-onset pre-school*  
127 *remitting wheeze* (18.6%), with high prevalence of wheeze during infancy, decreasing to 20% around  
128 early-childhood and to less than 10% afterwards; (3) *Early-onset middle-childhood remitting wheeze*  
129 (9.8%), with early-onset wheeze and peak prevalence in early-childhood (~70%), and diminishing by  
130 middle-childhood (<5%); (4) *Early-onset persistent wheeze* (10.4%) with 58% wheeze prevalence  
131 during infancy, and prevalence between 70- 80% thereafter; (5) *Late-onset wheeze* (8.8%) with very  
132 low prevalence until middle childhood, increasing rapidly to 55% in adolescence. These latent classes  
133 were used in the subsequent GWAS.

#### 134 *Minimising Bias and missing data effects*

135 Extracted from reference Oskel et al. (10): "*One of the advantages of our multicohort approach is that*  
136 *individual studies that might not provide conclusive evidence to make inference about the general*  
137 *population because of cohort specific effects and biases can contribute to revealing a more accurate*  
138 *picture when integrated together. The integration of five cohorts and their pooled analysis enhanced*  
139 *the credibility and generalizability of the phenotyping results to the U.K. population. A further*  
140 *advantage is to minimize the study-specific biases (including cohort specific effects, attrition effects,*  
141 *different recruitment strategies, and geographic factors) affecting the certainty of allocation of*  
142 *individuals to each latent class, while maximizing the benefits of individual cohort studies (e.g.,*  
143 *potentially important risk factors and outcomes are captured in some, but not all cohorts)."*

144 *“Another strength of pooling cohort data is that a multicohort design allowed us to analyze a large*  
145 *sample with complete data on wheeze from birth to adolescence, thus increasing statistical power to*  
146 *detect less prevalent phenotypes.”* However, *“The optimal solution in the model using 15,941 children*  
147 *(allowing for missing data) remained five classes (see Table E3, Figure E1), and was very similar to that*  
148 *derived from a complete data set.”* We used results from the larger sample, that is individuals with at  
149 least 2 observations of wheezing, to assign individuals to their most likely wheezing phenotype and  
150 used this as our primary outcome in this study.

## 151 **Genotyping and imputation**

### 152 ALSPAC

153 Participants were genotyped using the Illumina HumanHap550 quad genome-wide SNP genotyping  
154 platform (Illumina Inc., San Diego, CA, USA) by the Wellcome Trust Sanger Institute (WTSI; Cambridge,  
155 UK) and the Laboratory Corporation of America (LCA, Burlington, NC, USA), using support from  
156 23andMe. Haplotypes were estimated using ShapeIT (v2.r644) which uses relationship information to  
157 improve phasing accuracy. The phased haplotypes were then imputed to the Haplotype Reference  
158 Consortium (HRCr1.1, 2016) panel<sup>11</sup> of approximately 31,000 phased whole genomes. The HRC panel  
159 was phased using Shapelt v2, and the imputation was performed using the Michigan imputation  
160 server.

### 161 MAAS

162 In MAAS, we used the Illumina 610 quad genome-wide SNP genotyping platform (Illumina Inc., San  
163 Diego, CA, USA). Prior to imputation samples were excluded on the basis of gender mismatches;  
164 minimal or excessive heterozygosity, genotyping call rates of <97%. SNPs were excluded if they had  
165 call rates of < 95%, minor allele frequencies of < 0.5% and HWE  $p < 3 \times 10^{-8}$ . Prior to imputation each  
166 chromosome was pre-phased using EAGLE2 (v2.0.5)<sup>11</sup> as recommended by the Sanger imputation  
167 server<sup>12</sup>. We then imputed with PBWT<sup>13</sup> with the Haplotype Reference Consortium (release 1.1) of  
168 32,470 reference genomes<sup>12</sup> using the Sanger Imputation Server.

### 169 IOW, SEATON and ASHFORD

170 IOW, SEATON and ASHFORD were genotyped using the illumina Infinium Omni2.5-8 v1.3 BeadChip  
171 genotyping platform (Illumina Inc., San Diego, CA, USA). Genotype QC and imputation was carried out  
172 as described for MAAS.

### 173 Exclusions

174 Individuals were excluded on the basis of gender mismatches; minimal or excessive heterozygosity;  
175 disproportionate levels of individual missingness (>3%), insufficient sample replication (IBD < 0.8) or  
176 evidence of cryptic relatedness (IBD > 0.1). Following imputation, single nucleotide polymorphisms  
177 (SNPs) with a minor allele frequency of <1%, a call rate of <95%, evidence for violations of Hardy-  
178 Weinberg equilibrium ( $P < 5 \times 10^{-7}$ ) or imputation quality measure (MaCH-Rsq or IMPUTE-info score) <0.40  
179 were excluded. All individuals with non-European ancestry and siblings were removed.

## 180 **GWAS Meta-analysis**

181 GWAS of the joint wheezing phenotypes were performed independently in ALSPAC, MAAS and the  
182 combined IOW-SEATON-ASHFORD (combined as they were genotyped on the same platform, at the  
183 same time, and quality-controlled and imputed together). All genetic data were imputed to a new  
184 Haplotype Reference Consortium panel. This comprises around 31,000 sequenced individuals (mostly  
185 European), so the coverage of European haplotypes is much greater than in other panels. As a  
186 consequence, we expect to improve imputation accuracy, particularly at lower frequencies.



187 We used SNPTTEST v2.5.2<sup>14</sup> with a multinomial regression model (-method newml, never/infrequent  
188 wheeze as the reference) to investigate the association between SNPs and wheezing phenotypes. A  
189 meta-analysis of the three GWASs including 5,887 controls and 943 cases for early-onset persistent,  
190 1482 cases for early-onset remitting, 603 cases for mid-childhood-onset remitting and 652 cases for  
191 late-onset wheeze, was performed using METAL<sup>15</sup> with a total of 8,057,852 SNPs present. We used the  
192 option SCHEME STDERR in METAL to implement an effect-size based method weighted by each study-  
193 specific standard error in a fixed-effects model. We performed clumping to keep only one  
194 representative SNP per Linkage disequilibrium (LD) block and used locus zoom plots to short-list  
195 independent SNPs for further annotation.

## 196 **LD clumping, pre-Selection and Gene Annotation**

197 LD clumping was performed for all SNPs with  $p\text{-value} < 10^{-5}$  for at least one wheezing phenotype. In  
198 order to avoid redundancy between SNPs and to ensure associations are independent, we used  
199 significance thresholds of 0.05 for index and clumped SNPs (--clump-p1 0.05,--clump-p2 0.05), LD  
200 threshold of 0.80 (--clump-r2 0.80) and physical distance threshold of 250kb (--clump-kb 250).  
201 European 1000 Genome data were used to infer LD structure.

202 Locus Zoom plots (<http://locuszoom.org/>)<sup>16</sup> were used for close inspection of all independent signals.  
203 Loci showing a peak with different colour dots (possibly indicating more than one causal variant) were  
204 short-listed for further annotation. SNPnexus database (<https://www.snp-nexus.org/v4/>)<sup>17</sup> was used  
205 to annotate the overlapping, upstream and downstream genes; the GWAS catalogue (by SNP and then  
206 gene) (<https://www.ebi.ac.uk/gwas/search>), GeneCards (<https://www.genecards.org/>)<sup>18</sup> database and  
207 phenoscanner (<http://www.phenoscaner.medschl.cam.ac.uk/>) were used to further explore  
208 previously associated relevant phenotypes and gene function. Lead SNPs were looked in  
209 <https://www.regulomedb.org/> to assess potential functionality.

## 210 **Gene expression in whole blood and lung tissues**

211 The top independent SNPs associated with each of the wheeze phenotypes were assessed for their  
212 association with cis- and trans-acting gene expression (mRNA) in whole blood and lung tissues. We  
213 identified potential eQTL signals using Genotype-Tissue Expression database (<https://gtexportal.org>)  
214 using the European reference panel.

## 215 **Post-GWAS: rs75260654 (ANXA1)**

### 216 Annotation & distribution

217 Information including chromosome, strand, clinical significance was retrieved from ENSEMBL using the  
218 R package biomaRt<sup>19,20</sup>. The effects of rs75260654 on genomic features were predicted by querying  
219 the using Ensembl Variant Effect Predictor (VEP)<sup>21</sup> web tool.

220 rs75260654 distribution in the GRCh38.p13 build of the human genome across African, Asian and  
221 European populations of the 1000 Genomes Project Phase 3 were accessed by querying the Ensembl  
222 ([www.ensembl.org](http://www.ensembl.org)) web browser on 24 May 2021.

### 223 Promoter Capture

224 The Hi-C libraries were prepared from CD4+ T-cells isolated from 7 healthy individuals (2 libraries per  
225 individual) from the MAAS cohort using the Arima-HiC kit (Arima Genomics). Promoter Capture Hi-C  
226 (PCHi-C) libraries were generated by capturing the restriction fragments (RF) overlapping the TSS of  
227 18775 protein coding genes using the Agilent SureSelectXT HS Target Enrichment System according to  
228 the manufacturers' protocols. The final design included 305419 probes covering 13.476Mb and 18630  
229 protein-coding genes. The restriction fragments (RF) overlapping the TSS (+/-1 RF, 3 RF per promoter)

230 were captured with custom-designed biotinylated RNA baits. Libraries were sequenced to ~300M  
231 2x150bp reads each (~600M reads/individual). The 3'-end of the reads was quality trimmed with  
232 Sickle. The sequencing data were processed with the HiCUP pipeline to map the sequencing reads and  
233 eliminate experimental artefacts and PCR duplicates<sup>22</sup>. The BAM files from technical replicates were  
234 merged. Promoter interactions were called using the CHiCAGO pipeline<sup>23</sup>, which calls statistically  
235 significant interactions in PChi-C data while accounting for noise and PChi-C specific bias. A CHiCAGO  
236 score > 5 (soft-thresholded -log weighted p-value) was considered significant. To gain information  
237 from all the available data, the BAM files from all 7 individuals were supplied as biological replicates in  
238 the analysis with CHiCAGO. Moreover, to increase power, restriction fragments were binned as  
239 follows: 10 consecutive RF that were not covered by the baits were binned together; the 3 baited  
240 fragments for each promoter were binned with 1 RF upstream and 1 downstream, totaling 5  
241 fragments per promoter. If the bins for two consecutive promoters overlapped, these were binned  
242 together into a single larger bin. Publicly available ENCODE ATAC-seq (A549 cell line) and ChIP-seq  
243 (A549 cell line ENCFF900GVO) and POLR2A ChIP-seq (A549 cell line, ENCFF737ZKN) data and 18-state  
244 ChromHMM from the EpiMap Project (BSS00007)<sup>24</sup> for A549 cell line were downloaded. The PChi-C  
245 interactions of interest and their overlap with ATAC-seq and ChIP-seq peaks, and putative enhancers  
246 from the 18-state ChromHMM model were visualised using the Sushi R package.

#### 247 eQTL catalogue lookup

248 We queried the eQTL catalogue (<https://www.ebi.ac.uk/eql/>; accessed 6 May 2021) using tabix-0.2.6  
249 to assess if *rs75260654*, *rs116849664* or *rs78320984* are eQTLs in studies that utilised the following  
250 cell types: lung, T cells, blood, monocytes, neutrophils, NK cells, fibroblasts, B cells, CD4+ T cells, CD8+  
251 T cells, Th17 cells, Th1 cells, Th2 cells, Treg naive, Treg memory, CD16+ monocytes, Cultured  
252 fibroblasts, EBV-transformed lymphocytes. We defined nominal significance as  $p \leq 0.05$ .

#### 253 Variant effect

254 Variant effect on tissue-specific gene expression, which is based on GTEx eQTL, was retrieved on May  
255 24 from eQTL Ensembl database (<https://www.ensembl.org/>) and eQTLGene Consortium  
256 (<https://www.eqtlgen.org/cis-eqtls.html>). Using downloaded correlation of variant on tissue specific  
257 gene expression from Ensembl, the relative effect of T allele on the *ANXA1* expression across 86 tissue  
258 types was presented in scatter plot using R version 3.6.1.<sup>25</sup> To get information on the functional role of  
259 *ANXA1*, the top 30 interacting proteins and enrichment were retrieved from STRING database<sup>26</sup> into  
260 cystoscape for visualization.<sup>27</sup>

#### 261 **Functional mouse experiments**

##### 262 Mice

263 In accordance with the Animals (scientific procedures) act 1986, all animal experiments were  
264 conducted under the approved UK Home Office Project License No: PPL 70/7643, reviewed by Imperial  
265 College's Animal Welfare and Ethical Review body. Female WT BALB/c and annexin A1 knockout (KO)  
266 mice were purchased from Charles River (Bicester, UK). Animals aged 6-8 weeks of age received 25ug  
267 intranasal instillation of either HDM (Greer Laboratories, Lenoir, NC, USA; Cat: XPB70D3A25), or PBS  
268 3x a week for 3 weeks. Mice were sacrificed 4 hours post-final HDM challenge. Mice were housed  
269 under specific pathogen-free conditions and a 12:12 light:dark cycle. Food and water were supplied ad  
270 libitum. All animal experiments were completed twice, with N=4-6 per group.

##### 271 Airway hyperresponsiveness

272 Airway hyperreactivity was measured using Flexivent™. Lung resistance was measured in response to  
273 increasing doses of methacholine (3-100mg/ml, Sigma, Poole, UK, Cat: [A2251](#)) in tracheotomised  
274 anaesthetised mice using an EMMS system (Electro-Medical Measurement Systems, UK).

## 275 Flow Cytometry Analysis

276 Bronchoalveolar lavage (BAL) was collected. BAL cells were restimulated with ionomycin and phorbol  
277 12-myristate 13-acetate in the presence of brefeldin (Sigma), as previously described<sup>28</sup>. Specific  
278 antibodies for T1/ST2 staining were purchased from Morwell Diagnostics (Zurich, Switzerland). Cells  
279 were also stained for lineage negative cocktail, Ly6G, CD45, CD11b, CD11c, SiglecF. Labelled cells were  
280 acquired on a BD Fortessa (BD Biosciences, Oxford, UK) and analysed using FlowJo software (Treestar,  
281 Ashland, Oregon, USA). Details of antibodies used can be found in the table below.

282

Molecule	Manufacturer	Isotype	Conjugated dye	Clone
T1/ST2	Morwell Diagnostics GMBH, Switzerland	Rat IgG1	FITC	DJ8
CD45	e-Bioscience Ltd, Hatfield, UK	Rat IgG2b	PerCP-CY5.5	30-F11
CD11b	BD Biosciences, Oxford, UK	Rat IgG2b	e450	M1/70
CD11c	e-Bioscience Ltd, Hatfield, UK	Hamster IgG1	PerCP-CY5.5	N418
Siglec F	BD Biosciences, Oxford, UK	Rat IgG2a	PE	E50-2440

283

## 284 Analysis of cytokines and chemokines

285 Murine lung tissue homogenate supernatants were processed as previously described<sup>28</sup>. Cytokine  
286 levels were analysed by ELISA: IL-4, IL-5 (PharMingen, Oxford, UK), IL-13 Ready-Set-Go kits  
287 (eBioscience).

## 288 Real time-PCR

289 Total RNA was extracted from murine lung tissue using an RNeasy Mini Kit (Qiagen). Total RNA (1µg)  
290 was reverse transcribed into cDNA using a High-Capacity cDNA Reverse Transcription Kit (Life  
291 Technologies, UK). Real-time PCR reactions were performed using TaqMan Gene Expression Master  
292 Mix and TaqMan Gene Expression probes, annexin A1, and HPRT (Applied Biosystems). Values were  
293 normalised to HPRT and gene expression was analysed using the change-in-threshold 2- $\Delta$ CT method.

## 294 Annexin A1 Immunohistochemistry

295 Paraffin-embedded mouse lung sections were stained with annexin A1 (R&D Systems, MAB3770).  
296 Annexin A1 primary antibody was followed by a secondary detection antibody (donkey anti-goat 488,  
297 Thermofisher, A11055). Annexin A1+ cells were quantified by manual counting under microscope and  
298 numbers averaged over four fields, from five biological replicates per group.

## 299 Statistical analysis

300 Data are expressed as median  $\pm$  IQR. Statistical differences between groups were calculated using  
301 Mann Whitney U test, unless otherwise specified. p-values are indicated in figures.

## 302 **Replication of ANXA1 top hits in PIAMA cohort**

### 303 PIAMA cohort description

304 PIAMA (Prevention and Incidence of Asthma and Mite Allergy) is an ongoing birth cohort study. Details  
305 of the study design have been published previously<sup>29,30</sup>. In brief, pregnant women were recruited from  
306 the general population through antenatal clinics in the north, west and centre of the Netherlands in  
307 1996-1997. The baseline study population consisted of 3963 new-borns. Questionnaires were  
308 completed by the parents during pregnancy when the child was 3 months old, and then annually from

309 1 up to 8 years; at ages 11, 14 and 17 years, questionnaires were completed by the parents as well as  
310 the participants themselves.

311 LCA wheezing phenotypes

312 A 6 class LCA model was identified including 3,832 individuals with at least 2 observations of wheeze  
313 between 1 and 11-12 years of age. The identified classes were labelled: never/Infrequent  
314 (2909,75.91%), pre-school onset remitting (571, 14.90%), mid-childhood school remitting (108, 2.82%),  
315 intermediate onset remitting (106, 2.77%), school-age onset persisting (74, 1.93%) and continuous  
316 wheeze (64, 1.67%).

317 Replication analyses

318 We analyzed associations between SNPs downstream of *ANXA1* (**Table E7, Figure E6**) and continuous  
319 wheezing in PIAMA, using the never/infrequent wheezing as the baseline category. Analyses were  
320 carried out in SPSS using a logistic regression model.

## 321 ONLINE RESULTS

### 322 Participants and descriptive data

323 Demographic characteristics of the participants in STELAR cohorts included in this analysis are shown  
324 in **Table E3**. Cohorts contain similar proportions of males (range 48%-54%), maternal history of asthma  
325 (11%-14%), maternal smoking (14%-23%), (doctor-diagnosed) asthma ever during mid-childhood  
326 (16%-24%) and adolescence (20%-30%), current wheeze (12%-20% mid-childhood, 9%-25%  
327 adolescence) and current use of asthma medication (12%-17% mid-childhood, 11%-17% adolescence).

### 328 Comparison of participants included and excluded from this analysis

329 In ALSPAC, 11,176 individuals had data on wheezing phenotypes, of these 6,833 were white unrelated  
330 and had genetic data. We found more children from mothers who smoked during pregnancy in the  
331 excluded sample compared to the included sample; no difference in gender, maternal history of  
332 asthma, current wheezing at 8 or 15 years, and small evidence for more asthma ever and current  
333 medication at 8 years in the excluded sample (**Table E4**).

334 In MAAS, 1150 individuals had data on wheezing phenotypes, of these 887 were white unrelated and  
335 had genetic data. We found no difference in children from mothers who smoked during pregnancy in  
336 the excluded sample compared to the included sample; no difference in gender, maternal history of  
337 asthma or current wheezing at both 8 and 16 years. There was small evidence for more asthma ever  
338 and current medication at 8 years in the excluded sample (**Table E4**).

339 In SEATON, 1535 individuals had data on joint wheezing phenotypes, of these 548 were white  
340 unrelated and had genetic data. We found evidence for more children from mothers who smoked  
341 during pregnancy in the excluded sample compared to the included sample; and more males in the  
342 excluded sample. There was no difference in maternal history of asthma or current wheezing, asthma  
343 ever or current medication at both 10 and 15 years in the excluded sample compared to the included  
344 sample (**Table E4**).

345 In ASHFORD, 620 individuals had data on joint wheezing phenotypes, of these 348 were white  
346 unrelated and had genetic data. We found evidence for more children from mothers who smoked  
347 during pregnancy in the excluded sample compared to the included sample; no difference in gender,  
348 maternal history of asthma or asthma ever. There was small evidence for less current wheezing at 8  
349 years, or current medication at 8 years in the excluded sample compared to the included sample  
350 (**Table E4**).

351 In IOW, 1460 individuals had data on joint wheezing phenotypes, of these 952 were white unrelated  
352 and had genetic data. We found evidence for more children from mothers who smoked during  
353 pregnancy in the excluded sample compared to the included sample; no difference in gender,  
354 maternal history of asthma, asthma ever at 10 and 18 years in the excluded sample compared to the  
355 included sample. There was small evidence for more children with current wheeze and medication at  
356 8 years in the included sample compared to the included sample (**Table E4**).

### 357 GWAS Meta-Analysis: Main Results

358 The distribution of the minor allele frequencies was consistent across genotyped datasets (mean SD  
359 0.01). We assessed the deviation of the observed p-values from the null hypothesis by plotting QQ-  
360 plots for each wheezing phenotype (**Figure E1**). Some observed p-values were clearly more significant  
361 than expected under the null hypothesis, particularly for early-onset persistent wheeze, without an  
362 early separation of the expected from the observed (low evidence of population stratification;  
363 genomic control  $\lambda \geq 0.93$  for all four phenotypes).

### 364 Results of GWAS meta-analysis for 85 SNPs with main associations across wheeze phenotypes

365 Previously associated traits for each region/gene associated with different wheezing phenotypes are  
366 presented in **Table E6**.

367 Persistent Wheeze

368 We identified two GWAS-significant loci: 17q21,  $p < 5.5 \times 10^{-9}$ , and a novel locus on 9q21.13 (*ANXA1*),  
369  $p < 6.7 \times 10^{-9}$ . The remaining 17 loci ( $4.0 \times 10^{-7} \leq p\text{-values} \leq 9.8 \times 10^{-6}$ ) included regions previously associated  
370 with childhood asthma (1q43, 4p16.3, 4q31.21, 5p15.31, 7q22.3, 17q12), asthma and rhinitis (2p25.1),  
371 eosinophil count (3q21.3, 10q24.2, 11q23.3, 22q11.1), bronchial hyper-responsiveness (2q12.2), lung  
372 function (1q43, 3q21.3, 5q13.3, 15q25.2, 19p13.2), triglycerides measurement and/or glucose  
373 metabolism (11q23.3, 19p13.2 and 22q11.1), severe asthma (14q22.1) and severe asthma and insulin  
374 resistance (11p15.4).

375 Early-onset Preschool Remitting Wheeze

376 Among the regions associated with early-onset preschool remitting wheeze, we identified loci  
377 previously associated with smoking (3q26.31, 7q21.11 and 15q21.1), waist circumference and obesity  
378 (1q32.3), asthma and/or BMI (5q13.3, 6q27, 7q21.11), allergic disease and atopy (7q21.11) and airway  
379 repair (2p24.2 and 9p13.3).

380 Early-onset Mid-childhood Remitting Wheeze

381 Loci associated with this phenotype were previously associated with neutrophil counts (1q23.2, 3q29,  
382 20p12.3-p12.2), eosinophil counts and allergic rhinitis (4q24), pollution and DNA methylation (2p22.3),  
383 atopy (3p25.3), food allergy (13q31.1) and BMI (3q29, 5q14.1).

384 Late-onset Wheeze

385 Regions associated with late-onset wheeze were previously associated with adult-onset non-allergic  
386 asthma (3q13.2), asthma/allergic disease and allergy/atopic sensitization (3q23, 7q36.1), asthma  
387 and/or allergy in adolescence (9p22.3, 16p12.1), late-onset asthma and obesity (2q13), lung function  
388 or body height (2q14.3, 3p24.2, 3q13.2, 15q25.2), lymphocyte count and asthma susceptibility  
389 (1p32.2), obesity and/or metabolic syndrome/dysfunction (1p36.13 and 22q13.32), eczema (7q36.3),  
390 insulin resistance (10q23.31), type 1 diabetes (8q21.3), alcohol drinking (15q15.3-q21.1) and sex  
391 hormone-binding globulin levels (11q23.3).

392 **ONLINE TABLES**

393 **Table E1.** Definition of variables in each of the five STELAR birth cohorts

<b>Variable</b>	<b>Definition</b>
<b>Cohort: ALSPAC</b>	
<b>Mother-asthma</b>	Have you ever had asthma? (recruitment)
<b>Mother smoking</b>	Mother smoked when expecting (recruitment)
<b>Doctor-diagnosed asthma ever</b>	Has a doctor ever said that your child has asthma? (year 8 and 14)
<b>Current wheezing</b>	Two questions combined: Occurrence of wheezing and/or wheezing with whistling on the chest in the last 12 months (year ½, 2 <sup>1/2</sup> , 4 <sup>3/4</sup> , 8 <sup>1/2</sup> and 14)
<b>Current asthma medication</b>	Asthma medication in the last 12 months (year 8 <sup>1/2</sup> and 14)
<b>Current rhinitis</b>	Child had problem with sneezing/runny nose without cold/flu in last 12 months (year 7 and 16 <sup>1/2</sup> )
<b>Current hay-fever</b>	Child had hay-fever in last 12 months (year 10 <sup>1/2</sup> and 14)
<b>Cohort: MAAS</b>	
<b>Mother-asthma</b>	Has a doctor ever told you that you had asthma? (recruitment)
<b>Mother smoking</b>	Do you smoke- mother (recruitment)
<b>Doctor-diagnosed asthma ever</b>	Has your doctor ever told you that your child has or had asthma? (year 8 & 16)
<b>Asthma ever</b>	Has your child ever suffered from asthma (year 8 and 16)
<b>Current wheezing</b>	Has your child had wheezing or whistling in the chest in the last 6/12 months (year 1, 3, 5, 8 and 16)
<b>Current asthma medication</b>	Asthma medication in the last 12 months (year 8 and year 16)
<b>Current rhinitis</b>	Has your child ever had a problem with sneezing, or a runny nose, or a blocked nose when he /she did not have a cold or the flu? (year 8 and year 16)
<b>Current hay-fever</b>	Does your child have hay-fever now? (year 8 and year 16)
<b>Cohort: SEATON</b>	
<b>Mother-asthma</b>	Do you suffer from asthma? (recruitment)
<b>Mother smoking</b>	Which of the following best describes your smoking status? (recruitment)
<b>Doctor-diagnosed asthma ever</b>	Has your child ever suffered from asthma? If yes, has this been confirmed by a doctor? (year 10 & 15)
<b>Asthma ever</b>	Has your child ever suffered from asthma? (year 10) ; Have you ever suffered from asthma? (year 15)
<b>Current wheezing</b>	Has your child had wheezing in the chest in the last 12 months (year 1, 2, 5, 10 and 15)
<b>Current asthma medication</b>	Has your child been prescribed medicines/inhalers for asthma in the last 12 months? (year 10) Have you been prescribed medicines/inhalers for asthma in the last 12 months? (year 15)

<b>Current hay-fever</b>	Has your child suffered from hay-fever last 12 months? (year 10 & 15)
<b>Cohort: ASHFORD</b>	
<b>Mother-asthma</b>	Do you have or have you ever been told you have asthma? (recruitment)
<b>Mother smoking</b>	Do you smoke cigarettes? (recruitment)
<b>Doctor-diagnosed asthma ever</b>	Has your doctor ever told you that your child has or had asthma? (year 8 & 14)
<b>Asthma ever</b>	In the past 12 months has your daughter suffered from asthma? (year 8); Has she/he ever suffered from asthma? (year 14)
<b>Current wheezing</b>	Which one best describes your child's wheeze in past 12 months? 'Yes' (B:1-6, C:7+), 'No' (A:0) (year 1, 2, 5, 8 and 14)
<b>Current asthma medication</b>	Over the last 12 months has your daughter taken any of the following treatments (preventer, reliever, nebulizer, steroids) for asthma? (year 8 and year 14)
<b>Current rhinitis</b>	In the last twelve months has your child had a problem with sneezing or a runny or blocked nose? (year 8 & 14)
<b>Current hay-fever</b>	In your opinion does your child have hay fever now? year 8 Has your child ever had hay fever? year 14
<b>Cohort: IOW</b>	
<b>Mother-asthma</b>	Do you or have you suffered from asthma or wheezing (recruitment)
<b>Mother smoking</b>	Do you smoke in the house? (recruitment)
<b>Doctor-diagnosed asthma ever</b>	Asthma cared for by hospital specialist/ GP or nurse (year 10, 18 & 26)
<b>Asthma ever</b>	Child ever had asthma (year 10 and 18)
<b>Current wheezing</b>	Presence of wheeze since previous review (year 1, 2, 4, 10 and 18)
<b>Asthma medication ever</b>	Child ever had asthma treatment (year 18 year) combined with asthma treatment questions being asked at year 1, 2, 4, 10 and 18
<b>Current rhinitis</b>	In the past 12 months have you had a problem with sneezing, or a runny or blocked nose when you did not have a cold or the flu? (year 10, 18 & 26)

394

395



396 **Table E2.** The cohort-specific time points and sample size used to ascertain wheeze  
 397 phenotypes

398

<b>Birth Cohort:</b>	<b>IOW</b>	<b>MAAS</b>	<b>SEATON</b>	<b>ASHFORD</b>	<b>ALSPAC</b>
<i>Year of birth</i>	1989	1995	1997	1992	1991
<i>Questionnaire</i>	Interviewer -administered	Interviewer -administered	Postal	Interviewer -administered	Postal
<i>Data collection age (years)</i>	1, 2, 4, 10, 18	1, 3, 5, 8, 16	1, 2, 5, 10, 15	1, 2, 5, 8, 14	½, 2 <sup>1/2</sup> , 4 <sup>3/4</sup> , 8 <sup>1/2</sup> , 14
<i>No. of children with ≥2 observations on wheezing at five selected time points</i>	1460	1150	1535	620	11176

399

400 **Table E3.** Characteristics of the participants in STELAR cohorts included in this analysis (restricted to individuals with genetic data). Numbers  
 401 are N (%) except for age, where we report Mean (SD).

402

	<b>ALSPAC</b>	<b>MAAS</b>		<b>SEATON</b>		<b>ASHFORD</b>		<b>IOW</b>		
	N=6833 (71.4%)		N=887 (9.3%)		N=548 (5.7%)		N=348 (3.6%)		N=952 (9.9%)	
Males	3492 (51.1)		475 (53.6)		260 (47.5)		179 (51.4)		466 (49.0)	
Maternal history of asthma	748 (11.5)		120 (13.5)		77 (14.1)		49 (14.1)		106 (11.2)	
Maternal smoking	1423 (22.1)		122 (13.8)		107 (19.5)		52 (14.9)		217 (23.1)	
<b>Wheeze Phenotypes</b>										
<i>Never/Infrequent</i>	4331 (63.4)		506 (57.1)		332 (60.6)		145 (41.7)		573 (60.2)	
<i>Early-onset persistent</i>	656 (9.6)		133 (15.0)		36 (6.6)		41 (11.8)		77 (8.1)	
<i>Early-onset pre-school remitting</i>	1076 (15.8)		145 (16.4)		117 (21.4)		145 (41.7)		0	
<i>Early-onset mid-childhood remitting</i>	474 (6.9)		48 (5.4)		13 (2.4)		13 (3.7)		55 (5.8)	
<i>Late-onset</i>	296 (4.3)		55 (6.2)		50 (9.1)		4 (1.2)		247 (26.0)	
	<b>7-8 years</b>	<b>14-15 years</b>	<b>8 years</b>	<b>16 years</b>	<b>10 years</b>	<b>15 years</b>	<b>8 years</b>	<b>14 years</b>	<b>10 years</b>	<b>18 years</b>
Age Mean (SD) in years	8.7 (0.3)	15.4 (0.3)	7.98 (0.16)	16.09 (0.62)	10.15 (0.18)	15.09 (0.28)	7.97 (NA)	13.95 (NA)	9.98 (0.27)	17.87 (0.59)
Doctor-Diagnosed Asthma ever*	1060 (19.7)	796 (23.2)	198 (23.9)	198 (30.0)	86 (16.0)	80 (19.5)	75 (21.6)	83 (23.9)	350 (40.9)	255 (28.6)
Asthma ever	NA	NA	193 (22.8)	192 (29.5)	87 (16.2)	66 (21.9)	54 (15.6)	65 (18.7)	194 (20.9)	264 (29.3)
Current wheeze	683 (12.5)	306 (9.0)	150 (17.6)	112 (16.9)	67 (12.4)	63 (15.5)	54 (15.6)	54 (15.5)	190 (20.4)	227 (25.1)
Current asthma medication	695 (12.9)	361 (10.6)	141 (16.5)	114 (17.1)	68 (12.6)	58 (14.0)	50 (14.41)	49 (14.1)	41 (11.81)	38 (10.9)

403 \* DDA ever not available in IOW, we used Asthma cared for by hospital specialist/ GP or nurse as proxy

404

405 **Table E4.** Comparison of included vs. excluded participants in the five cohorts at different ages

406

<b>ALSPAC</b>	<b>N</b>	<b>Included</b>	<b>N</b>	<b>Excluded</b>	<b>P-value</b>					
Males (%)	6833	3492 (51.1)	4343	2269 (52.2)	0.24					
Maternal history asthma (%)	6497	748 (11.5)	4038	453 (11.2)	0.64					
Maternal smoking-pregnancy (%)	6438	1423 (22.1)	4019	1167 (29.0)	<0.001					
	<b>At 7.5-8.5 years</b>					<b>At 14-15 years</b>				
<b>ALSPAC</b>	<b>N</b>	<b>Included</b>	<b>N</b>	<b>Excluded</b>	<b>P-value</b>	<b>N</b>	<b>Included</b>	<b>N</b>	<b>Excluded</b>	<b>P-value</b>
Age Mean (SD) years	5139	8.7 (0.3)	1872	8.7 (0.3)	<0.001	3885	15.4 (0.3)	1237	15.5 (0.4)	<0.001
Current wheeze (%)	5453	683 (12.5)	2579	344 (13.3)	0.308	3419	306 (9.0)	1078	105 (9.7)	0.432
Asthma ever (%)	5377	1060 (19.7)	2605	562 (21.6)	0.053	3425	796 (23.2)	1079	279 (25.9)	0.079
Current asthma medication (%)	5379	695 (12.9)	2529	368 (14.6)	0.047	3400	361 (10.6)	1077	134 (12.4)	0.096
<b>MAAS</b>	<b>N</b>	<b>Included</b>	<b>N</b>	<b>Excluded</b>	<b>P-value</b>					
Males (%)	887	475 (53.6)	263	149 (56.7)	0.38					
Maternal history asthma (%)	886	120 (13.5)	259	45 (17.4)	0.12					
Maternal smoking* (%)	884	122 (13.8)	260	47 (18.1)	0.09					
	<b>At 8 years</b>					<b>At 16 years</b>				
<b>MAAS</b>	<b>N</b>	<b>Included</b>	<b>N</b>	<b>Excluded</b>	<b>P-value</b>	<b>N</b>	<b>Included</b>	<b>N</b>	<b>Excluded</b>	<b>P-value</b>
Age Mean (SD) years	827	7.98 (0.16)	149	8.00 (0.21)	0.31	605	16.09 (0.62)	59	15.98 (0.60)	0.20
Current wheeze (%)	853	150 (17.6)	172	35 (20.4)	0.39	664	112 (16.9)	82	15 (18.3)	0.11

Asthma ever (%)	845	193 (22.8)	173	52 (30.1)	0.043	651	192 (29.5)	79	28 (35.4)	0.28
Current asthma medication (%)	855	141 (16.5)	173	43 (24.9)	0.009	666	114 (17.1)	83	14 (16.9)	0.96

<b>SEATON</b>	<b>N</b>	<b>Included</b>	<b>N</b>	<b>Excluded</b>	<b>P-value</b>
Males (%)	548	260 (47.5)	987	525 (53.2)	0.031
Maternal history asthma (%)	548	77 (14.1)	985	161 (16.4)	0.24
Maternal smoking* (%)	548	107 (19.5)	987	276 (28.0)	<0.001

**At 10 years**

**At 15 years**

<b>SEATON</b>	<b>N</b>	<b>Included</b>	<b>N</b>	<b>Excluded</b>	<b>P-value</b>	<b>N</b>	<b>Included</b>	<b>N</b>	<b>Excluded</b>	<b>P-value</b>
Age Mean (SD) years	548	10.15 (0.18)	987	10.23 (0.16)	<0.001	545	15.09 (0.28)	916	15.11 (0.26)	0.20
Current wheeze (%)	541	67 (12.4)	376	42 (11.2)	0.58	407	63 (15.5)	310	48 (15.5)	0.99
Asthma ever (%)	537	87 (16.2)	374	53 (14.2)	0.40	409	66 (21.9)	302	85 (20.8)	0.73
Current asthma medication (%)	542	68 (12.6)	378	39 (10.3)	0.30	414	58 (14.0)	309	34 (11.0)	0.23

<b>ASHFORD</b>	<b>N</b>	<b>Included</b>	<b>N</b>	<b>Excluded</b>	<b>P-value</b>
Males (%)	348	179 (51.4)	272	153 (56.3)	0.23
Maternal history asthma (%)	348	49 (14.1)	272	38 (14.0)	0.97
Maternal smoking* (%)	348	52 (14.9)	270	61 (22.6)	0.015

**At 8 years**

**At 14 years**

<b>ASHFORD</b>	<b>N</b>	<b>Included</b>	<b>N</b>	<b>Excluded</b>	<b>P-value</b>	<b>N</b>	<b>Included</b>	<b>N</b>	<b>Excluded</b>	<b>P-value</b>
Age Mean (SD) years	348	NA	272	NA	NA	348	NA	272	NA	NA
Current wheeze (%)	347	54 (15.6)	246	25 (10.2)	0.06	348	54 (15.5)	150	18 (12.00)	0.31
Asthma ever (%)	347	54 (15.6)	246	38 (15.5)	0.97	348	65 (18.7)	150	25 (16.7)	0.59

Current asthma medication (%)	347	50 (14.41)	246	22 (8.9)	0.05	348	49 (14.1)	150	16 (10.7)	0.30
-------------------------------	-----	------------	-----	----------	------	-----	-----------	-----	-----------	------

<b>Isle Of Wight</b>	<b>N</b>	<b>Included</b>	<b>N</b>	<b>Excluded</b>	<b>P-value</b>
----------------------	----------	-----------------	----------	-----------------	----------------

Males (%)	952	466 (49.0)	508	275 (54.1)	0.06
Maternal history asthma (%)	946	106 (11.2)	505	52 (10.3)	0.60
Maternal smoking* (%)	941	217 (23.1)	502	147 (29.3)	0.01

**At 10 years**

**At 18 years**

<b>IOW</b>	<b>N</b>	<b>Included</b>	<b>N</b>	<b>Excluded</b>	<b>P-value</b>	<b>N</b>	<b>Included</b>	<b>N</b>	<b>Excluded</b>	<b>P-value</b>
Age Mean (SD) years	932	9.98 (0.27)	426	10.04 (0.31)	<0.001	914	17.87 (0.59)	389	18.14 (0.67)	<0.001
Current wheeze (%)	932	190 (20.4)	426	69 (16.2)	0.07	903	227 (25.1)	377	58 (15.4)	<0.002
Asthma ever (%)	930	194 (20.9)	425	80 (18.8)	0.39	900	264 (29.3)	385	108 (28.1)	0.64
Current asthma medication (%)	347	41 (11.81)	246	15 (6.10)	0.02	348	38 (10.9)	150	13(8.7)	0.45

407

408 **Table E5.** List of 134 independent SNPs identified after clumping and associated with at least one  
 409 wheezing phenotype ( $p < 10 \times 10^{-5}$ )

CHR	SNP	BP	short-listed after inspection of locus zoom plot
<b>Persistent Wheezing</b>			
1	rs4620530	240063821	yes
2	rs13398488	7142199	yes
2	rs77062323	53049017	no
2	rs6543291	106011626	yes
3	rs77655717	128737320	yes
4	rs77822621	1008212	yes
4	rs7680608	1050437	yes
4	rs115228498	142969757	yes
4	rs145937716	143192224	no
5	rs116494115	7736317	yes
5	rs78701483	95680422	no
6	rs138099941	7654240	no
6	rs9346404	71606613	no
6	rs143979498	151040328	no
7	rs76871421	105676144	yes
8	rs59670576	128555771	no
9	rs116933120	27458652	no
9	rs75260654	75788108	yes
9	rs116849664	75820902	yes
9	rs143481506	139515723	no
10	rs7088157	100038964	yes
11	rs112474574	5885773	yes
11	rs116861530	116962661	yes
13	rs7982350	73106322	no
13	rs17461573	106711373	no
14	rs1105683	56213787	yes
15	rs2202714	49811991	yes
15	rs117540214	84338642	yes
17	rs17676191	37949924	yes
17	rs79026872	37965932	yes
17	rs4795400	38067020	yes
17	rs1031460	38072247	yes
17	rs56199421	38090808	yes
17	rs4795406	38100134	yes
17	rs72832972	38110575	yes
17	rs4794821	38124203	yes
17	rs59843584	38124892	yes
18	rs111812993	30353181	no
19	rs4804311	8615589	yes
19	rs2013694	8616392	yes
19	rs73501545	8620823	yes
19	rs111644945	8625081	yes
22	rs5994170	17615213	yes

22 rs34902370 17632194 yes

**Early-onset Remitting Wheezing**

1 rs12730098 212427488 yes  
1 rs75639566 233019116 no  
2 rs2880066 17107219 yes  
2 rs10180268 17126699 yes  
3 rs115031796 86691640 no  
3 rs3861377 173317378 yes  
3 rs10513743 176022304 yes  
5 rs10075253 75548246 yes  
5 rs12520884 84406634 no  
6 rs117477297 92565052 no  
6 rs2453395 166286532 yes  
7 rs56027869 50072919 no  
7 rs4730561 78531705 yes  
7 rs73144976 78586112 yes  
7 rs67259321 78686582 yes  
7 rs146771277 154438861 no  
9 rs10758259 34392908 yes  
11 rs7128994 71242209 no  
11 rs72994149 106837223 yes  
12 rs117367256 93508478 no  
13 rs2872948 57442480 yes  
13 rs73527654 57447994 yes  
13 rs2151504 82291577 no  
15 rs116966886 47043587 yes  
15 rs117565527 47342882 yes

**Mid-childhood onset Remitting Wheezing**

1 rs35725789 159207367 yes  
1 rs146141555 159227423 yes  
1 rs146575092 159374228 yes  
1 rs140877050 220848829 no  
1 rs72745905 223451086 no  
2 rs7595553 36127878 yes  
2 rs145007503 50688324 no  
2 rs6546068 64583398 no  
2 rs17387431 206651315 no  
2 rs144791928 236963432 no  
3 rs34315999 8969653 yes  
3 rs115245770 99209128 no  
3 rs146961758 194285978 yes  
4 rs138794367 103859545 yes  
5 rs115719402 77538102 yes  
6 rs76026399 47531792 no  
7 rs73172838 154842348 no  
8 rs112631708 134500083 no  
9 rs72752356 98094970 no  
13 rs113195384 46333770 no  
13 rs9602218 84139813 yes

13	rs61960366	84144202	yes
13	rs74589927	84208697	yes
13	rs2210726	84492936	yes
13	rs4390476	84598570	yes
14	rs117443464	73460284	yes
16	rs72820814	81916262	no
17	rs190526697	12274299	no
18	rs75286534	26206826	no
18	rs138888086	63591085	no
18	rs76551535	71879807	no
19	rs77496444	19192132	no
20	rs6077514	9302948	yes

#### Late-onset Wheezing

1	rs9439669	18859049	yes
1	rs2051039	57067560	yes
1	rs72673642	80727443	yes
2	rs147557117	19778063	no
2	rs140983998	111402871	yes
2	rs117617447	123387601	no
2	rs13025116	127505482	no
2	rs148008098	128633620	yes
3	rs4072729	24780393	yes
3	rs143960666	31227943	no
3	rs4677102	72193991	no
3	rs145629570	113422516	yes
3	rs113643470	141728174	yes
4	rs17472015	48467594	yes
7	rs117660982	149438923	yes
7	rs118027705	150456728	yes
7	rs139489493	150481499	yes
7	rs144271668	157934780	yes
8	rs990182	89976447	yes
9	rs79110962	14432953	yes
10	rs9325460	82492323	no
10	rs7896106	91196402	yes
10	rs115465993	109372900	no
11	rs16935643	41395746	no
11	rs141958628	119083284	yes
14	rs113363660	69410278	no
15	rs139134265	44923960	yes
15	rs143862030	84922146	yes
16	rs113390367	1118849	yes
16	rs4788025	28003221	yes
18	rs72918264	51009510	no
22	rs133498	48913809	yes

410

411



412 **Table E6.** References to previous relevant associated traits for each region/gene identified by wheezing phenotype

Early Onset Persistent Wheezing			
Gene(s)	Locus	Previous Associated Trait	Reference or Source
<i>CHRM3</i>	1q43	<b>FEV<sub>1</sub>, FEV<sub>1</sub>/FVC, asthma</b> - high priority drug target	Patel, K.R. <i>et al.</i> Targeting acetylcholine receptor M3 prevents the progression of airway hyperreactivity in a mouse model of childhood asthma. <i>FASEB J</i> <b>31</b> , 4335-4346 (2017).
<i>RNF144A</i>	2p25.1	<b>asthma</b> , allergy, childhood onset <b>asthma</b> , <b>allergic rhinitis</b>	Schoettler, N. <i>et al.</i> Advances in asthma and allergic disease genetics: Is bigger always better? <i>J Allergy Clin Immunol</i> <b>144</b> , 1495-1506 (2019).
<i>FHL2</i>	2q12.2	<b>bronchial hyper-responsiveness</b> , airway inflammation, novel gene associated with <b>asthma</b> severity in human	Kurakula, K. <i>et al.</i> Deficiency of FHL2 attenuates airway inflammation in mice and genetic variation associates with human bronchial hyper-responsiveness. <i>Allergy</i> <b>70</b> , 1531-44 (2015).
<i>RAB7A</i>	3q21.3	<b>eosinophil count</b>	GeneCards
<i>RAB43</i>	3q21.3	response to <b>bronchodilator</b> , <b>FEV<sub>1</sub>/FEC ratio</b>	GWAS Catalog
<i>RNF212, IDUA, DGKQ, SLC26A1</i>	4p16.3	<b>asthma</b>	Gautam, Y. <i>et al.</i> Comprehensive functional annotation of susceptibility variants associated with asthma. <i>Hum Genet</i> <b>139</b> , 1037-1053 (2020).
<i>INPP4B</i>	4q31.21	atopic <b>asthma</b>	Sharma, M. <i>et al.</i> A genetic variation in inositol polyphosphate 4 phosphatase a enhances susceptibility to asthma. <i>Am J Respir Crit Care Med</i> <b>177</b> , 712-9 (2008).
<i>ADCY2</i>	5p15.31	<b>asthma</b> × air pollution, childhood <b>asthma</b>	Gref, A. <i>et al.</i> Genome-Wide Interaction Analysis of Air Pollution Exposure and Childhood Asthma with Functional Follow-up. <i>Am J Respir Crit Care Med</i> <b>195</b> , 1373-1383 (2017).
<i>CDHR3</i>	7q22.3	childhood <b>asthma</b>	Everman, J.L. <i>et al.</i> Functional genomics of CDHR3 confirms its role in HRV-C infection and childhood asthma exacerbations. <i>J Allergy Clin Immunol</i> <b>144</b> , 962-971 (2019).
<i>ANXA1</i>	9q21.13	<b>FEV<sub>1</sub>/FVC</b> , response to <b>bronchodilators</b> in smokers	Lutz, S.M. <i>et al.</i> A genome-wide association study identifies risk loci for spirometric measures among smokers of European and African ancestry. <i>BMC Genet</i> <b>16</b> , 138 (2015).
<i>ANXA1</i>	9q21.13	anti-inflammatory properties, strongly expressed in <b>bronchial mast cells</b>	Vieira Braga FA <i>et al.</i> A cellular census of human lungs identifies novel cell states in health and in asthma. (2019).

<i>ANXA1</i>	9q21.13	potentially involved in epithelial <b>airway repair</b>	Leoni, G. <i>et al.</i> Annexin A1-containing extracellular vesicles and polymeric nanoparticles promote epithelial wound repair. <i>J Clin Invest</i> <b>125</b> , 1215-27 (2015).
<i>R3HCC1L</i>	10q24.2	atopic <b>eczema</b> , psoriasis	GWAS Catalog
<i>R3HCC1L</i>	10q24.2	<b>eosinophil count</b> , BMI	GeneCards
<i>TRIM5, TRIM6, TRIM22</i>	11p15.4	severe <b>asthma</b> and insulin resistance	Kimura, T. <i>et al.</i> Precision autophagy directed by receptor regulators - emerging examples within the TRIM family. <i>J Cell Sci</i> <b>129</b> , 881-91 (2016).
<i>SIK3</i>	11q23.3	triglycerides, glucose metabolism, <b>eosinophil count</b>	Sun, Z. <i>et al.</i> The potent roles of salt-inducible kinases (SIKs) in metabolic homeostasis and tumorigenesis. <i>Sig Transduct Target Ther</i> <b>5</b> (2020).
<i>KTN1</i>	14q22.1	severe <b>asthma</b>	Bigler, J. <i>et al.</i> A Severe Asthma Disease Signature from Gene Expression Profiling of Peripheral Blood from U-BIOPRED Cohorts. <i>Am J Respir Crit Care Med</i> <b>195</b> , 1311-1320 (2017).
<i>FAM227B</i>	5q13.3	rs35251997 and <b>FEV<sub>1</sub>, FEV<sub>1</sub>/FVC</b>	Shrine, N. <i>et al.</i> New genetic signals for lung function highlight pathways and chronic obstructive pulmonary disease associations across multiple ancestries. <i>Nat Genet</i> <b>51</b> , 481-493 (2019).
<i>ADAMTSL3</i>	15q25.2	<b>FEV<sub>1</sub>/FVC</b>	Sakornsakolpat, P. <i>et al.</i> Genetic landscape of chronic obstructive pulmonary disease identifies heterogeneous cell-type and phenotype associations. <i>Nat Genet</i> <b>51</b> , 494-505 (2019).
<i>IKZF3, GSDMB, LRRC3C, GSDMA</i>	17q12	early-onset <b>asthma, persistent wheezing</b> (chr17q12-q21)	Granell R <i>et al.</i> Examination of the relationship between variation at 17q21 and childhood wheeze phenotypes. <i>J Allergy Clin Immunol.</i> 2013 Mar;131(3):685-94.
<i>MARCH2, HNRNPM, MYO1F</i>	19p13.2	triglycerides, HDL-cholesterol, metabolic syndrome	Sajuthi, S.P. <i>et al.</i> Genetic regulation of adipose tissue transcript expression is involved in modulating serum triglyceride and HDL-cholesterol. <i>Gene</i> <b>632</b> , 50-58 (2017).
<i>MYO1F</i>	19p13.2	<b>FEV<sub>1</sub> and FVC</b>	GeneCards
<i>CECR5</i>	22q11.1	triglycerides, <b>eosinophil count</b> and body height	Liu, D.J. <i>et al.</i> Exome-wide association study of plasma lipids in >300,000 individuals. <i>Nat Genet</i> <b>49</b> , 1758-1766 (2017).

Early-onset Pre-school Remitting Wheezing			
Gene(s)	Locus	Previous Associated Trait	Reference or source
<i>PPP2R5A</i>	1q32.3	<b>waist circumference &amp; obesity</b>	Kim, H.J. <i>et al.</i> Combined linkage and association analyses identify a novel locus for obesity near PROX1 in Asians. <i>Obesity (Silver Spring)</i> <b>21</b> , 2405-12 (2013).
<i>FAM49A</i> or <i>CYRIA</i>	2p24.2	<b>airway repair in non-atopic asthma</b>	Hoang, T.T. <i>et al.</i> Epigenome-wide association study of DNA methylation and adult asthma in the Agricultural Lung Health Study. <i>Eur Respir J</i> <b>56</b> (2020).
<i>NLGN1</i>	3q26.31	<b>smoking</b>	Drgon, T. <i>et al.</i> Genome-wide association for nicotine dependence and smoking cessation success in NIH research volunteers. <i>Mol Med</i> <b>15</b> , 21-7 (2009).
<i>NAALADL2</i>	3q26.31	<i>suggestive association with severe asthma exacerbations</i>	Herrera-Luis E <i>et al.</i> Genome-wide association study reveals a novel locus for asthma with severe exacerbations in diverse populations. <i>Pediatr Allergy Immunol.</i> 2021;32(1):106-115.
<i>SV2C</i>	5q13.3	<b>BMI, diastolic blood pressure</b>	GeneCards
<i>PDE10A</i>	6q27	Birthweight, <b>asthma</b> and <b>BMI</b>	Melen, E. <i>et al.</i> Analyses of shared genetic factors between asthma and obesity in children. <i>J Allergy Clin Immunol</i> <b>126</b> , 631-7 e1-8 (2010).
<i>MAGI2</i>	7q21.11	<b>allergic diseases &amp; atopy</b>	Freidin, M.B. <i>et al.</i> [Genome-wide association study of allergic diseases in Russians of Western Siberia]. <i>Mol Biol (Mosk)</i> <b>45</b> , 464-72 (2011).
<i>MAGI2</i>	7q21.11	<b>smoking</b>	Quach, B.C. <i>et al.</i> Expanding the genetic architecture of nicotine dependence and its shared genetics with multiple traits. <i>Nat Commun</i> <b>11</b> , 5562 (2020).
<i>MAGI2</i>	7q21.11	<b>BMI</b>	GeneCards
<i>MAGI2</i>	7q21.11	<b>airway wall thickness</b>	GWAS Catalog
<i>C9orf24</i>	9p13.3	<b>airway repair</b>	Yoshisue, H. <i>et al.</i> Characterization of ciliated bronchial epithelium 1, a ciliated cell-associated gene induced during mucociliary differentiation. <i>Am J Respir Cell Mol Biol</i> <b>31</b> , 491-500 (2004).
<i>GUCY1A2</i>	11q22.3	<b>systolic/diastolic blood pressure</b>	GeneCards

PRR20A/B/C/D/E	13q21.1	systolic <b>blood pressure</b>	GeneCards
SEMA6D	15q21.1	<b>smoking</b>	Minica, C.C. et al. Pathways to smoking behaviours: biological insights from the Tobacco and Genetics Consortium meta-analysis. <i>Mol Psychiatry</i> 22, 82-88 (2017).

415

<b>Early-onset Mid-childhood Remitting Wheezing</b>			
<b>Gene(s)</b>	<b>Locus</b>	<b>Previous Associated Trait</b>	<b>Reference</b>
<i>CADM3, FCER1A, MPTX1, OR10J1</i>	1q23.2	<b>neutrophil count, CRP</b>	Barreto, M. et al. Duffy phenotyping and FY*B-67T/C genotyping as screening test for benign constitutional neutropenia. <i>Hematol Transfus Cell Ther</i> (2020).
<i>MRPL50P1</i>	2p22.3	<b>PM 2.5</b> exposure level and global <b>DNA methylation</b> level	Liu, J. et al. Genetic variants, PM2.5 exposure level and global DNA methylation level: A multi-center population-based study in Chinese. <i>Toxicol Lett</i> <b>269</b> , 77-82 (2017).
<i>RAD18</i>	3p25.3	<b>atopy/SPT</b>	Bouzigon, E. et al. Meta-analysis of 20 genome-wide linkage studies evidenced new regions linked to asthma and atopy. <i>Eur J Hum Genet</i> <b>18</b> , 700-6 (2010).
<i>MRPL50P1</i>	3q29	<b>BMI</b>	Kettunen, J. et al. Multicenter dizygotic twin cohort study confirms two linkage susceptibility loci for body mass index at 3q29 and 7q36 and identifies three further potential novel loci. <i>Int J Obes (Lond)</i> <b>33</b> , 1235-42 (2009).
<i>LSG1</i>	3q29	<b>BMI, eosinophil &amp; neutrophil count</b>	GeneCards
<i>TMEM44-AS1, TMEM44, ATP13A3</i>	3q29	diastolic blood pressure	GeneCards
<i>SLC9B1</i>	4q24	<b>eosinophil count</b>	Aschard, H. et al. Sex-specific effect of IL9 polymorphisms on lung function and polysensitization. <i>Genes Immun</i> <b>10</b> , 559-65 (2009).

<i>SLC9B1</i>	4q24	<b>allergic rhinitis</b>	Haagerup, A. <i>et al.</i> Allergic rhinitis--a total genome-scan for susceptibility genes suggests a locus on chromosome 4q24-q27. <i>Eur J Hum Genet</i> <b>9</b> , 945-52 (2001).
<i>AP3B1</i>	5q14.1	<b>vital capacity, BMI</b>	GeneCards, GWAS Catalog
<i>RNU6-67P</i> , <i>SLITRK1</i> , <i>VENTXP2</i> , <i>UBE2D3P4</i> , <i>MTND4P1</i>	13q31.1	RNU6-67P/ rs976078: <b>food allergy</b>	Liu, X. <i>et al.</i> Genome-wide association study of maternal genetic effects and parent-of-origin effects on food allergy. <i>Medicine (Baltimore)</i> <b>97</b> , e0043 (2018).
<i>ZFYVE1</i>	14q24.2	LDL <b>cholesterol</b> and systolic blood pressure	GWAS Catalog
<i>PLCB4</i>	20p12.3-p12.2	<b>neutrophil count</b>	Okada, Y. <i>et al.</i> Common variations in PSMD3-CSF3 and PLCB4 are associated with neutrophil count. <i>Hum Mol Genet</i> <b>19</b> , 2079-85 (2010).

416

<b>Late-onset Wheezing</b>			
<b>Gene(s)</b>	<b>Locus</b>	<b>Previous Associated Trait</b>	<b>Reference</b>
<i>KLHDC7A</i>	1p36.13	1p36.13: <b>metabolic syndrome</b>	Hoffmann, K. <i>et al.</i> A German genome-wide linkage scan for type 2 diabetes supports the existence of a metabolic syndrome locus on chromosome 1p36.13 and a type 2 diabetes locus on chromosome 16p12.2. <i>Diabetologia</i> <b>50</b> , 1418-22 (2007).
<i>PPAP2B</i> , <i>PRKAA2</i>	1p32.2	PRKAA2: lymphocyte count and <b>asthma</b> susceptibility	Cusanovich, D.A. <i>et al.</i> The combination of a genome-wide association study of lymphocyte count and analysis of gene expression data reveals novel asthma candidate genes. <i>Hum Mol Genet</i> <b>21</b> , 2111-23 (2012).
<i>HMGB1P18</i>	1p31.1	<b>smoking, BMI</b>	GeneCards
<i>ACOXL</i> , <i>BUB1</i>	2q13	ACOXL: <b>later onset asthma and obesity</b>	Zhu, Z. <i>et al.</i> Shared genetic and experimental links between obesity-related traits and asthma subtypes in UK Biobank. <i>J Allergy Clin Immunol</i> <b>145</b> , 537-549 (2020).

AMMECR1L	2q14.3	<b>body height</b> , blood protein, growth, bone, and heart alterations	Moyses-Oliveira, M. <i>et al.</i> Inactivation of AMMECR1 is associated with growth, bone, and heart alterations. <i>Hum Mutat</i> <b>39</b> , 281-291 (2018).
RARB	3p24.2	<b>FEV<sub>1</sub>/FVC</b> , adult lung function	Collins, S.A. <i>et al.</i> HHIP, HDAC4, NCR3 and RARB polymorphisms affect fetal, childhood and adult lung function. <i>Eur Respir J</i> <b>41</b> , 756-7 (2013).
KIAA2018, NAA50, SIDT1, CD200	3q13.2	SIDT1: <b>FEV<sub>1</sub>/FVC</b> , CD200: <b>adult-onset non-allergic asthma</b>	Siroux, V. <i>et al.</i> Genetic heterogeneity of asthma phenotypes identified by a clustering approach. <i>Eur Respir J</i> <b>43</b> , 439-52 (2014).
TFDP2, XRN1	3q23	XRN1: <b>eosinophil count</b> , 3q23: <b>allergic disease and atopic sensitisation</b>	Freidin, M.B. <i>et al.</i> [Genome-wide association study of allergic diseases in Russians of Western Siberia]. <i>Mol Biol (Mosk)</i> <b>45</b> , 464-72 (2011).
SLAIN2, SLC10A4, FRYL	4p11	FRYL: body height, age at menopause	GeneCards
KRBA1, ZNF467	7q36.1	systolic blood pressure	GWAS Catalog
GIMAP family, AOC1, LOC105375566	7q36.1	AOC1: CV disease, <b>smoking</b> , GIMAP family: autoimmune diabetes, <b>asthma and allergy</b>	Heinonen, M.T. <i>et al.</i> GIMAP GTPase family genes: potential modifiers in autoimmune diabetes, asthma, and allergy. <i>J Immunol</i> <b>194</b> , 5885-94 (2015).
PTPRN2	7q36.3	<b>eczema</b>	Bogari, N.M. <i>et al.</i> Whole exome sequencing detects novel variants in Saudi children diagnosed with eczema. <i>J Infect Public Health</i> <b>13</b> , 27-33 (2020).
LOC105375631	8q21.3	8q21.3: <b>type 1 diabetes</b>	Mukhopadhyay, N., Noble, J.A., Govil, M., Marazita, M.L. & Greenberg, D.A. Identifying genetic risk loci for diabetic complications and showing evidence for heterogeneity of type 1 diabetes based on complications risk. <i>PLoS One</i> <b>13</b> , e0192696 (2018).
NFIB, ZDHHC21	9p22.3	9p22.3: <b>asthma</b> (mean age<16 years)	Denham, S. <i>et al.</i> Meta-analysis of genome-wide linkage studies of asthma and related traits. <i>Respir Res</i> <b>9</b> , 38 (2008).
SLC16A12, IFIT family, PANK1	10q23.31	SLC16A12: Body height, PANK1: <b>insulin resistance</b>	Yang, L. <i>et al.</i> P53/PANK1/miR-107 signalling pathway spans the gap between metabolic reprogramming and insulin resistance induced by high-fat diet. <i>J Cell Mol Med</i> <b>24</b> , 3611-3624 (2020).

<i>CBL, CCDC84, MCAM</i>	11q23.3	CBL: <b>Sex hormone-binding globulin levels</b> ; MCAM: Blood protein levels	GWAS Catalog
<i>SPG11, CTDSPL2</i>	15q15.3-q21.1	CTDSPL2: <b>alcohol drinking</b>	GWAS Catalog
<i>ADAMTSL3, GOLGA6L4, UBE2Q2P8</i>	15q25.2	ADAMTSL3: <b>FEV<sub>1</sub>/FVC</b> , lean mass	Karasik, D. <i>et al.</i> Disentangling the genetics of lean mass. <i>Am J Clin Nutr</i> <b>109</b> , 276-287 (2019).
<i>SSTR5-AS1, CACNA1H</i>	16p13.3	CACNA1H: <b>eosinophil count</b>	GWAS Catalog
<i>GSG1L</i>	16p12.1	16p12.1: current <b>asthma</b> and rhino-conjunctivitis at 10-15 years	Sottile, G. <i>et al.</i> An association analysis to identify genetic variants linked to asthma and rhino-conjunctivitis in a cohort of Sicilian children. <i>Ital J Pediatr</i> <b>45</b> , 16 (2019).
<i>FAM19A5 or TAF5</i>	22q13.32	<b>Obesity and Metabolic Dysfunction</b>	Recinella L. <i>et al.</i> Adipokines: New Potential Therapeutic Target for Obesity and Metabolic, Rheumatic, and Cardiovascular Diseases. <i>Front Physiol.</i> 2020 Oct 30;11:578966

417

418 **Table E7.** SNPs near *ANXA1* associated with persistent wheeze

Chr	Rsid	position	A1	A2	freqA2	beta	se	P value	Direction (3 GWAS)
9	rs75260654	75788108	t	c	0.02	0.90	0.16	6.66e-09	---
9	rs116849664	75820902	t	c	0.02	0.89	0.16	1.99e-08	---
9	rs78320984	75844302	t	g	0.02	0.81	0.15	6.41e-08	---

425 A1 is the effect allele, A2 is the reference allele.

426 **Table E8.** Allele Frequencies of *rs75260654* across different wheeze phenotypes

Phenotype	CC	CT	TT
Never/infrequent	5641 (97.2)	161 (2.8)	0 (0)
Early-onset pre-school remitting	1409 (97.1)	42 (2.9)	0 (0)
Early-onset mid-childhood remitting	572 (96.1)	23 (3.9)	0 (0)
Late-onset	613 (95.2)	31 (4.8)	0 (0)
Early-onset persistent	867 (94.2)	51 (5.5)	2 (0.2)

427 **Table E9.** Selected immune eQTLs of *rs75260654*

Rsid	P value	beta	se	an	symbol	Study
rs75260654	0.014	-0.65	0.26	382	<i>ANXA1</i>	Quach_2016_monocyte_R848
rs75260654	0.015	-1.02	0.41	396	<i>ANXA1</i>	Quach_2016_monocyte_IAV

428 **Table E10.** Lung eQTLs of *rs75260654*

Rsid	P value	beta	se	an	symbol	Study
rs116849664	0.0489	0.22	0.11	620	<i>ANXA1</i>	GTEx_exon_lung
rs78320984	0.0489	0.22	0.11	620	<i>ANXA1</i>	GTEx_exon_lung

429



430 **Table E11:** Functional enrichment for ANXA1: Top 10 GO terms

Term name	Description	FDR value
GO.0007186	G protein-coupled receptor signaling pathway	$3.57 \times 10^{-18}$
GO.0006954	inflammatory response	$1.13 \times 10^{-16}$
GO.0006874	cellular calcium ion homeostasis	$5.55 \times 10^{-15}$
GO.0007204	positive regulation of cytosolic calcium ion concentration	$1.65 \times 10^{-14}$
GO.0060326	cell chemotaxis	$1.95 \times 10^{-14}$
GO.0006955	immune response	$4.23 \times 10^{-14}$
GO.0006935	Chemotaxis	$4.93 \times 10^{-14}$
GO.0006952	defense response	$1.68 \times 10^{-13}$
GO.0050801	ion homeostasis	$2.23 \times 10^{-13}$
GO.0002376	immune system process	$2.87 \times 10^{-13}$

431  
432 **Table E12.** Replication of associations between SNPs downstream of ANXA1 and early-onset  
433 persistent wheezing in PIAMA

Rsid	chr:position	R2	A2/freqA2	CW (40) vs NI (1557)		
				beta	se	p-value
rs75260654	9:75788108	0.60	c/0.02	-0.287	0.91	0.75 <sup>434</sup> 435 436
rs116849664	9:75820902	0.61	c/0.02	0.119	1.08	0.94 <sup>437</sup> 438
rs78320984	9:75844302	0.59	g/0.02	0.125	1.04	0.96 <sup>439</sup> 440

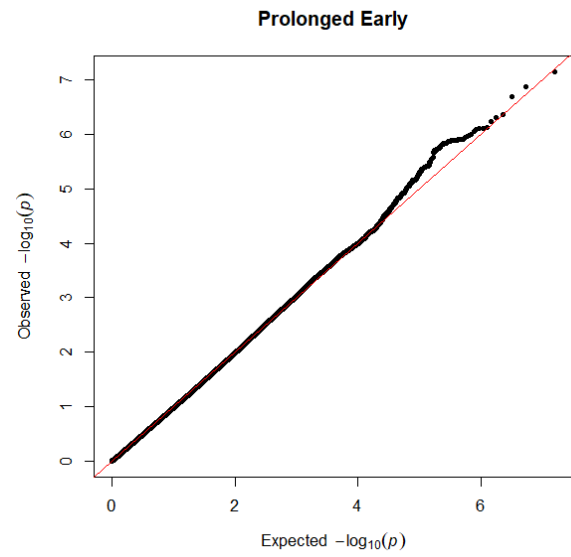
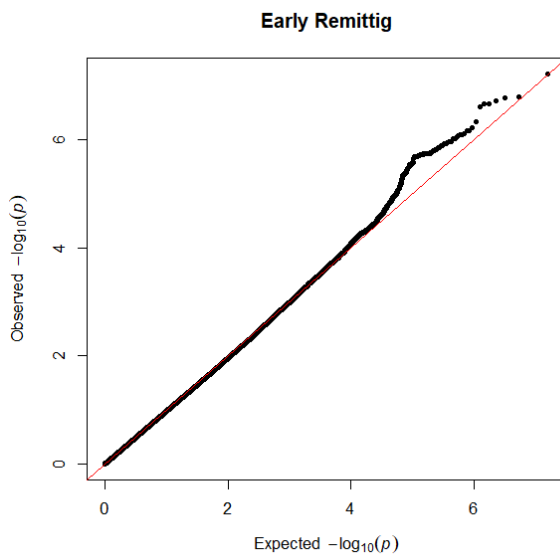
441 A1=T is the effect allele, A2 is the reference allele. CW=continuous wheezing, IR=intermediate wheezing  
442 derived from LCA 1 to 12 years in PIAMA.

443

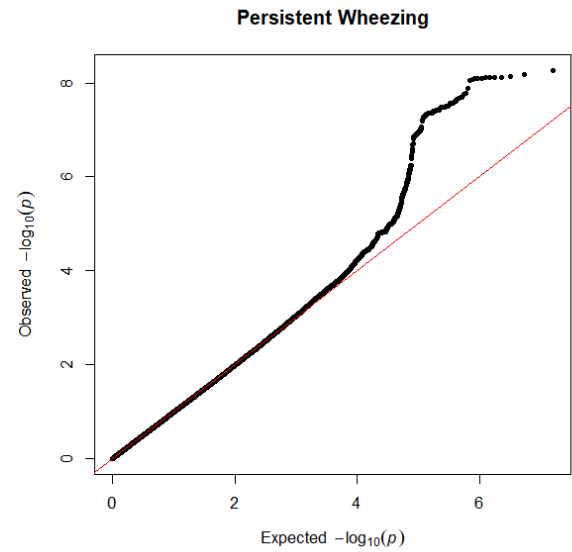
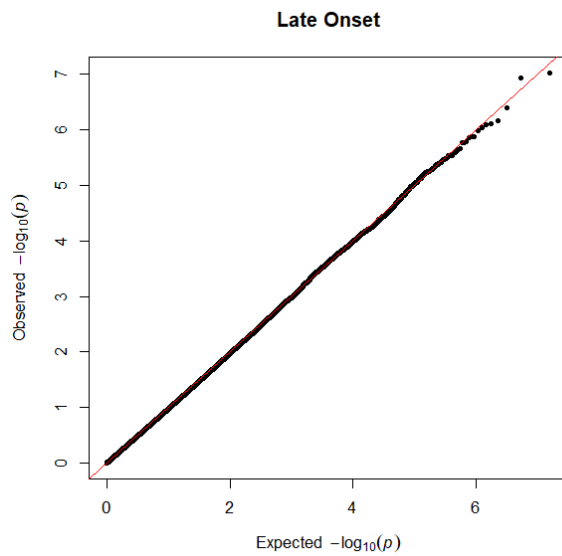
444 **SUPPLEMENTARY FIGURES**

445 **Figure E1.** QQ-plots for each wheezing phenotype

446

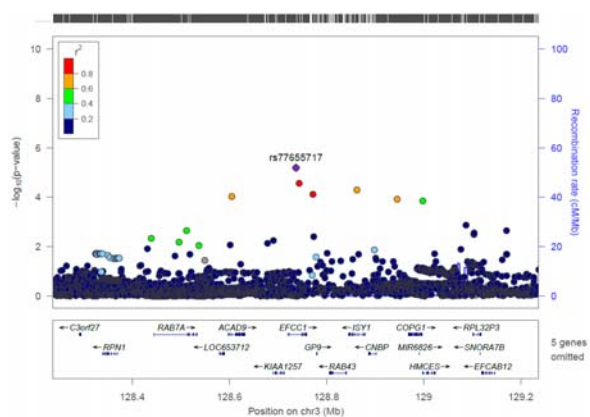
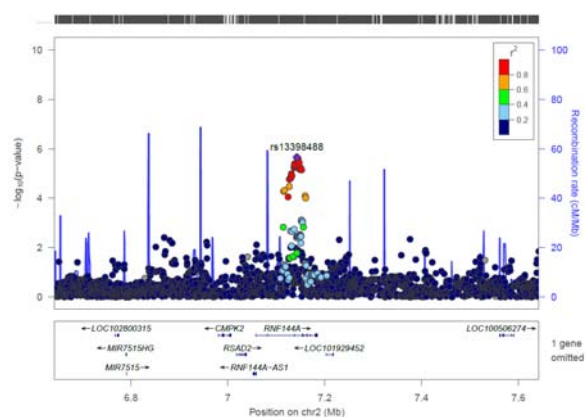
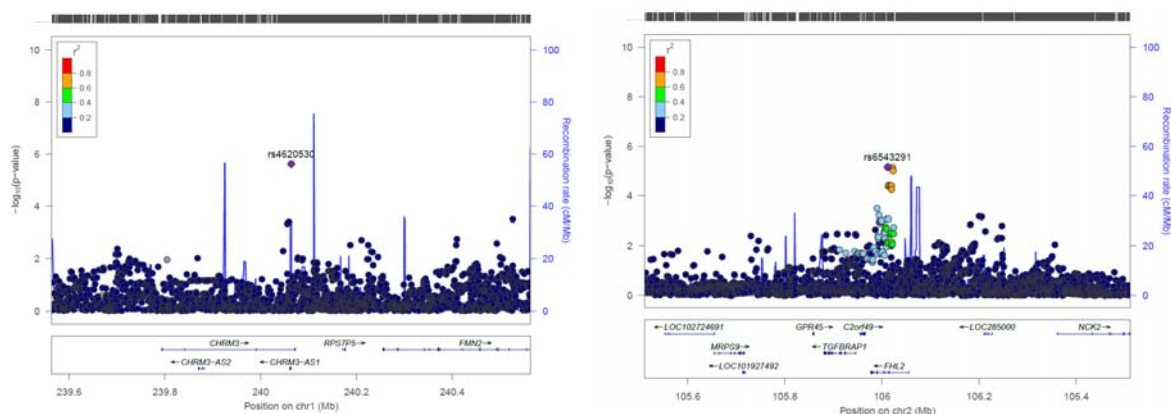


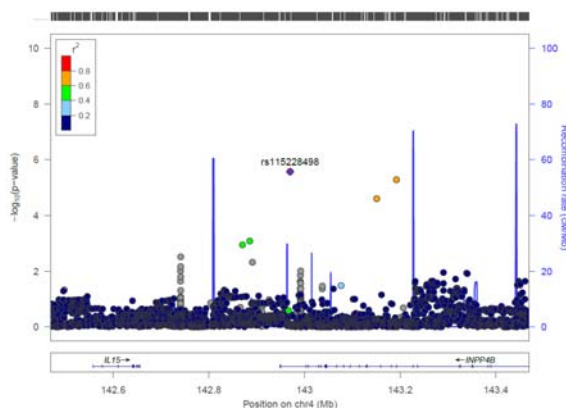
447



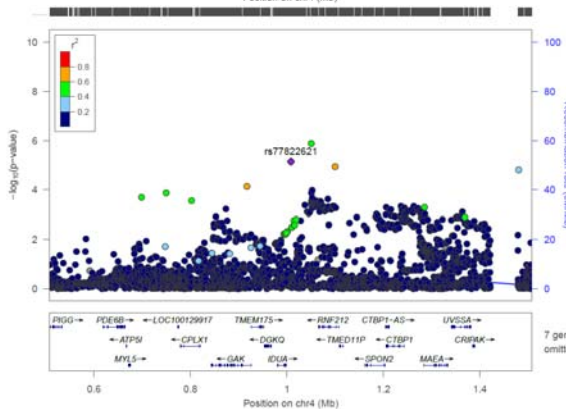
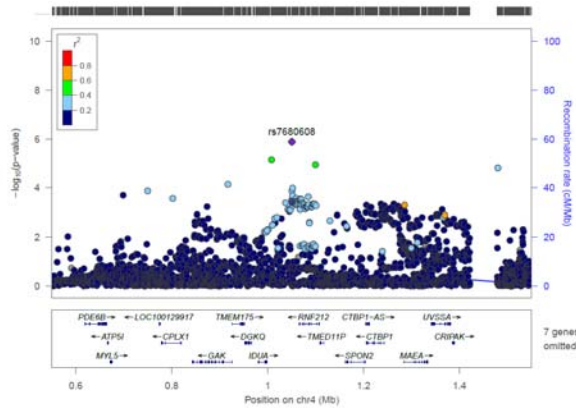
448

449 **Figure E2.** Zoom locus plots for short-listed independent top hits for Persistent Wheezing

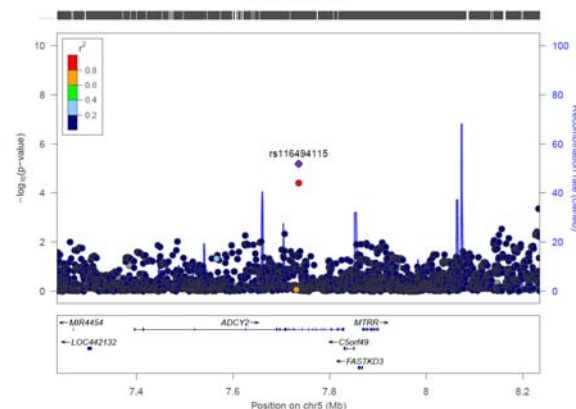


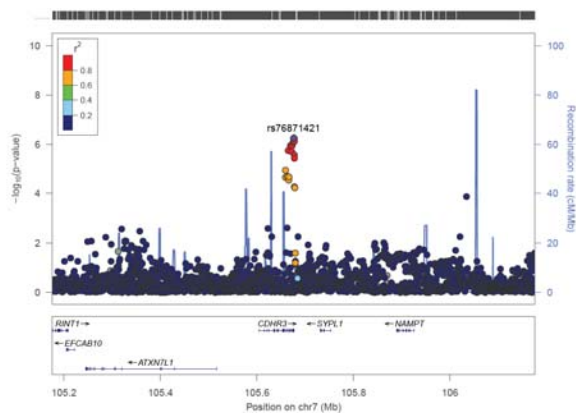


453

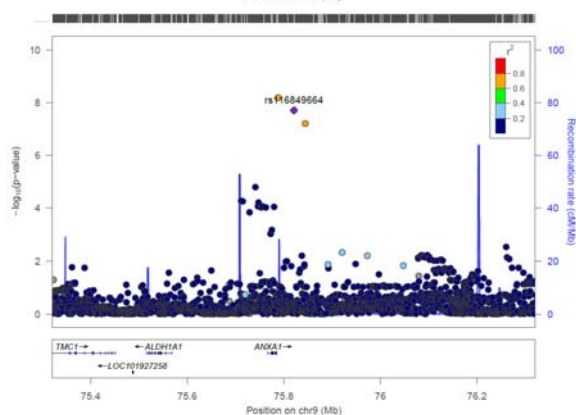
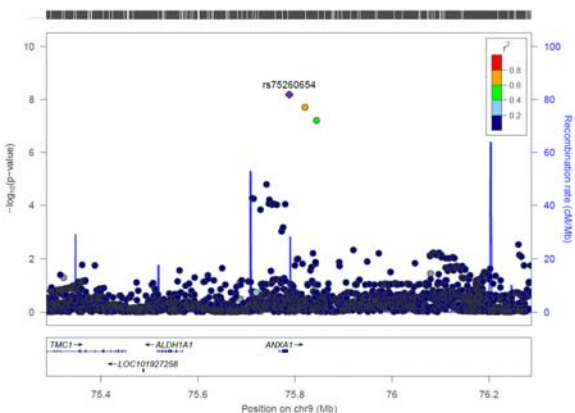


454

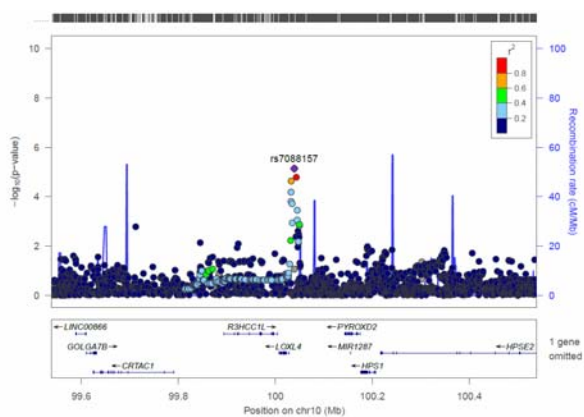




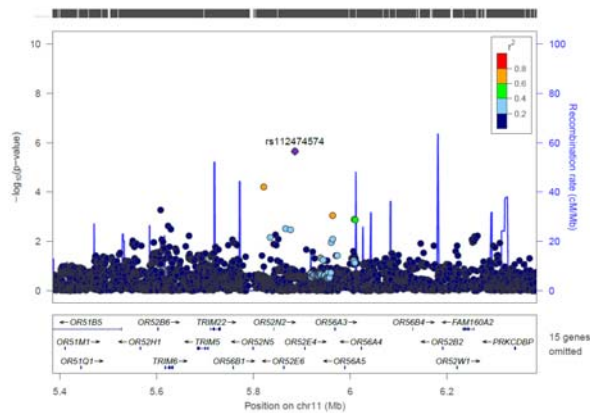
455



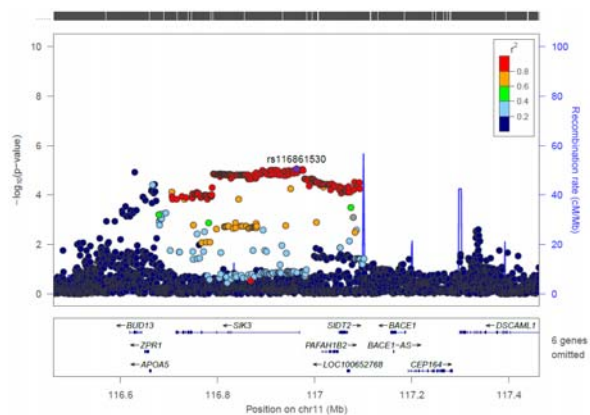
456



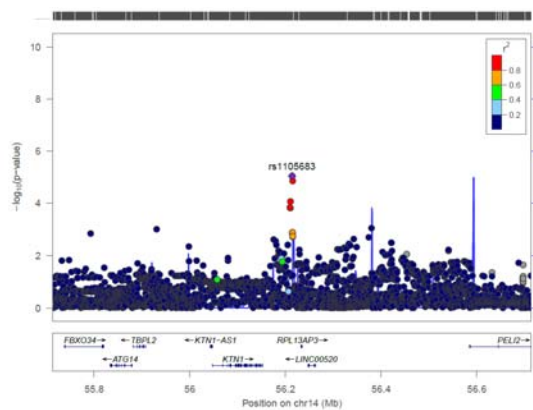
457



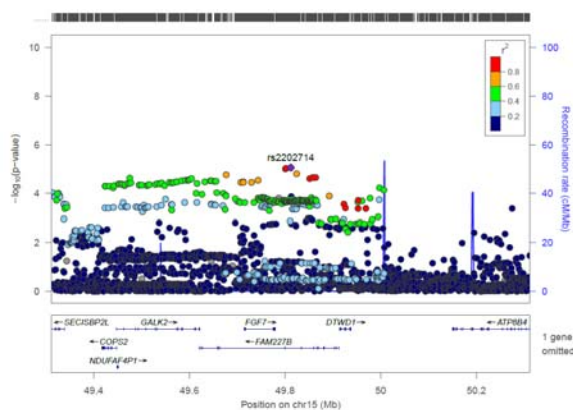
458



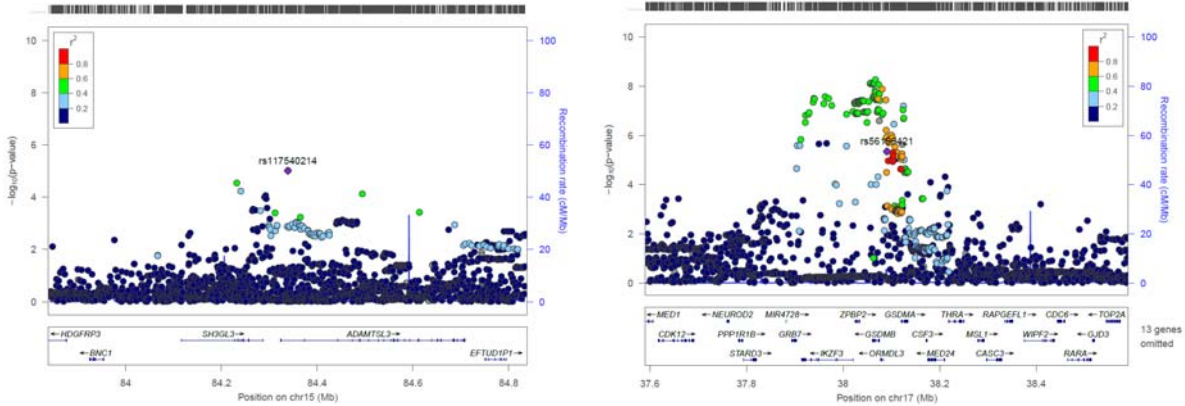
459



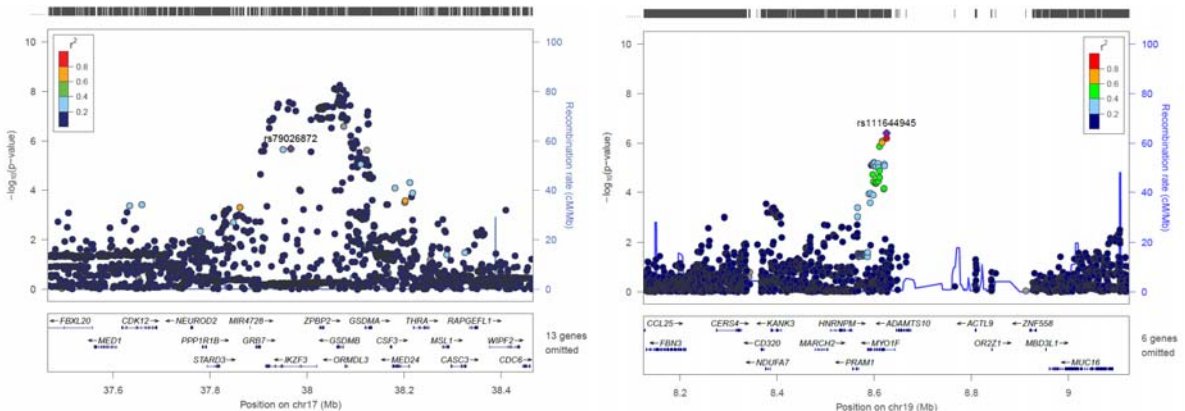
460



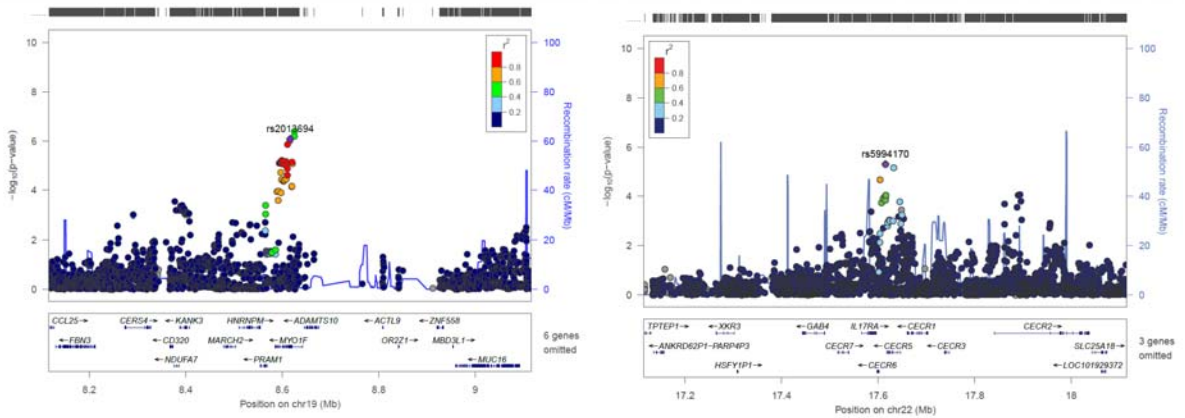
461



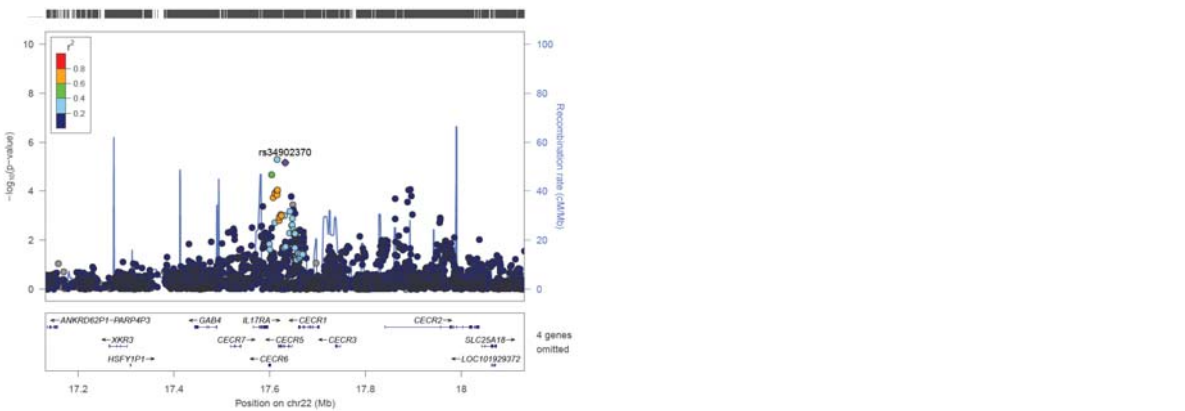
462



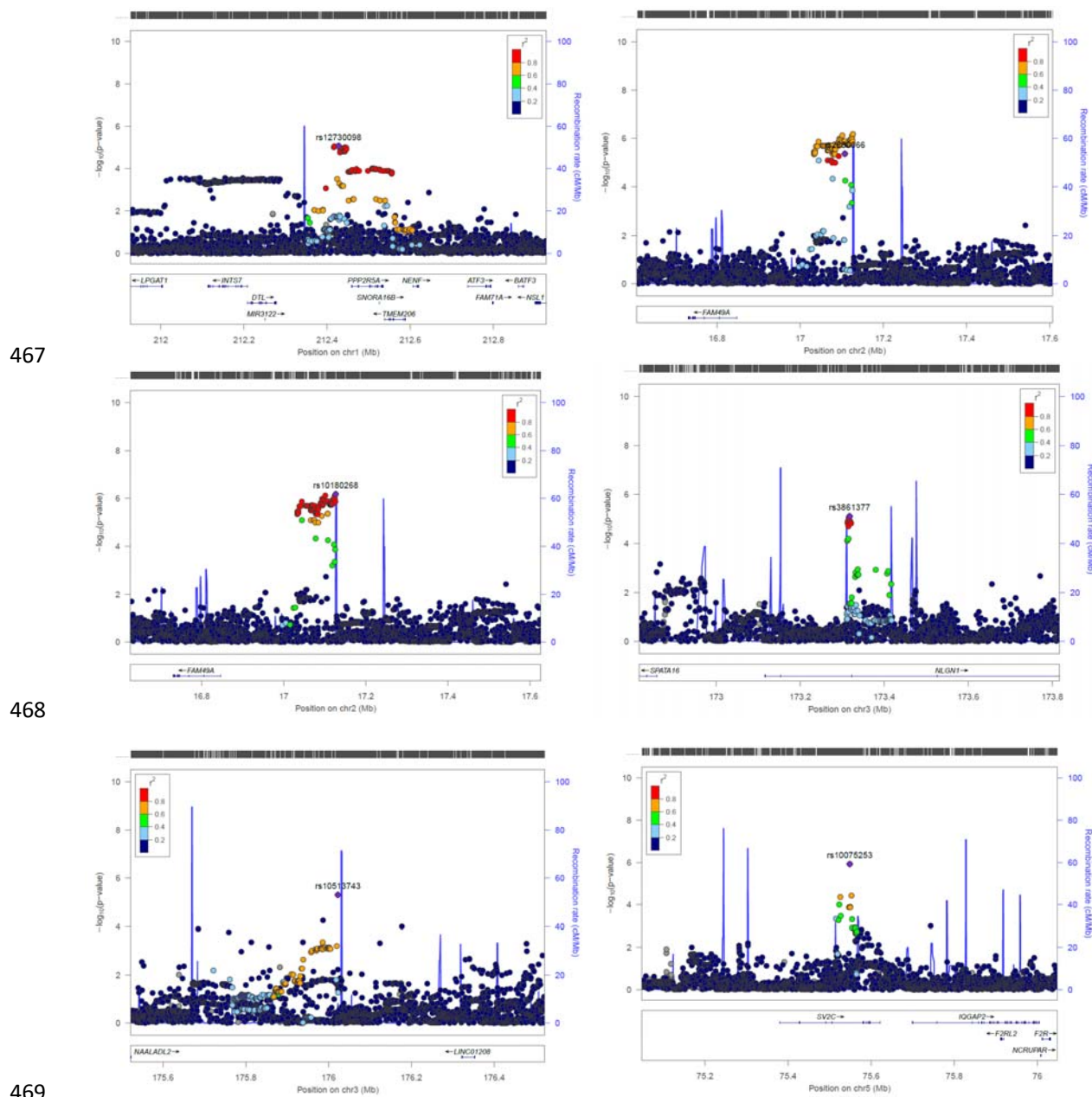
463



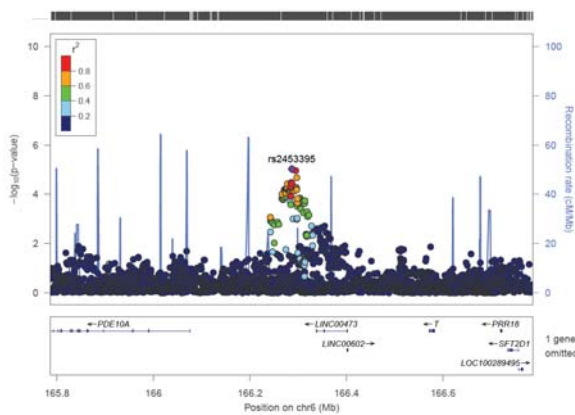
464



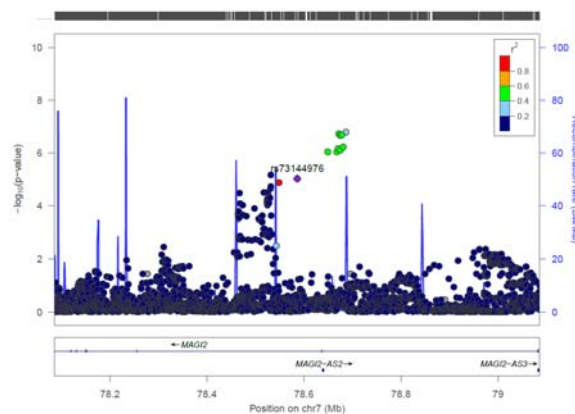
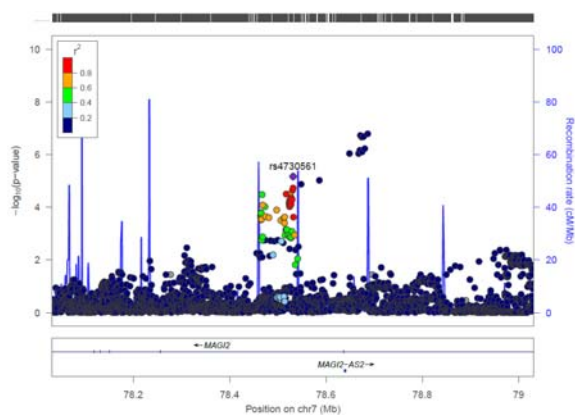
465 **Figure E3.** Zoom locus plots for short-listed independent top hits for Early-onset Pre-school  
 466 Remitting Wheezing



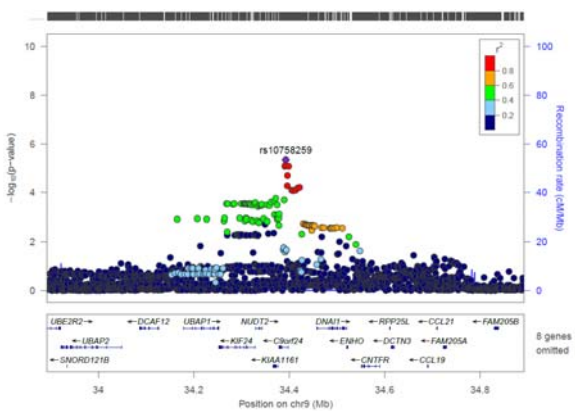




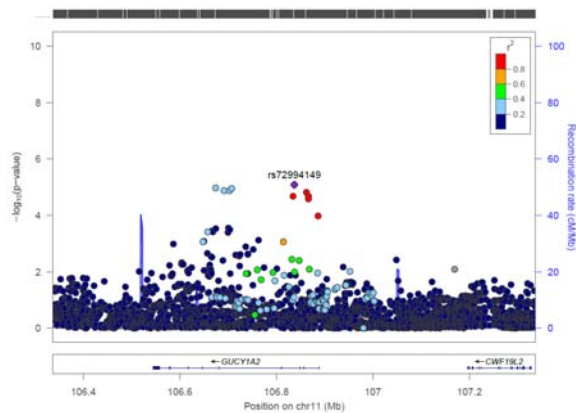
470



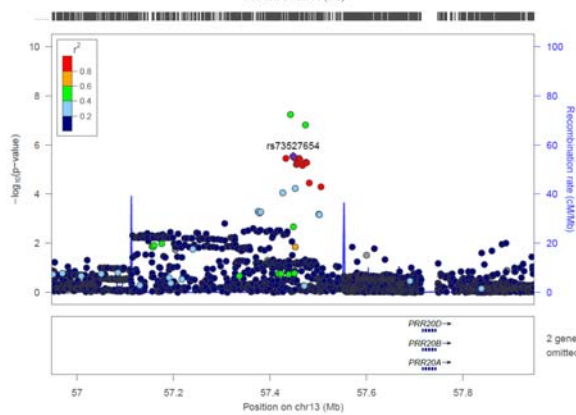
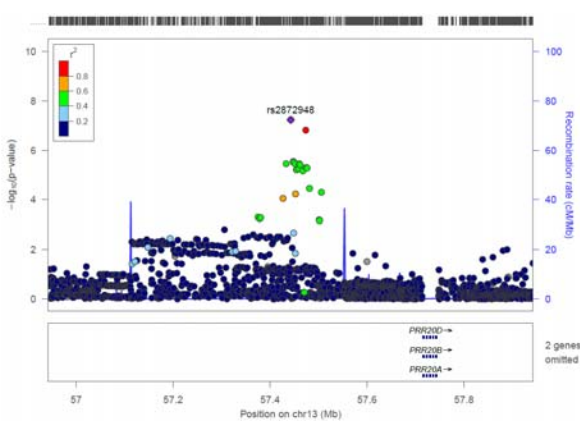
471



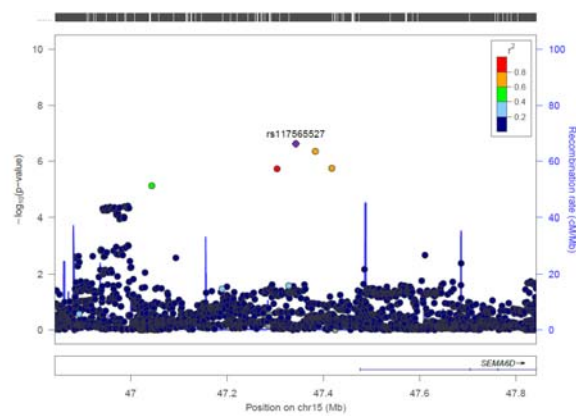
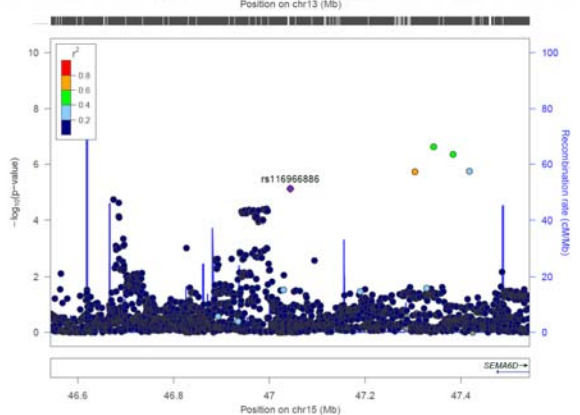
472



473



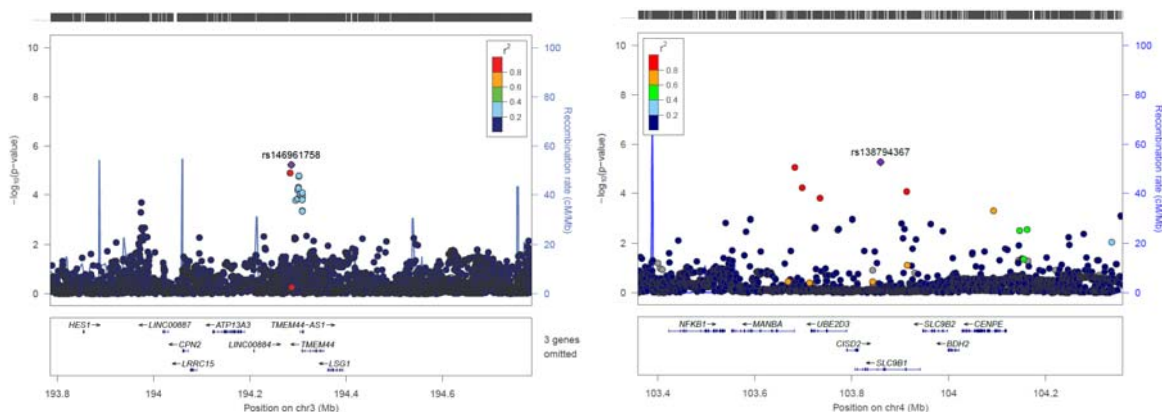
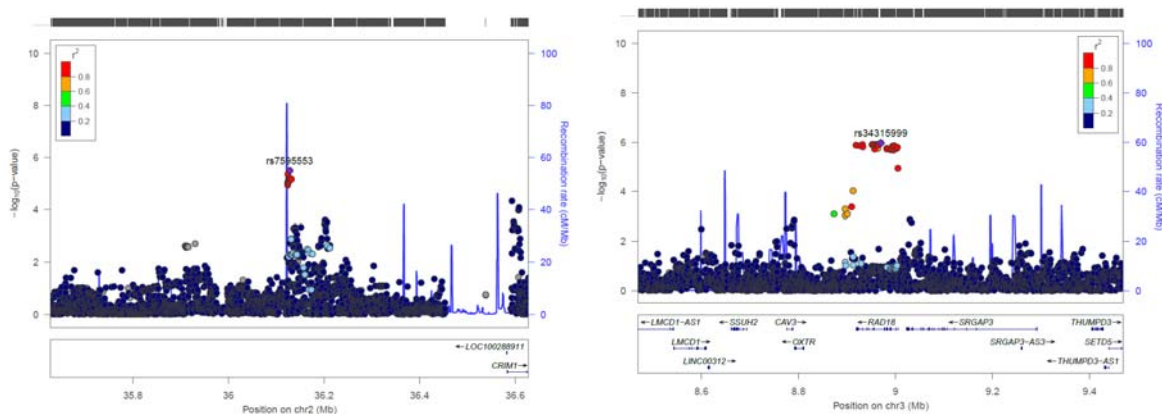
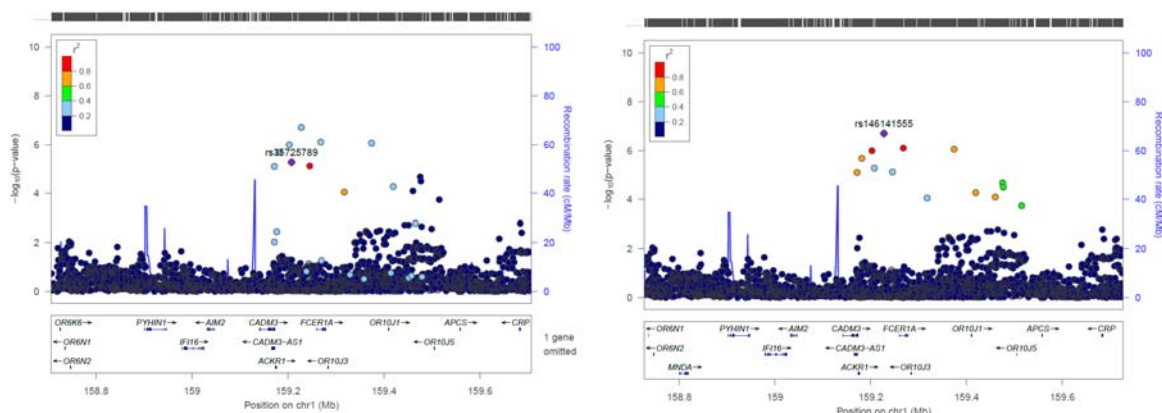
474



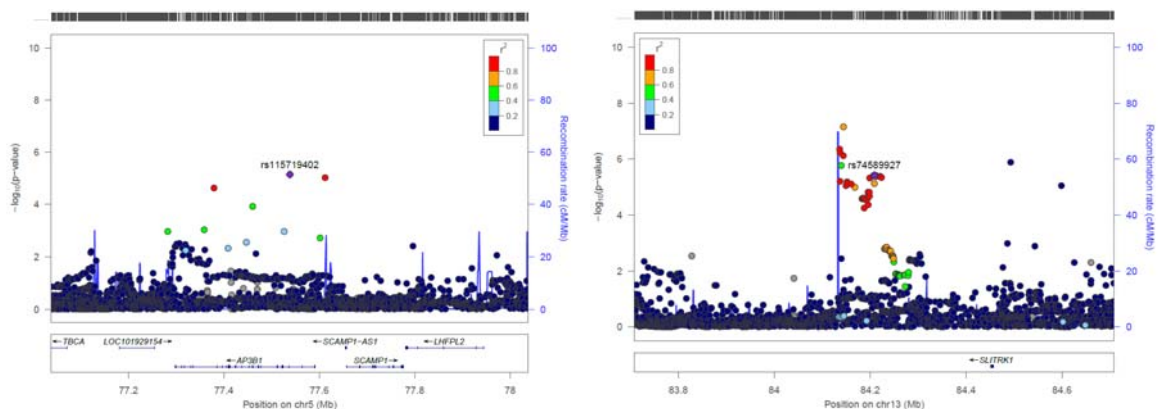
475

476

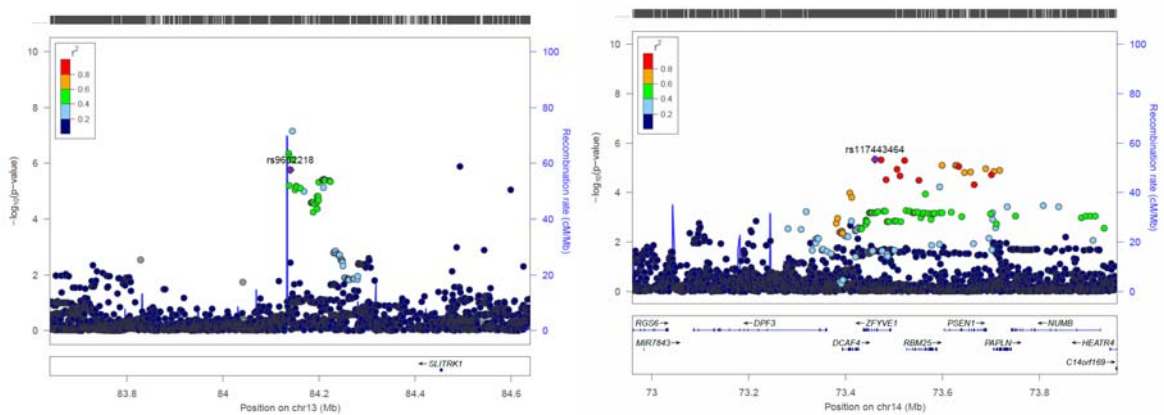
477 **Figure E4.** Zoom locus plots for short-listed independent top hits for Early-onset Mid-childhood  
 478 Remitting Wheezing



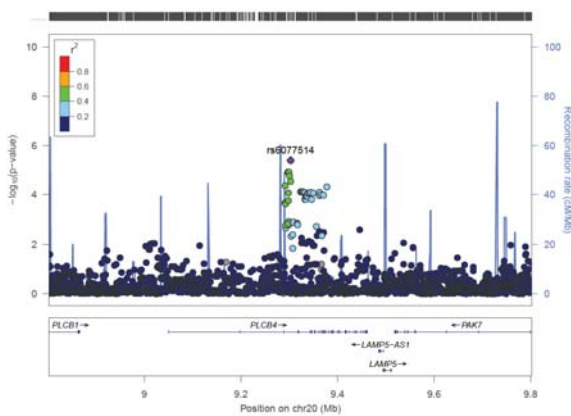
482



483

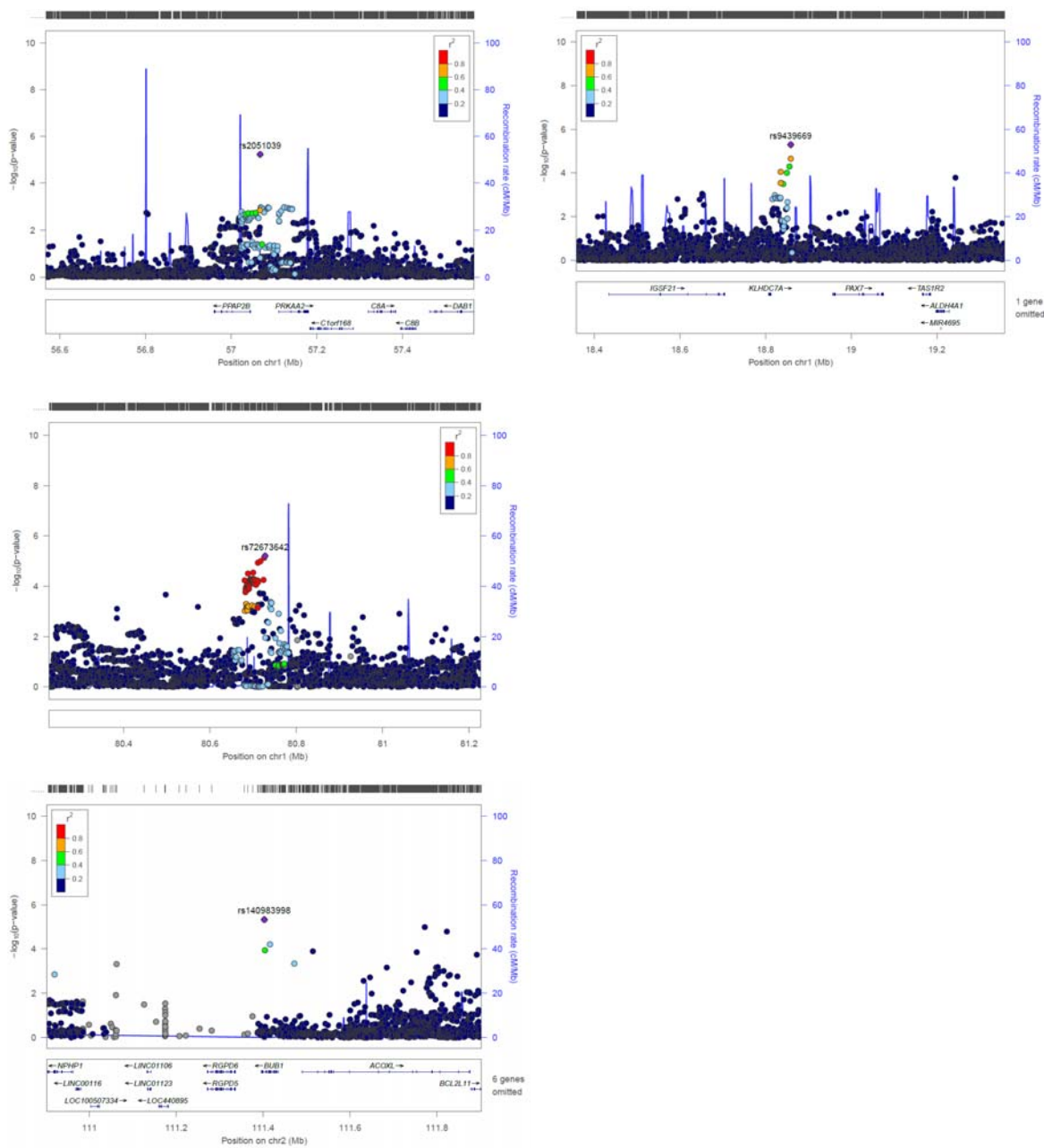


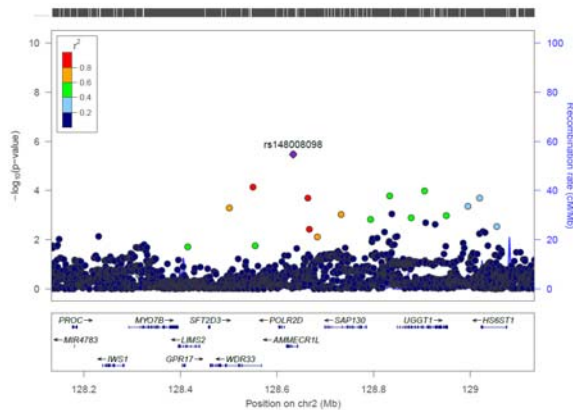
484



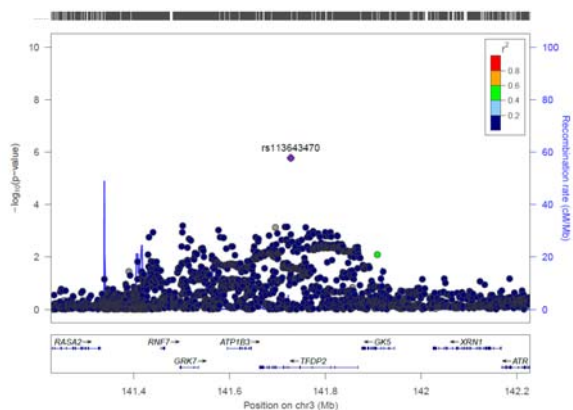
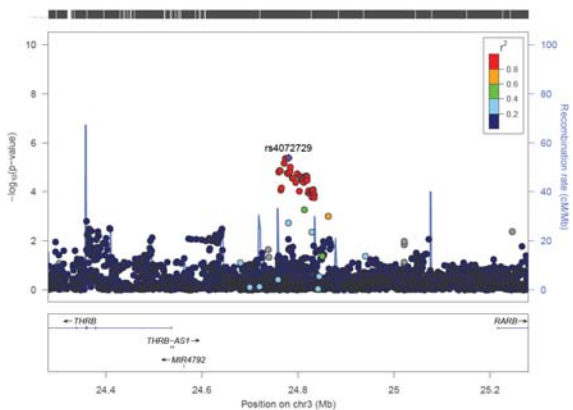
485

486 **Figure E5.** Zoom locus plots for short-listed independent top hits for Late-onset Wheezing

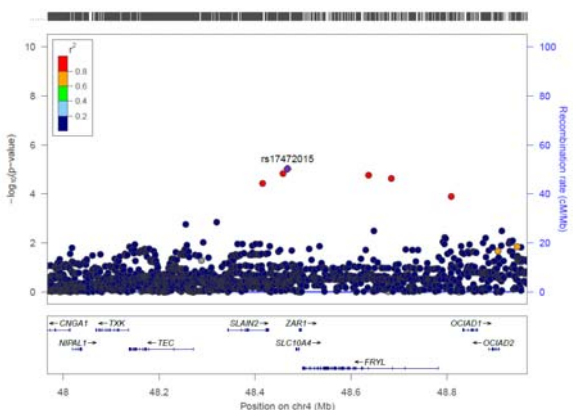
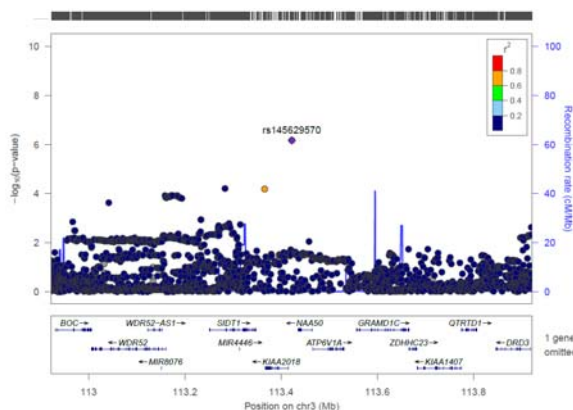




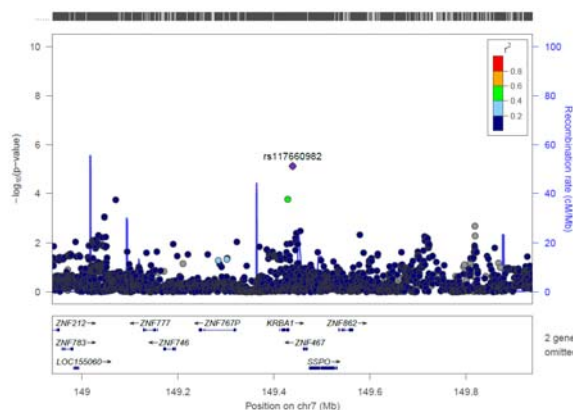
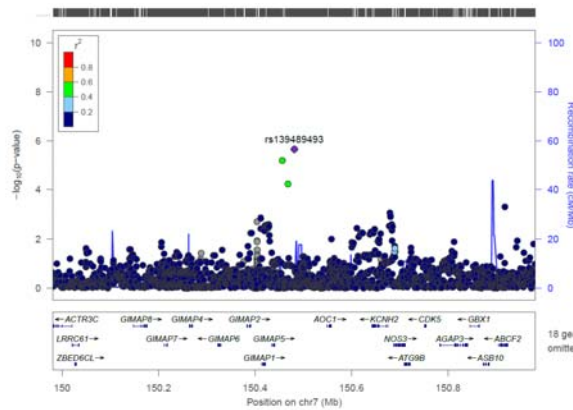
490



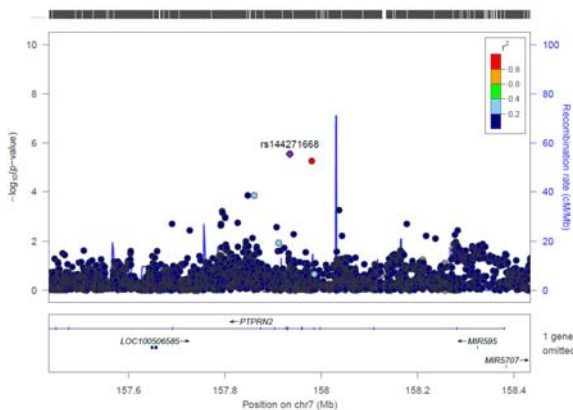
491

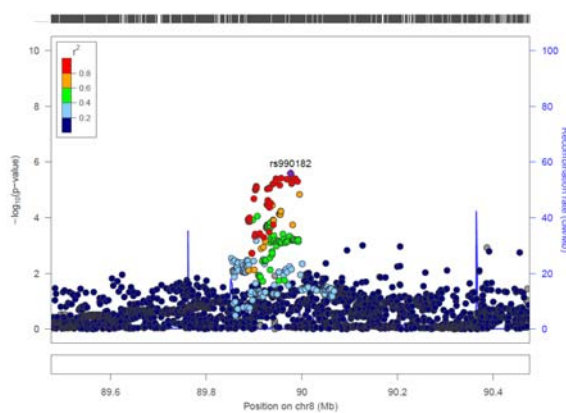


492

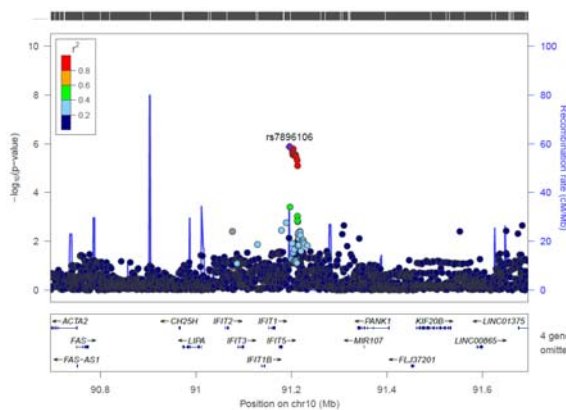
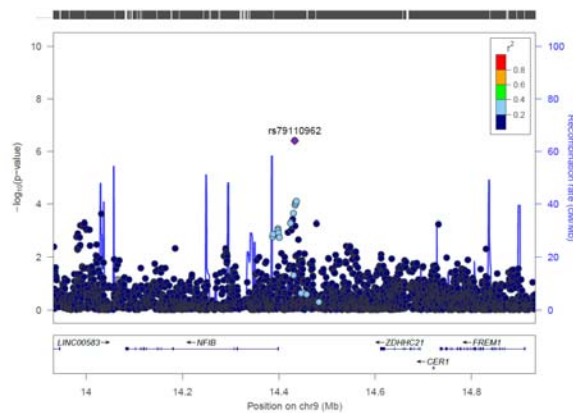


493

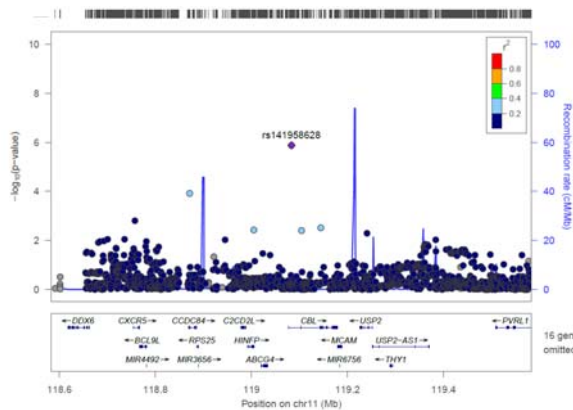




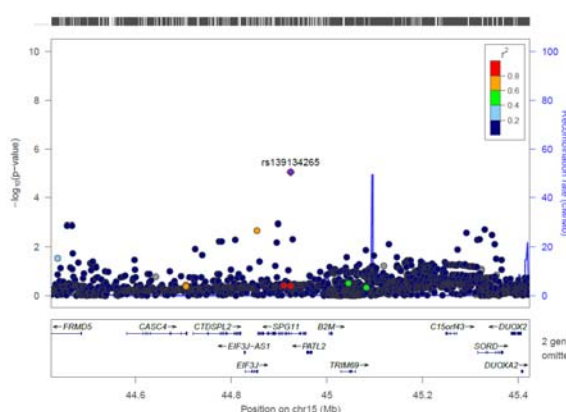
494



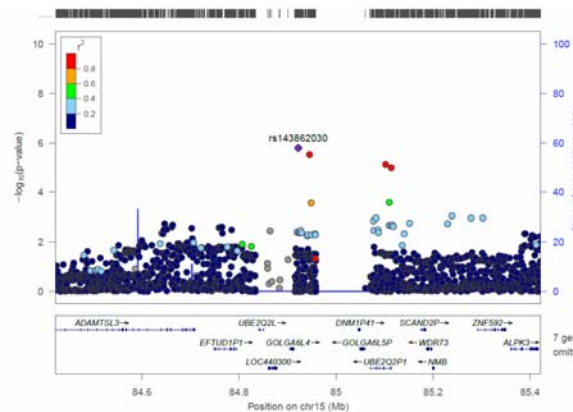
495

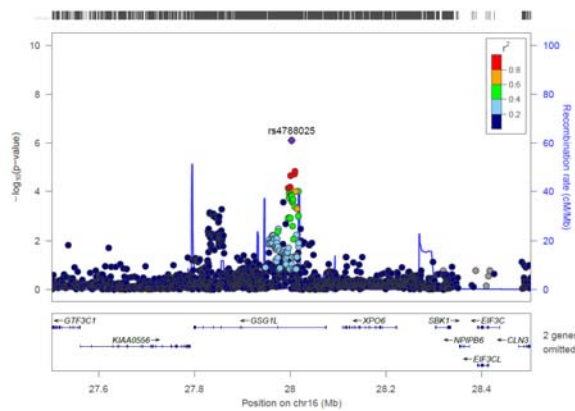


496

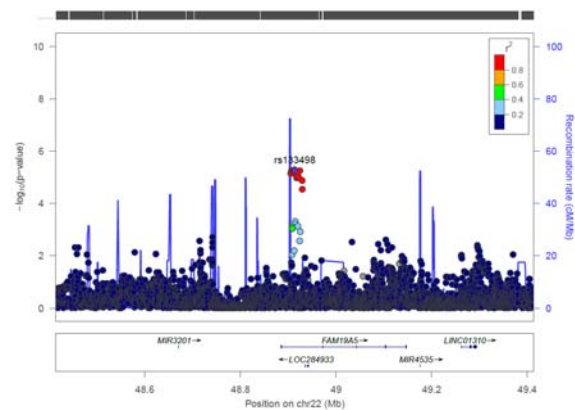
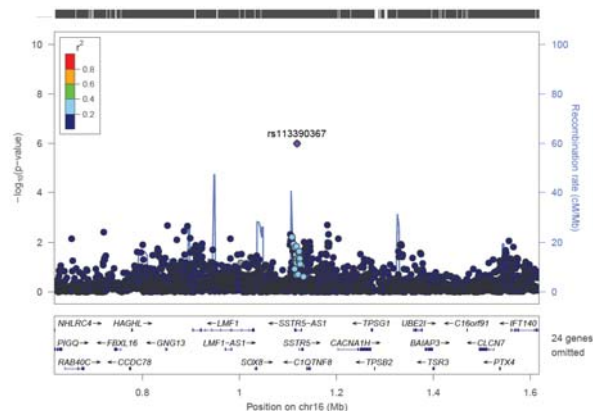


497





498



499



500 **Figure E6.** Linkage disequilibrium between SNPs downstream of *ANXA1* that were associated  
501 with persistent wheeze.



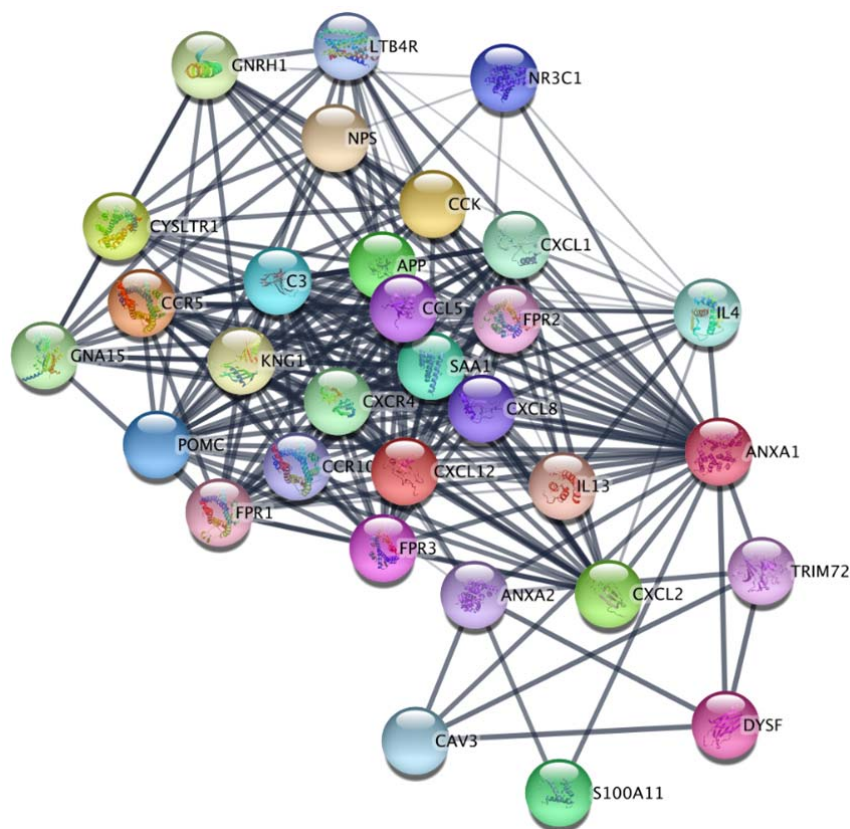
502

503

504

505 **Figure E7.** *ANXA1* interactors. Protein-protein interaction of *ANXA1* including *IL-4*, *IL-13* and  
506 *NR3C1*

507



508

## 509 REFERENCES

- 510 1. Custovic, A. *et al.* The Study Team for Early Life Asthma Research (STELAR) consortium 'Asthma e-  
511 lab': team science bringing data, methods and investigators together. *Thorax* **70**, 799-801 (2015).
- 512 2. Boyd A, Golding J, Macleod J, Lawlor DA, Fraser A, Henderson J, Molloy L, Ness A, Ring S, Davey  
513 Smith G. Cohort Profile: The 'Children of the 90s'--the index offspring of the Avon Longitudinal  
514 Study of Parents and Children. *Int J Epidemiol.* **42**, 111-27 (2013).
- 515 3. Fraser, A. *et al.* Cohort Profile: the Avon Longitudinal Study of Parents and Children: ALSPAC  
516 mothers cohort. *Int J Epidemiol* **42**, 97-110 (2013).
- 517 4. Custovic, A. *et al.* The National Asthma Campaign Manchester Asthma and Allergy Study. *Pediatr*  
518 *Allergy Immunol* **13**, 32-7 (2002).
- 519 5. Martindale, S. *et al.* Antioxidant intake in pregnancy in relation to wheeze and eczema in the first  
520 two years of life. *Am J Respir Crit Care Med* **171**, 121-8 (2005).
- 521 6. Atkinson, W. *et al.* Domestic aeroallergen exposures among infants in an English town. *Eur Respir*  
522 *J* **13**, 583-9 (1999).
- 523 7. Kurukulaaratchy, R.J., Fenn, M., Twiselton, R., Matthews, S. & Arshad, S.H. The prevalence of  
524 asthma and wheezing illnesses amongst 10-year-old schoolchildren. *Respir Med* **96**, 163-9 (2002).
- 525 8. Kurukulaaratchy, R.J. *et al.* Characterization of wheezing phenotypes in the first 10 years of life.  
526 *Clin Exp Allergy* **33**, 573-8 (2003).
- 527 9. Arshad, S.H. *et al.* Cohort Profile: The Isle Of Wight Whole Population Birth Cohort (IOWBC). *Int J*  
528 *Epidemiol* **47**, 1043-1044i (2018).
- 529 10. Oksel, C. *et al.* Distinguishing Wheezing Phenotypes from Infancy to Adolescence. A Pooled  
530 Analysis of Five Birth Cohorts. *Ann Am Thorac Soc* **16**, 868-876 (2019).
- 531 11. Loh, P.R. *et al.* Reference-based phasing using the Haplotype Reference Consortium panel. *Nat*  
532 *Genet* **48**, 1443-1448 (2016).
- 533 12. McCarthy, S. *et al.* A reference panel of 64,976 haplotypes for genotype imputation. *Nat Genet*  
534 **48**, 1279-83 (2016).
- 535 13. Durbin, R. Efficient haplotype matching and storage using the positional Burrows-Wheeler  
536 transform (PBWT). *Bioinformatics* **30**, 1266-72 (2014).
- 537 14. Marchini, J. & Howie, B. Genotype imputation for genome-wide association studies. *Nat Rev*  
538 *Genet* **11**, 499-511 (2010).
- 539 15. Willer, C.J., Li, Y. & Abecasis, G.R. METAL: fast and efficient meta-analysis of genomewide  
540 association scans. *Bioinformatics* **26**, 2190-1 (2010).
- 541 16. Pruim, R.J. *et al.* LocusZoom: regional visualization of genome-wide association scan results.  
542 *Bioinformatics* **26**, 2336-7 (2010).
- 543 17. Dayem Ullah, A.Z. *et al.* SNPnexus: assessing the functional relevance of genetic variation to  
544 facilitate the promise of precision medicine. *Nucleic Acids Res* **46**, W109-W113 (2018).
- 545 18. Stelzer, G. *et al.* The GeneCards Suite: From Gene Data Mining to Disease Genome Sequence  
546 Analyses. *Curr Protoc Bioinformatics* **54**, 1 30 1-1 30 33 (2016).
- 547 19. Durinck, S. *et al.* BioMart and Bioconductor: a powerful link between biological databases and  
548 microarray data analysis. *Bioinformatics* **21**, 3439-3440 (2005).
- 549 20. Durinck, S., Spellman, P.T., Birney, E. & Huber, W. Mapping identifiers for the integration of  
550 genomic datasets with the R/Bioconductor package biomaRt. *Nat Protoc* **4**, 1184-91 (2009).
- 551 21. McLaren, W. *et al.* The Ensembl Variant Effect Predictor. *Genome Biol* **17**, 122 (2016).
- 552 22. Wingett, S. *et al.* HiCUP: pipeline for mapping and processing Hi-C data. *F1000Res* **4**, 1310 (2015).
- 553 23. Cairns, J. *et al.* CHICAGO: robust detection of DNA looping interactions in Capture Hi-C data.  
554 *Genome Biol* **17**, 127 (2016).
- 555 24. Adsera, C.B., Park, Y.P., Meuleman, W. & Kellis, M. Integrative analysis of 10,000 epigenomic  
556 maps across 800 samples for regulatory genomics and disease dissection. 810291 (2019).
- 557 25. Ihaka, R. & Gentleman, R. R: A Language for Data Analysis and Graphics. *Journal of Computational*  
558 *and Graphical Statistics* **5**, 299-314 (1996).

- 559 26. Szklarczyk, D. *et al.* STRING v11: protein-protein association networks with increased coverage,  
560 supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res* **47**,  
561 D607-D613 (2019).
- 562 27. Shannon, P. *et al.* Cytoscape: a software environment for integrated models of biomolecular  
563 interaction networks. *Genome Res* **13**, 2498-504 (2003).
- 564 28. Branchett, W.J., Walker, S.A. & Lloyd, C.M. Experimental Mouse Models of Asthma and Analysis of  
565 CD4 T Cells. *Methods Mol Biol* **2285**, 329-348 (2021).
- 566 29. Brunekreef, B. *et al.* The prevention and incidence of asthma and mite allergy (PIAMA) birth  
567 cohort study: design and first results. *Pediatr Allergy Immunol* **13**, 55-60 (2002).
- 568 30. Wijga, A.H. *et al.* Cohort profile: the prevention and incidence of asthma and mite allergy (PIAMA)  
569 birth cohort. *Int J Epidemiol* **43**, 527-35 (2014).
- 570