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

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

on behalf of UNICORN[13] and Breathing Together[14] Investigators

[§]Equal contribution, joint senior authors

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

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

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

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

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

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

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

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

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

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

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

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

*Correspondence and requests for reprints should be addressed to Dr Raquel Granell, Population Health Sciences, Bristol Medical School, University of Bristol, BRISTOL BS8 2BN, UK

E-mail Raquel.Granell@bristol.ac.uk

This article has an online appendix.

This article is dedicated to the memory of our wonderful colleague and friend Professor John Henderson (1958-2019), whose contribution to the understanding of the heterogeneity of childhood asthma cannot be overstated. Rainbow-chasers and UNICORN riders forever.

[13]STELAR/UNICORN investigators

Professor Graham Devereux

Dr Dimitrios Charalampopoulos

[41]Breathing Together investigators

Prof Andrew Bush

Prof Sejal Saglani

Prof Benjamin Marsland

Prof Jonathan Grigg

Prof Jurgen Schwarze

Prof Mike Shields

Prof Peter Ghazal

Prof Ultan Power

Dr Ceyda Oksel

Word count: 3,500

Keywords: GWAS, asthma, wheeze phenotypes, Annexin 1

Data availability: The informed consent obtained from all included participants does not allow the data to be made freely available through any third party maintained public repository. However,

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

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

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

Author Contributions: R.G., J.C., S.H. and A.C. conceived and planned the study. R.G., A.C., J.W.H., L.L.Y. and C.M.L. wrote the manuscript. R.G. and J.C. analysed the data. J.C. and N.K. provided eQTL analysis for *ANXA1*, J.W.H. supervised eQTL analyses. J.H. and M.T. performed the PCHi-C experiments and analysis. J.C. and M.T. interpreted the genetic architecture results for *ANXA1*. 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 designed murine experiments. S.A.M and L.G.G. conducted experiments in mouse models. M.P. provided annexin$^{-/-}$ mice and advice on annexin. All authors provided critical feedback and helped shape the research, analysis, and manuscript. This publication is the work of the authors and Raquel Granell and Adnan Custovic will serve as guarantors for the contents of this paper.

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

(Lecture fees). Clare Lloyd: Wellcome Trust 107059/Z/15/Z. John Holloway: Medial Research Council grant MR/S025340/1 (to institution) American Academy of Allergy Asthma and Immunology (AAAI) (Support for speaker travel to AAAAI annual congress). Adnan Custovic: MRC (research grants) EPSRC (research grant) Wellcome Trust (research grant) Worg Pharmaceoticals (Personal payment <US$5000) GSK Honorarium for lecture, personal, <US$ 5000 AstraZeneca Honorarium for lecture, personal, <US$ 5000 Sanofi Honorarium for lecture, personal, <US$ 5000 Stallergens-Greer Honorarium for lecture, personal, <US$ 5000 WAO (Board of officers, unpaid).

Funding: Supported by the UK Medical Research Council (MRC) Programme Grant MR/S025340/1, the Wellcome Trust (WT) Strategic Award (108,818/15/Z) and a WT Senior Fellowship to CML (107059/Z/15/Z). The MRC and Wellcome (Grant ref: 217065/Z/19/Z) and the University of Bristol provide core support for ALSPAC. ALSPAC GWAS data was generated by Sample Logistics and Genotyping Facilities at Wellcome Sanger Institute and LabCorp (Laboratory Corporation of America) using support from 23andMe. A comprehensive list of grants funding is available on the ALSPAC website (http://www.bristol.ac.uk/alspac/external/documents/grant-acknowledgements.pdf).

PIAMA was funded by the Netherlands Lung Foundation (grant 3.4.01.26, 3.2.06.022, 3.4.09.081 and 3.2.10.085CO), the ZON-MW Netherlands Organization for Health Research and Development (grant 912-03-031), the Stichting Astmabestrijding and the Ministry of the Environment. Genome-wide genotyping was funded by the European Commission as part of GABRIEL (grant number 018996) and a grant from BBMRI-NL (CP 29). GHK is supported by a ZON-MW VICI grant.

This research was funded in part by the Wellcome Trust [217065/Z/19/Z]. For the purpose of Open Access, the author has applied a CC BY public copyright license to any Author Accepted Manuscript 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*), $p<6.7\times10^{-9}$), 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$^{-/-}$
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

INTRODUCTION

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

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

One way of disaggregating asthma diagnosis is to use data-driven methods to derive subtypes in a hypothesis-neutral way(18). For example, we jointly modelled data on wheezing from birth to 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.

MATERIALS & METHODS

Study design, setting, participants and data sources/measurement

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

Validated questionnaires were completed on multiple occasions from infancy to adolescence(19). A list of variables, per cohort, is shown in **Table E1**, and the cohort−specific time points and sample sizes in **Table E2**. Data were harmonised and imported into Asthma eLab web−based knowledge management platform to facilitate joint analyses(20).

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

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

Genotyping, imputation and GWAS Meta-Analysis

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

Post-GWAS studies

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

RESULTS

Participants and descriptive data

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

GWAS Meta-Analysis

We conducted three GWASs (ALSPAC, MAAS, IOW-SEATON-Ashford) in parallel and results were meta-analyzed. QQ-plots are shown in **Figure E1**. A circular Manhattan plot showing an overview of the GWAS results by wheeze phenotype is shown in **Figure 1**. A total of 589 SNPs were associated with at least one phenotype with $p<10^{-5}$. After clumping, we identified 134 independent SNPs uniquely associated with different phenotypes ($p<10^{-5}$): of these, 44 were associated with early-onset persistent, 25 with early-onset preschool remitting, 33 with early-onset mid-childhood remitting and 32 with late-onset wheeze (**Table E5**). For some SNPs there were nominal associations with more than one phenotype. For example, chr17q21 was identified as a top locus for early-onset persistent wheeze ($p=5.42\times10^{-9}$), but some of the SNPs in this region were also associated with the early-onset mid-childhood remitting phenotype ($p<0.0001$).

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

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

We identified two GWAS-significant loci for early-onset persistent wheeze: 17q21, $p<5.5\times10^{-9}$, and a novel locus on 9q21.13 (*ANXA1*), $p<6.7\times10^{-9}$. The *ANXA1* locus was the only GWAS-significant locus that had not previously been associated with asthma or atopic traits, with one previous study showing an association with FEV_1/FVC and bronchodilator response in smokers(36). *ANXA1* is strongly expressed in bronchial mast cells and has anti-inflammatory properties(37), and may be involved in epithelial airway repair(38) **(Table E6**). We therefore followed up top SNPs from this locus.

***ANXA1* locus and persistent wheeze**

Two SNPs (rs75260654, the lead SNP, and rs116849664 located downstream of *ANXA1*) were 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 1st 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.

rs75260654: Effect on gene expression

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

Additional supporting evidence regarding the significance of the T-allele on the expression of these genes was provided using eQTLGene Consortium meta-analysis of 24 Cohorts and 24331 samples(40). This method reproduced the previous modest results showing a cis-eQTL effect of rs75260654 on both the *ANXA1* ($p=6.02\times10^{-23}$) and *ALDH1A1* ($p=1.11\times10^{-19}$) at FDR=0. No significant trans-eQTLs were observed.

Potential biological function of ANXA1 in asthma

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

Functional studies of *anxa1* in a murine model

Pulmonary expression of anxa1 is modulated by aeroallergen exposure

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

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

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

DISCUSSION

Herein, we present a comprehensive description of the genetic architecture of childhood wheezing disorders. Using a novel approach applied to a unique dataset from five UK birth cohorts, we identified subsets of SNPs differentially associated across four wheezing phenotypes: early-onset persistent (44 SNPs, 19 loci), early-onset preschool remitting (25 SNPs, 10 loci), early-onset mid-childhood remitting (33 SNPs, 9 loci) and late-onset (32 SNPs, 20 loci). We found little evidence of genetic associations spanning across different phenotypes. This suggests that genetic architecture of different wheeze phenotypes comprises a limited number of variants likely underpinning mechanisms which are shared across phenotypes, but that each phenotype is also characterized by unique phenotype-specific genetic associations. Importantly, we identified a novel locus in chr9q21 nearby *ANXA1* exclusively associated with early-onset persistent wheeze ($p<6.7\times10^{-9}$). To identify the most likely causative variant, we investigated the overlap of the associated SNPs with Promoter Capture Hi-C interactions to demonstrate that SNP rs75260654 overlapped a region interacting with the *ANXA1* promoter. Using eQTL data, we identified that the risk allele (T) of rs75260654 associated with early-onset persistent wheeze is also associated with *ANXA1* expression. Further investigation of the biological function of *ANXA1* revealed that it interacts with genes enriched for asthma (including *IL4* and *IL13*) and inflammatory regulation (*NR3C1*, Glucocorticoid receptor). In functional mouse experiments, anxa1 protein expression increased and anxa1 mRNA was significantly induced in lung tissue following HDM exposure, suggesting that the pro-resolving anxa1 may play a role in regulating the pulmonary immune response to allergen. Concurrently, by utilizing *anxa1*$^{-/-}$ deficient mice we demonstrated that loss of anxa1 results in heightened AHR and Th2 inflammation upon allergen challenge, providing important *in vivo* functional data to support our GWAS-finding.

ANXA1 is a 37-kDa glycoprotein with potent anti-inflammatory and pro-resolving properties that are mediated by interaction with a specific G-protein coupled receptor FPR2(42). This axis represents an important resolution pathway in chronic inflammatory settings such as those of rheumatoid arthritis(43) and ulcerative colitis(44). *ANXA1* belongs to the annexin family of Ca^{2+}-dependent phospholipid-binding proteins, and through inhibition of phospholipase A2, it reduces eicosanoid production, which also contributes to its anti-inflammatory activities. Modulation of M2 macrophage phenotype is also promoted by *ANXA1* to attenuate tissue inflammation(45). Corticosteroids (a mainstay of asthma treatment) increase the synthesis of *ANXA1*(46). Plasma *ANXA1* levels are significantly lower in asthmatic patients with frequent exacerbations compared to those with stable disease, suggesting a link between this mediator and disease state(47). Furthermore, children with wheeze have reduced airway levels of *ANXA1*(48).

Previous functional studies using *anxa1*$^{-/-}$ deficient mice challenged with ovalbumin showed *anxa1*-deficient mice to have elevated AHR compared to WT mice(49). Ng *et al.* demonstrated that untreated *anxa1*-deficient mice have spontaneous AHR that predisposes them to exacerbated response to allergen(49). In the current study, we demonstrated in the murine lung the induction of Anxa1 in response to HDM exposure. In addition, genetic deletion of *anxa1* potentiated the development of AHR and enhanced eosinophilia and markers of Th2 inflammation in mice treated with HDM, which is consistent with and extends previous reports. Of interest, in mice, *anxa1* expression was recently found to be characteristic of a novel cell-type called the Hillock cell, which may be involved in squamous barrier function and immunomodulation(50). These data identify the ANXA1/FPR2 signaling axis as an important regulator of allergic disease, that could be 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

identify asthma-susceptibility loci. Examples include the GABRIEL(6), EVE(60) and TAGC(7) consortia, and the value of diverse, multi-ethnic participants in large-scale genomic studies has recently been shown(61). However, such consortia do not have the depth of longitudinal data to allow the type of analyses which we carried out to derive a multivariable primary outcome.

One strength of our approach is that we used data from five birth cohorts with detailed and lifelong phenotyping, which were harmonised in a common knowledge management platform(20), allowing joint analyses. We performed three parallel GWASs that produced estimates with remarkably consistent directions of effects.

In conclusion, using unique data from five UK birth cohorts, we identified subsets of SNPs differentially associated across four wheezing phenotypes from infancy to adolescence. We found little evidence of genetic associations spanning across different phenotypes. We discovered a novel locus in chr9q21 uniquely associated with early-onset persistent wheeze ($p<6.7\times10^{-9}$), identified SNP rs75260654 as the most likely causative variant, and demonstrated that the risk allele (T) confers a reduction in *ANXA1* expression. In mouse experiments, *ANXA1* expression increased in lung tissue following allergen exposure, suggesting that the pro-resolving ANXA1 may play a role in regulating the pulmonary immune response to allergen. Using *ANXA1*-deficient mice, we demonstrated that loss of *ANXA1* results in heightened AHR and Th2 inflammation upon allergen challenge, providing important *in vivo* functional data to support our GWAS finding. Targeting these pathways to promote the clearance of chronic inflammation in persistent disease may represents an exciting therapeutic prospect.

REFERENCES

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

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

locus for asthma with severe exacerbations in diverse populations. *Pediatr Allergy Immunol* 2021; 32: 106-115.

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

Downes K, Stegle O, Battle A, Yang J, Visscher PM, Scholz M, Gibson G, Esko T, Franke L. Unraveling the polygenic architecture of complex traits using blood eQTL metaanalysis. 2018: 447367.

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

653 **TABLES AND FIGURES**

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	*CHRM3*	g(0.56)/t	0.25	0.05	2.45E-06	+++	0.79	FEV_1, FEV_1/FVC, asthma- high priority drug target
2p25.1	rs13398488	*RNF144A*	g(0.29)/a	0.25	0.05	2.18E-06	+--	0.13	asthma, allergy, childhood onset asthma, allergic rhinitis
2q12.2	rs6543291	*FHL2*	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	*EFCC1, RAB43, RAB7A*	c(0.05)/t	0.47	0.10	6.40E-06	+++	0.39	*RAB43*: response to bronchodilator, FEV/FEC ratio; *RAB7A*: eosinophil count
4p16.3	rs7680608[eQTL]	*RNF212, IDUA, DGKQ, SLC26A1*	g(0.93)/c	-0.42	0.09	1.31E-06	---	0.15	*4p16*: asthma
	rs77822621[eQTL]		c(0.96)/t	-0.50	0.11	7.16E-06	---	0.01	
4q31.21	rs115228498	*INPP4B*	c(0.02)/t	0.79	0.17	2.70E-06	+++	0.02	atopic asthma
5p15.31	rs116494115	*ADCY2*	g(0.01)/a	0.75	0.17	6.49E-06	+++	0.09	asthma × air pollution, childhood asthma
7q22.3	rs76871421	*CDHR3*	c(0.12)/t	0.37	0.07	5.71E-07	+++	0.22	childhood asthma
9q21.13	rs75260654	***ANXA1**, TMC1, LOC101927258, ALDH1A1*	c(0.98)/t	-0.90	0.16	**6.66E-09**	---	0.05	***ANXA1***: FEV_1/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	**1.99E-08**	---	0.06	

									repair
10q24.2	rs7088157	*LOXL4, R3HCC1L*	g(0.5)/a	-0.23	0.05	7.34E-06	---	0.26	*R3HCC1L*: eosinophil count, atopic eczema, psoriasis, BMI
11p15.4	rs112474574	*TRIM5, TRIM6, TRIM22*	c(0.96)/t	-0.55	0.12	2.29E-06	---	0.14	severe asthma and insulin resistance
11q23.3	rs116861530[eQTL]	*SIK3*	g(0.94)/a	-0.42	0.09	9.07E-06	---	0.01	triglycerides, glucose metabolism, eosinophil count
14q22.1	rs1105683	*KTN1*	c(0.07)/t	0.41	0.09	9.15E-06	+++	0.24	severe asthma
5q13.3	rs2202714[eQTL]	*FAM227B*	g(0.36)/a	0.23	0.05	8.71E-06	+++	0.01	rs35251997 and FEV_1; FEV_1/FVC
15q25.2	rs117540214[eQTL]	*ADAMTSL3*	g(0.06)/a	0.42	0.10	9.82E-06	+++	3.91E-03	FEV_1/FVC
17q12	rs17676191	*IKZF3*	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	**5.42E-09**	+++	1.96E-04	
	rs1031460	*GSDMB*	g(0.50)/t	0.27	0.05	8.71E-08	+++	1.87E-04	
	rs56199421		c(0.45)/t	-0.23	0.05	4.50E-06	---	9.61E-04	
17q21	rs4795406	*LRRC3C*	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	*GSDMA*	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	
19p13.2	rs4804311	*MARCH2, HNRNPM, MYO1F*	g(0.08)/a	0.42	0.09	9.65E-07	+-+	0.05	triglycerides, HDL-cholesterol, metabolic syndrome; *MYO1F*: FEV_1 and FVC
	rs2013694		c(0.89)/t	-0.38	0.08	8.29E-07	--+	0.39	
	rs73501545		g(0.16)/a	0.31	0.07	8.39E-06	+++	0.29	
	rs111644945		g(0.9)/a	-0.41	0.08	4.01E-07	---	0.02	
22q11.1	rs5994170	*CECR5*	g(0.4)/a	0.23	0.05	4.95E-06	+++	0.58	triglycerides, eosinophil count and body height
	rs34902370		c(0.75)/t	-0.25	0.06	6.80E-06	---	0.41	

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[eQTL]	*PPP2R5A*	c(0.79)/t	-0.22	0.05	8.44E-06	---	0.53	waist circumference & obesity
2p24.2	rs2880066	*FAM49A or CYRIA*	t(0.09)/a	0.32	0.07	4.34E-06	+++	0.20	airway repair in non-atopic asthma
	rs10180268		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[eQTL]	*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	*CADM3, FCER1A, MPTX1, OR10J1*	c(0.95)/a	-0.56	0.12	5.42E-06	-+-	0.01	neutrophil count, CRP
	rs146141555		c(0.98)/t	-0.89	0.17	2.04E-07	-+-	0.08	
	rs146575092		g(0.98)/a	-0.85	0.17	8.73E-07	-+-	0.07	
2p22.3	rs7595553	*MRPL50P1*	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[eQTL]	*RAD18*	c(0.03)/t	0.69	0.14	1.11E-06	+++	0.14	atopy/SPT
3q29	rs146961758	*MRPL50P1, LSG1, TMEM44-AS1, TMEM44, ATP13A3*	t(0.05)/a	0.57	0.12	6.01E-06	+-+	0.11	*3q29:* BMI *TMEM44-AS1, TMEM44, ATP13A3*: diastolic blood pressure; *LSG1*: BMI, eosinophil count
4q24	rs138794367	*SLC9B1*	c(0.99)/t	-1.02	0.22	5.47E-06	---	0.13	eosinophil count, allergic rhinitis
5q14.1	rs115719402	*AP3B1*	g(0.96)/a	-0.60	0.13	7.20E-06	---	0.06	Vital capacity, BMI
13q31.1	rs9602218	*RNU6-67P, SLITRK1*	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		g(0.02)/a	0.73	0.16	3.78E-06	+-+	0.02	
	rs2210726	*VENTXP2, UBE2D3P4, MTND4P1*	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	*ZFYVE1*	g(0.95)/a	-0.57	0.12	4.68E-06	--+	0.19	LDL cholesterol and systolic blood pressure
20p12.3-p12.2	rs6077514	*PLCB4*	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

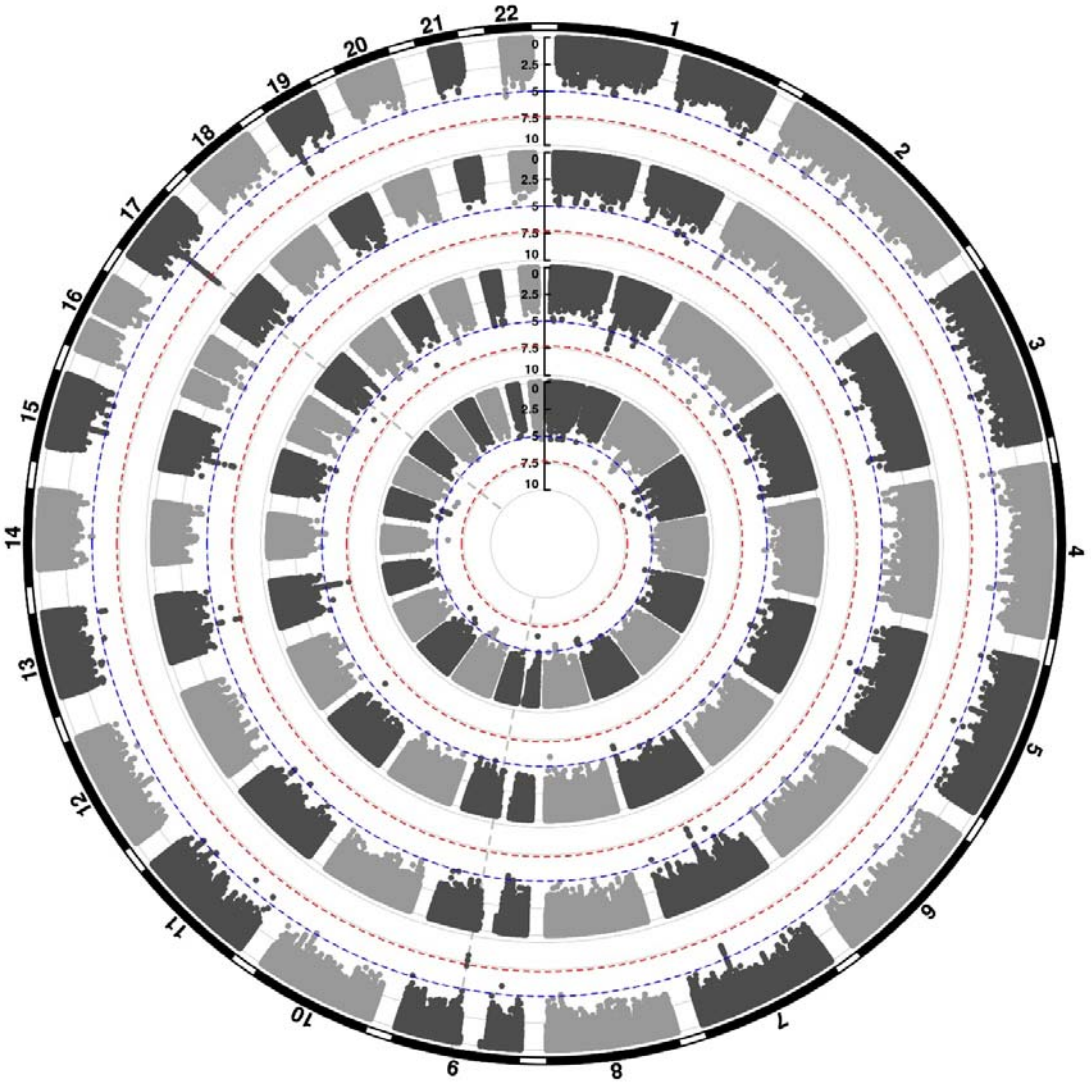
	rs139489493	*LOC105375566*	c(0.98)/t	-0.95	0.20	2.28E-06	---	0.03	
7q36.3	rs144271668	*PTPRN2*	c(0.01)/a	0.88	0.19	2.91E-06	+++	0.28	eczema
8q21.3	rs990182	*LOC105375631*	t(0.42)/a	0.28	0.06	2.57E-06	+++	0.46	*8q21.3*: type 1 diabetes
9p22.3	rs79110962	*NFIB, ZDHHC21*	c(0.08)/t	0.51	0.10	3.98E-07	+++	0.05	*9p22.3*: asthma (mean age<16 years)
10q23.31	rs7896106	*SLC16A12, IFIT family, PANK1*	g(0.35)/t	0.30	0.06	1.35E-06	+++	0.05	*SLC16A12*: Body height; *PANK1*: insulin
11q23.3	rs141958628	*CBL, CCDC84, MCAM*	c(0.98)/t	-0.98	0.20	1.33E-06	-+-	0.27	*CCDC84*: asthma, allergy
15q15.3-q21.1	rs139134265	*SPG11, CTDSPL2*	g(0.02)/c	0.87	0.20	9.11E-06	+-+	0.13	*CTDSPL2*: alcohol drinking
15q25.2	rs143862030	*ADAMTSL3, GOLGA6L4, UBE2Q2P8*	c(0.04)/t	0.64	0.13	1.65E-06	+++	0.08	*ADAMTSL3*: FEV1/FVC; lean mass
16p13.3	rs113390367	*SSTR5-AS1, CACNA1H*	g(0.86)/a	-0.40	0.08	1.04E-06	---	0.16	*CACNA1H*: eosinophil count
16p12.1	rs4788025	*GSG1L*	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	*FAM19A5 or TAFA5*	g(0.94)/a	-0.48	0.11	5.35E-06	---	0.84	Obesity and Metabolic Dysfunction

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

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

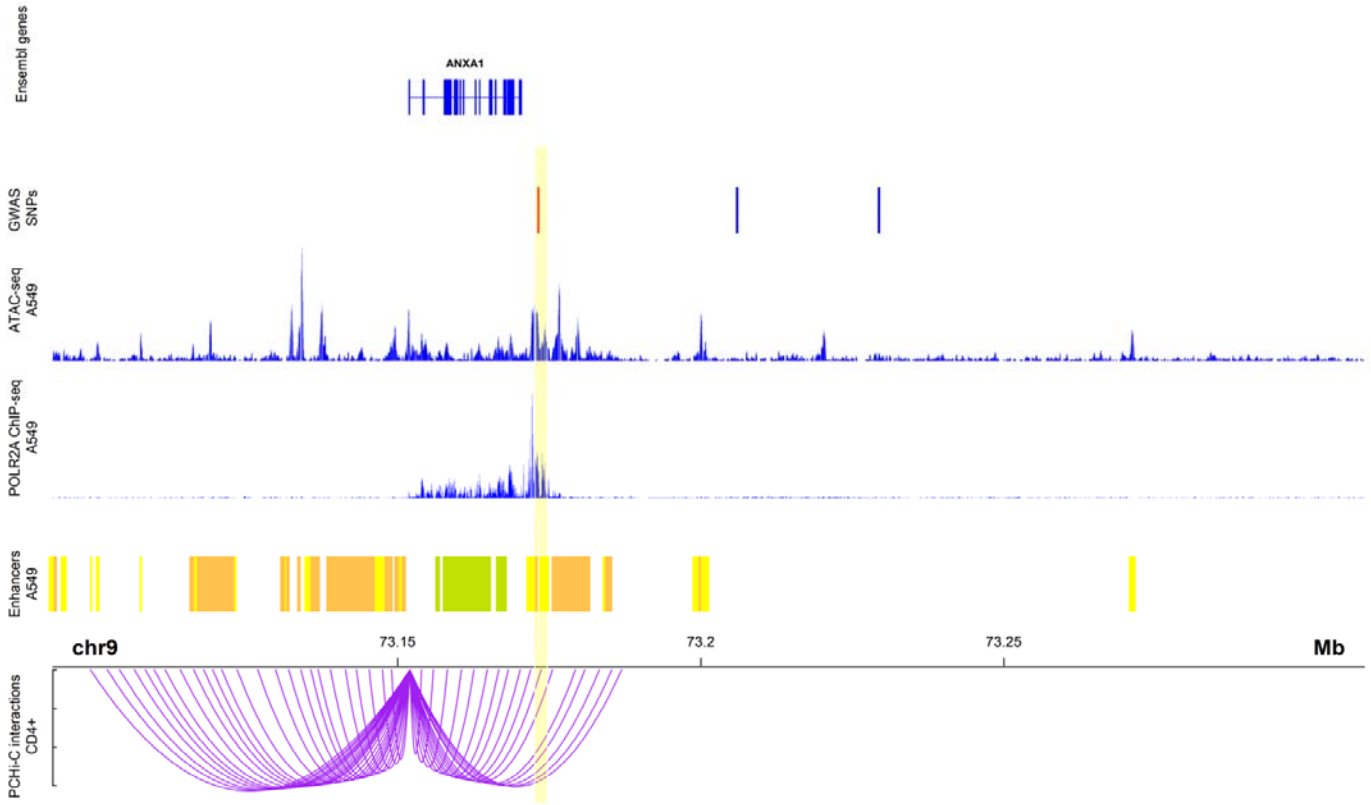
† List of references or sources (GeneCards, GWAS Catalog, PhenoScanner) available in **Table E6.**

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 ($P < 10^{-5}$).



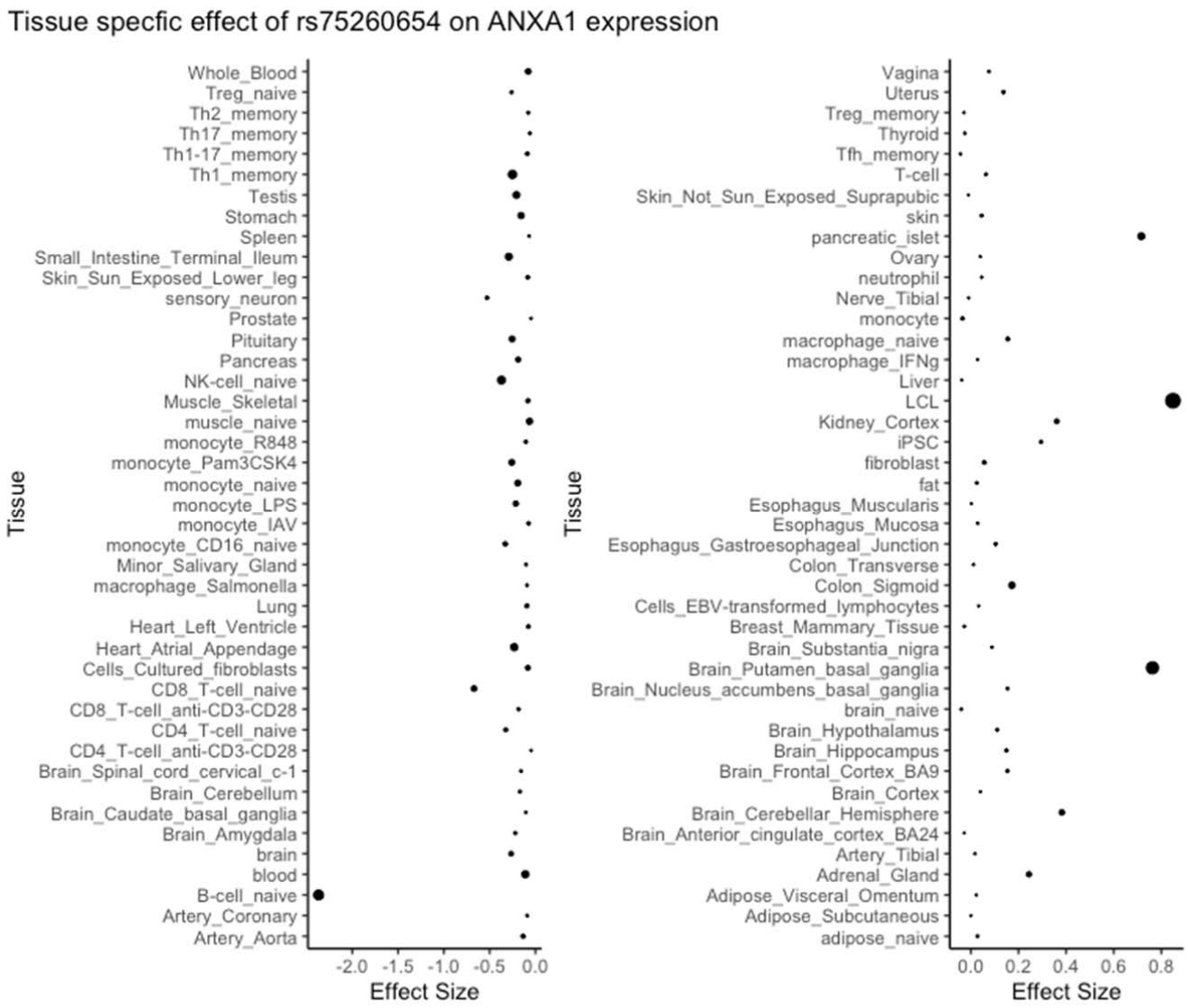
665
666
667
668

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



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

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

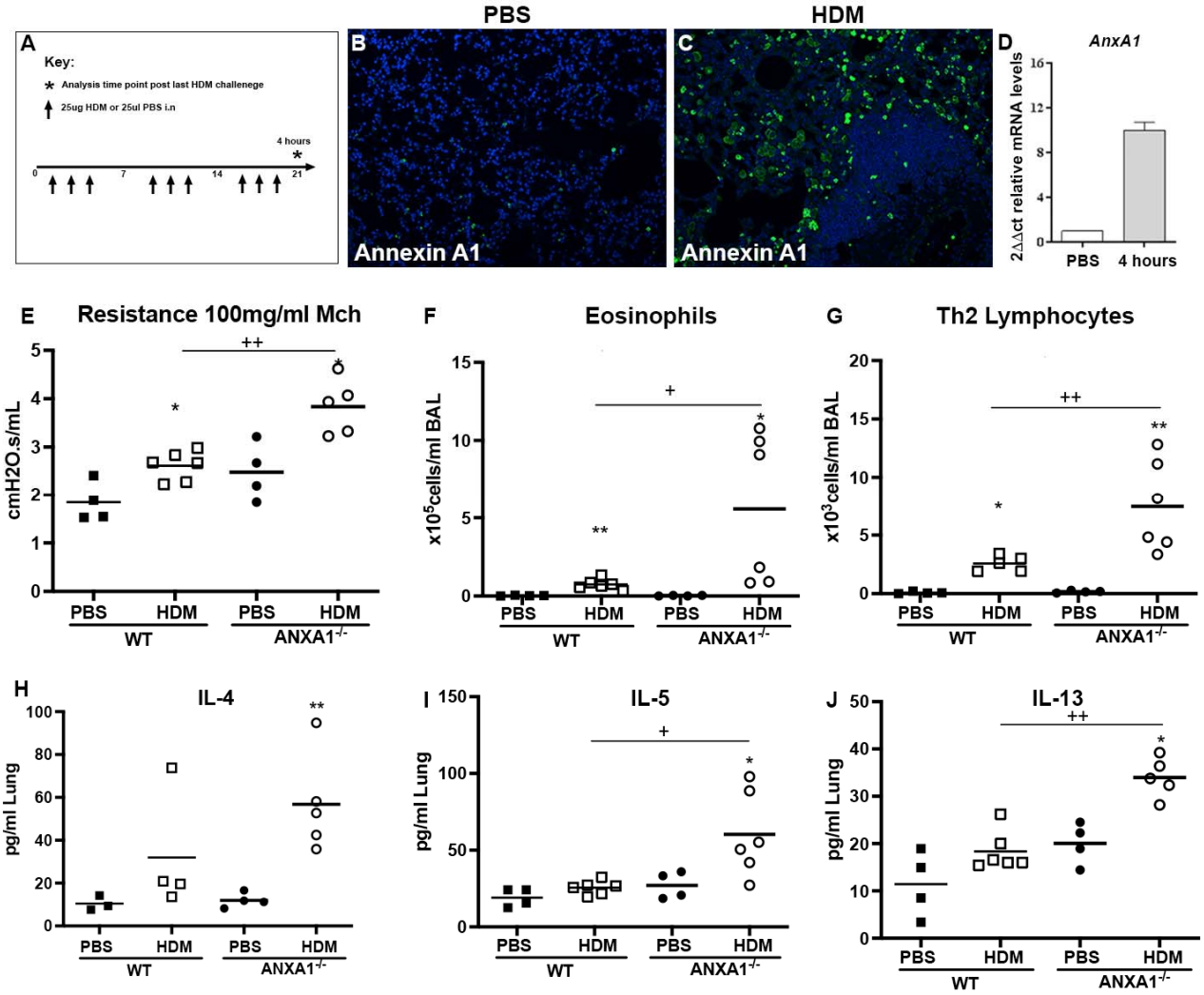


676 to -log10 p-value.

677

678

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



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

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

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

on behalf of UNICORN and Breathing Together Investigators

SUPPLEMENTARY INFORMATION

ONLINE METHODS

Description of cohorts

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

The Avon Longitudinal Study of Parents and Children (ALSPAC)

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

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

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

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

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

The Manchester Asthma and Allergy Study (MAAS)

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

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

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

ASHFORD

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

The Isle of Wight (IOW) cohort

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

Definition of variables

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

Demographic, exposures and outcomes

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

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

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

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

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

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

Primary outcome: Joint wheeze phenotypes

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

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

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

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

Minimising Bias and missing data effects

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

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

Genotyping and imputation

ALSPAC

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

MAAS

In MAAS, we used the Illumina 610 quad genome-wide SNP genotyping platform (Illumina Inc., San Diego, CA, USA). Prior to imputation samples were excluded on the basis of gender mismatches; minimal or excessive heterozygosity, genotyping call rates of <97%. SNPS were excluded if they had call rates of < 95%, minor allele frequencies of < 0.5% and HWE p<3x10-8. Prior to imputation each chromosome was pre-phased using EAGLE2 (v2.0.5)[11] as recommended by the Sanger imputation server[12]. We then imputed with PBWT[13] with the Haplotype Reference Consortium (release 1.1) of 32,470 reference genomes[12] using the Sanger Imputation Server.

IOW, SEATON and ASHFORD

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

Exclusions

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

GWAS Meta-analysis

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

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

LD clumping, pre-Selection and Gene Annotation

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

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

Gene expression in whole blood and lung tissues

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

Post-GWAS: rs75260654 (*ANXA1*)

Annotation & distribution

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

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

Promoter Capture

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

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

eQTL catalogue lookup

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

Variant effect

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

Functional mouse experiments

Mice

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

Airway hyperresponsiveness

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

Flow Cytometry Analysis

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

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

Analysis of cytokines and chemokines

Murine lung tissue homogenate supernatants were processed as previously described[28]. Cytokine levels were analysed by ELISA: IL-4, IL-5 (PharMingen, Oxford, UK), IL-13 Ready-Set-Go kits (eBioscience).

Real time-PCR

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

Annexin A1 Immunohistochemistry

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

Statistical analysis

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

Replication of ANXA1 top hits in PIAMA cohort

PIAMA cohort description

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

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

LCA wheezing phenotypes

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

Replication analyses

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

ONLINE RESULTS

Participants and descriptive data

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

Comparison of participants included and excluded from this analysis

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

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

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

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

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

GWAS Meta-Analysis: Main Results

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

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

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

Persistent Wheeze

We identified two GWAS-significant loci: 17q21, $p<5.5\times10^{-9}$, and a novel locus on 9q21.13 (*ANXA1*), $p<6.7\times10^{-9}$. The remaining 17 loci ($4.0\times10^{-7}\leq$p-values$\leq9.8\times10^{-6}$) included regions previously associated with childhood asthma (1q43, 4p16.3, 4q31.21, 5p15.31, 7q22.3, 17q12), asthma and rhinitis (2p25.1), eosinophil count (3q21.3, 10q24.2, 11q23.3, 22q11.1), bronchial hyper-responsiveness (2q12.2), lung function (1q43, 3q21.3, 5q13.3, 15q25.2, 19p13.2), triglycerides measurement and/or glucose metabolism (11q23.3, 19p13.2 and 22q11.1), severe asthma (14q22.1) and severe asthma and insulin resistance (11p15.4).

Early-onset Preschool Remitting Wheeze

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

Early-onset Mid-childhood Remitting Wheeze

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

Late-onset Wheeze

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

ONLINE TABLES

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

Variable	Definition
Cohort: ALSPAC	
Mother–asthma	Have you ever had asthma? (recruitment)
Mother smoking	Mother smoked when expecting (recruitment)
Doctor-diagnosed asthma ever	Has a doctor ever said that your child has asthma? (year 8 and 14)
Current wheezing	Two questions combined: Occurrence of wheezing and/or wheezing with whistling on the chest in the last 12 months (year ½, $2^{1/2}$, $4^{3/4}$, $8^{1/2}$ and 14)
Current asthma medication	Asthma medication in the last 12 months (year $8^{1/2}$ and 14)
Current rhinitis	Child had problem with sneezing/runny nose without cold/flu in last 12 months (year 7 and $16^{1/2}$)
Current hay-fever	Child had hay-fever in last 12 months (year $10^{1/2}$ and 14)
Cohort: MAAS	
Mother–asthma	Has a doctor ever told you that you had asthma? (recruitment)
Mother smoking	Do you smoke– mother (recruitment)
Doctor-diagnosed asthma ever	Has your doctor ever told you that your child has or had asthma? (year 8 & 16)
Asthma ever	Has your child ever suffered from asthma (year 8 and 16)
Current wheezing	Has your child had wheezing or whistling in the chest in the last 6/12 months (year 1, 3, 5, 8 and 16)
Current asthma medication	Asthma medication in the last 12 months (year 8 and year 16)
Current rhinitis	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)
Current hay-fever	Does your child have hay-fever now? (year 8 and year 16)
Cohort: SEATON	
Mother–asthma	Do you suffer from asthma? (recruitment)
Mother smoking	Which of the following best describes your smoking status? (recruitment)
Doctor-diagnosed asthma ever	Has your child ever suffered from asthma? If yes, has this been confirmed by a doctor? (year 10 & 15)
Asthma ever	Has your child ever suffered from asthma? (year 10) ; Have you ever suffered from asthma? (year 15)
Current wheezing	Has your child had wheezing in the chest in the last 12 months (year 1, 2, 5, 10 and 15)
Current asthma medication	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)

Current hay-fever	Has your child suffered from hay-fever last 12 months? (year 10 & 15)
Cohort: ASHFORD	
Mother–asthma	Do you have or have you ever been told you have asthma? (recruitment)
Mother smoking	Do you smoke cigarettes? (recruitment)
Doctor-diagnosed asthma ever	Has your doctor ever told you that your child has or had asthma? (year 8 & 14)
Asthma ever	In the past 12 months has your daughter suffered from asthma? (year 8); Has she/he ever suffered from asthma? (year 14)
Current wheezing	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)
Current asthma medication	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)
Current rhinitis	In the last twelve months has your child had a problem with sneezing or a runny or blocked nose? (year 8 & 14)
Current hay-fever	In your opinion does your child have hay fever now? year 8 Has your child ever had hay fever? year 14
Cohort: IOW	
Mother–asthma	Do you or have you suffered from asthma or wheezing (recruitment)
Mother smoking	Do you smoke in the house? (recruitment)
Doctor-diagnosed asthma ever	Asthma cared for by hospital specialist/ GP or nurse (year 10, 18 & 26)
Asthma ever	Child ever had asthma (year 10 and 18)
Current wheezing	Presence of wheeze since previous review (year 1, 2, 4, 10 and 18)
Asthma medication ever	Child ever had asthma treatment (year 18 year) combined with asthma treatment questions being asked at year 1, 2, 4, 10 and 18
Current rhinitis	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

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

399

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

	ALSPAC		**MAAS**		**SEATON**		**ASHFORD**		**IOW**	
	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)	
Wheeze Phenotypes										
Never/Infrequent	4331 (63.4)		506 (57.1)		332 (60.6)		145 (41.7)		573 (60.2)	
Early-onset persistent	656 (9.6)		133 (15.0)		36 (6.6)		41 (11.8)		77 (8.1)	
Early-onset pre-school remitting	1076 (15.8)		145 (16.4)		117 (21.4)		145 (41.7)		0	
Early-onset mid-childhood remitting	474 (6.9)		48 (5.4)		13 (2.4)		13 (3.7)		55 (5.8)	
Late-onset	296 (4.3)		55 (6.2)		50 (9.1)		4 (1.2)		247 (26.0)	
	7-8 years	**14-15 years**	**8 years**	**16 years**	**10 years**	**15 years**	**8 years**	**14 years**	**10 years**	**18 years**
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)

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

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

406

ALSPAC	**N**	**Included**	**N**	**Excluded**	**P-value**					
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					
	At 7.5-8.5 years					**At 14-15 years**				
ALSPAC	**N**	**Included**	**N**	**Excluded**	**P-value**	**N**	**Included**	**N**	**Excluded**	**P-value**
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
MAAS	**N**	**Included**	**N**	**Excluded**	**P-value**					
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					
	At 8 years					**At 16 years**				
MAAS	**N**	**Included**	**N**	**Excluded**	**P-value**	**N**	**Included**	**N**	**Excluded**	**P-value**
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
SEATON	**N**	**Included**	**N**	**Excluded**	**P-value**					
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**				
SEATON	**N**	**Included**	**N**	**Excluded**	**P-value**	**N**	**Included**	**N**	**Excluded**	**P-value**
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
ASHFORD	**N**	**Included**	**N**	**Excluded**	**P-value**					
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**				
ASHFORD	**N**	**Included**	**N**	**Excluded**	**P-value**	**N**	**Included**	**N**	**Excluded**	**P-value**
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
Isle Of Wight	**N**	**Included**	**N**	**Excluded**	**P-value**					
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**				
IOW	**N**	**Included**	**N**	**Excluded**	**P-value**	**N**	**Included**	**N**	**Excluded**	**P-value**
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

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

CHR	SNP	BP	short-listed after inspection of locus zoom plot
Persistent Wheezing			
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**
CHRM3	1q43	**FEV_1, FEV_1/FVC**, **asthma**- high priority drug target	Patel, K.R. *et al.* Targeting acetylcholine receptor M3 prevents the progression of airway hyperreactivity in a mouse model of childhood asthma. *FASEB J* **31**, 4335-4346 (2017).
RNF144A	2p25.1	**asthma**, allergy, childhood onset **asthma**, **allergic rhinitis**	Schoettler, N. *et al.* Advances in asthma and allergic disease genetics: Is bigger always better? *J Allergy Clin Immunol* **144**, 1495-1506 (2019).
FHL2	2q12.2	**bronchial hyper-responsiveness**, airway inflammation, novel gene associated with **asthma** severity in human	Kurakula, K. *et al.* Deficiency of FHL2 attenuates airway inflammation in mice and genetic variation associates with human bronchial hyper-responsiveness. *Allergy* **70**, 1531-44 (2015).
RAB7A	3q21.3	**eosinophil count**	GeneCards
RAB43	3q21.3	response to **bronchodilator**, **FEV_1/FEC ratio**	GWAS Catalog
RNF212, IDUA, DGKQ, SLC26A1	4p16.3	**asthma**	Gautam, Y. *et al.* Comprehensive functional annotation of susceptibility variants associated with asthma. *Hum Genet* **139**, 1037-1053 (2020).
INPP4B	4q31.21	atopic **asthma**	Sharma, M. *et al.* A genetic variation in inositol polyphosphate 4 phosphatase a enhances susceptibility to asthma. *Am J Respir Crit Care Med* **177**, 712-9 (2008).
ADCY2	5p15.31	**asthma** × air pollution, childhood **asthma**	Gref, A. *et al.* Genome-Wide Interaction Analysis of Air Pollution Exposure and Childhood Asthma with Functional Follow-up. *Am J Respir Crit Care Med* **195**, 1373-1383 (2017).
CDHR3	7q22.3	childhood **asthma**	Everman, J.L. *et al.* Functional genomics of CDHR3 confirms its role in HRV-C infection and childhood asthma exacerbations. *J Allergy Clin Immunol* **144**, 962-971 (2019).
ANXA1	9q21.13	**FEV_1/FVC**, response to **bronchodilators** in smokers	Lutz, S.M. *et al.* A genome-wide association study identifies risk loci for spirometric measures among smokers of European and African ancestry. *BMC Genet* **16**, 138 (2015).
ANXA1	9q21.13	anti-inflammatory properties, strongly expressed in **bronchial mast cells**	Vieira Braga FA *et al.* A cellular census of human lungs identifies novel cell states in health and in asthma. (2019).

ANXA1	9q21.13	potentially involved in epithelial **airway repair**	Leoni, G. *et al.* Annexin A1-containing extracellular vesicles and polymeric nanoparticles promote epithelial wound repair. *J Clin Invest* **125**, 1215-27 (2015).
R3HCC1L	10q24.2	atopic **eczema**, psoriasis	GWAS Catalog
R3HCC1L	10q24.2	**eosinophil count**, BMI	GeneCards
TRIM5, TRIM6, TRIM22	11p15.4	severe **asthma** and insulin resistance	Kimura, T. *et al.* Precision autophagy directed by receptor regulators - emerging examples within the TRIM family. *J Cell Sci* **129**, 881-91 (2016).
SIK3	11q23.3	triglycerides, glucose metabolism, **eosinophil count**	Sun, Z. et al. The potent roles of salt-inducible kinases (SIKs) in metabolic homeostasis and tumorigenesis. *Sig Transduct Target Ther* **5**(2020).
KTN1	14q22.1	severe **asthma**	Bigler, J. *et al.* A Severe Asthma Disease Signature from Gene Expression Profiling of Peripheral Blood from U-BIOPRED Cohorts. *Am J Respir Crit Care Med* **195**, 1311-1320 (2017).
FAM227B	5q13.3	rs35251997 and **FEV_1, FEV_1/FVC**	Shrine, N. *et al.* New genetic signals for lung function highlight pathways and chronic obstructive pulmonary disease associations across multiple ancestries. *Nat Genet* **51**, 481-493 (2019).
ADAMTSL3	15q25.2	**FEV1/FVC**	Sakornsakolpat, P. *et al.* Genetic landscape of chronic obstructive pulmonary disease identifies heterogeneous cell-type and phenotype associations. *Nat Genet* **51**, 494-505 (2019).
IKZF3, GSDMB, LRRC3C, GSDMA	17q12	early-onset **asthma**, **persistent wheezing** (chr17q12-q21)	Granell R *et al.* Examination of the relationship between variation at 17q21 and childhood wheeze phenotypes. J Allergy Clin Immunol. 2013 Mar;131(3):685-94.
MARCH2, HNRNPM, MYO1F	19p13.2	triglycerides, HDL-cholesterol, metabolic syndrome	Sajuthi, S.P. *et al.* Genetic regulation of adipose tissue transcript expression is involved in modulating serum triglyceride and HDL-cholesterol. *Gene* **632**, 50-58 (2017).
MYO1F	19p13.2	**FEV_1 and FVC**	GeneCards
CECR5	22q11.1	triglycerides, **eosinophil count** and body height	Liu, D.J. *et al.* Exome-wide association study of plasma lipids in >300,000 individuals. *Nat Genet* **49**, 1758-1766 (2017).

413

414

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

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

415

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

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

416

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

AMMECR1L	2q14.3	**body height**, blood protein, growth, bone, and heart alterations	Moyses-Oliveira, M. *et al.* Inactivation of AMMECR1 is associated with growth, bone, and heart alterations. *Hum Mutat* **39**, 281-291 (2018).
RARB	3p24.2	**FEV_1/FVC**, adult lung function	Collins, S.A. *et al.* HHIP, HDAC4, NCR3 and RARB polymorphisms affect fetal, childhood and adult lung function. *Eur Respir J* **41**, 756-7 (2013).
KIAA2018, NAA50, SIDT1, CD200	3q13.2	SIDT1: **FEV_1/FVC**, CD200: **adult-onset non-allergic asthma**	Siroux, V. *et al.* Genetic heterogeneity of asthma phenotypes identified by a clustering approach. *Eur Respir J* **43**, 439-52 (2014).
TFDP2, XRN1	3q23	XRN1: **eosinophil count**, 3q23: **allergic disease and atopic sensitisation**	Freidin, M.B. *et al.* [Genome-wide association study of allergic diseases in Russians of Western Siberia]. *Mol Biol (Mosk)* **45**, 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, **smoking**, GIMAP family: autoimmune diabetes, **asthma and allergy**	Heinonen, M.T. *et al.* GIMAP GTPase family genes: potential modifiers in autoimmune diabetes, asthma, and allergy. *J Immunol* **194**, 5885-94 (2015).
PTPRN2	7q36.3	**eczema**	Bogari, N.M. *et al.* Whole exome sequencing detects novel variants in Saudi children diagnosed with eczema. *J Infect Public Health* **13**, 27-33 (2020).
LOC105375631	8q21.3	8q21.3: **type 1 diabetes**	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. *PLoS One* **13**, e0192696 (2018).
NFIB, ZDHHC21	9p22.3	9p22.3: **asthma** (mean age<16 years)	Denham, S. *et al.* Meta-analysis of genome-wide linkage studies of asthma and related traits. *Respir Res* **9**, 38 (2008).
SLC16A12, IFIT family, PANK1	10q23.31	SLC16A12: Body height, PANK1: **insulin resistance**	Yang, L. *et al.* P53/PANK1/miR-107 signalling pathway spans the gap between metabolic reprogramming and insulin resistance induced by high-fat diet. *J Cell Mol Med* **24**, 3611-3624 (2020).

CBL, CCDC84, MCAM	11q23.3	CBL: **Sex hormone-binding globulin levels**; MCAM: Blood protein levels	GWAS Catalog
SPG11, CTDSPL2	15q15.3-q21.1	CTDSPL2: **alcohol drinking**	GWAS Catalog
ADAMTSL3, GOLGA6L4, UBE2Q2P8	15q25.2	ADAMTSL3: **FEV_1/FVC**, lean mass	Karasik, D. *et al.* Disentangling the genetics of lean mass. *Am J Clin Nutr* **109**, 276-287 (2019).
SSTR5-AS1, CACNA1H	16p13.3	CACNA1H: **eosinophil count**	GWAS Catalog
GSG1L	16p12.1	16p12.1: current **asthma** and rhino-conjunctivitis at 10-15 years	Sottile, G. *et al.* An association analysis to identify genetic variants linked to asthma and rhino-conjunctivitis in a cohort of Sicilian children. *Ital J Pediatr* **45**, 16 (2019).
FAM19A5 or TAFA5	22q13.32	**Obesity and Metabolic Dysfunction**	Recinella L. et al. Adipokines: New Potential Therapeutic Target for Obesity and Metabolic, Rheumatic, and Cardiovascular Diseases. Front Physiol. 2020 Oct 30;11:578966

417

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	---

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

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)

Table E9. Selected immune eQTLs of rs75260654

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

Table E10. Lung eQTLs of rs75260654

Rsid	P value	beta	se	an	symbol	Study
rs116849664	0.0489	0.22	0.11	620	*ANXA1*	GTEx_exon_lung
rs78320984	0.0489	0.22	0.11	620	*ANXA1*	GTEx_exon_lung

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

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

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

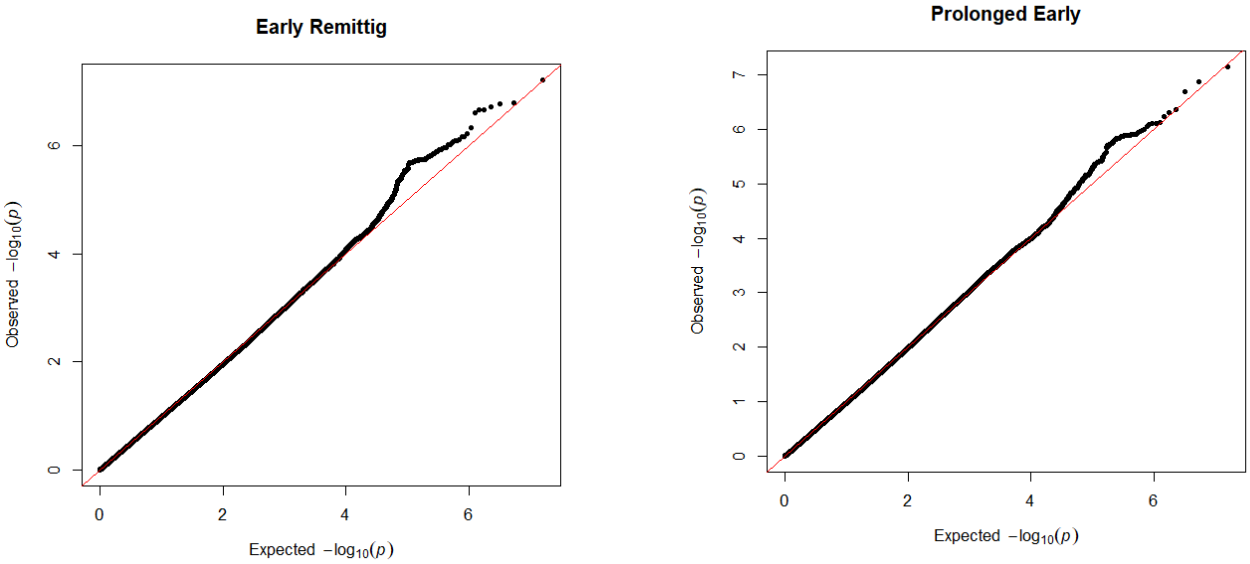
				CW (40) vs NI (1557)		
Rsid	chr:position	R2	A2/freqA2	beta	se	p-value
rs75260654	9:75788108	0.60	c/0.02	-0.287	0.91	0.75
rs116849664	9:75820902	0.61	c/0.02	0.119	1.08	0.91
rs78320984	9:75844302	0.59	g/0.02	0.125	1.04	0.90

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

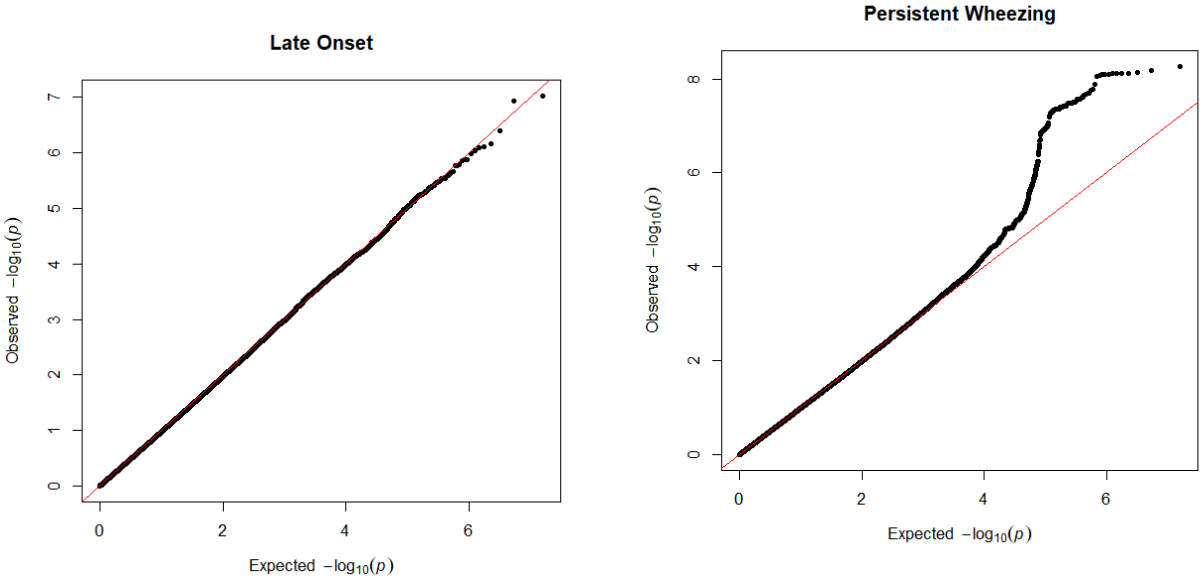
444 **SUPPLEMENTARY FIGURES**

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

446

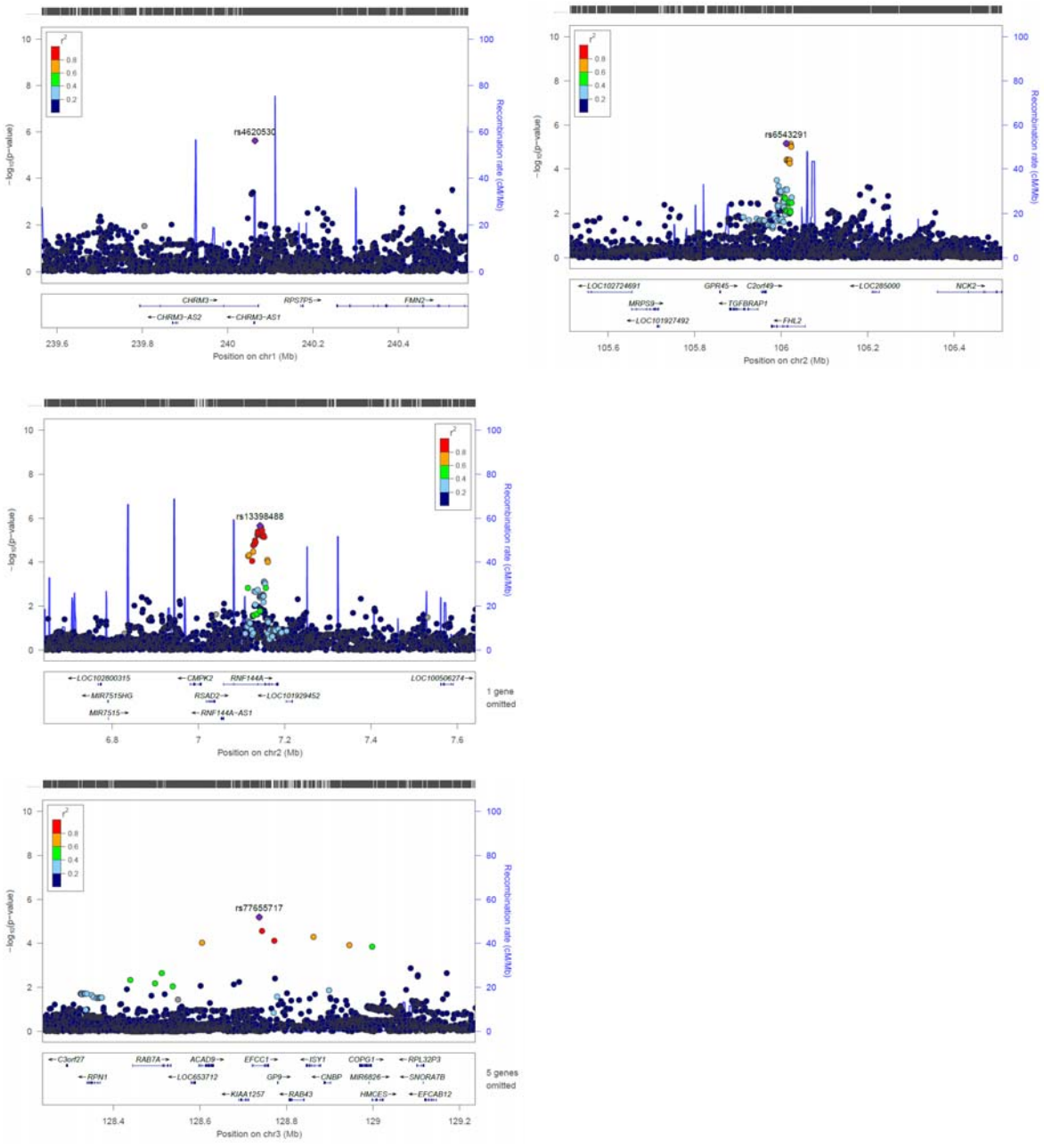


447



448

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

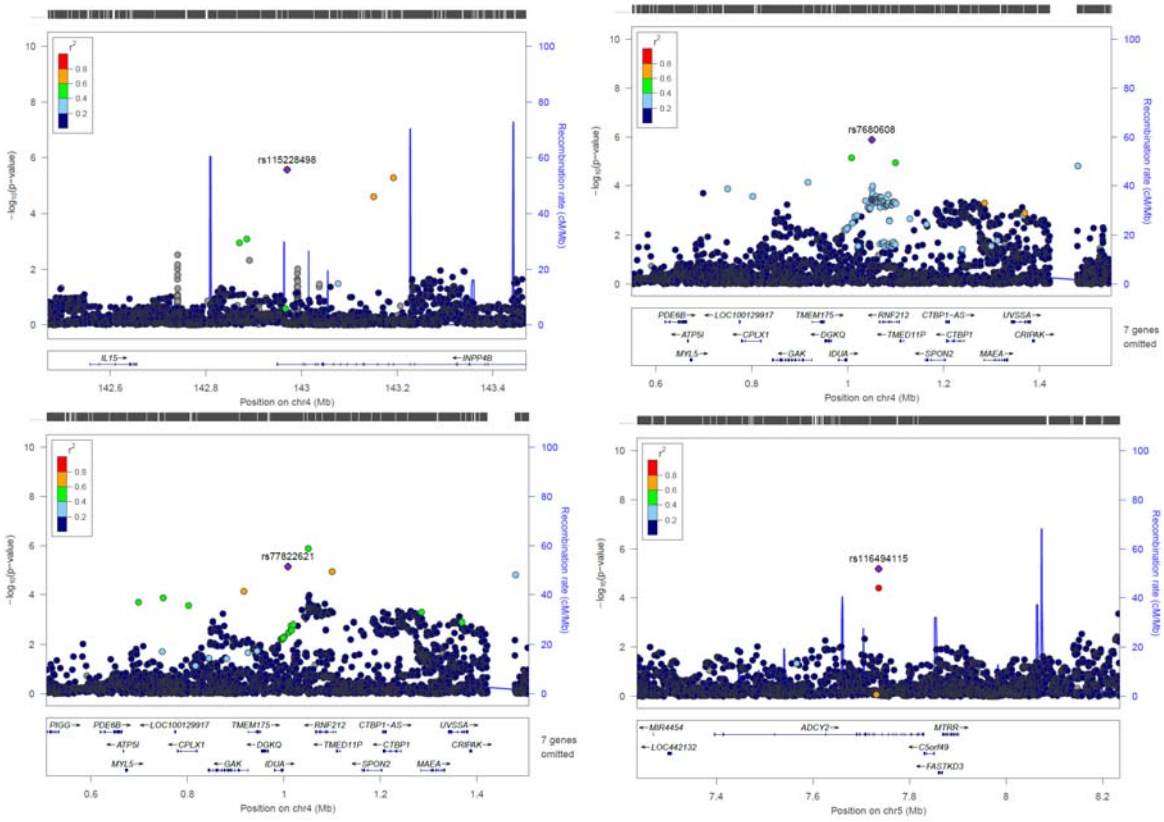


450

451

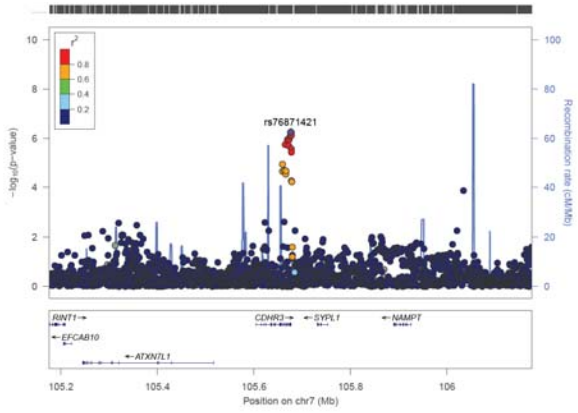
452

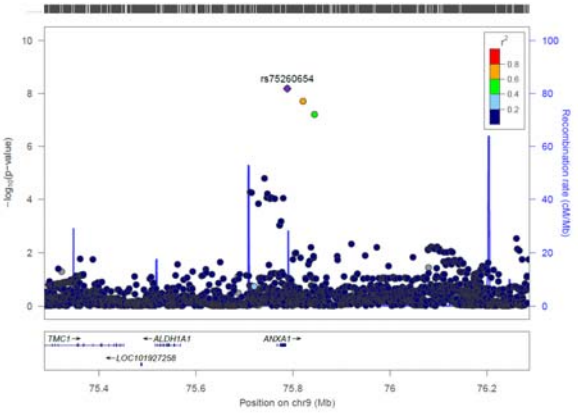
453



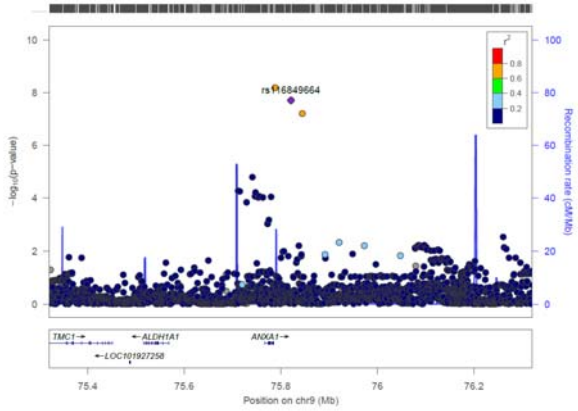
454

455

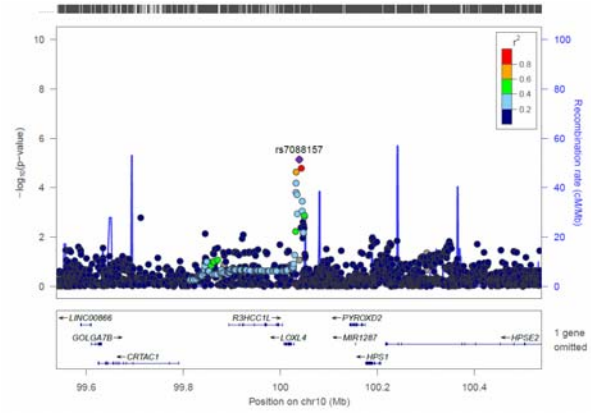




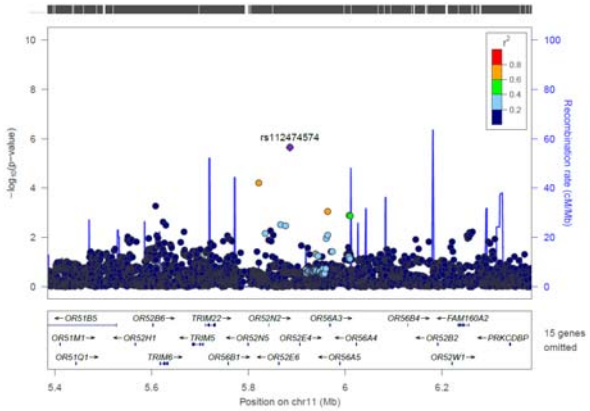
456



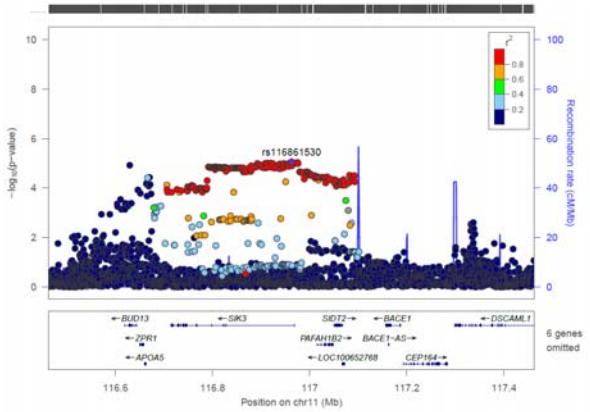
457



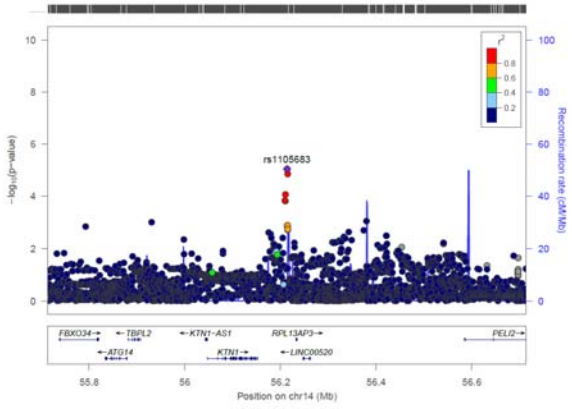
458

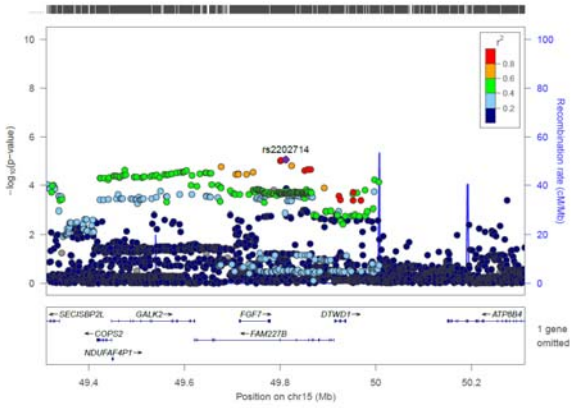


459

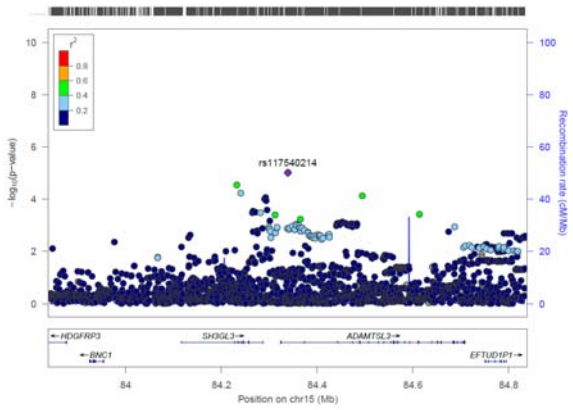


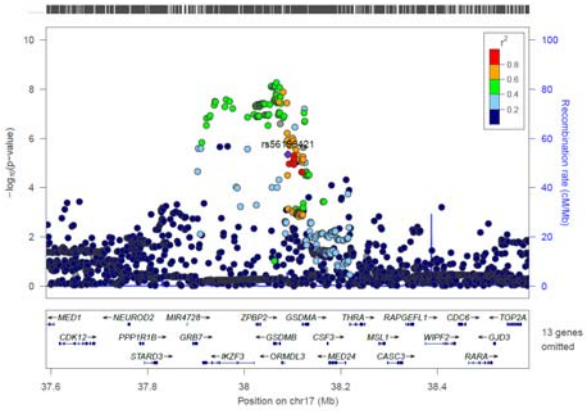
460



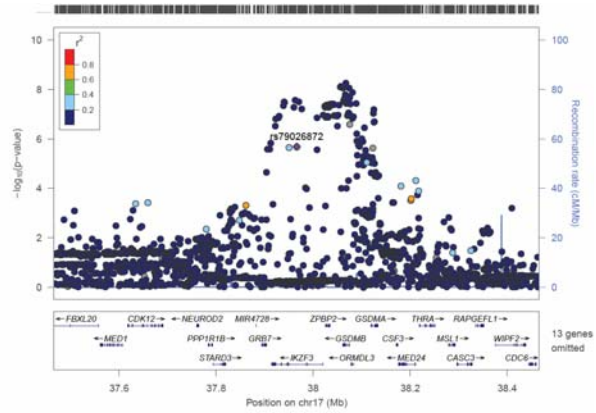


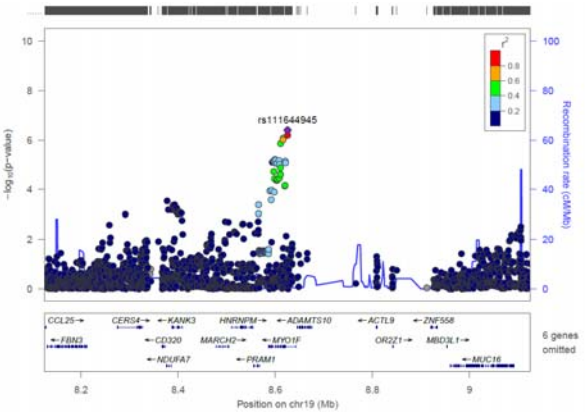
461



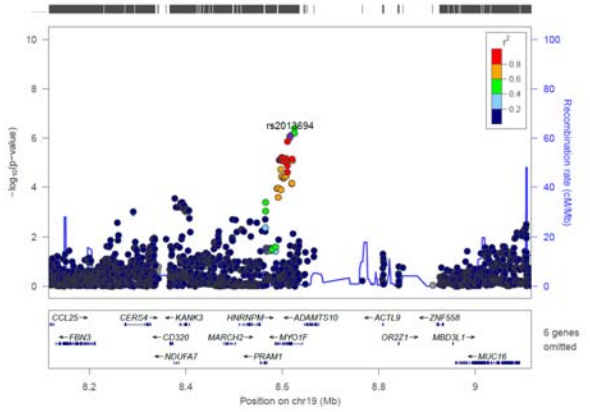


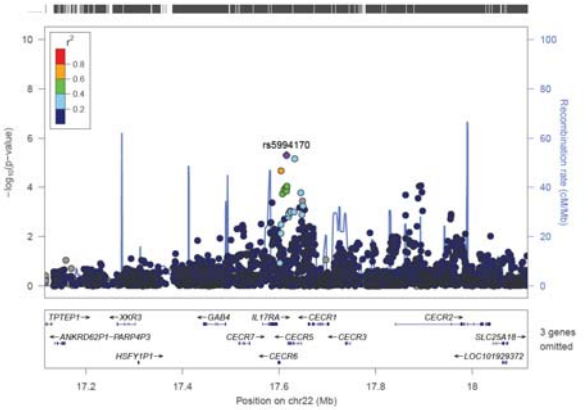
462





463





464

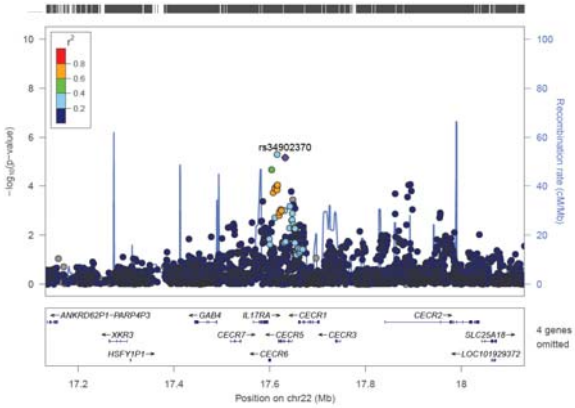
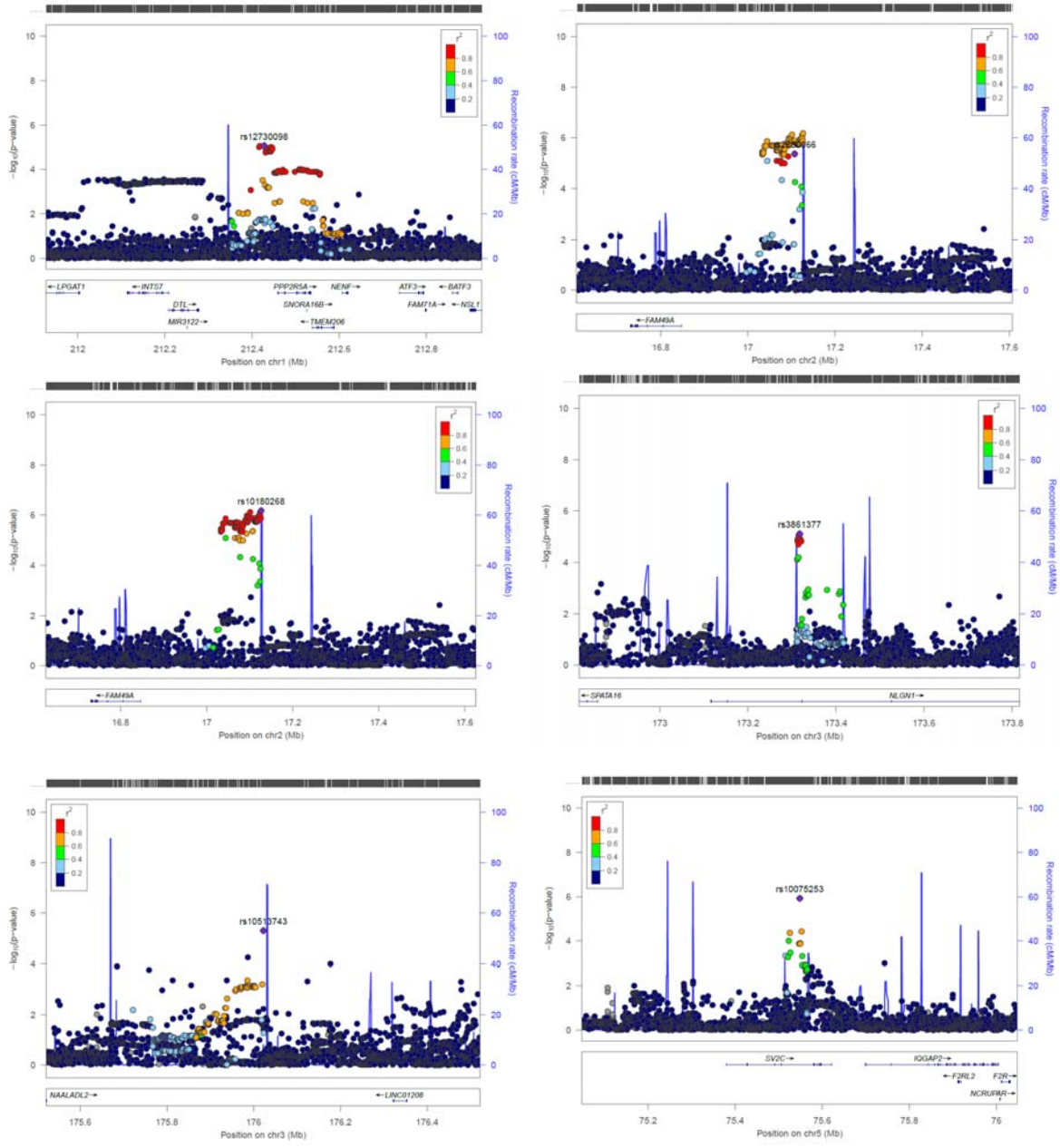
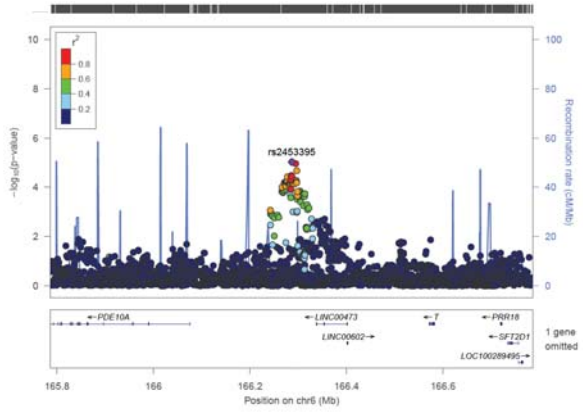
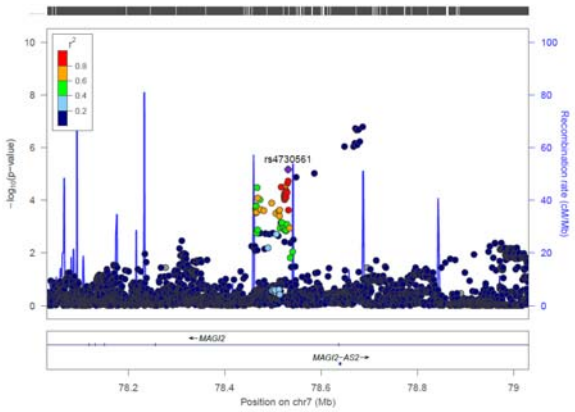


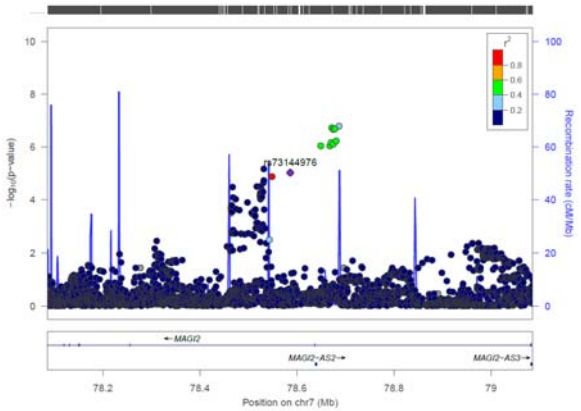
Figure E3. Zoom locus plots for short-listed independent top hits for Early-onset Pre-school Remitting Wheezing



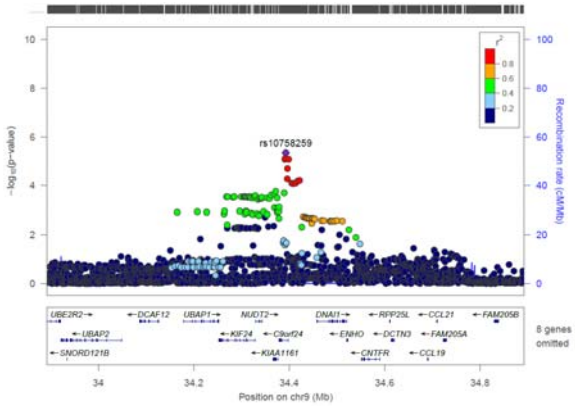




470

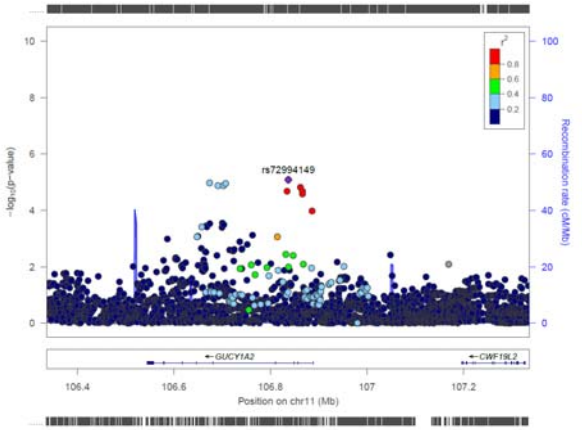


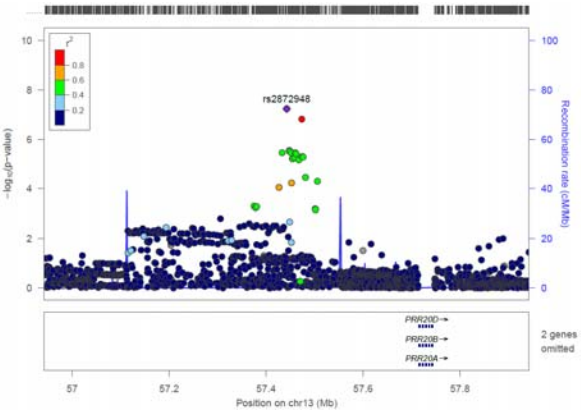
471



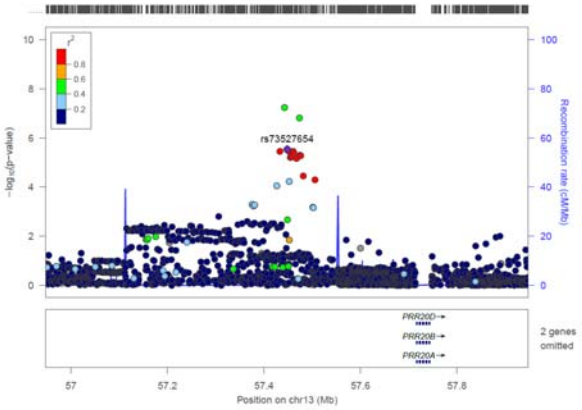
472

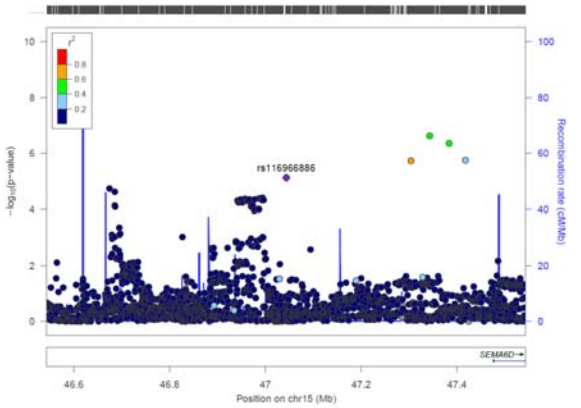
473



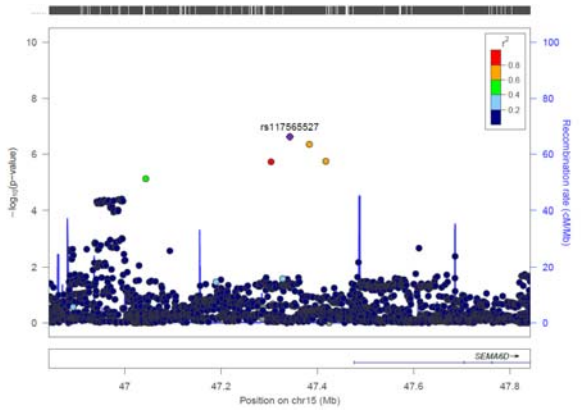


474



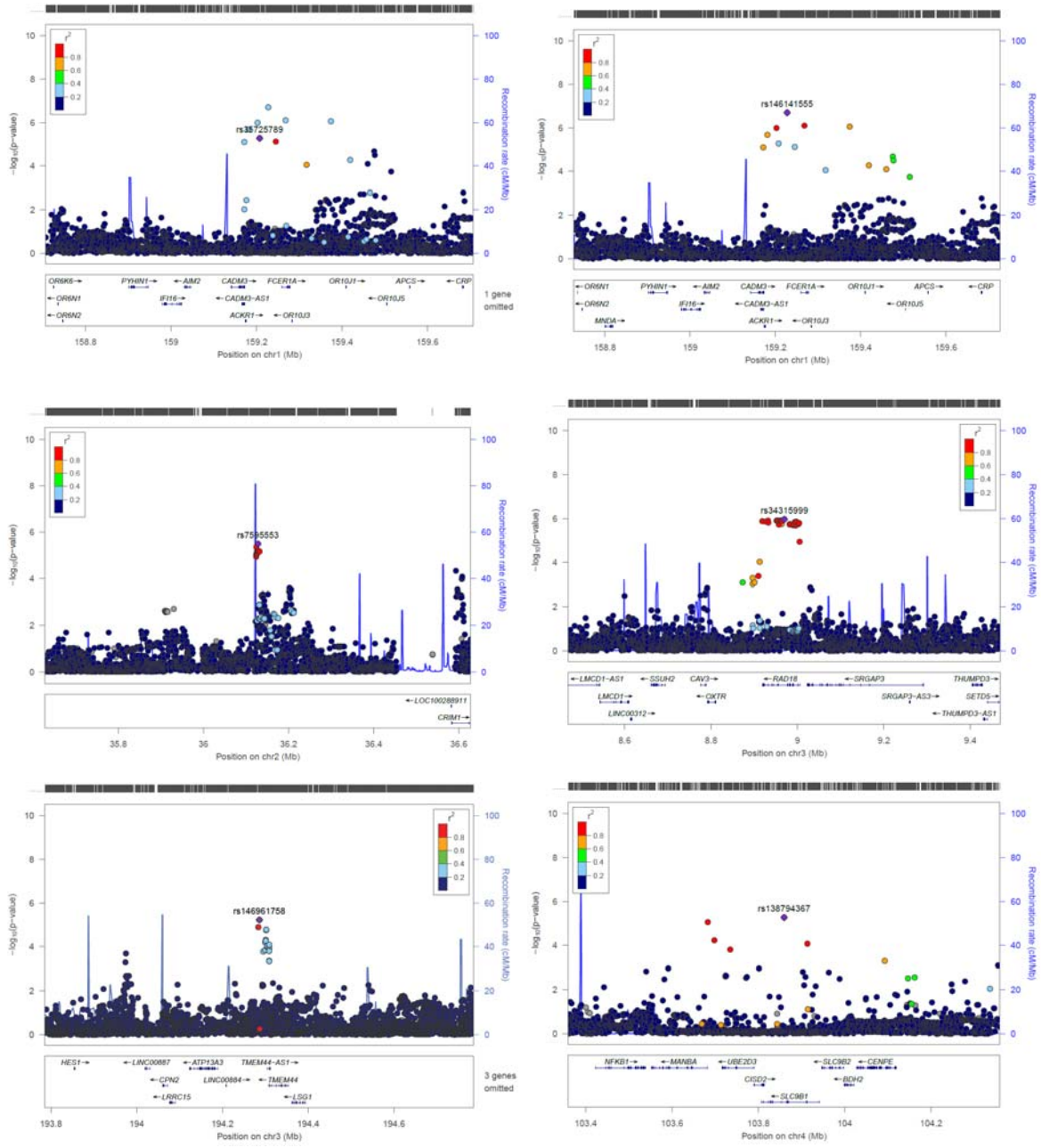


475

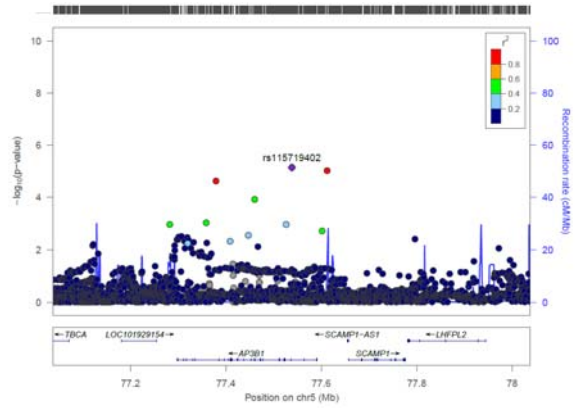


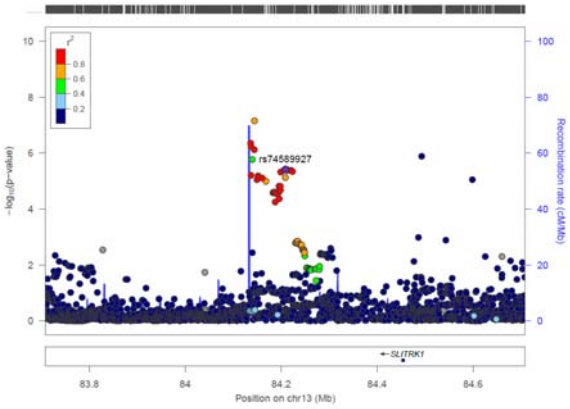
476

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

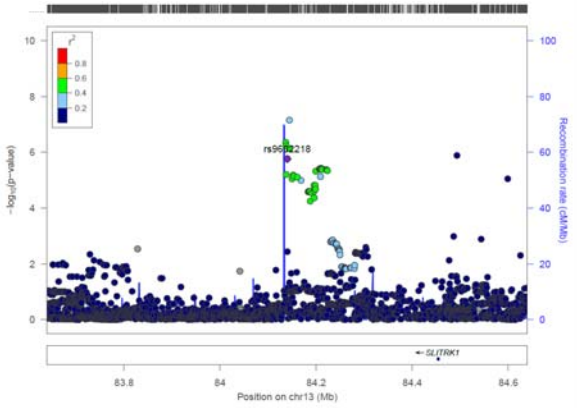


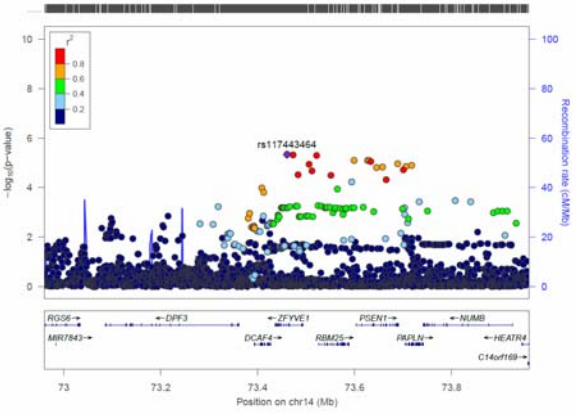
482



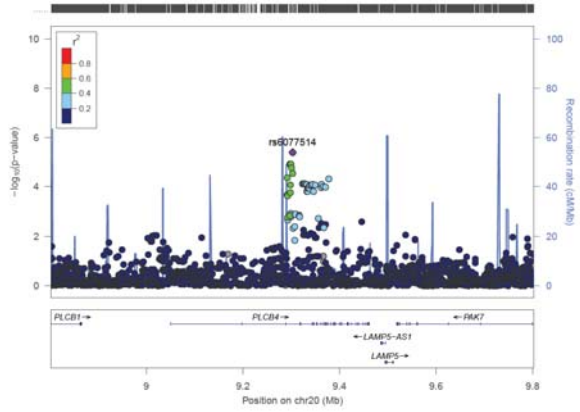


483



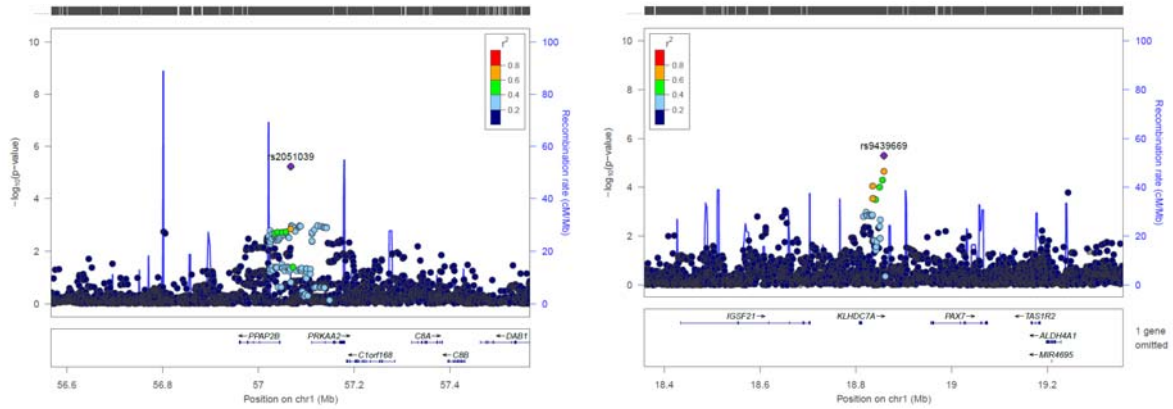


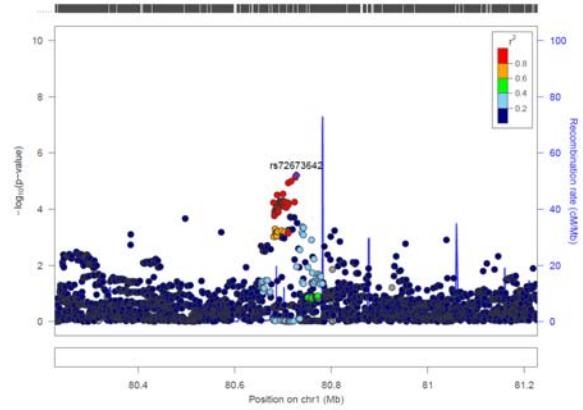
484

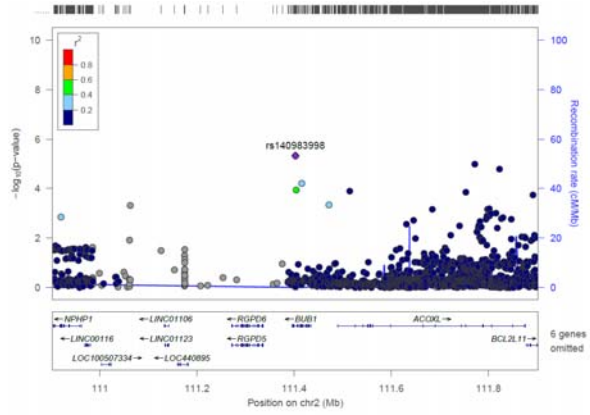


485

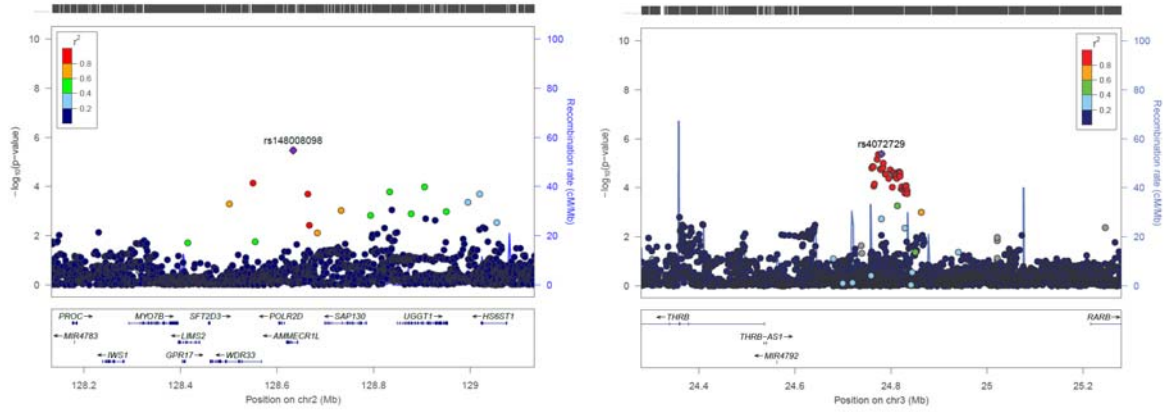
Figure E5. Zoom locus plots for short-listed independent top hits for Late-onset Wheezing



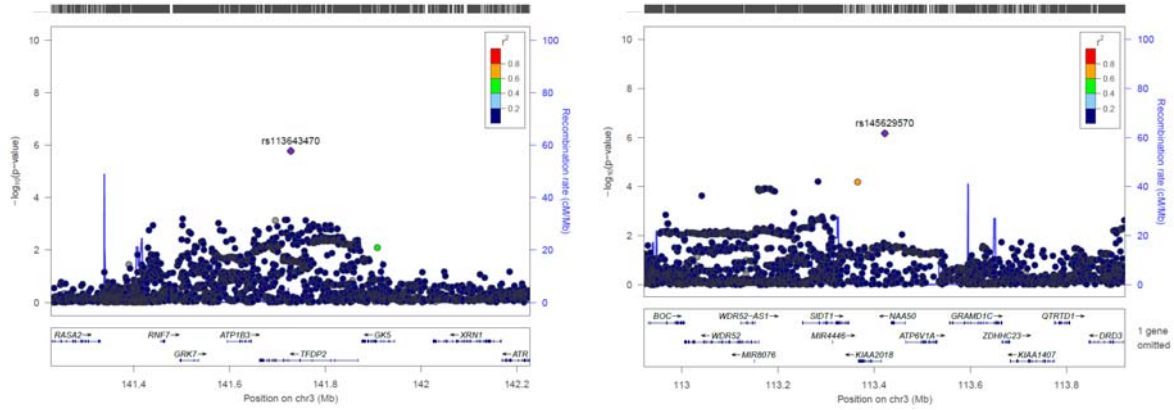




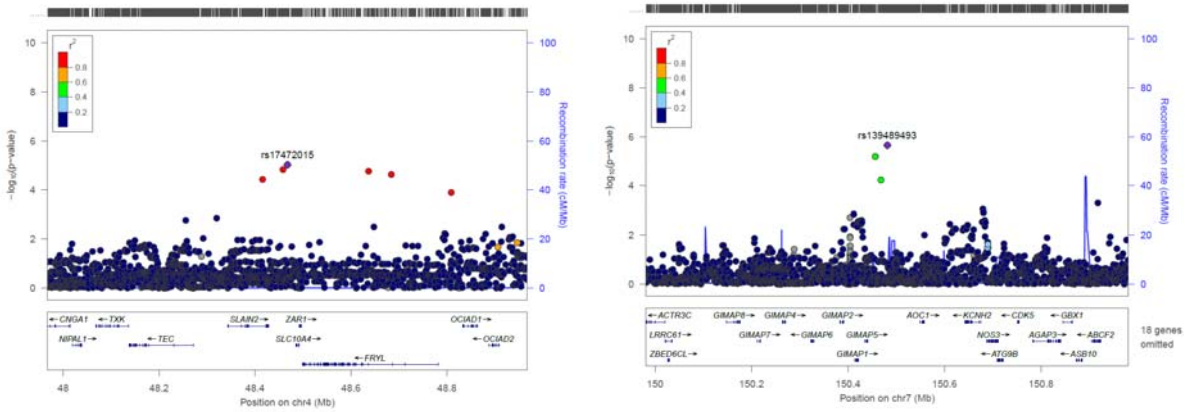
490



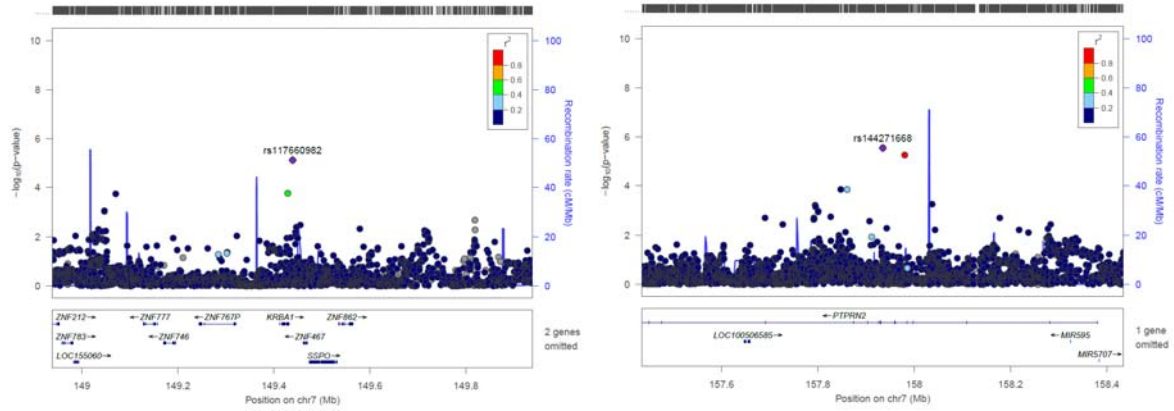
491



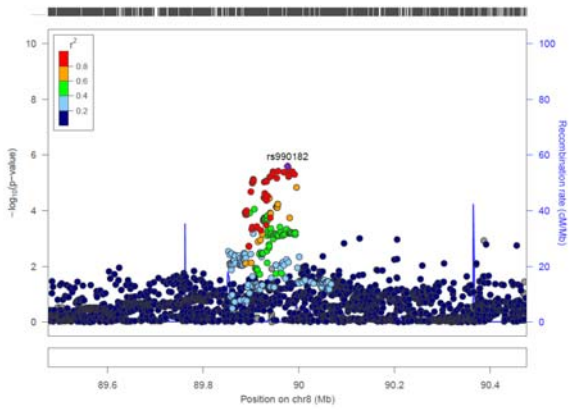
492

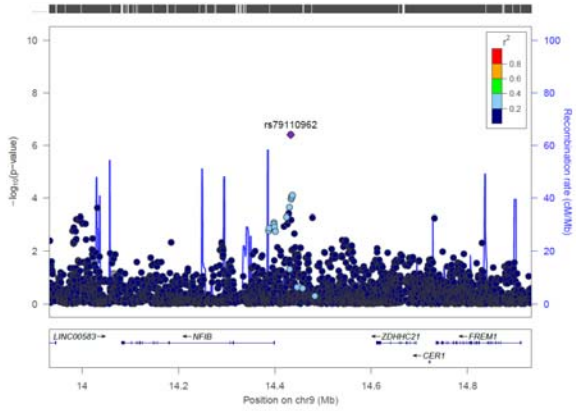


493

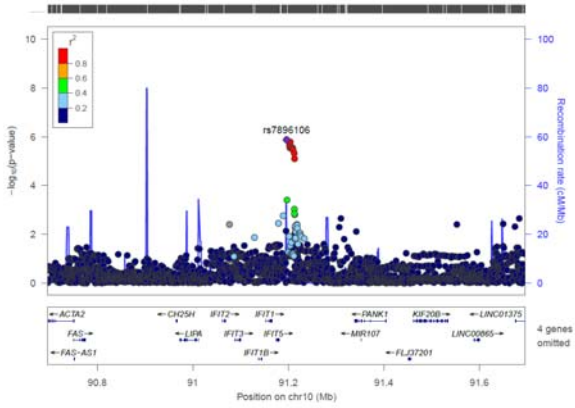


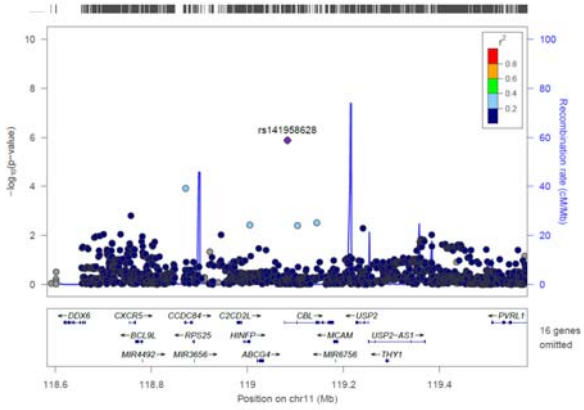
494



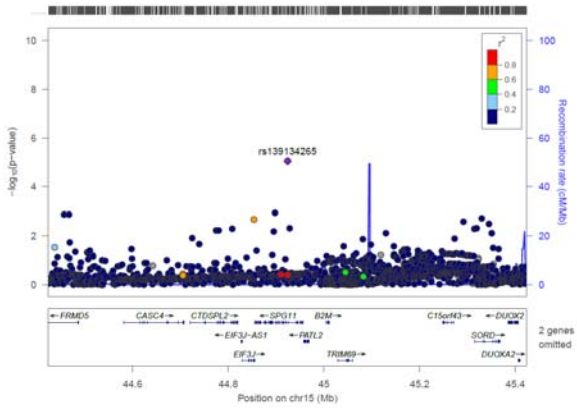


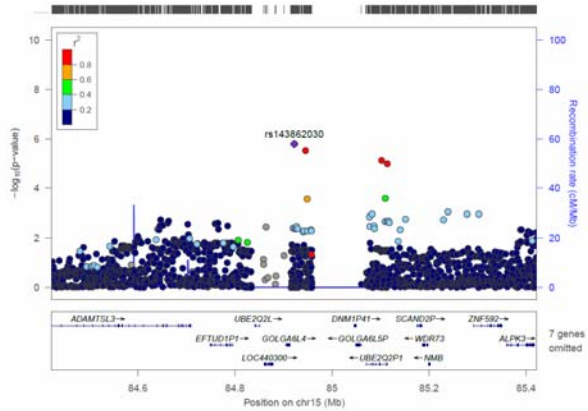
495





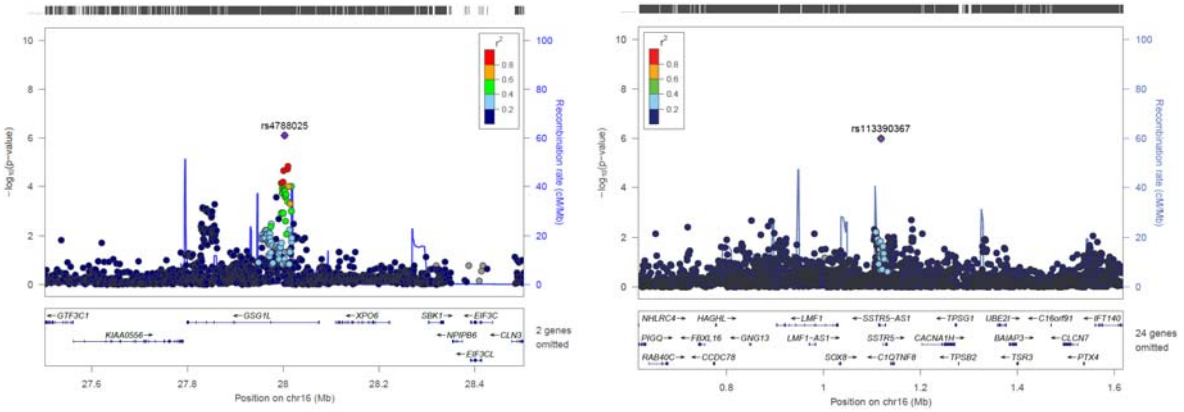
496





497

498



499

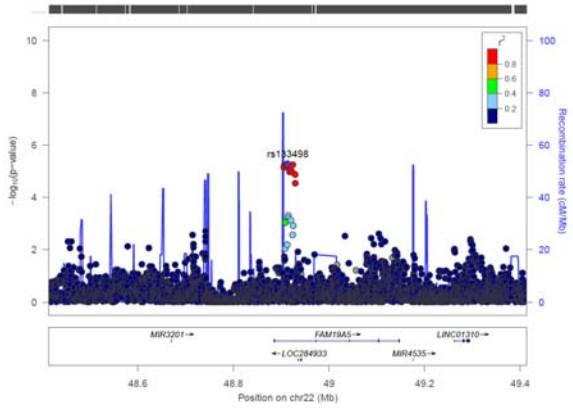


Figure E6. Linkage disequilibrium between SNPs downstream of *ANXA1* that were associated with persistent wheeze.

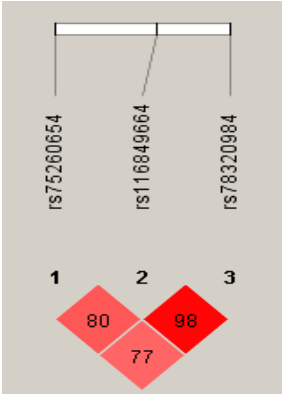
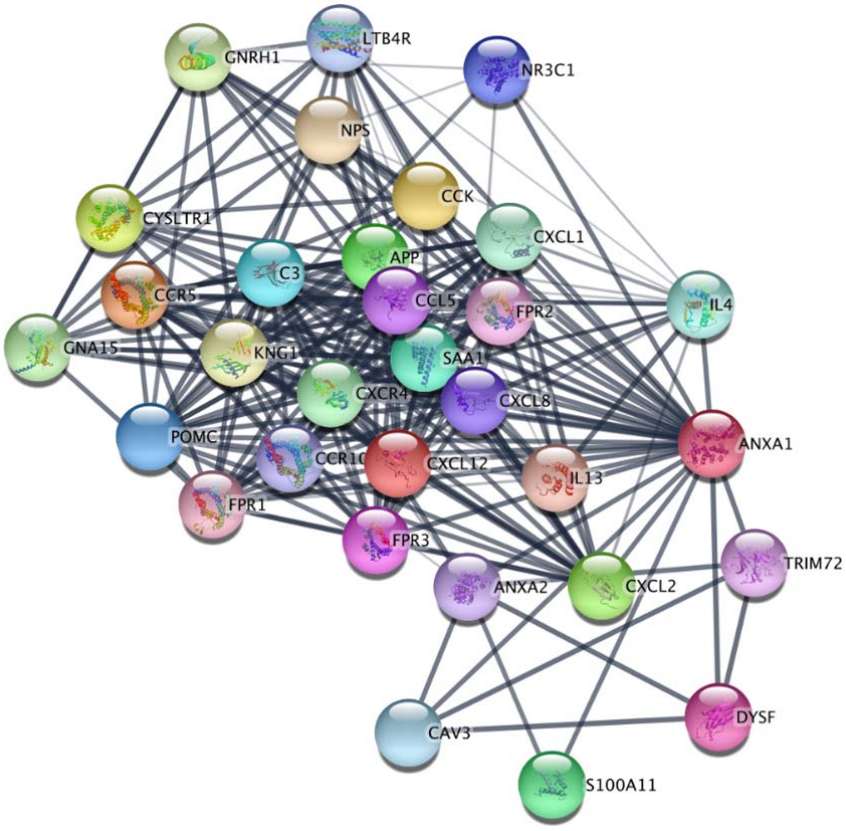


Figure E7. *ANXA1* interactors. Protein-protein interaction of *ANXA1* including *IL-4, IL-13* and *NR3C1*



REFERENCES

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