Phenotypic and functional translation of *IL33* genetics in asthma

Maria E. Ketelaar[1ab,2,] MSc MD, Michael A. Portelli[2,*] PhD, F. Nicole Dijk[1a,*] MD, PhD, N. Shrine[3,4] PhD, A. Faiz[1c] PhD, C.J. Vermeulen[1c] PhD, C.J. Xu PhD[1a,5], J. Hankinson[6] PhD, S. Bhaker[2] PhD, A.P. Henry[2] PhD, C.K. Billington[2] PhD, D.E. Shaw[2] MD PhD, Prof. S.R. Johnson[2] MBBS DM, A.V. Benest[7] PhD, V. Pang[7] PhD, Prof. D. Bates[7], Z.E.K. Pogson[8,9]MD PhD, A. Fogarty[9] PhD, Prof. T. M. McKeever[9] PhD, A. Singapuri[10] BSc, Prof. L. Heaney[11] MD PhD, A.H. Mansur[12] PhD, R. Chaudhuri[13] MD, Prof. N.C. Thomson[13], Prof. J.W. Holloway[14], G.A. Lockett[14] PhD, Prof. P.H. Howarth[14] MD PhD, R. Niven[6] MD, Prof. A. Simpson[6] MD PhD, Prof. M.D. Tobin[3,4], Prof. I.P. Hall[2] FMedSci, Prof. L.V. Wain[3,4], J.D. Blakey[15] PhD, Prof. C.E. Brightling[4,10] MD PhD, M. Obeidat[16] PhD, Prof. D.D. Sin[16,17] MD PhD, C. Nickle[18] PhD, Y. Bossé[19] PhD, J.M. Vonk[1d] PhD, M. van den Berge[1c] MD PhD, Prof. G.H. Koppelman[1a] MD PhD, Prof. Ian Sayers[2#], Martijn C. Nawijn[1b,#] PhD*

[1]University of Groningen, University Medical Center Groningen (UMCG), Groningen Research Institute for Asthma and COPD, Groningen, The Netherlands
[a]University of Groningen, UMCG, Department of Pediatric Pulmonology and Pediatric Allergology, Beatrix Children's Hospital, Groningen, the Netherlands
[b]University of Groningen, UMCG, Department of Pathology and Medical Biology, Groningen, The Netherlands
[c]University of Groningen, UMCG, Department of Pulmonary Diseases, Groningen, The Netherlands
[d]University of Groningen, UMCG, Department of Epidemiology, Groningen, The Netherlands
[2]Division of Respiratory Medicine, National Institute for Health Research Nottingham Biomedical Research Centre, Nottingham University Biodiscovery Institute, University of Nottingham, Nottingham, UK
[3]Department of Health Sciences, University of Leicester, Leicester, UK
[4]National Institute for Health Research Leicester Respiratory Biomedical Research Centre, University of Leicester, Leicester, UK
[5]CiiM & TWINCORE , joint ventures between the Helmholtz-Centre for Infection Research (HZI) and the Hannover Medical School (MHH), Hannover, Germany
[6]Manchester Academic Health Science Centre, University of Manchester, Manchester, UK
[7]Tumour Vascular Biology Laboratories, Division of Cancer and Stem Cells, School of Medicine, Queens Medical Centre, Nottingham, UK and COMPARE University of Birmingham and University of Nottingham, Nottingham, UK.
[8]Department of Respiratory Medicine, Lincoln County Hospital, Lincoln, UK
[9]Division of Epidemiology and Public Health, University of Nottingham, Nottingham, UK
[10]Institute for Lung Health, Department of Respiratory Sciences, University of Leicester, Glenfield Hospital, Leicester, UK
[11]Centre for Infection and Immunity, Queen's University of Belfast, Belfast, UK
[12]Respiratory Medicine, Birmingham Heartlands Hospital and University of Birmingham, Birmingham, UK
[13]Institute of Infection, Immunity and Inflammation, University of Glasgow, Glasgow, UK
[14]Human Development & Health & Clinical and Experimental Sciences, Faculty of Medicine and National Institute of Health Biomedical Research Centre, University of Southampton, Southampton, UK
[15]Respiratory Medicine, Sir Charles Gairdner Hospital, Perth, Australia
[16]The University of British Columbia Centre for Heart Lung Innovation, St Paul's Hospital, Vancouver, BC, Canada.
[17]Division of Respiratory Medicine, Department of Medicine, University of British Columbia, Vancouver, BC, Canada.
[18]Departments of Genetics and Pharmacogenomics, Merck Research Laboratories, Boston, Massachusetts, USA
[19]Institut universitaire de cardiologie et de pneumologie de Québec, Department of Molecular Medicine, Laval University, Québec, Canada
[*]shared first authors, [#]shared last authors

Corresponding author: Maria E. Ketelaar; University Medical Center Groningen, Groningen Research Institute for Asthma and COPD, Department of Pediatric Pulmonology and Pediatric Allergology, CA43, PO Box 30.001, 9700 RB Groningen, the Netherlands; m.e.ketelaar@umcg.nl; Tel: + 31 50 3611036; Fax: + 31 50 3614235

Sources of funding:

This study was supported by Lung Foundation of the Netherlands grants no. AF 95.05 (GHK), AF 98.48 (GHK) and no.AF3.2.09.081JU, (GHK, MCN), the University Medical Center Groningen (GHK), Dutch TerMeulen Fund (MEK) and the Ubbo Emmius Foundation (GHK), and a grant from GSK (IS, IH, MCN, GHK). The Lifelines Biobank initiative has been made possible by subsidy from the Dutch Ministry of Health, Welfare and Sport, the Dutch Ministry of Economic Affairs, the University Medical Center Groningen (UMCG the Netherlands), University Groningen and the Northern Provinces of the Netherlands. The generation of the lung tissue dataset was funded by Merck. This study was also funded by an Asthma UK Grant to IS, IPH, DES, CEB (AUK-PG-2013-188) and additional Asthma UK funding to IS and DES (Grants 10/006 and 11/031). Genotyping in GASP was additionally supported by Rosetrees Trust (Grant to IS), and AirPROM (CEB, MT and IS). This work was supported by the Medical Research Council [grant number MC_PC_12010], a Strategic Award to IPH, MDT, and LVW, and an MRC project grant to SRJ (G1100163). LVW holds a GSK/ British Lung Foundation Chair in Respiratory Research. Asthma UK funded the GASP initiative (AUK-PG-2013-188). This work was part funded by the NIHR Leicester Respiratory Biomedical Centre. AS is supported by the Manchester Biomedical Research Centre.

Declaration of potential conflict of interests:

GHK, MCN, MEK, CJX, MAP, IS and IH report research funding from Glaxo Smith Kline relating to this manuscript. IS has had research funding relating to this manuscript from AnaptysBio Inc. JDB reports personal fees and non-financial support from Napp, personal fees from Novartis, personal fees and non-financial support from Astra Zeneca, personal fees and non-financial support from Boehringer Ingelheim, personal fees from Teva, personal fees from Innovate UK, outside the submitted work; SRJ reports grants from Medical Research Council, during the conduct of the study; non-financial support from Boehringer-Ingelheim, outside the submitted work; CEB reports grants from AirPROM FP7, grants from Asthma UK, grants from NIHR Biomedical Research Centre, during the conduct of the study; DSP reports grants from Glaxo Smith Kline, during the conduct of the study; grants from Glaxo Smith Kline, outside the submitted work; GHK reports grants from TEVA the Netherlands, Vertex, and Stichting Astma Bestrijding, outside the submitted work; and advisory board fees from GSK and PureIMS, outside the submitted work; MCN reports grants from Glaxo Smith Kline, outside the submitted work; IS reports grants from Glaxo Smith Kline, grants from Anaptsbio Inc, outside the submitted work; RC reports personal fees and non-financial support from AstraZeneca, personal fees from Glaxo Smith Kline, personal fees from Teva Pharmaceuticals, personal fees and non-financial support from Novartis, outside the submitted work; the remaining authors have declared that no conflict of interest exists. AVB and DOB are supported by British Heart Foundation grant and AVB is supported by a Royal Society Project grant RGS\R1\191221.

Abstract (249):

Background: Asthma is a complex disease with multiple phenotypes that may differ in disease pathobiology and treatment response. Interleukin 33 (*IL33*) single nucleotide polymorphisms (SNPs) have been reproducibly associated with asthma. IL33 levels are elevated in sputum, and bronchial biopsies of asthma patients. The functional consequences of *IL33* asthma SNPs remain unknown.

Objective: We studied whether *IL33* SNPs associate with asthma-related phenotypes and with *IL33* expression in lung or bronchial epithelium. We investigated the effect of increased *IL33* expression on human bronchial epithelial cell (HBEC) function.

Methods: Association between *IL33* SNPs (Chr9: 5,815,786–6,657,983) and asthma phenotypes (Lifelines/DAG/GASP cohorts) and between SNPs and expression (lung tissue, bronchial brushes, HBECs) was done using regression modelling. Lentiviral overexpression was used to study *IL33* effects on HBECs.

Results: 161 SNPs spanning the *IL33* region associated with one or more asthma phenotypes after correction for multiple testing. We report one main independent signal tagged by rs992969 associating with blood eosinophil levels, asthma and eosinophilic asthma. A second, independent signal tagged by rs4008366 presented modest association with eosinophilic asthma. Neither signal associated with FEV_1, FEV_1/FVC, atopy, and age of asthma onset. The two *IL33* signals are expression quantitative loci (eQTLs) in bronchial brushes and cultured HBECs, but not in lung tissue. *IL33* overexpression *in vitro* resulted in reduced viability and ROS-capturing of HBECs, without influencing epithelial cell count, metabolic activity or barrier function.

Conclusion: We identify *IL33* as an epithelial susceptibility gene for eosinophilia and asthma, provide mechanistic insight, and implicate targeting of the IL33 pathway specifically in eosinophilic asthma.

Key Messages:

- Genetic signals at the *IL33* locus predominantly associate with blood eosinophil counts in the general population and with an eosinophilic asthma phenotype.
- These genetic signals influence *IL33* levels in the airway epithelium, with the disease risk allele associating with elevated *IL33 in vivo*.
- Elevated *IL33* has modest paracrine effects on bronchial epithelial cell function *in vitro*, implicating that epithelial derived IL33 may more likely affect other effector cell types such as type 2 immune cells, eosinophils or mast cells.

Capsule Summary (30, max 35 words):

This study identifies *IL33* as an epithelial susceptibility gene for eosinophilia and eosinophilic asthma, supporting the *IL33* pathway as a likely candidate for targeted treatment strategies specifically in eosinophilic asthma.

Keywords:

asthma phenotypes, *IL33* SNPs, eQTL, bronchial epithelium, functional translation

Abbreviations:

AHBEC: Asthma Human Bronchial Epithelial Cell
ALI: Air Liquid Interface
AOO: Age Of Onset
AUC: Area Under the Curve
BEGM: Bronchial Epithelial Growth Medium
BHR: Bronchial Hyper Responsiveness
CMV: CytoMegaloVirus
ECIS: Electric Cell Substrate Impedance Sensing
EUR: European
DAG: Dutch Asthma GWAS
FDR: False Discovery Rate
FeNO: Fraction of exhaled Nitric Oxide
FEV_1: Forced Expiratory Volume 1st second
FVC: Forced Vital Capacity
GASP: Genetics of Severe Asthma Phenotypes
GSH: Glutathione
GWAS: Genome Wide Association Study
HBEC: Human Bronchial Epithelial Cell
HDM: House Dust Mite
HWE: Hardy Weinberg Equilibrium
IF: ImmunoFluorescence
IgE: Immunoglobulin E
IL1RL1: Interleukin 1 Receptor Like 1
IL1RAP: Interleukin 1 Receptor Accessory Protein
IL33: Interleukin 33
kU: kiloUnit
LD: Linkage Disequilibrium
MAF: Minor Allele Frequency
MOI: Multiplicity Of Infection
NGS: Next Generation Sequencing
OR: Odds Ratio
P.adj: Adjusted p-value (FDR)
PI: Propidium Iodide
qPCR: quantitative Polymerase Chain Reaction
QTL: Quantitative Trait Locus
ROS: Reactive Oxygen Species
RV: Rhinovirus
SNP: Single Nucleotide Polymorphism
SPT: Skin Prick Test
TF: Transcription Factor
Th2: Type 2 T helper

Introduction:

Asthma is a common, complex, heterogeneous disease that results from the interaction between genetic and environmental factors. It is a chronic inflammatory condition of the airways, characterized by bronchial hyperresponsiveness and reversible airway obstruction. Asthma may consist of several endotypes characterized by differences in specific phenotypes, underlying pathobiology, and (treatment) outcomes in individual patients(1). Genome-wide association studies (GWAS) have identified a large number of asthma loci(2-9), including single nucleotide polymorphisms (SNPs) in *Interleukin (IL)33* and the gene encoding its receptor *Interleukin 1 Receptor Like 1 (IL1RL1)*(9). Both loci were originally discovered as associated with blood eosinophils in general population cohorts(10,11). Next to these common SNPs, a rare *IL33* loss of function mutation has been shown to reduce blood eosinophil counts and protect from asthma(12).

IL33 is an alarmin released upon cellular damage from e.g. epithelial cells. Extracellular IL33 induces signalling via the heterodimeric receptor complex IL1RL1/IL1RAP. Airway IL33 levels have been associated with type 2 cytokines levels and a positive correlation with eosinophil numbers in asthma patients was recently reported(13). High IL33 levels have been found in induced sputum and bronchial biopsies of asthma patients compared to non-asthmatic controls(14-16). Moreover, IL33 may have a paracrine effect on the airway epithelium, as this epithelium has been shown responsive to IL33(17,18). These data suggest a connection between epithelium-derived IL33, eosinophilic inflammation and asthma.

Nevertheless, the functional relevance of common asthma-associated SNPs in *IL33* remains largely unknown. Moreover, genetic association studies thus far have focused on asthma diagnosis, while the contribution of genetic variants to distinct phenotypes of asthma has not been addressed. We hypothesize that genetic variants at the *IL33* locus drive specific phenotypes of asthma by activating a type-2 cytokine dominated immune response, characterized by eosinophilic lung inflammation. Therefore, this study aimed to investigate 1) whether SNPs in the *IL33* region associate with specific asthma phenotypes; 2) whether these *IL33* SNPs form quantitative trait loci (QTL) for *IL33* expression in lung tissue and/or bronchial epithelial samples *in vivo* and *in vitro;* and 3) whether increased *IL33* expression alters human bronchial epithelial cell function.

Materials/Methods:

Detailed methods are described in the online supplement.
Codes available on https://git.web.rug.nl/P252222/IL33_Ketelaaretal_JACI2020. ('wiki')

Study design (see also figure 1)

SNPs in the region of *IL33* (Chr9: 5,815,786–6,657,983, GRCh37/hg19) were tested for association with asthma phenotypes using regression modelling. Briefly, we tested association of the *IL33* SNPs in a Dutch general population cohort (Lifelines(19); n=13,395) with eosinophil counts, FEV_1 and FEV_1/FVC. From this general population cohort we subsequently took the asthma subpopulation (n=1,066, doctor's diagnosed asthma) and investigated genetic association with eosinophil counts, eosinophilic asthma (asthma and blood eosinophils >150 cells/µL, n=707, as this cut-off is a good predictor for airway eosinophilia (>2% sputum eos (20), non-eosinophilic asthma (asthma and blood eosinophils <150 cells/uL, n=359), FEV_1, FEV_1/FVC and asthma with airway obstruction (asthma and FEV_1<80% of predicted (n=258) or FEV_1/FVC<70% (n=324)). In a meta-analysis of two independent asthma cohorts of n=2,536 moderate-severe asthma patients (GASP, UK(21))) and n=909 asthma patients of mild-moderate severity (DAG, the Netherlands(22)), we then evaluated association of *IL33* SNPs with atopy, blood eosinophils, total serum IgE, age of asthma onset, and lung function (FEV_1, FEV_1/FVC).
We then selected independent genetic signals based on LD (r^2<0.1), followed by conditional analyses on the most significantly associated SNP. Functional investigations of selected independent genetic signals included expression and protein quantitative loci studies in lung tissue (n=1,111), bronchial brushes (n=139) and primary asthma derived human bronchial epithelial cells (AHBECs, n=35). Potential function was investigated using ENCODE, PredictSNP, Meta-SNP, and Polyphen-2 data(23-25). We tested for inducible expression (e)QTL and protein (p)QTL by exposing AHBECs (n=18) of various *IL33* genotypes to asthma-relevant stimuli (HDM, RV16). Finally, we overexpressed *IL33* in (healthy-derived) HBECs (n=5) to investigate effects on cell count, metabolic activity, viability, ROS-capturing and epithelial barrier.

Genotype-phenotype analysis

A total of 1,970 imputed SNPs (Lifelines, all overlapping with DAG/GASP) and 2,457 imputed SNPs (DAG/GASP) were available for the association analyses based on a MAF≥0.01 and chromosomal location of 400kb up- and downstream *IL33* (Chr9: 5,815,786–6,657,983). This region encompasses all known asthma-associated SNPs (table I, table S1). Associations of SNPs with asthma phenotypes were performed with PLINK v1.90b6.7(26) (Lifelines) or SNPtest v2.5β(27) (DAG/GASP) using an additive genetic model. DAG/GASP were meta-analysed in METAL(28) using a fixed model, (table S2-S3). An adjusted p-value (P.adj)<0.05 (FDR) was considered statistically significant.

Functional genetics

QTL and ENCODE investigations

We tested for expression quantitative trait loci (eQTL) in lung tissue (n=1,111) and bronchial brushes (n=139) (table S5) using a linear regression model to investigate the association between SNPs and log-transformed *IL33* expression data. We employed an additive genetic model with age, gender, smoking status and the PCs explaining >1% of expression variance as covariates using R statistics(29). We did not have data on medication use for these cohorts, so could not correct for this covariate, but the currently used covariates are thought to reflect main confounders in eQTL analyses (22,30).We also tested for (inducible) QTLs in cultured bronchial epithelial cells (n=18-35) obtained from bronchial brushes/biopsies from asthma patients as described(31). AHBECs were stimulated with 50µg/ml house dust mite (HDM) or rhinovirus (RV16, MOI=1) for 24 hours and RNA lysates collected(32). Cells were genotype-stratified and expression compared using Kruskal-Wallis tests. A p-value<0.05 was considered statistically significant. ENCODE was used to identify potential functional effects of tagSNPs and SNPs in LD (r^2>0.3). SNPs were functionally checked for DNase I hypersensitive sites, histone mark sites, binding motifs, and regulatory motifs using RegulomeDB, HaploReg, ChromHMM and Segway(23,24).

Functional bronchial epithelial cell studies

In order to investigate the functional consequences of increased IL33 in bronchial epithelial cells, we stably overexpressed human full-length *IL33* (aa1-270) in primary HBECs isolated from 5 healthy individuals (Lonza, #CC-2540). IL33 mRNA and protein expression was quantified by qPCR and immunofluorescence respectively. We analysed cell count, viability and metabolic activity, as well as ROS-capturing ability (glutathione assay) and barrier function (ECIS) in these cultures. We used Kruskal-Wallis for all parameters except for longitudinal area under the curves (AUCs) comparisons of ECIS data, which were compared using a Z-test. A p-value<0.05 was considered statistically significant.

Results:

Genetic association with phenotypes of asthma:

The IL33 locus particularly associates with eosinophilia and eosinophilic asthma

Overall in DAG/GASP and Lifelines (S2-S3), 161 SNPs significantly associated with one or more asthma phenotypes (P.adj<0.05 (FDR); tables S11-15), mainly derived from the Lifelines cohort. From these, 144 SNPs comprised five LD blocks (A-E, r^2>0.1). Markedly, these five LD blocks all associated with an eosinophilic phenotype; either with blood eosinophil counts, eosinophilic asthma and/or asthma (table I, tables S11-S17, figure E2/E3). LD block A shows a significant association with blood eosinophil counts in the general population (tagSNP rs992969[allele A] beta=0.058+/- SE=0.0089, P.adj=7.09E-08, AF=0.25), whilst three other LD blocks were modestly associated with this phenotype (block B-D, table I). Block E showed association with eosinophilic asthma (tagSNP rs4008366 [allele T], OR=1.26+/-SE=0.0704, P.adj=0.045, AF=0.67) only.

Outside these five LD blocks, seven SNPs significantly associated with other phenotypes (age of onset or FEV_1/FVC, table S15) and ten significant SNPs were identified in the case-control analyses of resequencing data; these were relatively rare (MAF~0.03, ~~table S17~~) and hence were not followed-up functionally. We performed conditional analyses on the LD blocks associated with eosinophilic phenotypes to determine independent signals. A summary description of association results can be found in the supplementary section M1.

Conditional and sensitivity analyses show one main genetic signal associated with blood eosinophil counts in the general population

Four LD blocks (A-D, figure 2) showed association with blood eosinophil counts in the Lifelines general population. Thereby, block A (tagSNP rs992969) shows the largest effect size and statistical significance (table I); rs992969 explaining 1.6% (R^2 regression model=0.016) of the variance in blood eosinophil counts (corrected for age/gender). Therefore, we conditioned the association analysis for blood eosinophils on rs992969 to test whether block A-D are independent signals. Conditioning removed the association of signals B-D with blood eosinophil counts in the general population (see figure 3 and table II). Signal E was not significantly associated with eosinophil counts, regardless of conditioning. Sensitivity analysis for the main signal A showed that rs992969 still associated with eosinophil counts in the general population when removing asthma patients (figure 3a-2) (n=12,329; rs992969 [allele A] beta=0.055, SE=0.009, R^2=0.017, P.adj=1.04E-06) or both asthma/allergic patients (figure 3a-3) (n=6,227; rs992969 [allele A] beta=0.046, SE=0.012, R^2=0.020, P.adj=0.02). These analyses show the presence of one main genetic signal (A) at the *IL33* locus associated with blood eosinophil counts in the general population, independent of the presence of asthma/allergy phenotypes.

Signal A and E associate with eosinophilic asthma

Signal A, driving the association with blood eosinophil counts in the general population, also showed a significant association with asthma (rs992969[allele A], OR=1.22, SE=0.05, P.adj=0.03) and with eosinophilic asthma (rs992969[allele A], OR= 1.32+/-SE=0.0618, P.adj=4.73E-03), (figure 3+4). Signal A contains several SNPs previously associated with asthma (table I). The genetic effect of this main signal on asthma risk remained of similar size, even when correcting for blood eosinophil counts (OR[A] from 1.22 (P.adj=0.03) to 1.19 (P.adj=0.08), figure 4b).

Signal E was the other LD block associated with eosinophilic asthma (tagSNP rs4008366, figure 4d), with a significantly large effect size. Lack of power precluded conditional analyses for the eosinophilic asthma phenotype, so (in)dependency of block E could not be confirmed. However, this block represents a genetically independent signal in other cohorts (table II), underscoring it may be a distinct signal and may represent a distinct mechanism underlying asthma pathogenesis. Therefore, two signals (A and E) were selected for functional follow-up.

To assess whether our definition of eosinophilic asthma based on the cut-off for blood eosinophils at 150 cells/uL impacted on the associations observed, we repeated the analysis at a cut-off of 300 cells/uL as a definition for eosinophilic asthma. These additional analyses of eosinophilic asthma, including a higher cut-off of eosinophil counts, identify the same associations with slightly higher effect sizes (see supplemental table S8/S9), but not FDR (<0.05) significant anymore, likely explained by the more refined phenotype resulting in smaller group sizes.

309 **Table I- Five LD blocks (r^2>0.1) with phenotype associations could be distinguished**

Tag SNP (genetic signal)	Location	Gene context	Pheno risk allele (AF)	Alt allele	Associated feature(s)	Cohort	Effect size risk allele (B or OR)	SE	P.adj (FDR)	Lit. asthma GWAS SNP in block	Ref GWAS
A-rs992969	9:6209697	~6kb 5' of *IL33*	A (0.25)	G	*eos levels in GenPop,* *eos asthma vs HC,* *asthma case control*	Lifelines Lifelines Lifelines	0.058 (B) 1.321 (OR) 1.230 (OR)	0.009 0.062 0.053	7.09E-08 4.73E-03 0.034	rs1888909 rs7848215 rs992969 rs144829310 rs72699186 rs928413 rs1342326 rs2381416 rs2066362	(2-7,21,33-37)
B-rs1342327	9:6189874	~25kb 5'of *IL33*	G (0.15)	C	*eos levels in GenPop,* *eos levels in asthma subjects*	Lifelines, DAG/GASP	0.035 (B) 0.057 (B)	0.011 0.018	0.027 0.039	-	-
C-rs74438701	9:6282794	~25kb 3'of *IL33*	T (0.83)	C	*eos levels in GenPop*	Lifelines	0.035 (B)	0.011	0.041	-	-
D-rs2282162	9:6534466	intronic of *GLDC*	G (0.56)	A	*eos levels in GenPop*	Lifelines	0.029 (B)	0.008	0.011	-	-
E-rs4008366	9:6116407	intergenic	T (0.69)	C	*eos asthma vs HC*	Lifelines	1.264 (OR)	0.070	0.045	rs343478	(36,37)

310 The table shows the tagSNPs representing 5 LD blocks/signals (r^2>0.1) from the SNPs significantly (FDR<0.05) associated with asthma features in the Lifelines general population,
311 Lifelines asthma population and DAG/GASP asthma population. In the last columns these are put into context of previously reported genome-wide significant (5*10E-8) SNPs
312 associated with asthma, displaying SNPs part of the LD block at r^2>0.1. **Underlined**: the two genetic signals (A and E) taken forward in functional assesment in this study. Because of
313 its association with eosinophilic asthma, lack of LD with signal A, as well as this LD block also represented an independent signal in multiple studies, we took signal E forward as an
314 independent phenotype-associated signal in our functional analyses. *AF=frequency (EUR 1000G); Alt allele=alternative allele; B=beta; eos=eosinophils/eosinophilic; FDR=false*
315 *discovery rate value at alpha 0.05; GWAS= genomewide association study; GenPop =general population; HC=healthy control; kb= kilo basepairs; Lit.=literature; OR=odds ratio;*
316 *P.adj= FDR adjusted p-value; Pheno risk allele=phenotype associated allele; Ref=Reference; SE=standard error*. For complete analyses of all eosinophilic phenotypes in each cohort,
317 please see supplemental table S7.

318

Table II- Conditioning on rs992969 in the association of *IL33* with blood eosinophils in the general population of Lifelines removed signals B-D

LD Block (tagSNP)	Location	Gene context	Pheno risk allele (AF)	Alt allele	Effect size risk allele (beta)**	SE	P.adj (FDR)	Independent signal Ubiobank/ INTERVAL^ (blood eos GenPop)(10)	Independent signal UK biobank only^ (asthma) (29)	Independent signal SHARE^ (asthma/allergy) (31)
A-rs992969	9:62096 97	~6kb 5' of *IL33*	*A (0.25)*	*G*	**0.058** *n/a*	**0.009** *n/a*	**7.09E-08***** *n/a*	rs2381416 (r^2= 0.95)	rs7848215 (r^2=0.93)	rs144829310 (r^2= 0.59)
B-rs1342327	9:61898 74	~25kb 5'of *IL33*	*G (0.15)*	*C*	**0.035** *0.012*	**0.011** *0.011*	**0.027**** *0.877*	-	-	-
C-rs74438701	9:62827 94	~25kb 3'of *IL33*	*T (0.83)*	*C*	**0.035** *0.017*	**0.011** *0.011*	**0.041**** *0.722*	-	-	-
D-rs2282162	9:65344 66	intronic of *GLDC*	*G (0.56)*	*A*	**0.029** *0.012*	**0.008** *0.009*	**0.011**** *0.722*	-	-	-
E-rs4008366*	9:61164 07	intergenic	*T (0.69)*	*C*	**0.010** *0.002*	**0.009** *0.009*	**0.647** *0.974*	-	rs343478 (r^2= 0.17)	rs343478 (r^2= 0.17)

Conditional analyses were performed in n=13,395 subjects from the Lifelines general population, studying the effect of *IL33* SNPs on level of blood eosinophils, by taking rs992969 (=lowest p-value SNP associated with level of blood eos) as covariate in the regression model. These were put into the context of independent SNPs as determined in other large cohorts. r^2= relative to tagSNP of LD block A/B/C/D/E respectively.

* Signal E was not significantly associated with level of blood eosinophils in the general population before conditional analyses, nor after conditional analyses, but has only been included in this table to show it is in modest LD with rs343478 (an independent signal in other studies). Because of its association with eosinophilic asthma, lack of LD with signal A, as well as this LD block also represented an independent signal in multiple studies, we took this signal forward as an independent phenotype-associated signal in our functional analyses.

In **bold the unconditioned results, in *italics* the results conditioned on rs992969.

***Adjusted p-value (FDR) statistically significant <0.05.

^ Independent phenotype-associated SNPs at the *IL33* locus determined based on conditional analyses in other large population cohorts: the phenotype studied in the UK Biobank/INTERVAL was blood eosinophil levels in the general population (n=173,480)(10), in the UK biobank only was asthma (n=41,926 cases vs. n=239,773 controls) (36), whilst the SHARE study examined a combined asthma/allergy phenotype (n=180,129 cases vs 180,709 controls)(37).

Underlined: the two genetic signals taken forward in functional assessment in this study

QTL/functional investigation of *IL33* genetic variation:

After conditional analyses, two independent signals A and E remained for functional follow-up, each with a tagSNP (rs992969 and rs4008366, table II). These tagSNPs were chosen based on smallest p-value/largest effect size, largest number of associated phenotypes and if applicable known association with asthma from literature. In case the tagSNP was not available for functional look-up, a proxySNP at r^2>0.5 with the tagSNP of the original association signal was chosen (table S7).

Signal A and E are IL33 eQTL in bronchial epithelium

To investigate potential functionality of signal A and E, QTL analyses were performed in lung tissue, bronchial epithelial brushes and cultured bronchial epithelial cells (see table III, S5 and figure 5).

In lung tissue samples, no eQTLs for *IL33* were found (table S19 and figure E4). In bronchial brushes, the tagSNP of signal A was a significant and strong eQTL for *IL33*, with the disease associated allele correlating with higher mRNA levels, (rs992969[A] beta=0.331, SD=0.043, P=8.30E-12, AF=0.25). No significant eQTLs were found for signal E in bronchial brushes. In cultured primary human bronchial epithelial cells (HBECs), the disease associated allele of signal E (proxy SNP rs442246) associated with lower *IL33* mRNA (P=0.029, see table III and figure E6b). No significant pQTLs were found for IL33 in HBECs for both signal A and E (figure E7).

Signal A and E harbour potential functional elements related to expression regulation of IL33

ENCODE revealed several putative regulatory elements for SNPs in both genetic signals A and E relevant for *IL33* transcription (table IIIb). Signal A contained 27 SNPs (LD r^2>0.3 with tagSNP) with potential functionality. Among these is a SNP (rs928413) in strong LD with the phenotype and expression associated tagSNP rs992969 (r^2=0.96), forming a CREB1 binding site activating the *IL33* promotor. In signal E, 7 SNPs were potential functional elements, including specific transcription factor binding sites relevant to the regulation of the cellular oxidative state (e.g. Nrf2) in lung derived cells. Thus, the genetic signals A and E contain likely functional elements related to expression, forming a potential mechanistic link between phenotype and expression association.

Asthma stimuli induce differential IL33 expression, regardless of genetic background for signal A and E

Next, we tested for the presence of inducible QTLs for *IL33* in primary AHBECs after exposure to RV16 (MOI:1) or house dust mite (HDM, 50µg/mL) and analysed for effects on *IL33* mRNA and extracellular protein levels in an unstratified way or stratified for genetic signals A and E. RV16 induced a decrease in *IL33* mRNA levels in AHBEC (P=0.048), and a marked increase of IL33 protein in the cellular supernatant (P=0.0001). HDM exposure induced an increase in *IL33* RNA and had no significant effects on IL33 protein levels, measured 24h post stimulation (figure E7). When stratified on signals A and E, no significant differences on the RV16 or HDM induced effects on IL33 mRNA or protein levels were observed (figures E8/E9).

IL33 overexpression modestly impairs bronchial epithelial cell homeostasis:

In order to investigate the effect of increased *IL33* expression, we overexpressed full length *IL33* in primary bronchial epithelial cells using lentiviral delivery (figure E11/E12). We confirmed increased expression of *IL33* at the mRNA level and presence of IL33 protein in engineered cells (figure 5 and figures E13/E14). We found that overexpression of *IL33* does not significantly influence cell number or metabolic activity (figure E15). Viability was 15-20% lower (P=0.04, figure 5c) and ROS-capturing capacity (presence of free glutathione) was ~20% lower (P=0.03, figure 5d) in cells that overexpressed *IL33* under submerged culture condition. No effect of *IL33* overexpression was seen on spreading or formation of an epithelial barrier using ECIS (figure E14).

391

392 **Table III- QTL function and functional ENCODE annotation of the phenotype associated signals A and E:**

III A *Tag SNP (genetic signal)*	*Pheno risk allele (AF)*	*Alt allele*	*Associated phenotype(s)*	*QTL cohort*	*Effect size pheno risk allele*	*SD*	*P-value*	*Direction pheno risk allele*	*Literature Reported QTL function*
A- *rs992969*	***A* (0.25)**	***G***	*eos levels in GenPop, eos asthma vs HC, asthma case control*	Bronchial brushes	0.326 (B)	0.043	8.30E-12	*++IL33 RNA*	cisQTL IL33 bronchial biopsies/blood/brain; (30,39,40)
E- *rs442246 (proxy for: rs4008366)*	***T* (0.69)**	***G***	*eos asthma vs HC*	Cultured HBEC	-2.377 (fold change TT)		0.0298	*--IL33 RNA*	-

III B *Tag SNP (genetic signal)*	*Location*	*Gene context*	*Associated phenotypes*	*Functional annotation of genetic signal, SNPs r^2>0.3 with tagSNP*					
				Promotor cRE (lung)	**Enhancer cRE (lung)**	**DNase I site cRE (lung)**	**Protein-binding (lung)**	**PredictSNP/DANN**	**Experimental functionality**
A-rs992969	9:6209697	**~6kb 5' of *IL33***	*eos levels in GenPop, eos asthma vs HC, asthma case control*	Y- H3K4me3	Y-H3K27ac	Y	CTCF, SETDB1, CFOS, PRDM1, STAT3	neutral	r^2=0.96 with rs928413(G) forming CREB1 binding site, activating IL33 promotor lung epithelial cells (41)
E-rs4008366	9:6116407	**Intergenic**	*eos asthma vs HC*	-	Y- H3K27ac	Y	Nrf-2, TCF11, MafG, ZID, Hmbox1, Hoxd8	Deleterious (0.85 accuracy)	-

393 **Panel IIIA:** The table shows quantitative trait loci (QTL) function of the two genetic signals in the *IL33* region associated with eosinophilic asthma features in our cohorts. In case the tagSNP
394 was not available, a proxy at r^2>0.3 was used for QTL look-up. Expression (e)QTLs were studied in lung tissue (lung surgery patients) and bronchial brushes (healthy subjects); eQTL and
395 protein(p)QTL function were studied in cultured primary human bronchial epithelial cells (AHBECs) from asthma patients. Of note: in lung tissue no significant eQTLs for *IL33* were found in
396 the *IL33* region, and in HBECs no significant pQTLs were found for these 2 genetic signals (alpha=0.05). In bronchial brushes, signal A was an eQTL for *IL33*, with the phenotype risk allele
397 associating with higher *IL33* mRNA levels. In cultured HBECs signal E has potential QTL function; the eosinophilic asthma risk allele associating with lower *IL33* RNA. More details can be
398 found in figure E3-E5 (supplemental). *Pheno Risk allele=phenotype associated allele; Alt allele=alternative allele; AF=allele frequency (EUR 1000G); B=beta; SD=standard deviation;*
399 *eos=eosinophils/eosinophilic; GenPop=general population; HC=healthy control; ++=increased expression, --=decreased expression.* **Panel IIIB:** The table shows the functional ENCODE
400 and PredictSNP, Meta-SNP, Polyphen-2 (23,24) look-up of the two genetic signals that were selected from the SNPs significantly (P.adj (FDR)<0.05) associated with asthma features in
401 Lifelines general population, Lifelines asthma population and DAG/GASP asthma population. SNPs in LD r^2>0.3 with the tagSNP of the applicable genetic signal were included in the functional

look-up. ENCODE and Polyphen-2 retrieved functional annotation for the signals, while PredictSNP and Meta-SNP did not. *cRE=candidate regulatory element; eos=eosinophils/eosinophilic; DANN – Deleterious Annotation of Genetic Variants using Neural Networks, HC=healthy control; GenPop=general population; kb= kilo basepairs; Y=yes*

Discussion

We set out to determine if SNPs in the *IL33* region associate with specific phenotypes of asthma, whether these regulate *IL33* expression in lung tissue or bronchial epithelial samples, and whether increased *IL33* expression alters human bronchial epithelial cell biology. Genetic signals at the *IL33* locus predominantly associate with an eosinophilic phenotype in the general population and asthma subjects, whereby the *IL33* risk allele is associated with higher *IL33* expression *in vivo*. Using conditional analyses we observed one major genetic signal and one secondary signal. The major signal associates with blood eosinophil counts and (eosinophilic) asthma, while the secondary signal associates with eosinophilic asthma but not with eosinophil counts in the general population. Importantly, no association with other asthma-related phenotypes including lung function, atopy, serum IgE levels and asthma age of onset was observed. Studying the effects of these two genetic signals on *IL33* transcription, we report eQTLs in bronchial brushes and cultured bronchial epithelial cells, but not in lung tissue. Overexpression of *IL33* in bronchial epithelial cells resulted in modest paracrine effect on epithelial cell homeostasis, including reduction in cell viability and ROS-capturing capacity. With this approach we identify *IL33* as an epithelial susceptibility gene for eosinophilia and asthma, provide mechanistic insight, and support targeting of the IL33 pathway specifically in eosinophilic asthma.

Two genetic IL33 signals associate with eosinophilia in health and disease

The *IL33* gene, and the *IL1RL1* gene encoding its receptor, have consistently been associated with asthma and allergy(3-8,21,33-37). Both loci were originally discovered as regions associating with blood eosinophils in the Icelandic population(11,33), and a strong association with blood eosinophil counts was recently confirmed in a large general population cohort (n=173,480), combining UK Biobank and INTERVAL studies(10). Also, a rare loss-of-function *IL33* mutation was shown to both reduce eosinophil counts and to protect from asthma(12). These observations suggest a shared genetic effect of this locus for eosinophilia and asthma. However, it remained unknown whether these are the same or distinct genetic signals and what additional asthma related phenotypes these signals may be associated with.

We report five LD blocks that were associated with either blood eosinophil counts and/or eosinophilic asthma, which after conditional analysis correcting for the strongest signal (rs992969), were reduced to two independent signals. The fifth signal (E) was not associated with blood eosinophil counts in the general population, but with eosinophilic asthma. The available subjects (n=707) for the eosinophilic asthma phenotype did not allow conditional analyses for signal E. However, previous analyses in two very large cohorts (SHARE(37) and UK-biobank(36) supported the independence of signal E, representing a second signal associated with an eosinophilic phenotype in our cohorts. This left us with two genetic signals for further study.

We observe a strong association for signal A with both asthma and blood eosinophil counts in our Lifelines general population cohort, the tagSNP rs992969 explaining 1.6% (R^2=0.016) of the variance in eosinophil counts (corrected for age/gender). SNPs within this signal have previously been reported to associate with asthma in the UK Biobank, SHARE, and TAGC study, as well as in earlier asthma meta-analyses(2,4,5,7,21,35-37) and with blood eosinophil counts in the UK biobank/INTERVAL study (rs992969 in LD r^2=0.95 with rs2381416 from UK Biobank/INTERVAL)(10). Using a sensitivity analysis in Lifelines by removing asthmatic and allergic subjects from the general population, we show that the association with blood eosinophils remained present with a similar effect size, indicating that the association between this signal A and blood eosinophils is not fully driven by the presence of asthma or allergy. We find that the association of signal A with asthma is of similar effect size when correcting for blood eosinophil counts, suggesting that this IL33 genetic signal- in addition to its effect on blood eosinophil counts- may have an effect on asthma. However, we do find that the effect of signal A on asthma after correcting for blood eosinophils is no longer FDR significant. (Figure 4.) Therefore, a better powered study is required to conclusively investigate an effect of this signal on asthma independent from eosinophil counts. Interestingly, we observed an association of signal A with eosinophilic asthma, but not with non-eosinophilic asthma (figure 4), indicating that patients with this *IL33* genetic make-up would be enriched in the high-eosinophil group. A note of caution is the relatively limited number of subjects in our non-eosinophilic asthma group (n=359).

An intriguing implication could be that in asthma patients with this particular genetic background (signal A), treatment targeting the IL33 pathway could have additional effects over treatments targeting eosinophils(42,43). Notwithstanding, whether the association of *IL33* SNPs with asthma and eosinophils are (in)dependent from each

other remains to be conclusively determined in larger cohorts, allowing causal inference/mediating approaches like Mendelian randomization(44). Ideally, such an analysis would also take into account *IL1RL1* genotypes, which are likely to interact with *IL33* variants on outcomes such as eosinophilic inflammation; as well as a more direct measure of eosinophilic airway inflammation such as sputum eosinophil counts should be considered.

Functional effects of phenotype-associated IL33 polymorphisms and IL33 expression

Functionally, IL33 signalling has previously been linked to Th2-driven inflammation, contributing to eosinophilic inflammation(45-47). Moreover, levels of IL33 have been found elevated in induced sputum and bronchial biopsies of asthma patients compared to non-asthmatic controls(14-16), indicating a dysregulation of IL33 homeostasis in asthma. Therefore, specific genetic variation at the *IL33* locus might contribute to eosinophil numbers and asthma through regulation of *IL33* expression levels.

While we did not detect eQTLs for *IL33* in lung tissue samples, the tagSNP of signal A was a strong eQTL for *IL33* in bronchial epithelial brushes from healthy subjects (table III, figure 5), with the risk allele associating with increased *IL33* mRNA levels. Signal A harbours a reported *IL33* eQTL in a candidate eQTL study of bronchial biopsies(30), with the same direction of effect. This eQTL signal A also comprised a SNP (rs928413, in LD r^2=0.96 with rs992969) where the phenotype-risk allele was recently found to form a CREB1 binding site, functionally activating the *IL33* promotor in lung epithelial cells(41). This allele associates with higher level of eosinophils, higher risk of (eosinophilic) asthma and increased *IL33* expression in brushes in our cohorts. As lung tissue resection samples mainly consist of parenchymal lung tissue with minor contributions of airway epithelial cells, while bronchial brushes contain more than 90% bronchial epithelial cells(48), we interpret these data as evidence for regulation of *IL33* expression in bronchial epithelium.

The bronchial epithelium is the first barrier that the inhaled substances encounter when entering the lung, and serves to protect the body from potential threats from the environment. The airway epithelium is changed in asthma patients, with increased susceptibility to and altered repair responses after external damage(49,50), for example in response to respiratory viruses(51). A GWAS on exacerbation in asthma(2) found the *IL33* locus associated with frequent virus-induced exacerbations in severe childhood-onset asthma, their main *IL33* SNP in strong LD (r^2=0.96) with our eosinophilic signal rs992969. Therefore, we tested whether our two phenotype-associated signals are baseline and/or induced QTLs in cultured bronchial epithelium. Signal E is a modest, baseline *IL33* eQTL, with the eosinophilic risk allele associated with lower *IL33* mRNA levels (table III) in these cells. Although both RV16 and HDM regulated *IL33* expression, no effect of the two *IL33* signals on the RV16 and HDM induced IL33 response was observed in vitro in our samples. This could indicate that our two signals may specifically have effects on baseline changes of *IL33* expression in epithelium.

The opposite direction of effect in the cultured bronchial epithelial cells compared to the bronchial brushes might indicate that *IL33* gene regulation is different in asthmatic epithelium compared to healthy brushed cells, in agreement with recent data from Jurak *et al.* (2018)(52). Alternatively, it could reflect differences in epithelial cell state with cultured HBECs having a basal cell phenotype(48,53), while bronchial brushes contain mostly well-differentiated ciliated and secretory epithelial cells(48).

The cell-autonomous effects of increased *IL33* expression on cultured HBECs were modest. Nevertheless, the observed effect of sustained IL33 on reduction of glutathione levels in the epithelium is interesting, as Uchida *et al.*(18) showed that the balance between oxidative stress and antioxidant responses plays a key role in controlling IL33 release from airway epithelium. Our data indicate that the bronchial epithelium is the source of IL33, but that other cell types should be considered as the main IL33 responsive population, such as tissue-resident dendritic cells, eosinophils, type 2 innate lymphoid cells, Th2 cells, mast cells and basophils, but also lung mesenchymal, such as fibroblasts. This is also relevant in the context of steroid-resistant asthma patients. For example: elevated IL-33 and type 2 cells were still present in corticosteroid resistant pediatric asthma patients, contributing to airway remodelling via its effects on airway fibroblast.(54,55)

In conclusion, we have reduced the complex *IL33* locus into one major and one secondary genetic signal for eosinophilic asthma. The major *IL33* signal risk allele associates with increased *IL33* expression levels providing a putative mechanism. Importantly we have also shown a lack of genetic association of this main genetic signal with other studied asthma phenotypes. We identified the bronchial epithelial cell as the likely cellular source of *IL33* QTL signals, which is crucial to place the genetic effects on *IL33* expression in asthma pathophysiology. These data need confirmation by e.g. single-cell eQTL analyses in airway wall samples of asthma patients and healthy controls. This approach might also guide the identification of the main IL33 responding cells. Nevertheless, our data identifies *IL33* as an epithelial susceptibility gene for eosinophilia and asthma, and

supports the IL33 pathway as a likely candidate for targeted treatment strategies in specifically eosinophilic asthma, with the potential to affect both eosinophil counts and asthma independently.

Acknowledgements:
The authors wish to acknowledge the services of the Lifelines Cohort Study, the contributing research centres delivering data to Lifelines, and all the study participants. Also, we acknowledge the ENCODE Consortium for generating online accessible datasets to consult for functional regulatory elements. Furthermore, this work was supported in part by the NIHR Nottingham Biomedical Research Centre, which IH, IS and MAP would like to acknowledge here.

References:

(1) Kaur R, Chupp G. Phenotypes and endotypes of adult asthma: Moving toward precision medicine. J Allergy Clin Immunol 2019 Jul;144(1):1-12.

(2) Bonnelykke K, Sleiman P, Nielsen K, Kreiner-Moller E, Mercader JM, Belgrave D, et al. A genome-wide association study identifies CDHR3 as a susceptibility locus for early childhood asthma with severe exacerbations. Nat Genet 2014 Jan;46(1):51-55.

(3) Ferreira MA, Matheson MC, Tang CS, Granell R, Ang W, Hui J, et al. Genome-wide association analysis identifies 11 risk variants associated with the asthma with hay fever phenotype. J Allergy Clin Immunol 2014 Jun;133(6):1564-1571.

(4) Torgerson DG, Ampleford EJ, Chiu GY, Gauderman WJ, Gignoux CR, Graves PE, et al. Meta-analysis of genome-wide association studies of asthma in ethnically diverse North American populations. Nat Genet 2011 Jul 31;43(9):887-892.

(5) Moffatt MF, Gut IG, Demenais F, Strachan DP, Bouzigon E, Heath S, et al. A large-scale, consortium-based genomewide association study of asthma. N Engl J Med 2010 Sep 23;363(13):1211-1221.

(6) Pickrell JK, Berisa T, Liu JZ, Segurel L, Tung JY, Hinds DA. Detection and interpretation of shared genetic influences on 42 human traits. Nat Genet 2016 Jul;48(7):709-717.

(7) Ferreira MA, Matheson MC, Duffy DL, Marks GB, Hui J, Le Souef P, et al. Identification of IL6R and chromosome 11q13.5 as risk loci for asthma. Lancet 2011 Sep 10;378(9795):1006-1014.

(8) Vicente CT, Revez JA, Ferreira MAR. Lessons from ten years of genome-wide association studies of asthma. Clin Transl Immunology 2017 Dec 15;6(12):e165.

(9) Grotenboer NS, Ketelaar ME, Koppelman GH, Nawijn MC. Decoding asthma: translating genetic variation in IL33 and IL1RL1 into disease pathophysiology. J Allergy Clin Immunol 2013 Mar;131(3):856-865.

(10) Astle WJ, Elding H, Jiang T, Allen D, Ruklisa D, Mann AL, et al. The Allelic Landscape of Human Blood Cell Trait Variation and Links to Common Complex Disease. Cell 2016 Nov 17;167(5):1415-1429.e19.

(11) Gudbjartsson DF, Bjornsdottir US, Halapi E, Helgadottir A, Sulem P, Jonsdottir GM, et al. Sequence variants affecting eosinophil numbers associate with asthma and myocardial infarction. Nat Genet 2009 Mar;41(3):342-347.

(12) Smith D, Helgason H, Sulem P, Bjornsdottir US, Lim AC, Sveinbjornsson G, et al. A rare IL33 loss-of-function mutation reduces blood eosinophil counts and protects from asthma. PLoS Genet 2017 Mar 8;13(3):e1006659.

(13) Poulsen NN, Bjerregaard A, Khoo SK, Laing IA, Le Souef P, Backer V, et al. Airway Interleukin-33 and type 2 cytokines in adult patients with acute asthma. Respir Med 2018 Jul;140:50-56.

(14) Hamzaoui A, Berraies A, Kaabachi W, Haifa M, Ammar J, Kamel H. Induced sputum levels of IL-33 and soluble ST2 in young asthmatic children. J Asthma 2013 Oct;50(8):803-809.

(15) Gasiuniene E, Janulaityte I, Zemeckiene Z, Barkauskiene D, Sitkauskiene B. Elevated levels of interleukin-33 are associated with allergic and eosinophilic asthma. Scand J Immunol 2019 May;89(5):e12724.

(16) Wang W, Li Y, Lv Z, Chen Y, Li Y, Huang K, et al. Bronchial Allergen Challenge of Patients with Atopic Asthma Triggers an Alarmin (IL-33, TSLP, and IL-25) Response in the Airways Epithelium and Submucosa. J Immunol 2018 Oct 15;201(8):2221-2231.

(17) Altman MC, Lai Y, Nolin JD, Long S, Chen CC, Piliponsky AM, et al. Airway epithelium-shifted mast cell infiltration regulates asthmatic inflammation via IL-33 signaling. J Clin Invest 2019 Aug 22.

(18) Uchida M, Anderson EL, Squillace DL, Patil N, Maniak PJ, Iijima K, et al. Oxidative stress serves as a key checkpoint for IL-33 release by airway epithelium. Allergy 2017 Oct;72(10):1521-1531.

(19) Scholtens S, Smidt N, Swertz MA, Bakker SJ, Dotinga A, Vonk JM, et al. Cohort Profile: LifeLines, a three-generation cohort study and biobank. Int J Epidemiol 2015 Aug;44(4):1172-1180.

(20) Gao J, Wu F. Association between fractional exhaled nitric oxide, sputum induction and peripheral blood eosinophil in uncontrolled asthma. Allergy Asthma Clin Immunol 2018 May 23;14:21-018-0248-7. eCollection 2018.

(21) Shrine N, Portelli MA, John C, Soler Artigas M, Bennett N, Hall R, et al. Moderate-to-severe asthma in individuals of European ancestry: a genome-wide association study. Lancet Respir Med 2019 Jan;7(1):20-34.

(22) Nieuwenhuis MA, Siedlinski M, van den Berge M, Granell R, Li X, Niens M, et al. Combining genomewide association study and lung eQTL analysis provides evidence for novel genes associated with asthma. Allergy 2016 Dec;71(12):1712-1720.

(23) ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. Nature 2012 Sep 6;489(7414):57-74.

(24) Davis CA, Hitz BC, Sloan CA, Chan ET, Davidson JM, Gabdank I, et al. The Encyclopedia of DNA elements (ENCODE): data portal update. Nucleic Acids Res 2018 Jan 4;46(D1):D794-D801.

(25) Adzhubei I, Jordan DM, Sunyaev SR. Predicting functional effect of human missense mutations using PolyPhen-2. Curr Protoc Hum Genet 2013 Jan;Chapter 7:Unit7.20.

(26) Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007 Sep;81(3):559-575.

(27) Marchini J, Howie B, Myers S, McVean G, Donnelly P. A new multipoint method for genome-wide association studies by imputation of genotypes. Nat Genet 2007 Jul;39(7):906-913.

(28) Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics 2010 Sep 1;26(17):2190-2191.

(29) Core Team R. R: A language and environment for statistical computing. 2013; Available at: http://www.R-project.org/.

592 (30) Li X, Hastie AT, Hawkins GA, Moore WC, Ampleford EJ, Milosevic J, et al. eQTL of bronchial epithelial cells
593 and bronchial alveolar lavage deciphers GWAS-identified asthma genes. Allergy 2015 Oct;70(10):1309-1318.

594 (31) Slater M, Torr E, Harrison T, Forrester D, Knox A, Shaw D, et al. The differential effects of azithromycin on
595 the airway epithelium in vitro and in vivo. Physiol Rep 2016 Sep;4(18):10.14814/phy2.12960.

596 (32) Portelli MA, Stewart CE, Hall IP, Brightling CE, Sayers I. Cigarette Smoke and the Induction of Urokinase
597 Plasminogen Activator Receptor In Vivo: Selective Contribution of Isoforms to Bronchial Epithelial Phenotype.
598 Am J Respir Cell Mol Biol 2015 Aug;53(2):174-183.

599 (33) Kristjansson RP, Benonisdottir S, Davidsson OB, Oddsson A, Tragante V, Sigurdsson JK, et al. A loss-of-
600 function variant in ALOX15 protects against nasal polyps and chronic rhinosinusitis. Nat Genet 2019
601 Feb;51(2):267-276.

602 (34) Pividori M, Schoettler N, Nicolae DL, Ober C, Im HK. Shared and distinct genetic risk factors for childhood-
603 onset and adult-onset asthma: genome-wide and transcriptome-wide studies. Lancet Respir Med 2019
604 Jun;7(6):509-522.

605 (35) Demenais F, Margaritte-Jeannin P, Barnes KC, Cookson WOC, Altmuller J, Ang W, et al. Multiancestry
606 association study identifies new asthma risk loci that colocalize with immune-cell enhancer marks. Nat Genet
607 2018 Jan;50(1):42-53.

608 (36) Johansson A, Rask-Andersen M, Karlsson T, Ek WE. Genome-wide association analysis of 350 000
609 Caucasians from the UK Biobank identifies novel loci for asthma, hay fever and eczema. Hum Mol Genet 2019
610 Jul 30.

611 (37) Ferreira MA, Vonk JM, Baurecht H, Marenholz I, Tian C, Hoffman JD, et al. Shared genetic origin of asthma,
612 hay fever and eczema elucidates allergic disease biology. Nat Genet 2017 Dec;49(12):1752-1757.

613 (38) Pruim RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, Gliedt TP, et al. LocusZoom: regional visualization of
614 genome-wide association scan results. Bioinformatics 2010 Sep 15;26(18):2336-2337.

615 (39) GTEx Consortium. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue
616 gene regulation in humans. Science 2015 May 8;348(6235):648-660.

617 (40) Zeller T, Wild P, Szymczak S, Rotival M, Schillert A, Castagne R, et al. Genetics and beyond--the
618 transcriptome of human monocytes and disease susceptibility. PLoS One 2010 May 18;5(5):e10693.

619 (41) Gorbacheva AM, Korneev KV, Kuprash DV, Mitkin NA. The Risk G Allele of the Single-Nucleotide
620 Polymorphism rs928413 Creates a CREB1-Binding Site That Activates IL33 Promoter in Lung Epithelial Cells. Int J
621 Mol Sci 2018 Sep 25;19(10):10.3390/ijms19102911.

622 (42) Al Efraij K, FitzGerald JM. Benralizumab for the add-on maintenance treatment of patients with severe
623 asthma aged 12 years and older with an eosinophilic phenotype. Expert Rev Clin Pharmacol 2018 Jul;11(7):669-
624 676.

625 (43) Farne HA, Wilson A, Powell C, Bax L, Milan SJ. Anti-IL5 therapies for asthma. Cochrane Database Syst Rev
626 2017 Sep 21;9:CD010834.

627 (44) Emdin CA, Khera AV, Kathiresan S. Mendelian Randomization. JAMA 2017 Nov 21;318(19):1925-1926.

628 (45) Cayrol C, Girard JP. Interleukin-33 (IL-33): A nuclear cytokine from the IL-1 family. Immunol Rev 2018
629 Jan;281(1):154-168.

630 (46) Molofsky AB, Savage AK, Locksley RM. Interleukin-33 in Tissue Homeostasis, Injury, and Inflammation.
631 Immunity 2015 Jun 16;42(6):1005-1019.

632 (47) Lambrecht BN, Hammad H. The immunology of asthma. Nat Immunol 2015 Jan;16(1):45-56.

633 (48) Vieira Braga FA, Kar G, Berg M, Carpaij OA, Polanski K, Simon LM, et al. A cellular census of human lungs
634 identifies novel cell states in health and in asthma. Nat Med 2019 Jul;25(7):1153-1163.

635 (49) Heijink IH, Nawijn MC, Hackett TL. Airway epithelial barrier function regulates the pathogenesis of allergic
636 asthma. Clin Exp Allergy 2014;44(5):620-630.

637 (50) Lambrecht BN, Hammad H. The airway epithelium in asthma. Nat Med 2012 May 4;18(5):684-692.

638 (51) Wark PAB, Ramsahai JM, Pathinayake P, Malik B, Bartlett NW. Respiratory Viruses and Asthma. Semin
639 Respir Crit Care Med 2018 Feb;39(1):45-55.

640 (52) Jurak LM, Xi Y, Landgraf M, Carroll ML, Murray L, Upham JW. Interleukin 33 Selectively Augments
641 Rhinovirus-Induced Type 2 Immune Responses in Asthmatic but not Healthy People. Front Immunol 2018 Aug
642 16;9:1895.

643 (53) Schiller HB, Montoro DT, Simon LM, Rawlins EL, Meyer KB, Strunz M, et al. The Human Lung Cell Atlas: A
644 High-Resolution Reference Map of the Human Lung in Health and Disease. Am J Respir Cell Mol Biol 2019
645 Jul;61(1):31-41.

646 (54) Saglani S, Lui S, Ullmann N, Campbell GA, Sherburn RT, Mathie SA, et al. IL-33 promotes airway remodeling
647 in pediatric patients with severe steroid-resistant asthma. J Allergy Clin Immunol 2013 Sep;132(3):676-685.e13.

648 (55) Nagakumar P, Puttur F, Gregory LG, Denney L, Fleming L, Bush A, et al. Pulmonary type-2 innate lymphoid
649 cells in paediatric severe asthma: phenotype and response to steroids. Eur Respir J 2019 Aug
650 29;54(2):10.1183/13993003.01809-2018. Print 2019 Aug.

651

652

Figure legends:

Figure 1: Overview of the flow of the analyses

In above figure the flow of analysis of the current paper is shown. In the first phase, SNPs in a candidate region (400kb+/- *IL33*) were associated with asthma phenotypes in Lifelines (n=13,395) and GASP/DAG cohorts (1), including blood eosinophils, blood neutrophils, FEV_1, FEV_1/FVC, atopy, blood IgE and age of asthma onset. A total of 161 SNPs (MAF>0.01) were associated with one or more of these phenotypes; the majority of these associations were found in the Lifelines general population cohort. A total of 5 independent LD blocks (r2>0.1) were identified (2a). Conditional analyses on the most significantly associated SNP revealed 2 independent signals left for functional study in QTL cohorts (2b). eQTLs were studied in lung tissue (n=1,111) and bronchial brushes (n=139), eQTL and pQTL in cultured primary human bronchial epithelial cells (HBECS, n=35 (3)). Then, functional elements in the phenotype-associated genetic signals were investigated using ENCODE, PredictSNP, Meta-SNP, Polyphen-2 data (4). Further functional study was done by exposing HBECs (n=18) to asthma-relevant stimuli (HDM, RV16), investigating inducible eQTL and pQTL; as well as investigating the functional effects of elevated IL33 (n=5) *in vitro*, including cell count, metabolic activity, viability, ROS-capturing and resistance (5).

Figure 2- The LD pattern of the five LD blocks (r^2>0.1) with phenotype association

The panel shows the LD pattern of the 5 LD blocks/signals (r^2>0.1) from the 144 SNPs significantly (FDR<0.05) associated with asthma features in the Lifelines general population, Lifelines asthma population and DAG/GASP asthma population. Signal A and E were taken forward in functional assesment in this study. *Image generated using the EUR population of the Phase I cohort of the 1000 genomes study via the online software tool LDlink 3.0, available at: https://analysistools.nci.nih.gov/LDlink/?tab=home.*

Figure 3- Conditioning on the main genetic eosinophilic signal A (rs992969) removes three other signals (B,C,D) associated with blood eosinophils in the general population of Lifelines.

In figure 3 the association between *IL33* region SNPs and level of blood eosinophils in the general population is shown. Four LD blocks (r^2>0.1) could be distinguished for this phenotype (LD block A-D), with LD Block A representing a strong signal, and block B-D a modest signal. Indeed, conditioning on the tagSNP of LD block A (rs992969) removed signals B-D. Conditional analyses were performed in n=13,395 subjects from the Lifelines general population, studying the effect of *IL33* SNPs on level of blood eosinophils, by taking rs992969 (=lowest p-value SNP associated with level of blood eos) as covariate in the regression model. Statistical details can be found in table II. *Red line indicates the cut-off at which the adjusted p-value (FDR) is 0.05. Plots generated using LocusZoom.(38)*

Figure 4- The main genetic *IL33* signal (signal A) selected for functional follow-up associates with level of blood eosinophils (a), asthma diagnosis (b) and eosinophilic asthma (c), whilst signal E associates with eosinophilic asthma (d)

Fig 4a- Signal A (tagSNP rs992969) associates with level of blood eosinophils in the general population of Lifelines (a1), independent of the presence of asthma/allergy (a2 and a3). In panel ***a1*** the results of the association between *IL33* SNPs and blood eosinophil levels in the total general population (n=13,395) of Lifelines are shown, the reference SNP (purple) indicating the tagSNP of LD block A: rs992969, which was significantly associated with blood eosinophil (beta [A allele]= 0.058, SE=0.009, P.adj=7.09E-08). In panel ***a2*** this association was performed in the general population lacking asthma (n=1,066 asthma patients removed), rs992969 (purple) still associating with blood eosinophil levels at similar effect size (n=12,329; rs992969 [A] beta=0.055, SE=0.009, P.adj=1.04E-06). In panel ***a3*** individuals with asthma and allergies (n=6,227 asthma/allergic subjects) were removed, and also then rs992969 (purple dot) associated with blood eosinophil levels at similar effect size (n=7,168; rs992969 [A] beta=0.046, SE=0.012, P.adj=0.02). *Red line indicates the*

cut-off at which the adjusted p-value (FDR) is 0.05. Asthma was defined as self-reported doctor-diagnosed asthma. Allergy was defined based on at least one self-reported allergy, including eczema, rhinitis, food allergy, dust allergy, animal allergy, pollen allergy, medication allergy, contact allergy, and insect bite allergy. Plots generated using LocusZoom.(38)

Fig 4b- Signal A (tagSNP rs992969) associates with asthma diagnosis (Lifelines). Here the association between *IL33* locus SNPs and all asthma is shown, with panel ***b1*** showing the association model corrected for age and gender, whilst in panel ***b2*** the model in addition was corrected for level of blood eosinophils. b1- All asthma, uncorrected for blood eosinophils; asthma patients (n=1,066) vs healthy controls (n=6,863) (rs992969 [A], OR= 1.22, SE= 0.05, P.adj=0.03); b2- All asthma, corrected for blood eosinophils; asthma patients (n=1,066) vs healthy controls (n=6,863) (rs992969 [A], OR=1.19, SE= 0.05, P.adj=0.08). *Red line indicates the cut-off at which the adjusted p-value (FDR) is 0.05. Plots generated using LocusZoom.(38)*

Fig 4c- Signal A (tagSNP rs992969) also associates with eosinophilic asthma in Lifelines (c1), but this signal is not present in non-eosinophilic asthma (c2). In panel ***c1*** the results of the association between *IL33* SNPs and eosinophilic asthma in Lifelines is shown, rs992969 as tagSNP of LD block A significantly associated with this phenotype. Eosinophilic asthma (n=707) vs. healthy controls (n=6,863) (rs992969 [A] OR=1.32, SE=0.06, P.adj=4.73E-03). In panel ***c2*** the association with all asthma phenotypes lacking eosinophilic asthma ('non-eosinophilic asthma') is shown, to which rs992969 (purple) was not significantly associated. Non-eosinophilic asthma (n=359) vs healthy controls (n=6,863) (rs992969 [A] OR=1.09, SE=0.09, P.adj=0.62). *Red line indicates the cut-off at which the adjusted p-value (FDR) is 0.05. Plots generated using LocusZoom.(38)*

Fig 4d- Signal E (tagSNP rs4008366) associates with eosinophilic asthma in Lifelines. In panel ***d*** it is shown that a modest association for signal E exists for eosinophilic asthma. Eosinophilic asthma (n=707) vs. healthy controls (n=6,863) (rs4008366 [G] OR=1.26, SE=0.070, P.adj=0.045). *Red line indicates the cut-off at which the adjusted p-value (FDR) is 0.05. Plot generated using LocusZoom.(38)*

Figure 5- eQTL bronchial brushes in context of eosinophil associated signals

Figure 5- At the *IL33* locus, the phenotype association signals for blood eosinophil counts in the general population (n=13,395) is shown in the upper panel, and the eQTL signals for *IL33* expression in bronchial brushes shown in the lower panel (genotyped SNPs only, n=139 subjects). It becomes clear that the main eosinophil-associated genetic signal A, tagged by rs992969, is also a strong eQTL in bronchial brushes. The A allele associates with higher levels of *IL33* mRNA levels. Statistical details can be found in table II (phenotype) and table III (eQTL). Plots generated using LocusZoom.(38)

Figure 6: Elevated expression of *IL33* affects viability and ROS-capturing, but not barrier formation in bronchial epithelial cells

Panel 6a*:* Elevated *IL33* mRNA (qPCR) was confirmed in the five engineered donor HBEC which was titrated to result in a range around 10 times higher levels of *IL33* in the overexpression condition; matching the fold change in IL33 expression that we found in HBECs from asthmatic donors compared to HBECs from healthy controls (8-10 times higher in asthma HBECs, *not shown*). Data expressed as fold difference in *IL33* mRNA levels compared to no vector control. N=5 HBEC donors, data points represent mean +/-standard deviation for 2 technical replicates per donor.
Panel 6b*:* Protein expression of IL33 (red) was confirmed in HBECs transduced with lentivirus expressing human *IL33*. Cells were processed for immunofluorescent staining at passage 2, two weeks after the lentiviral transduction when cells were considered virus-free.
Panel 6c: Viability of HBECs overexpressing IL33 ('IL-33') was determined using propidium iodide staining in passage 2 cells and compared to empty vector (EV) controls (Kruskall Wallis, followed by MWU *posthoc* statistics). Data expressed relative to no vector (NV) control, mean +/- standard deviation of n=5 cell donors.
Panel 6d: Level of reduced glutathione was stained using a commercially available assay (VitaBright-48™, Chemometec) in passage 2 cells, and HBECs overexpressing *IL33* ('IL-33') compared to empty vector (EV)

controls (Kruskall Wallis, followed by Wilcoxon *posthoc* statistics). Data expressed relative to no vector (NV) control, mean +/- standard deviation of n=5 cell donors.

Phenotypic and functional translation of *IL33* genetics in asthma

Maria E. Ketelaar[1ab,2,] MSc MD, Michael A. Portelli[2,*] PhD, F. Nicole Dijk[1a,*] MD, PhD, N. Shrine[3,4] PhD, A. Faiz[1c] PhD, C.J. Vermeulen[1c] PhD, C.J. Xu PhD[1a,5], J. Hankinson[6] PhD, S. Bhaker[2] PhD, A.P. Henry[2] PhD, C.K. Billington[2] PhD, D.E. Shaw[2] MD PhD, Prof. S.R. Johnson[2] MBBS DM, A.V. Benest[7] PhD, V. Pang[7] PhD, Prof. D. Bates[7], Z.E.K. Pogson[8,9]MD PhD, A. Fogarty[9] PhD, Prof. T. M. McKeever[9] PhD, A. Singapuri[10] BSc, Prof. L. Heaney[11] MD PhD, A.H. Mansur[12] PhD, R. Chaudhuri[13] MD, Prof. N.C. Thomson[13], Prof. J.W. Holloway[14], G.A. Lockett[14] PhD, Prof. P.H. Howarth[14] MD PhD, R. Niven[6] MD, Prof. A. Simpson[6] MD PhD, Prof. M.D. Tobin[3,4], Prof. I.P. Hall[2] FMedSci, Prof. L.V. Wain[3,4], J.D. Blakey[15] PhD, Prof. C.E. Brightling[4,10] MD PhD, M. Obeidat[16] PhD, Prof. D.D. Sin[16,17] MD PhD, C. Nickle[18] PhD, Y. Bossé[19] PhD, J.M. Vonk[1d] PhD, M. van den Berge[1c] MD PhD, Prof. G.H. Koppelman[1a] MD PhD, Prof. Ian Sayers[2#], Martijn C. Nawijn[1b,#] PhD*

[1]University of Groningen, University Medical Center Groningen (UMCG), Groningen Research Institute for Asthma and COPD, Groningen, The Netherlands
[a]University of Groningen, UMCG, Department of Pediatric Pulmonology and Pediatric Allergology, Beatrix Children's Hospital, Groningen, the Netherlands
[b]University of Groningen, UMCG, Department of Pathology and Medical Biology, Groningen, The Netherlands
[c]University of Groningen, UMCG, Department of Pulmonary Diseases, Groningen, The Netherlands
[d]University of Groningen, UMCG, Department of Epidemiology, Groningen, The Netherlands
[2]Division of Respiratory Medicine, National Institute for Health Research Nottingham Biomedical Research Centre, Nottingham University Biodiscovery Institute, University of Nottingham, Nottingham, UK
[3]Department of Health Sciences, University of Leicester, Leicester, UK
[4]National Institute for Health Research Leicester Respiratory Biomedical Research Centre, University of Leicester, Leicester, UK
[5]CiiM & TWINCORE , joint ventures between the Helmholtz-Centre for Infection Research (HZI) and the Hannover Medical School (MHH), Hannover, Germany
[6]Manchester Academic Health Science Centre, University of Manchester, Manchester, UK
[7]Tumour Vascular Biology Laboratories, Division of Cancer and Stem Cells, School of Medicine, Queens Medical Centre, Nottingham, UK and COMPARE University of Birmingham and University of Nottingham, Nottingham, UK.
[8]Department of Respiratory Medicine, Lincoln County Hospital, Lincoln, UK
[9]Division of Epidemiology and Public Health, University of Nottingham, Nottingham, UK
[10]Institute for Lung Health, Department of Respiratory Sciences, University of Leicester, Glenfield Hospital, Leicester, UK
[11]Centre for Infection and Immunity, Queen's University of Belfast, Belfast, UK
[12]Respiratory Medicine, Birmingham Heartlands Hospital and University of Birmingham, Birmingham, UK
[13]Institute of Infection, Immunity and Inflammation, University of Glasgow, Glasgow, UK
[14]Human Development & Health & Clinical and Experimental Sciences, Faculty of Medicine and National Institute of Health Biomedical Research Centre, University of Southampton, Southampton, UK
[15]Respiratory Medicine, Sir Charles Gairdner Hospital, Perth, Australia
[16]The University of British Columbia Centre for Heart Lung Innovation, St Paul's Hospital, Vancouver, BC, Canada.
[17]Division of Respiratory Medicine, Department of Medicine, University of British Columbia, Vancouver, BC, Canada.
[18]Departments of Genetics and Pharmacogenomics, Merck Research Laboratories, Boston, Massachusetts, USA
[19]Institut universitaire de cardiologie et de pneumologie de Québec, Department of Molecular Medicine, Laval University, Québec, Canada
[*]shared first authors, [#]shared last authors

Corresponding author: Maria E. Ketelaar; University Medical Center Groningen, Groningen Research Institute for Asthma and COPD, Department of Pediatric Pulmonology and Pediatric Allergology, CA43, PO Box 30.001, 9700 RB Groningen, the Netherlands; m.e.ketelaar@umcg.nl; Tel: + 31 50 3611036; Fax: + 31 50 3614235

Sources of funding:

This study was supported by Lung Foundation of the Netherlands grants no. AF 95.05 (GHK), AF 98.48 (GHK) and no.AF3.2.09.081JU, (GHK, MCN), the University Medical Center Groningen (GHK), Dutch TerMeulen Fund (MEK) and the Ubbo Emmius Foundation (GHK), and a grant from GSK (IS, IH, MCN, GHK). The Lifelines Biobank initiative has been made possible by subsidy from the Dutch Ministry of Health, Welfare and Sport, the Dutch Ministry of Economic Affairs, the University Medical Center Groningen (UMCG the Netherlands), University Groningen and the Northern Provinces of the Netherlands. The generation of the lung tissue dataset was funded by Merck. This study was also funded by an Asthma UK Grant to IS, IPH, DES, CEB (AUK-PG-2013-188) and additional Asthma UK funding to IS and DES (Grants 10/006 and 11/031). Genotyping in GASP was additionally supported by Rosetrees Trust (Grant to IS), and AirPROM (CEB, MT and IS). This work was supported by the Medical Research Council [grant number MC_PC_12010], a Strategic Award to IPH, MDT, and LVW, and an MRC project grant to SRJ (G1100163). LVW holds a GSK/ British Lung Foundation Chair in Respiratory Research. Asthma UK funded the GASP initiative (AUK-PG-2013-188). This work was part funded by the NIHR Leicester Respiratory Biomedical Centre. AS is supported by the Manchester Biomedical Research Centre.

Declaration of potential conflict of interests:

GHK, MCN, MEK, CJX, MAP, IS and IH report research funding from Glaxo Smith Kline relating to this manuscript. IS has had research funding relating to this manuscript from AnaptysBio Inc. JDB reports personal fees and non-financial support from Napp, personal fees from Novartis, personal fees and non-financial support from Astra Zeneca, personal fees and non-financial support from Boehringer Ingelheim, personal fees from Teva, personal fees from Innovate UK, outside the submitted work; SRJ reports grants from Medical Research Council, during the conduct of the study; non-financial support from Boehringer-Ingelheim, outside the submitted work; CEB reports grants from AirPROM FP7, grants from Asthma UK, grants from NIHR Biomedical Research Centre, during the conduct of the study; DSP reports grants from Glaxo Smith Kline, during the conduct of the study; grants from Glaxo Smith Kline, outside the submitted work; GHK reports grants from TEVA the Netherlands, Vertex, and Stichting Astma Bestrijding, outside the submitted work; and advisory board fees from GSK and PureIMS, outside the submitted work; MCN reports grants from Glaxo Smith Kline, outside the submitted work; IS reports grants from Glaxo Smith Kline, grants from Anaptsbio Inc, outside the submitted work; RC reports personal fees and non-financial support from AstraZeneca, personal fees from Glaxo Smith Kline, personal fees from Teva Pharmaceuticals, personal fees and non-financial support from Novartis, outside the submitted work; the remaining authors have declared that no conflict of interest exists. AVB and DOB are supported by British Heart Foundation grant and AVB is supported by a Royal Society Project grant RGS\R1\191221.

Abstract (249):

Background: Asthma is a complex disease with multiple phenotypes that may differ in disease pathobiology and treatment response. Interleukin 33 (*IL33*) single nucleotide polymorphisms (SNPs) have been reproducibly associated with asthma. IL33 levels are elevated in sputum, and bronchial biopsies of asthma patients. The functional consequences of *IL33* asthma SNPs remain unknown.

Objective: We studied whether *IL33* SNPs associate with asthma-related phenotypes and with *IL33* expression in lung or bronchial epithelium. We investigated the effect of increased *IL33* expression on human bronchial epithelial cell (HBEC) function.

Methods: Association between *IL33* SNPs (Chr9: 5,815,786–6,657,983) and asthma phenotypes (Lifelines/DAG/GASP cohorts) and between SNPs and expression (lung tissue, bronchial brushes, HBECs) was done using regression modelling. Lentiviral overexpression was used to study *IL33* effects on HBECs.

Results: 161 SNPs spanning the *IL33* region associated with one or more asthma phenotypes after correction for multiple testing. We report one main independent signal tagged by rs992969 associating with blood eosinophil levels, asthma and eosinophilic asthma. A second, independent signal tagged by rs4008366 presented modest association with eosinophilic asthma. Neither signal associated with FEV_1, FEV_1/FVC, atopy, and age of asthma onset. The two *IL33* signals are expression quantitative loci (eQTLs) in bronchial brushes and cultured HBECs, but not in lung tissue. *IL33* overexpression *in vitro* resulted in reduced viability and ROS-capturing of HBECs, without influencing epithelial cell count, metabolic activity or barrier function.

Conclusion: We identify *IL33* as an epithelial susceptibility gene for eosinophilia and asthma, provide mechanistic insight, and implicate targeting of the IL33 pathway specifically in eosinophilic asthma.

Key Messages:

- Genetic signals at the *IL33* locus predominantly associate with blood eosinophil counts in the general population and with an eosinophilic asthma phenotype.
- These genetic signals influence *IL33* levels in the airway epithelium, with the disease risk allele associating with elevated *IL33 in vivo*.
- Elevated *IL33* has modest paracrine effects on bronchial epithelial cell function *in vitro*, implicating that epithelial derived IL33 may more likely affect other effector cell types such as type 2 immune cells, eosinophils or mast cells.

Capsule Summary (30, max 35 words):

This study identifies *IL33* as an epithelial susceptibility gene for eosinophilia and eosinophilic asthma, supporting the *IL33* pathway as a likely candidate for targeted treatment strategies specifically in eosinophilic asthma.

Keywords:

asthma phenotypes, *IL33* SNPs, eQTL, bronchial epithelium, functional translation

Abbreviations:

AHBEC: Asthma Human Bronchial Epithelial Cell
ALI: Air Liquid Interface
AOO: Age Of Onset
AUC: Area Under the Curve
BEGM: Bronchial Epithelial Growth Medium
BHR: Bronchial Hyper Responsiveness
CMV: CytoMegaloVirus
ECIS: Electric Cell Substrate Impedance Sensing
EUR: European
DAG: Dutch Asthma GWAS
FDR: False Discovery Rate
FeNO: Fraction of exhaled Nitric Oxide
FEV_1: Forced Expiratory Volume 1st second
FVC: Forced Vital Capacity
GASP: Genetics of Severe Asthma Phenotypes
GSH: Glutathione
GWAS: Genome Wide Association Study
HBEC: Human Bronchial Epithelial Cell
HDM: House Dust Mite
HWE: Hardy Weinberg Equilibrium
IF: ImmunoFluorescence
IgE: Immunoglobulin E
IL1RL1: Interleukin 1 Receptor Like 1
IL1RAP: Interleukin 1 Receptor Accessory Protein
IL33: Interleukin 33
kU: kiloUnit
LD: Linkage Disequilibrium
MAF: Minor Allele Frequency
MOI: Multiplicity Of Infection
NGS: Next Generation Sequencing
OR: Odds Ratio
P.adj: Adjusted p-value (FDR)
PI: Propidium Iodide
qPCR: quantitative Polymerase Chain Reaction
QTL: Quantitative Trait Locus
ROS: Reactive Oxygen Species
RV: Rhinovirus
SNP: Single Nucleotide Polymorphism
SPT: Skin Prick Test
TF: Transcription Factor
Th2: Type 2 T helper

Introduction:

Asthma is a common, complex, heterogeneous disease that results from the interaction between genetic and environmental factors. It is a chronic inflammatory condition of the airways, characterized by bronchial hyperresponsiveness and reversible airway obstruction. Asthma may consist of several endotypes characterized by differences in specific phenotypes, underlying pathobiology, and (treatment) outcomes in individual patients(1). Genome-wide association studies (GWAS) have identified a large number of asthma loci(2-9), including single nucleotide polymorphisms (SNPs) in *Interleukin (IL)33* and the gene encoding its receptor *Interleukin 1 Receptor Like 1 (IL1RL1)*(9). Both loci were originally discovered as associated with blood eosinophils in general population cohorts(10,11). Next to these common SNPs, a rare *IL33* loss of function mutation has been shown to reduce blood eosinophil counts and protect from asthma(12).

IL33 is an alarmin released upon cellular damage from e.g. epithelial cells. Extracellular IL33 induces signalling via the heterodimeric receptor complex IL1RL1/IL1RAP. Airway IL33 levels have been associated with type 2 cytokines levels and a positive correlation with eosinophil numbers in asthma patients was recently reported(13). High IL33 levels have been found in induced sputum and bronchial biopsies of asthma patients compared to non-asthmatic controls(14-16). Moreover, IL33 may have a paracrine effect on the airway epithelium, as this epithelium has been shown responsive to IL33(17,18). These data suggest a connection between epithelium-derived IL33, eosinophilic inflammation and asthma.

Nevertheless, the functional relevance of common asthma-associated SNPs in *IL33* remains largely unknown. Moreover, genetic association studies thus far have focused on asthma diagnosis, while the contribution of genetic variants to distinct phenotypes of asthma has not been addressed. We hypothesize that genetic variants at the *IL33* locus drive specific phenotypes of asthma by activating a type-2 cytokine dominated immune response, characterized by eosinophilic lung inflammation. Therefore, this study aimed to investigate 1) whether SNPs in the *IL33* region associate with specific asthma phenotypes; 2) whether these *IL33* SNPs form quantitative trait loci (QTL) for *IL33* expression in lung tissue and/or bronchial epithelial samples *in vivo* and *in vitro;* and 3) whether increased *IL33* expression alters human bronchial epithelial cell function.

Materials/Methods:

Detailed methods are described in the online supplement.
Codes available on https://git.web.rug.nl/P252222/IL33_Ketelaaretal_JACI2020. ('wiki')

Study design (see also figure 1)

SNPs in the region of *IL33* (Chr9: 5,815,786–6,657,983, GRCh37/hg19) were tested for association with asthma phenotypes using regression modelling. Briefly, we tested association of the *IL33* SNPs in a Dutch general population cohort (Lifelines(19); n=13,395) with eosinophil counts, FEV_1 and FEV_1/FVC. From this general population cohort we subsequently took the asthma subpopulation (n=1,066, doctor's diagnosed asthma) and investigated genetic association with eosinophil counts, eosinophilic asthma (asthma and blood eosinophils >150 cells/µL, n=707, as this cut-off is a good predictor for airway eosinophilia (>2% sputum eos (20), non-eosinophilic asthma (asthma and blood eosinophils <150 cells/uL, n=359), FEV_1, FEV_1/FVC and asthma with airway obstruction (asthma and FEV_1<80% of predicted (n=258) or FEV_1/FVC<70% (n=324)). In a meta-analysis of two independent asthma cohorts of n=2,536 moderate-severe asthma patients (GASP, UK(21))) and n=909 asthma patients of mild-moderate severity (DAG, the Netherlands(22)), we then evaluated association of *IL33* SNPs with atopy, blood eosinophils, total serum IgE, age of asthma onset, and lung function (FEV_1, FEV_1/FVC).
We then selected independent genetic signals based on LD (r^2<0.1), followed by conditional analyses on the most significantly associated SNP. Functional investigations of selected independent genetic signals included expression and protein quantitative loci studies in lung tissue (n=1,111), bronchial brushes (n=139) and primary asthma derived human bronchial epithelial cells (AHBECs, n=35). Potential function was investigated using ENCODE, PredictSNP, Meta-SNP, and Polyphen-2 data(23-25). We tested for inducible expression (e)QTL and protein (p)QTL by exposing AHBECs (n=18) of various *IL33* genotypes to asthma-relevant stimuli (HDM, RV16). Finally, we overexpressed *IL33* in (healthy-derived) HBECs (n=5) to investigate effects on cell count, metabolic activity, viability, ROS-capturing and epithelial barrier.

Genotype-phenotype analysis

A total of 1,970 imputed SNPs (Lifelines, all overlapping with DAG/GASP) and 2,457 imputed SNPs (DAG/GASP) were available for the association analyses based on a MAF≥0.01 and chromosomal location of 400kb up- and downstream *IL33* (Chr9: 5,815,786–6,657,983). This region encompasses all known asthma-associated SNPs (table I, table S1). Associations of SNPs with asthma phenotypes were performed with PLINK v1.90b6.7(26) (Lifelines) or SNPtest v2.5β(27) (DAG/GASP) using an additive genetic model. DAG/GASP were meta-analysed in METAL(28) using a fixed model, (table S2-S3). An adjusted p-value (P.adj)<0.05 (FDR) was considered statistically significant.

Functional genetics

QTL and ENCODE investigations

We tested for expression quantitative trait loci (eQTL) in lung tissue (n=1,111) and bronchial brushes (n=139) (table S5) using a linear regression model to investigate the association between SNPs and log-transformed *IL33* expression data. We employed an additive genetic model with age, gender, smoking status and the PCs explaining >1% of expression variance as covariates using R statistics(29). We did not have data on medication use for these cohorts, so could not correct for this covariate, but the currently used covariates are thought to reflect main confounders in eQTL analyses (22,30).We also tested for (inducible) QTLs in cultured bronchial epithelial cells (n=18-35) obtained from bronchial brushes/biopsies from asthma patients as described(31). AHBECs were stimulated with 50µg/ml house dust mite (HDM) or rhinovirus (RV16, MOI=1) for 24 hours and RNA lysates collected(32). Cells were genotype-stratified and expression compared using Kruskal-Wallis tests. A p-value<0.05 was considered statistically significant. ENCODE was used to identify potential functional effects of tagSNPs and SNPs in LD (r^2>0.3). SNPs were functionally checked for DNase I hypersensitive sites, histone mark sites, binding motifs, and regulatory motifs using RegulomeDB, HaploReg, ChromHMM and Segway(23,24).

Functional bronchial epithelial cell studies

In order to investigate the functional consequences of increased IL33 in bronchial epithelial cells, we stably overexpressed human full-length *IL33* (aa1-270) in primary HBECs isolated from 5 healthy individuals (Lonza, #CC-2540). IL33 mRNA and protein expression was quantified by qPCR and immunofluorescence respectively. We analysed cell count, viability and metabolic activity, as well as ROS-capturing ability (glutathione assay) and barrier function (ECIS) in these cultures. We used Kruskal-Wallis for all parameters except for longitudinal area under the curves (AUCs) comparisons of ECIS data, which were compared using a Z-test. A p-value<0.05 was considered statistically significant.

Results:

Genetic association with phenotypes of asthma:

The IL33 locus particularly associates with eosinophilia and eosinophilic asthma

Overall in DAG/GASP and Lifelines (S2-S3), 161 SNPs significantly associated with one or more asthma phenotypes (P.adj<0.05 (FDR); tables S11-15), mainly derived from the Lifelines cohort. From these, 144 SNPs comprised five LD blocks (A-E, r^2>0.1). Markedly, these five LD blocks all associated with an eosinophilic phenotype; either with blood eosinophil counts, eosinophilic asthma and/or asthma (table I, tables S11-S17, figure E2/E3). LD block A shows a significant association with blood eosinophil counts in the general population (tagSNP rs992969[allele A] beta=0.058+/- SE=0.0089, P.adj=7.09E-08, AF=0.25), whilst three other LD blocks were modestly associated with this phenotype (block B-D, table I). Block E showed association with eosinophilic asthma (tagSNP rs4008366 [allele T], OR=1.26+/-SE=0.0704, P.adj=0.045, AF=0.67) only.

Outside these five LD blocks, seven SNPs significantly associated with other phenotypes (age of onset or FEV_1/FVC, table S15) and ten significant SNPs were identified in the case-control analyses of resequencing data; these were relatively rare (MAF~0.03) and hence were not followed-up functionally. We performed conditional analyses on the LD blocks associated with eosinophilic phenotypes to determine independent signals. A summary description of association results can be found in the supplementary section M1.

Conditional and sensitivity analyses show one main genetic signal associated with blood eosinophil counts in the general population

Four LD blocks (A-D, figure 2) showed association with blood eosinophil counts in the Lifelines general population. Thereby, block A (tagSNP rs992969) shows the largest effect size and statistical significance (table I); rs992969 explaining 1.6% (R^2 regression model=0.016) of the variance in blood eosinophil counts (corrected for age/gender). Therefore, we conditioned the association analysis for blood eosinophils on rs992969 to test whether block A-D are independent signals. Conditioning removed the association of signals B-D with blood eosinophil counts in the general population (see figure 3 and table II). Signal E was not significantly associated with eosinophil counts, regardless of conditioning. Sensitivity analysis for the main signal A showed that rs992969 still associated with eosinophil counts in the general population when removing asthma patients (figure 3a-2) (n=12,329; rs992969 [allele A] beta=0.055, SE=0.009, R^2=0.017, P.adj=1.04E-06) or both asthma/allergic patients (figure 3a-3) (n=6,227; rs992969 [allele A] beta=0.046, SE=0.012, R^2=0.020, P.adj=0.02). These analyses show the presence of one main genetic signal (A) at the *IL33* locus associated with blood eosinophil counts in the general population, independent of the presence of asthma/allergy phenotypes.

Signal A and E associate with eosinophilic asthma

Signal A, driving the association with blood eosinophil counts in the general population, also showed a significant association with asthma (rs992969[allele A], OR=1.22, SE=0.05, P.adj=0.03) and with eosinophilic asthma (rs992969[allele A], OR= 1.32+/-SE=0.0618, P.adj=4.73E-03), (figure 3+4). Signal A contains several SNPs previously associated with asthma (table I). The genetic effect of this main signal on asthma risk remained of similar size, even when correcting for blood eosinophil counts (OR[A] from 1.22 (P.adj=0.03) to 1.19 (P.adj=0.08), figure 4b).

Signal E was the other LD block associated with eosinophilic asthma (tagSNP rs4008366, figure 4d), with a significantly large effect size. Lack of power precluded conditional analyses for the eosinophilic asthma phenotype, so (in)dependency of block E could not be confirmed. However, this block represents a genetically independent signal in other cohorts (table II), underscoring it may be a distinct signal and may represent a distinct mechanism underlying asthma pathogenesis. Therefore, two signals (A and E) were selected for functional follow-up.

To assess whether our definition of eosinophilic asthma based on the cut-off for blood eosinophils at 150 cells/uL impacted on the associations observed, we repeated the analysis at a cut-off of 300 cells/uL as a definition for eosinophilic asthma. These additional analyses of eosinophilic asthma, including a higher cut-off of eosinophil counts, identify the same associations with slightly higher effect sizes (see supplemental table S8/S9), but not FDR (<0.05) significant anymore, likely explained by the more refined phenotype resulting in smaller group sizes.

309 **Table I- Five LD blocks (r^2>0.1) with phenotype associations could be distinguished**

Tag SNP (genetic signal)	Location	Gene context	Pheno risk allele (AF)	Alt allele	Associated feature(s)	Cohort	Effect size risk allele (B or OR)	SE	P.adj (FDR)	Lit. asthma GWAS SNP in block	Ref GWAS
A-rs992969	9:6209697	~6kb 5' of *IL33*	A (0.25)	G	*eos levels in GenPop,* *eos asthma vs HC,* *asthma case control*	Lifelines Lifelines Lifelines	0.058 (B) 1.321 (OR) 1.230 (OR)	0.009 0.062 0.053	7.09E-08 4.73E-03 0.034	rs1888909 rs7848215 rs992969 rs144829310 rs72699186 rs928413 rs1342326 rs2381416 rs2066362	(2-7,21,33-37)
B-rs1342327	9:6189874	~25kb 5'of *IL33*	G (0.15)	C	*eos levels in GenPop,* *eos levels in asthma subjects*	Lifelines, DAG/GASP	0.035 (B) 0.057 (B)	0.011 0.018	0.027 0.039	-	-
C-rs74438701	9:6282794	~25kb 3'of *IL33*	T (0.83)	C	*eos levels in GenPop*	Lifelines	0.035 (B)	0.011	0.041	-	-
D-rs2282162	9:6534466	intronic of *GLDC*	G (0.56)	A	*eos levels in GenPop*	Lifelines	0.029 (B)	0.008	0.011	-	-
E-rs4008366	9:6116407	intergenic	T (0.69)	C	*eos asthma vs HC*	Lifelines	1.264 (OR)	0.070	0.045	rs343478	(36,37)

310 The table shows the tagSNPs representing 5 LD blocks/signals (r^2>0.1) from the SNPs significantly (FDR<0.05) associated with asthma features in the Lifelines general population,
311 Lifelines asthma population and DAG/GASP asthma population. In the last columns these are put into context of previously reported genome-wide significant (5*10E-8) SNPs
312 associated with asthma, displaying SNPs part of the LD block at r^2>0.1. **Underlined**: the two genetic signals (A and E) taken forward in functional assesment in this study. Because of
313 its association with eosinophilic asthma, lack of LD with signal A, as well as this LD block also represented an independent signal in multiple studies, we took signal E forward as an
314 independent phenotype-associated signal in our functional analyses. *AF=frequency (EUR 1000G); Alt allele=alternative allele; B=beta; eos=eosinophils/eosinophilic; FDR=false*
315 *discovery rate value at alpha 0.05; GWAS= genomewide association study; GenPop =general population; HC=healthy control; kb= kilo basepairs; Lit.=literature; OR=odds ratio;*
316 *P.adj= FDR adjusted p-value; Pheno risk allele=phenotype associated allele; Ref=Reference; SE=standard error*. For complete analyses of all eosinophilic phenotypes in each cohort,
317 please see supplemental table S7.

318

Table II- Conditioning on rs992969 in the association of *IL33* with blood eosinophils in the general population of Lifelines removed signals B-D

LD Block (tagSNP)	Location	Gene context	Pheno risk allele (AF)	Alt allele	Effect size risk allele (beta)**	SE	P.adj (FDR)	Independent signal Ubiobank/ INTERVAL^ (blood eos GenPop)(10)	Independent signal UK biobank only^ (asthma) (29)	Independent signal SHARE^ (asthma/allergy) (31)
A-rs992969	9:62096 97	~6kb 5' of *IL33*	*A (0.25)*	*G*	**0.058** *n/a*	**0.009** *n/a*	**7.09E-08***** *n/a*	rs2381416 (r^2= 0.95)	rs7848215 (r^2=0.93)	rs144829310 (r^2= 0.59)
B-rs1342327	9:61898 74	~25kb 5'of *IL33*	*G (0.15)*	*C*	**0.035** *0.012*	**0.011** *0.011*	**0.027**** *0.877*	-	-	-
C-rs74438701	9:62827 94	~25kb 3'of *IL33*	*T (0.83)*	*C*	**0.035** *0.017*	**0.011** *0.011*	**0.041**** *0.722*	-	-	-
D-rs2282162	9:65344 66	intronic of *GLDC*	*G (0.56)*	*A*	**0.029** *0.012*	**0.008** *0.009*	**0.011**** *0.722*	-	-	-
E-rs4008366*	9:61164 07	intergenic	*T (0.69)*	*C*	**0.010** *0.002*	**0.009** *0.009*	**0.647** *0.974*	-	rs343478 (r^2= 0.17)	rs343478 (r^2= 0.17)

Conditional analyses were performed in n=13,395 subjects from the Lifelines general population, studying the effect of *IL33* SNPs on level of blood eosinophils, by taking rs992969 (=lowest p-value SNP associated with level of blood eos) as covariate in the regression model. These were put into the context of independent SNPs as determined in other large cohorts. r^2= relative to tagSNP of LD block A/B/C/D/E respectively.

* Signal E was not significantly associated with level of blood eosinophils in the general population before conditional analyses, nor after conditional analyses, but has only been included in this table to show it is in modest LD with rs343478 (an independent signal in other studies). Because of its association with eosinophilic asthma, lack of LD with signal A, as well as this LD block also represented an independent signal in multiple studies, we took this signal forward as an independent phenotype-associated signal in our functional analyses.

In **bold the unconditioned results, in *italics* the results conditioned on rs992969.

***Adjusted p-value (FDR) statistically significant <0.05.

^ Independent phenotype-associated SNPs at the *IL33* locus determined based on conditional analyses in other large population cohorts: the phenotype studied in the UK Biobank/INTERVAL was blood eosinophil levels in the general population (n=173,480)(10), in the UK biobank only was asthma (n=41,926 cases vs. n=239,773 controls) (36), whilst the SHARE study examined a combined asthma/allergy phenotype (n=180,129 cases vs 180,709 controls)(37).

Underlined: the two genetic signals taken forward in functional assessment in this study

QTL/functional investigation of *IL33* genetic variation:

After conditional analyses, two independent signals A and E remained for functional follow-up, each with a tagSNP (rs992969 and rs4008366, table II). These tagSNPs were chosen based on smallest p-value/largest effect size, largest number of associated phenotypes and if applicable known association with asthma from literature. In case the tagSNP was not available for functional look-up, a proxySNP at r^2>0.5 with the tagSNP of the original association signal was chosen (table S7).

Signal A and E are IL33 eQTL in bronchial epithelium

To investigate potential functionality of signal A and E, QTL analyses were performed in lung tissue, bronchial epithelial brushes and cultured bronchial epithelial cells (see table III, S5 and figure 5).

In lung tissue samples, no eQTLs for *IL33* were found (table S19 and figure E4). In bronchial brushes, the tagSNP of signal A was a significant and strong eQTL for *IL33*, with the disease associated allele correlating with higher mRNA levels, (rs992969[A] beta=0.331, SD=0.043, P=8.30E-12, AF=0.25). No significant eQTLs were found for signal E in bronchial brushes. In cultured primary human bronchial epithelial cells (HBECs), the disease associated allele of signal E (proxy SNP rs442246) associated with lower *IL33* mRNA (P=0.029, see table III and figure E6b). No significant pQTLs were found for IL33 in HBECs for both signal A and E (figure E7).

Signal A and E harbour potential functional elements related to expression regulation of IL33

ENCODE revealed several putative regulatory elements for SNPs in both genetic signals A and E relevant for *IL33* transcription (table IIIb). Signal A contained 27 SNPs (LD r^2>0.3 with tagSNP) with potential functionality. Among these is a SNP (rs928413) in strong LD with the phenotype and expression associated tagSNP rs992969 (r^2=0.96), forming a CREB1 binding site activating the *IL33* promotor. In signal E, 7 SNPs were potential functional elements, including specific transcription factor binding sites relevant to the regulation of the cellular oxidative state (e.g. Nrf2) in lung derived cells. Thus, the genetic signals A and E contain likely functional elements related to expression, forming a potential mechanistic link between phenotype and expression association.

Asthma stimuli induce differential IL33 expression, regardless of genetic background for signal A and E

Next, we tested for the presence of inducible QTLs for *IL33* in primary AHBECs after exposure to RV16 (MOI:1) or house dust mite (HDM, 50µg/mL) and analysed for effects on *IL33* mRNA and extracellular protein levels in an unstratified way or stratified for genetic signals A and E. RV16 induced a decrease in *IL33* mRNA levels in AHBEC (P=0.048), and a marked increase of IL33 protein in the cellular supernatant (P=0.0001). HDM exposure induced an increase in *IL33* RNA and had no significant effects on IL33 protein levels, measured 24h post stimulation (figure E7). When stratified on signals A and E, no significant differences on the RV16 or HDM induced effects on IL33 mRNA or protein levels were observed (figures E8/E9).

IL33 overexpression modestly impairs bronchial epithelial cell homeostasis:

In order to investigate the effect of increased *IL33* expression, we overexpressed full length *IL33* in primary bronchial epithelial cells using lentiviral delivery (figure E11/E12). We confirmed increased expression of *IL33* at the mRNA level and presence of IL33 protein in engineered cells (figure 5 and figures E13/E14). We found that overexpression of *IL33* does not significantly influence cell number or metabolic activity (figure E15). Viability was 15-20% lower (P=0.04, figure 5c) and ROS-capturing capacity (presence of free glutathione) was ~20% lower (P=0.03, figure 5d) in cells that overexpressed *IL33* under submerged culture condition. No effect of *IL33* overexpression was seen on spreading or formation of an epithelial barrier using ECIS (figure E14).

391

392 **Table III- QTL function and functional ENCODE annotation of the phenotype associated signals A and E:**

III A *Tag SNP (genetic signal)*	*Pheno risk allele (AF)*	*Alt allele*	*Associated phenotype(s)*	*QTL cohort*	*Effect size pheno risk allele*	*SD*	*P-value*	*Direction pheno risk allele*	*Literature Reported QTL function*
A- *rs992969*	***A* (0.25)**	***G***	*eos levels in GenPop, eos asthma vs HC, asthma case control*	Bronchial brushes	0.326 (B)	0.043	8.30E-12	*++IL33 RNA*	cisQTL IL33 bronchial biopsies/blood/brain; (30,39,40)
E- *rs442246 (proxy for: rs4008366)*	***T* (0.69)**	***G***	*eos asthma vs HC*	Cultured HBEC	-2.377 (fold change TT)		0.0298	*--IL33 RNA*	-

III B *Tag SNP (genetic signal)*	*Location*	*Gene context*	*Associated phenotypes*	*Functional annotation of genetic signal, SNPs r^2>0.3 with tagSNP*					
				Promotor cRE (lung)	**Enhancer cRE (lung)**	**DNase I site cRE (lung)**	**Protein-binding (lung)**	**PredictSNP/DANN**	**Experimental functionality**
A-rs992969	9:6209697	**~6kb 5' of *IL33***	*eos levels in GenPop, eos asthma vs HC, asthma case control*	Y- H3K4me3	Y-H3K27ac	Y	CTCF, SETDB1, CFOS, PRDM1, STAT3	neutral	r^2=0.96 with rs928413(G) forming CREB1 binding site, activating IL33 promotor lung epithelial cells (41)
E-rs4008366	9:6116407	**Intergenic**	*eos asthma vs HC*	-	Y- H3K27ac	Y	Nrf-2, TCF11, MafG, ZID, Hmbox1, Hoxd8	Deleterious (0.85 accuracy)	-

393 **Panel IIIA:** The table shows quantitative trait loci (QTL) function of the two genetic signals in the *IL33* region associated with eosinophilic asthma features in our cohorts. In case the tagSNP
394 was not available, a proxy at r^2>0.3 was used for QTL look-up. Expression (e)QTLs were studied in lung tissue (lung surgery patients) and bronchial brushes (healthy subjects); eQTL and
395 protein(p)QTL function were studied in cultured primary human bronchial epithelial cells (AHBECs) from asthma patients. Of note: in lung tissue no significant eQTLs for *IL33* were found in
396 the *IL33* region, and in HBECs no significant pQTLs were found for these 2 genetic signals (alpha=0.05). In bronchial brushes, signal A was an eQTL for *IL33*, with the phenotype risk allele
397 associating with higher *IL33* mRNA levels. In cultured HBECs signal E has potential QTL function; the eosinophilic asthma risk allele associating with lower *IL33* RNA. More details can be
398 found in figure E3-E5 (supplemental). *Pheno Risk allele=phenotype associated allele; Alt allele=alternative allele; AF=allele frequency (EUR 1000G); B=beta; SD=standard deviation;*
399 *eos=eosinophils/eosinophilic; GenPop=general population; HC=healthy control; ++=increased expression, --=decreased expression.* **Panel IIIB:** The table shows the functional ENCODE
400 and PredictSNP, Meta-SNP, Polyphen-2 (23,24) look-up of the two genetic signals that were selected from the SNPs significantly (P.adj (FDR)<0.05) associated with asthma features in
401 Lifelines general population, Lifelines asthma population and DAG/GASP asthma population. SNPs in LD r^2>0.3 with the tagSNP of the applicable genetic signal were included in the functional

look-up. ENCODE and Polyphen-2 retrieved functional annotation for the signals, while PredictSNP and Meta-SNP did not. *cRE=candidate regulatory element; eos=eosinophils/eosinophilic; DANN – Deleterious Annotation of Genetic Variants using Neural Networks, HC=healthy control; GenPop=general population; kb= kilo basepairs; Y=yes*

Discussion

We set out to determine if SNPs in the *IL33* region associate with specific phenotypes of asthma, whether these regulate *IL33* expression in lung tissue or bronchial epithelial samples, and whether increased *IL33* expression alters human bronchial epithelial cell biology. Genetic signals at the *IL33* locus predominantly associate with an eosinophilic phenotype in the general population and asthma subjects, whereby the *IL33* risk allele is associated with higher *IL33* expression *in vivo*. Using conditional analyses we observed one major genetic signal and one secondary signal. The major signal associates with blood eosinophil counts and (eosinophilic) asthma, while the secondary signal associates with eosinophilic asthma but not with eosinophil counts in the general population. Importantly, no association with other asthma-related phenotypes including lung function, atopy, serum IgE levels and asthma age of onset was observed. Studying the effects of these two genetic signals on *IL33* transcription, we report eQTLs in bronchial brushes and cultured bronchial epithelial cells, but not in lung tissue. Overexpression of *IL33* in bronchial epithelial cells resulted in modest paracrine effect on epithelial cell homeostasis, including reduction in cell viability and ROS-capturing capacity. With this approach we identify *IL33* as an epithelial susceptibility gene for eosinophilia and asthma, provide mechanistic insight, and support targeting of the IL33 pathway specifically in eosinophilic asthma.

Two genetic IL33 signals associate with eosinophilia in health and disease

The *IL33* gene, and the *IL1RL1* gene encoding its receptor, have consistently been associated with asthma and allergy(3-8,21,33-37). Both loci were originally discovered as regions associating with blood eosinophils in the Icelandic population(11,33), and a strong association with blood eosinophil counts was recently confirmed in a large general population cohort (n=173,480), combining UK Biobank and INTERVAL studies(10). Also, a rare loss-of-function *IL33* mutation was shown to both reduce eosinophil counts and to protect from asthma(12). These observations suggest a shared genetic effect of this locus for eosinophilia and asthma. However, it remained unknown whether these are the same or distinct genetic signals and what additional asthma related phenotypes these signals may be associated with.

We report five LD blocks that were associated with either blood eosinophil counts and/or eosinophilic asthma, which after conditional analysis correcting for the strongest signal (rs992969), were reduced to two independent signals. The fifth signal (E) was not associated with blood eosinophil counts in the general population, but with eosinophilic asthma. The available subjects (n=707) for the eosinophilic asthma phenotype did not allow conditional analyses for signal E. However, previous analyses in two very large cohorts (SHARE(37) and UK-biobank(36) supported the independence of signal E, representing a second signal associated with an eosinophilic phenotype in our cohorts. This left us with two genetic signals for further study.

We observe a strong association for signal A with both asthma and blood eosinophil counts in our Lifelines general population cohort, the tagSNP rs992969 explaining 1.6% (R^2=0.016) of the variance in eosinophil counts (corrected for age/gender). SNPs within this signal have previously been reported to associate with asthma in the UK Biobank, SHARE, and TAGC study, as well as in earlier asthma meta-analyses(2,4,5,7,21,35-37) and with blood eosinophil counts in the UK biobank/INTERVAL study (rs992969 in LD r^2=0.95 with rs2381416 from UK Biobank/INTERVAL)(10). Using a sensitivity analysis in Lifelines by removing asthmatic and allergic subjects from the general population, we show that the association with blood eosinophils remained present with a similar effect size, indicating that the association between this signal A and blood eosinophils is not fully driven by the presence of asthma or allergy. We find that the association of signal A with asthma is of similar effect size when correcting for blood eosinophil counts, suggesting that this IL33 genetic signal- in addition to its effect on blood eosinophil counts- may have an effect on asthma. However, we do find that the effect of signal A on asthma after correcting for blood eosinophils is no longer FDR significant. (Figure 4.) Therefore, a better powered study is required to conclusively investigate an effect of this signal on asthma independent from eosinophil counts. Interestingly, we observed an association of signal A with eosinophilic asthma, but not with non-eosinophilic asthma (figure 4), indicating that patients with this *IL33* genetic make-up would be enriched in the high-eosinophil group. A note of caution is the relatively limited number of subjects in our non-eosinophilic asthma group (n=359).

An intriguing implication could be that in asthma patients with this particular genetic background (signal A), treatment targeting the IL33 pathway could have additional effects over treatments targeting eosinophils(42,43). Notwithstanding, whether the association of *IL33* SNPs with asthma and eosinophils are (in)dependent from each

other remains to be conclusively determined in larger cohorts, allowing causal inference/mediating approaches like Mendelian randomization(44). Ideally, such an analysis would also take into account *IL1RL1* genotypes, which are likely to interact with *IL33* variants on outcomes such as eosinophilic inflammation; as well as a more direct measure of eosinophilic airway inflammation such as sputum eosinophil counts should be considered.

Functional effects of phenotype-associated IL33 polymorphisms and IL33 expression

Functionally, IL33 signalling has previously been linked to Th2-driven inflammation, contributing to eosinophilic inflammation(45-47). Moreover, levels of IL33 have been found elevated in induced sputum and bronchial biopsies of asthma patients compared to non-asthmatic controls(14-16), indicating a dysregulation of IL33 homeostasis in asthma. Therefore, specific genetic variation at the *IL33* locus might contribute to eosinophil numbers and asthma through regulation of *IL33* expression levels.

While we did not detect eQTLs for *IL33* in lung tissue samples, the tagSNP of signal A was a strong eQTL for *IL33* in bronchial epithelial brushes from healthy subjects (table III, figure 5), with the risk allele associating with increased *IL33* mRNA levels. Signal A harbours a reported *IL33* eQTL in a candidate eQTL study of bronchial biopsies(30), with the same direction of effect. This eQTL signal A also comprised a SNP (rs928413, in LD r^2=0.96 with rs992969) where the phenotype-risk allele was recently found to form a CREB1 binding site, functionally activating the *IL33* promotor in lung epithelial cells(41). This allele associates with higher level of eosinophils, higher risk of (eosinophilic) asthma and increased *IL33* expression in brushes in our cohorts. As lung tissue resection samples mainly consist of parenchymal lung tissue with minor contributions of airway epithelial cells, while bronchial brushes contain more than 90% bronchial epithelial cells(48), we interpret these data as evidence for regulation of *IL33* expression in bronchial epithelium.

The bronchial epithelium is the first barrier that the inhaled substances encounter when entering the lung, and serves to protect the body from potential threats from the environment. The airway epithelium is changed in asthma patients, with increased susceptibility to and altered repair responses after external damage(49,50), for example in response to respiratory viruses(51). A GWAS on exacerbation in asthma(2) found the *IL33* locus associated with frequent virus-induced exacerbations in severe childhood-onset asthma, their main *IL33* SNP in strong LD (r^2=0.96) with our eosinophilic signal rs992969. Therefore, we tested whether our two phenotype-associated signals are baseline and/or induced QTLs in cultured bronchial epithelium. Signal E is a modest, baseline *IL33* eQTL, with the eosinophilic risk allele associated with lower *IL33* mRNA levels (table III) in these cells. Although both RV16 and HDM regulated *IL33* expression, no effect of the two *IL33* signals on the RV16 and HDM induced IL33 response was observed in vitro in our samples. This could indicate that our two signals may specifically have effects on baseline changes of *IL33* expression in epithelium.

The opposite direction of effect in the cultured bronchial epithelial cells compared to the bronchial brushes might indicate that *IL33* gene regulation is different in asthmatic epithelium compared to healthy brushed cells, in agreement with recent data from Jurak *et al.* (2018)(52). Alternatively, it could reflect differences in epithelial cell state with cultured HBECs having a basal cell phenotype(48,53), while bronchial brushes contain mostly well-differentiated ciliated and secretory epithelial cells(48).

The cell-autonomous effects of increased *IL33* expression on cultured HBECs were modest. Nevertheless, the observed effect of sustained IL33 on reduction of glutathione levels in the epithelium is interesting, as Uchida *et al.*(18) showed that the balance between oxidative stress and antioxidant responses plays a key role in controlling IL33 release from airway epithelium. Our data indicate that the bronchial epithelium is the source of IL33, but that other cell types should be considered as the main IL33 responsive population, such as tissue-resident dendritic cells, eosinophils, type 2 innate lymphoid cells, Th2 cells, mast cells and basophils, but also lung mesenchymal, such as fibroblasts. This is also relevant in the context of steroid-resistant asthma patients. For example: elevated IL-33 and type 2 cells were still present in corticosteroid resistant pediatric asthma patients, contributing to airway remodelling via its effects on airway fibroblast.(54,55)

In conclusion, we have reduced the complex *IL33* locus into one major and one secondary genetic signal for eosinophilic asthma. The major *IL33* signal risk allele associates with increased *IL33* expression levels providing a putative mechanism. Importantly we have also shown a lack of genetic association of this main genetic signal with other studied asthma phenotypes. We identified the bronchial epithelial cell as the likely cellular source of *IL33* QTL signals, which is crucial to place the genetic effects on *IL33* expression in asthma pathophysiology. These data need confirmation by e.g. single-cell eQTL analyses in airway wall samples of asthma patients and healthy controls. This approach might also guide the identification of the main IL33 responding cells. Nevertheless, our data identifies *IL33* as an epithelial susceptibility gene for eosinophilia and asthma, and

supports the IL33 pathway as a likely candidate for targeted treatment strategies in specifically eosinophilic asthma, with the potential to affect both eosinophil counts and asthma independently.

Acknowledgements:
The authors wish to acknowledge the services of the Lifelines Cohort Study, the contributing research centres delivering data to Lifelines, and all the study participants. Also, we acknowledge the ENCODE Consortium for generating online accessible datasets to consult for functional regulatory elements. Furthermore, this work was supported in part by the NIHR Nottingham Biomedical Research Centre, which IH, IS and MAP would like to acknowledge here.

References:

(1) Kaur R, Chupp G. Phenotypes and endotypes of adult asthma: Moving toward precision medicine. J Allergy Clin Immunol 2019 Jul;144(1):1-12.

(2) Bonnelykke K, Sleiman P, Nielsen K, Kreiner-Moller E, Mercader JM, Belgrave D, et al. A genome-wide association study identifies CDHR3 as a susceptibility locus for early childhood asthma with severe exacerbations. Nat Genet 2014 Jan;46(1):51-55.

(3) Ferreira MA, Matheson MC, Tang CS, Granell R, Ang W, Hui J, et al. Genome-wide association analysis identifies 11 risk variants associated with the asthma with hay fever phenotype. J Allergy Clin Immunol 2014 Jun;133(6):1564-1571.

(4) Torgerson DG, Ampleford EJ, Chiu GY, Gauderman WJ, Gignoux CR, Graves PE, et al. Meta-analysis of genome-wide association studies of asthma in ethnically diverse North American populations. Nat Genet 2011 Jul 31;43(9):887-892.

(5) Moffatt MF, Gut IG, Demenais F, Strachan DP, Bouzigon E, Heath S, et al. A large-scale, consortium-based genomewide association study of asthma. N Engl J Med 2010 Sep 23;363(13):1211-1221.

(6) Pickrell JK, Berisa T, Liu JZ, Segurel L, Tung JY, Hinds DA. Detection and interpretation of shared genetic influences on 42 human traits. Nat Genet 2016 Jul;48(7):709-717.

(7) Ferreira MA, Matheson MC, Duffy DL, Marks GB, Hui J, Le Souef P, et al. Identification of IL6R and chromosome 11q13.5 as risk loci for asthma. Lancet 2011 Sep 10;378(9795):1006-1014.

(8) Vicente CT, Revez JA, Ferreira MAR. Lessons from ten years of genome-wide association studies of asthma. Clin Transl Immunology 2017 Dec 15;6(12):e165.

(9) Grotenboer NS, Ketelaar ME, Koppelman GH, Nawijn MC. Decoding asthma: translating genetic variation in IL33 and IL1RL1 into disease pathophysiology. J Allergy Clin Immunol 2013 Mar;131(3):856-865.

(10) Astle WJ, Elding H, Jiang T, Allen D, Ruklisa D, Mann AL, et al. The Allelic Landscape of Human Blood Cell Trait Variation and Links to Common Complex Disease. Cell 2016 Nov 17;167(5):1415-1429.e19.

(11) Gudbjartsson DF, Bjornsdottir US, Halapi E, Helgadottir A, Sulem P, Jonsdottir GM, et al. Sequence variants affecting eosinophil numbers associate with asthma and myocardial infarction. Nat Genet 2009 Mar;41(3):342-347.

(12) Smith D, Helgason H, Sulem P, Bjornsdottir US, Lim AC, Sveinbjornsson G, et al. A rare IL33 loss-of-function mutation reduces blood eosinophil counts and protects from asthma. PLoS Genet 2017 Mar 8;13(3):e1006659.

(13) Poulsen NN, Bjerregaard A, Khoo SK, Laing IA, Le Souef P, Backer V, et al. Airway Interleukin-33 and type 2 cytokines in adult patients with acute asthma. Respir Med 2018 Jul;140:50-56.

(14) Hamzaoui A, Berraies A, Kaabachi W, Haifa M, Ammar J, Kamel H. Induced sputum levels of IL-33 and soluble ST2 in young asthmatic children. J Asthma 2013 Oct;50(8):803-809.

(15) Gasiuniene E, Janulaityte I, Zemeckiene Z, Barkauskiene D, Sitkauskiene B. Elevated levels of interleukin-33 are associated with allergic and eosinophilic asthma. Scand J Immunol 2019 May;89(5):e12724.

(16) Wang W, Li Y, Lv Z, Chen Y, Li Y, Huang K, et al. Bronchial Allergen Challenge of Patients with Atopic Asthma Triggers an Alarmin (IL-33, TSLP, and IL-25) Response in the Airways Epithelium and Submucosa. J Immunol 2018 Oct 15;201(8):2221-2231.

(17) Altman MC, Lai Y, Nolin JD, Long S, Chen CC, Piliponsky AM, et al. Airway epithelium-shifted mast cell infiltration regulates asthmatic inflammation via IL-33 signaling. J Clin Invest 2019 Aug 22.

(18) Uchida M, Anderson EL, Squillace DL, Patil N, Maniak PJ, Iijima K, et al. Oxidative stress serves as a key checkpoint for IL-33 release by airway epithelium. Allergy 2017 Oct;72(10):1521-1531.

(19) Scholtens S, Smidt N, Swertz MA, Bakker SJ, Dotinga A, Vonk JM, et al. Cohort Profile: LifeLines, a three-generation cohort study and biobank. Int J Epidemiol 2015 Aug;44(4):1172-1180.

(20) Gao J, Wu F. Association between fractional exhaled nitric oxide, sputum induction and peripheral blood eosinophil in uncontrolled asthma. Allergy Asthma Clin Immunol 2018 May 23;14:21-018-0248-7. eCollection 2018.

(21) Shrine N, Portelli MA, John C, Soler Artigas M, Bennett N, Hall R, et al. Moderate-to-severe asthma in individuals of European ancestry: a genome-wide association study. Lancet Respir Med 2019 Jan;7(1):20-34.

(22) Nieuwenhuis MA, Siedlinski M, van den Berge M, Granell R, Li X, Niens M, et al. Combining genomewide association study and lung eQTL analysis provides evidence for novel genes associated with asthma. Allergy 2016 Dec;71(12):1712-1720.

(23) ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. Nature 2012 Sep 6;489(7414):57-74.

(24) Davis CA, Hitz BC, Sloan CA, Chan ET, Davidson JM, Gabdank I, et al. The Encyclopedia of DNA elements (ENCODE): data portal update. Nucleic Acids Res 2018 Jan 4;46(D1):D794-D801.

(25) Adzhubei I, Jordan DM, Sunyaev SR. Predicting functional effect of human missense mutations using PolyPhen-2. Curr Protoc Hum Genet 2013 Jan;Chapter 7:Unit7.20.

(26) Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007 Sep;81(3):559-575.

(27) Marchini J, Howie B, Myers S, McVean G, Donnelly P. A new multipoint method for genome-wide association studies by imputation of genotypes. Nat Genet 2007 Jul;39(7):906-913.

(28) Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics 2010 Sep 1;26(17):2190-2191.

(29) Core Team R. R: A language and environment for statistical computing. 2013; Available at: http://www.R-project.org/.

592 (30) Li X, Hastie AT, Hawkins GA, Moore WC, Ampleford EJ, Milosevic J, et al. eQTL of bronchial epithelial cells
593 and bronchial alveolar lavage deciphers GWAS-identified asthma genes. Allergy 2015 Oct;70(10):1309-1318.

594 (31) Slater M, Torr E, Harrison T, Forrester D, Knox A, Shaw D, et al. The differential effects of azithromycin on
595 the airway epithelium in vitro and in vivo. Physiol Rep 2016 Sep;4(18):10.14814/phy2.12960.

596 (32) Portelli MA, Stewart CE, Hall IP, Brightling CE, Sayers I. Cigarette Smoke and the Induction of Urokinase
597 Plasminogen Activator Receptor In Vivo: Selective Contribution of Isoforms to Bronchial Epithelial Phenotype.
598 Am J Respir Cell Mol Biol 2015 Aug;53(2):174-183.

599 (33) Kristjansson RP, Benonisdottir S, Davidsson OB, Oddsson A, Tragante V, Sigurdsson JK, et al. A loss-of-
600 function variant in ALOX15 protects against nasal polyps and chronic rhinosinusitis. Nat Genet 2019
601 Feb;51(2):267-276.

602 (34) Pividori M, Schoettler N, Nicolae DL, Ober C, Im HK. Shared and distinct genetic risk factors for childhood-
603 onset and adult-onset asthma: genome-wide and transcriptome-wide studies. Lancet Respir Med 2019
604 Jun;7(6):509-522.

605 (35) Demenais F, Margaritte-Jeannin P, Barnes KC, Cookson WOC, Altmuller J, Ang W, et al. Multiancestry
606 association study identifies new asthma risk loci that colocalize with immune-cell enhancer marks. Nat Genet
607 2018 Jan;50(1):42-53.

608 (36) Johansson A, Rask-Andersen M, Karlsson T, Ek WE. Genome-wide association analysis of 350 000
609 Caucasians from the UK Biobank identifies novel loci for asthma, hay fever and eczema. Hum Mol Genet 2019
610 Jul 30.

611 (37) Ferreira MA, Vonk JM, Baurecht H, Marenholz I, Tian C, Hoffman JD, et al. Shared genetic origin of asthma,
612 hay fever and eczema elucidates allergic disease biology. Nat Genet 2017 Dec;49(12):1752-1757.

613 (38) Pruim RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, Gliedt TP, et al. LocusZoom: regional visualization of
614 genome-wide association scan results. Bioinformatics 2010 Sep 15;26(18):2336-2337.

615 (39) GTEx Consortium. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue
616 gene regulation in humans. Science 2015 May 8;348(6235):648-660.

617 (40) Zeller T, Wild P, Szymczak S, Rotival M, Schillert A, Castagne R, et al. Genetics and beyond--the
618 transcriptome of human monocytes and disease susceptibility. PLoS One 2010 May 18;5(5):e10693.

619 (41) Gorbacheva AM, Korneev KV, Kuprash DV, Mitkin NA. The Risk G Allele of the Single-Nucleotide
620 Polymorphism rs928413 Creates a CREB1-Binding Site That Activates IL33 Promoter in Lung Epithelial Cells. Int J
621 Mol Sci 2018 Sep 25;19(10):10.3390/ijms19102911.

622 (42) Al Efraij K, FitzGerald JM. Benralizumab for the add-on maintenance treatment of patients with severe
623 asthma aged 12 years and older with an eosinophilic phenotype. Expert Rev Clin Pharmacol 2018 Jul;11(7):669-
624 676.

625 (43) Farne HA, Wilson A, Powell C, Bax L, Milan SJ. Anti-IL5 therapies for asthma. Cochrane Database Syst Rev
626 2017 Sep 21;9:CD010834.

627 (44) Emdin CA, Khera AV, Kathiresan S. Mendelian Randomization. JAMA 2017 Nov 21;318(19):1925-1926.

628 (45) Cayrol C, Girard JP. Interleukin-33 (IL-33): A nuclear cytokine from the IL-1 family. Immunol Rev 2018
629 Jan;281(1):154-168.

630 (46) Molofsky AB, Savage AK, Locksley RM. Interleukin-33 in Tissue Homeostasis, Injury, and Inflammation.
631 Immunity 2015 Jun 16;42(6):1005-1019.

632 (47) Lambrecht BN, Hammad H. The immunology of asthma. Nat Immunol 2015 Jan;16(1):45-56.

633 (48) Vieira Braga FA, Kar G, Berg M, Carpaij OA, Polanski K, Simon LM, et al. A cellular census of human lungs
634 identifies novel cell states in health and in asthma. Nat Med 2019 Jul;25(7):1153-1163.

635 (49) Heijink IH, Nawijn MC, Hackett TL. Airway epithelial barrier function regulates the pathogenesis of allergic
636 asthma. Clin Exp Allergy 2014;44(5):620-630.

637 (50) Lambrecht BN, Hammad H. The airway epithelium in asthma. Nat Med 2012 May 4;18(5):684-692.

638 (51) Wark PAB, Ramsahai JM, Pathinayake P, Malik B, Bartlett NW. Respiratory Viruses and Asthma. Semin
639 Respir Crit Care Med 2018 Feb;39(1):45-55.

640 (52) Jurak LM, Xi Y, Landgraf M, Carroll ML, Murray L, Upham JW. Interleukin 33 Selectively Augments
641 Rhinovirus-Induced Type 2 Immune Responses in Asthmatic but not Healthy People. Front Immunol 2018 Aug
642 16;9:1895.

643 (53) Schiller HB, Montoro DT, Simon LM, Rawlins EL, Meyer KB, Strunz M, et al. The Human Lung Cell Atlas: A
644 High-Resolution Reference Map of the Human Lung in Health and Disease. Am J Respir Cell Mol Biol 2019
645 Jul;61(1):31-41.

646 (54) Saglani S, Lui S, Ullmann N, Campbell GA, Sherburn RT, Mathie SA, et al. IL-33 promotes airway remodeling
647 in pediatric patients with severe steroid-resistant asthma. J Allergy Clin Immunol 2013 Sep;132(3):676-685.e13.

648 (55) Nagakumar P, Puttur F, Gregory LG, Denney L, Fleming L, Bush A, et al. Pulmonary type-2 innate lymphoid
649 cells in paediatric severe asthma: phenotype and response to steroids. Eur Respir J 2019 Aug
650 29;54(2):10.1183/13993003.01809-2018. Print 2019 Aug.

651

652

Figure legends:

Figure 1: Overview of the flow of the analyses

In above figure the flow of analysis of the current paper is shown. In the first phase, SNPs in a candidate region (400kb+/- *IL33*) were associated with asthma phenotypes in Lifelines (n=13,395) and GASP/DAG cohorts (1), including blood eosinophils, blood neutrophils, FEV_1, FEV_1/FVC, atopy, blood IgE and age of asthma onset. A total of 161 SNPs (MAF>0.01) were associated with one or more of these phenotypes; the majority of these associations were found in the Lifelines general population cohort. A total of 5 independent LD blocks (r2>0.1) were identified (2a). Conditional analyses on the most significantly associated SNP revealed 2 independent signals left for functional study in QTL cohorts (2b). eQTLs were studied in lung tissue (n=1,111) and bronchial brushes (n=139), eQTL and pQTL in cultured primary human bronchial epithelial cells (HBECS, n=35 (3)). Then, functional elements in the phenotype-associated genetic signals were investigated using ENCODE, PredictSNP, Meta-SNP, Polyphen-2 data (4). Further functional study was done by exposing HBECs (n=18) to asthma-relevant stimuli (HDM, RV16), investigating inducible eQTL and pQTL; as well as investigating the functional effects of elevated IL33 (n=5) *in vitro*, including cell count, metabolic activity, viability, ROS-capturing and resistance (5).

Figure 2- The LD pattern of the five LD blocks (r^2>0.1) with phenotype association

The panel shows the LD pattern of the 5 LD blocks/signals (r^2>0.1) from the 144 SNPs significantly (FDR<0.05) associated with asthma features in the Lifelines general population, Lifelines asthma population and DAG/GASP asthma population. Signal A and E were taken forward in functional assesment in this study. *Image generated using the EUR population of the Phase I cohort of the 1000 genomes study via the online software tool LDlink 3.0, available at: https://analysistools.nci.nih.gov/LDlink/?tab=home.*

Figure 3- Conditioning on the main genetic eosinophilic signal A (rs992969) removes three other signals (B,C,D) associated with blood eosinophils in the general population of Lifelines.

In figure 3 the association between *IL33* region SNPs and level of blood eosinophils in the general population is shown. Four LD blocks (r^2>0.1) could be distinguished for this phenotype (LD block A-D), with LD Block A representing a strong signal, and block B-D a modest signal. Indeed, conditioning on the tagSNP of LD block A (rs992969) removed signals B-D. Conditional analyses were performed in n=13,395 subjects from the Lifelines general population, studying the effect of *IL33* SNPs on level of blood eosinophils, by taking rs992969 (=lowest p-value SNP associated with level of blood eos) as covariate in the regression model. Statistical details can be found in table II. *Red line indicates the cut-off at which the adjusted p-value (FDR) is 0.05. Plots generated using LocusZoom.(38)*

Figure 4- The main genetic *IL33* signal (signal A) selected for functional follow-up associates with level of blood eosinophils (a), asthma diagnosis (b) and eosinophilic asthma (c), whilst signal E associates with eosinophilic asthma (d)

Fig 4a- Signal A (tagSNP rs992969) associates with level of blood eosinophils in the general population of Lifelines (a1), independent of the presence of asthma/allergy (a2 and a3). In panel ***a1*** the results of the association between *IL33* SNPs and blood eosinophil levels in the total general population (n=13,395) of Lifelines are shown, the reference SNP (purple) indicating the tagSNP of LD block A: rs992969, which was significantly associated with blood eosinophil (beta [A allele]= 0.058, SE=0.009, P.adj=7.09E-08). In panel ***a2*** this association was performed in the general population lacking asthma (n=1,066 asthma patients removed), rs992969 (purple) still associating with blood eosinophil levels at similar effect size (n=12,329; rs992969 [A] beta=0.055, SE=0.009, P.adj=1.04E-06). In panel ***a3*** individuals with asthma and allergies (n=6,227 asthma/allergic subjects) were removed, and also then rs992969 (purple dot) associated with blood eosinophil levels at similar effect size (n=7,168; rs992969 [A] beta=0.046, SE=0.012, P.adj=0.02). *Red line indicates the*

cut-off at which the adjusted p-value (FDR) is 0.05. Asthma was defined as self-reported doctor-diagnosed asthma. Allergy was defined based on at least one self-reported allergy, including eczema, rhinitis, food allergy, dust allergy, animal allergy, pollen allergy, medication allergy, contact allergy, and insect bite allergy. Plots generated using LocusZoom.(38)

Fig 4b- Signal A (tagSNP rs992969) associates with asthma diagnosis (Lifelines). Here the association between *IL33* locus SNPs and all asthma is shown, with panel ***b1*** showing the association model corrected for age and gender, whilst in panel ***b2*** the model in addition was corrected for level of blood eosinophils. b1- All asthma, uncorrected for blood eosinophils; asthma patients (n=1,066) vs healthy controls (n=6,863) (rs992969 [A], OR= 1.22, SE= 0.05, P.adj=0.03); b2- All asthma, corrected for blood eosinophils; asthma patients (n=1,066) vs healthy controls (n=6,863) (rs992969 [A], OR=1.19, SE= 0.05, P.adj=0.08). *Red line indicates the cut-off at which the adjusted p-value (FDR) is 0.05. Plots generated using LocusZoom.(38)*

Fig 4c- Signal A (tagSNP rs992969) also associates with eosinophilic asthma in Lifelines (c1), but this signal is not present in non-eosinophilic asthma (c2). In panel ***c1*** the results of the association between *IL33* SNPs and eosinophilic asthma in Lifelines is shown, rs992969 as tagSNP of LD block A significantly associated with this phenotype. Eosinophilic asthma (n=707) vs. healthy controls (n=6,863) (rs992969 [A] OR=1.32, SE=0.06, P.adj=4.73E-03). In panel ***c2*** the association with all asthma phenotypes lacking eosinophilic asthma ('non-eosinophilic asthma') is shown, to which rs992969 (purple) was not significantly associated. Non-eosinophilic asthma (n=359) vs healthy controls (n=6,863) (rs992969 [A] OR=1.09, SE=0.09, P.adj=0.62). *Red line indicates the cut-off at which the adjusted p-value (FDR) is 0.05. Plots generated using LocusZoom.(38)*

Fig 4d- Signal E (tagSNP rs4008366) associates with eosinophilic asthma in Lifelines. In panel ***d*** it is shown that a modest association for signal E exists for eosinophilic asthma. Eosinophilic asthma (n=707) vs. healthy controls (n=6,863) (rs4008366 [G] OR=1.26, SE=0.070, P.adj=0.045). *Red line indicates the cut-off at which the adjusted p-value (FDR) is 0.05. Plot generated using LocusZoom.(38)*

Figure 5- eQTL bronchial brushes in context of eosinophil associated signals

Figure 5- At the *IL33* locus, the phenotype association signals for blood eosinophil counts in the general population (n=13,395) is shown in the upper panel, and the eQTL signals for *IL33* expression in bronchial brushes shown in the lower panel (genotyped SNPs only, n=139 subjects). It becomes clear that the main eosinophil-associated genetic signal A, tagged by rs992969, is also a strong eQTL in bronchial brushes. The A allele associates with higher levels of *IL33* mRNA levels. Statistical details can be found in table II (phenotype) and table III (eQTL). Plots generated using LocusZoom.(38)

Figure 6: Elevated expression of *IL33* affects viability and ROS-capturing, but not barrier formation in bronchial epithelial cells

Panel 6a*:* Elevated *IL33* mRNA (qPCR) was confirmed in the five engineered donor HBEC which was titrated to result in a range around 10 times higher levels of *IL33* in the overexpression condition; matching the fold change in IL33 expression that we found in HBECs from asthmatic donors compared to HBECs from healthy controls (8-10 times higher in asthma HBECs, *not shown*). Data expressed as fold difference in *IL33* mRNA levels compared to no vector control. N=5 HBEC donors, data points represent mean +/-standard deviation for 2 technical replicates per donor.
Panel 6b*:* Protein expression of IL33 (red) was confirmed in HBECs transduced with lentivirus expressing human *IL33*. Cells were processed for immunofluorescent staining at passage 2, two weeks after the lentiviral transduction when cells were considered virus-free.
Panel 6c: Viability of HBECs overexpressing IL33 ('IL-33') was determined using propidium iodide staining in passage 2 cells and compared to empty vector (EV) controls (Kruskall Wallis, followed by MWU *posthoc* statistics). Data expressed relative to no vector (NV) control, mean +/- standard deviation of n=5 cell donors.
Panel 6d: Level of reduced glutathione was stained using a commercially available assay (VitaBright-48™, Chemometec) in passage 2 cells, and HBECs overexpressing *IL33* ('IL-33') compared to empty vector (EV)

controls (Kruskall Wallis, followed by Wilcoxon *posthoc* statistics). Data expressed relative to no vector (NV) control, mean +/- standard deviation of n=5 cell donors.

Figure No.1

Click here to access/download;Figure No.;Figure 1_Overview of the flow of the analyses.jpg

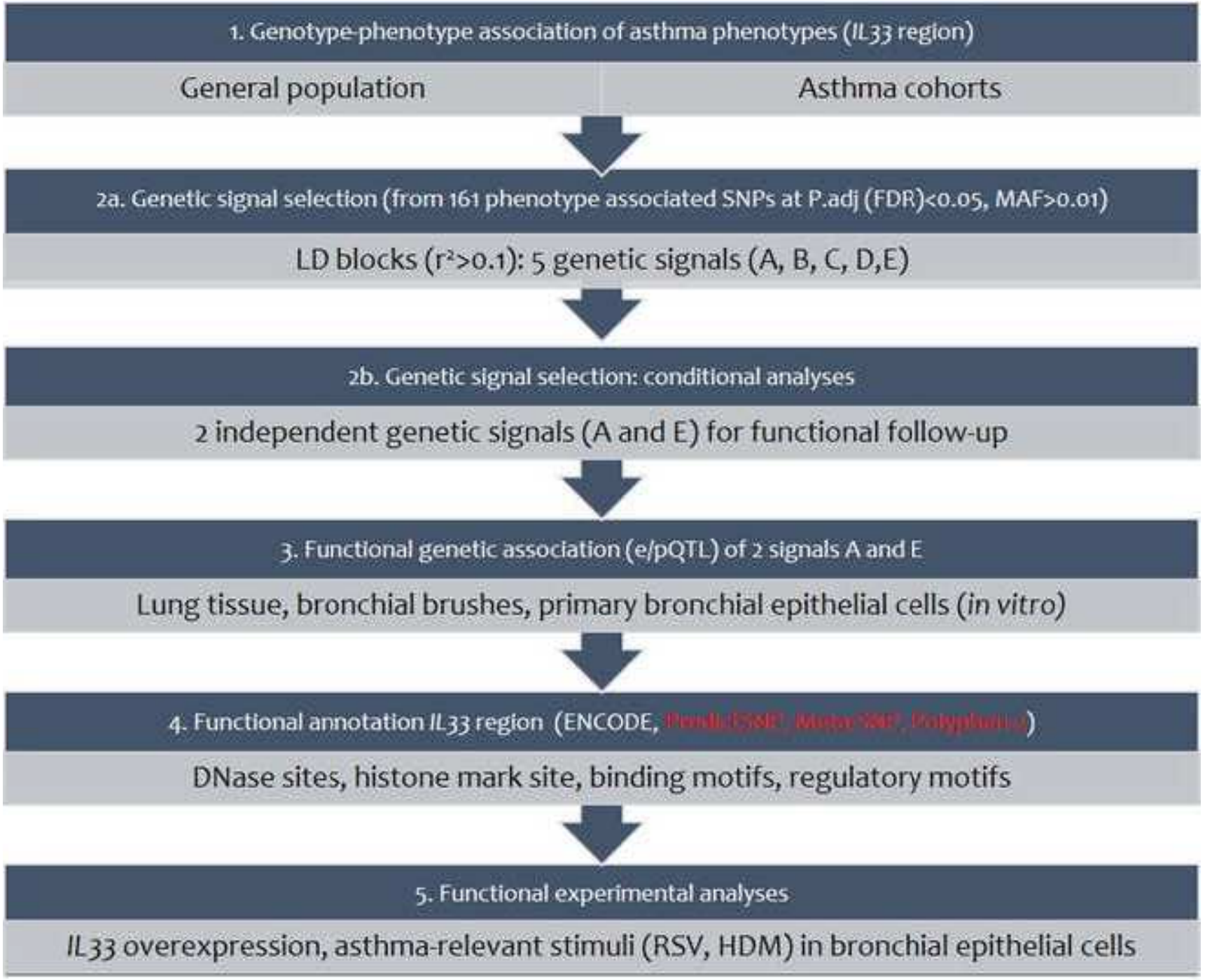


Figure No.2

Click here to access/download;Figure No.;Figure 2_The LD pattern of the five LD blocks.jpg

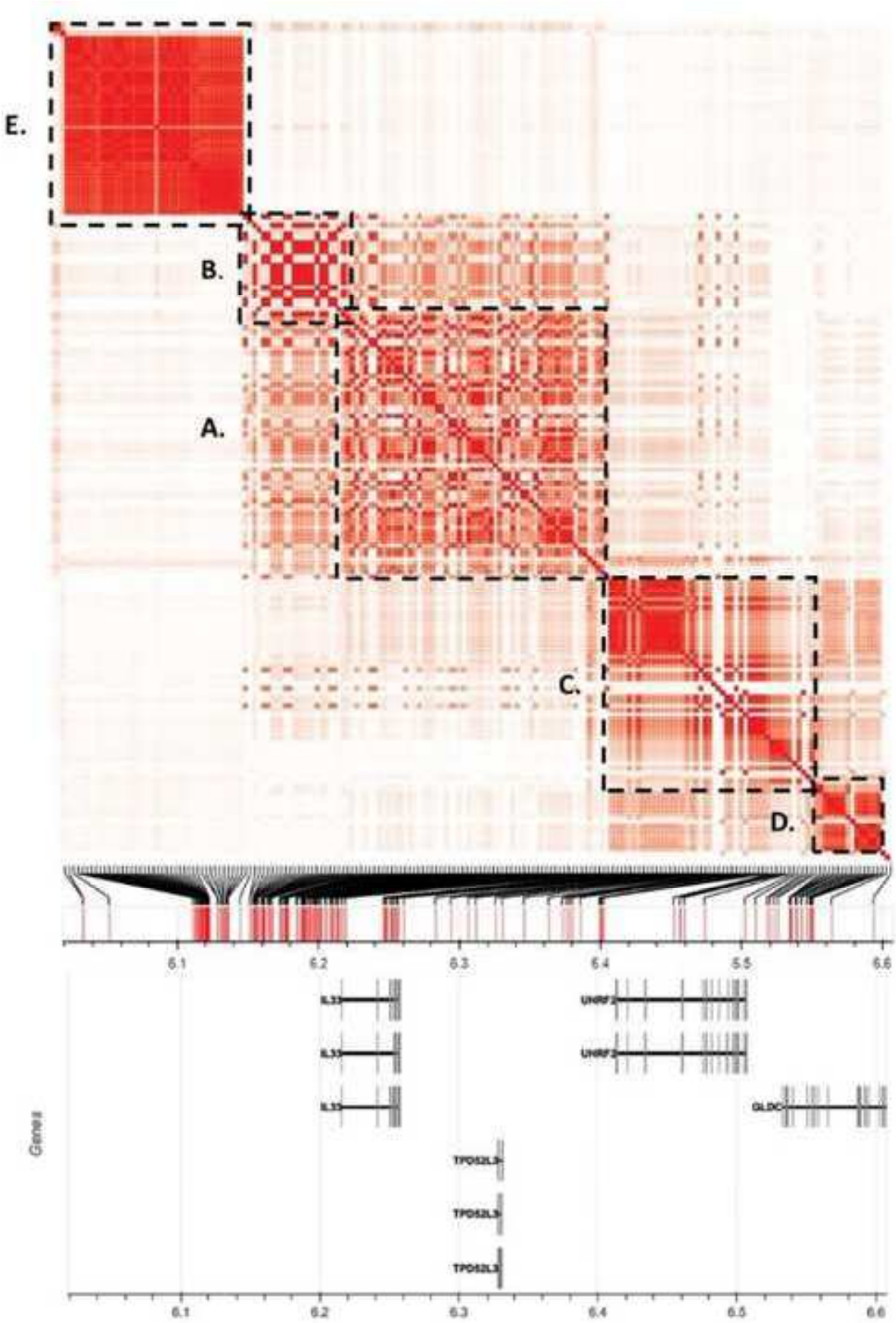


Figure No.3

Click here to access/download;Figure No.;Figure 3_ConditioningonLDblockA.jpg

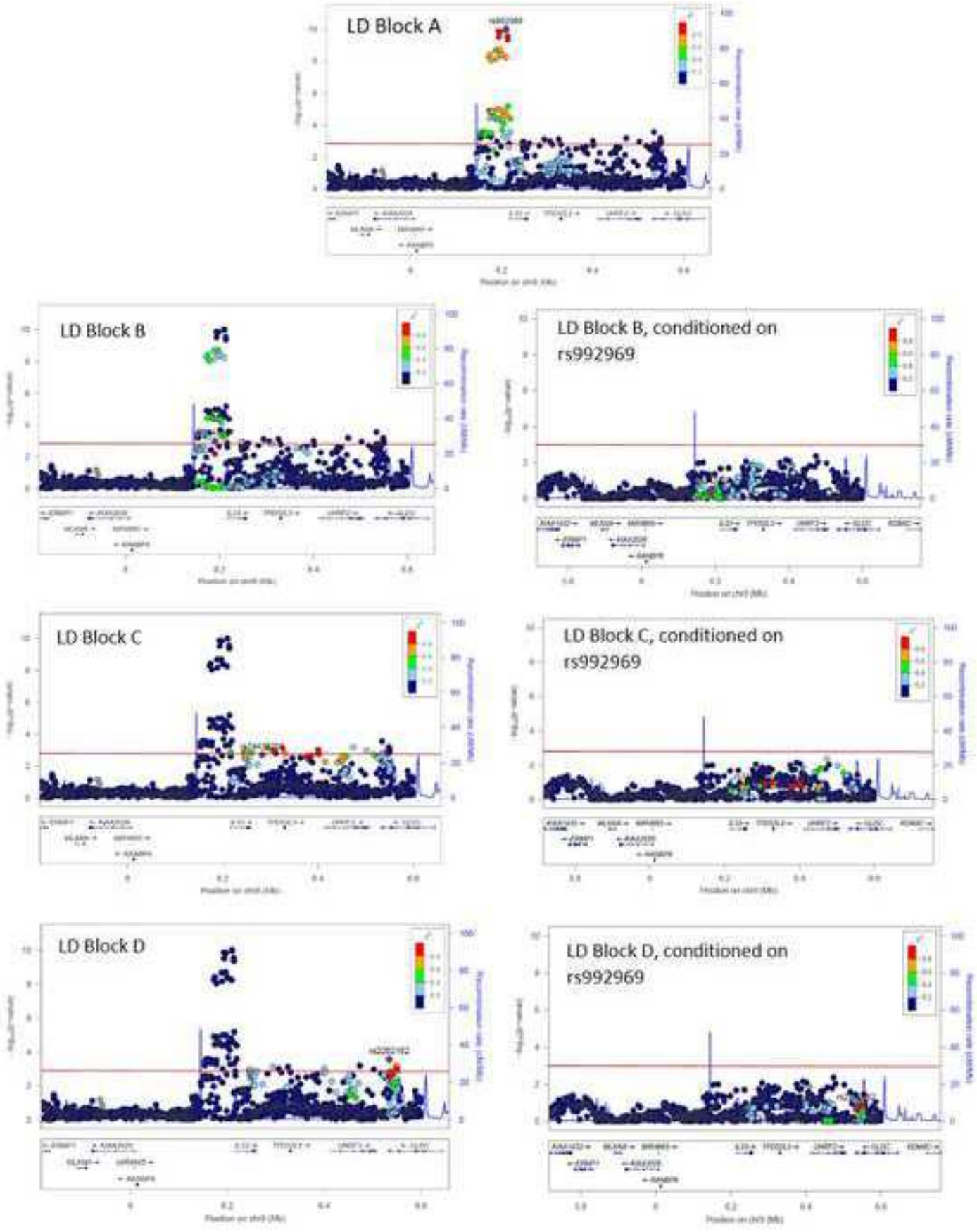


Figure No.4

Click here to access/download;Figure No.;Figure 4_SignalAandEPhenotypeAssociations.jpg

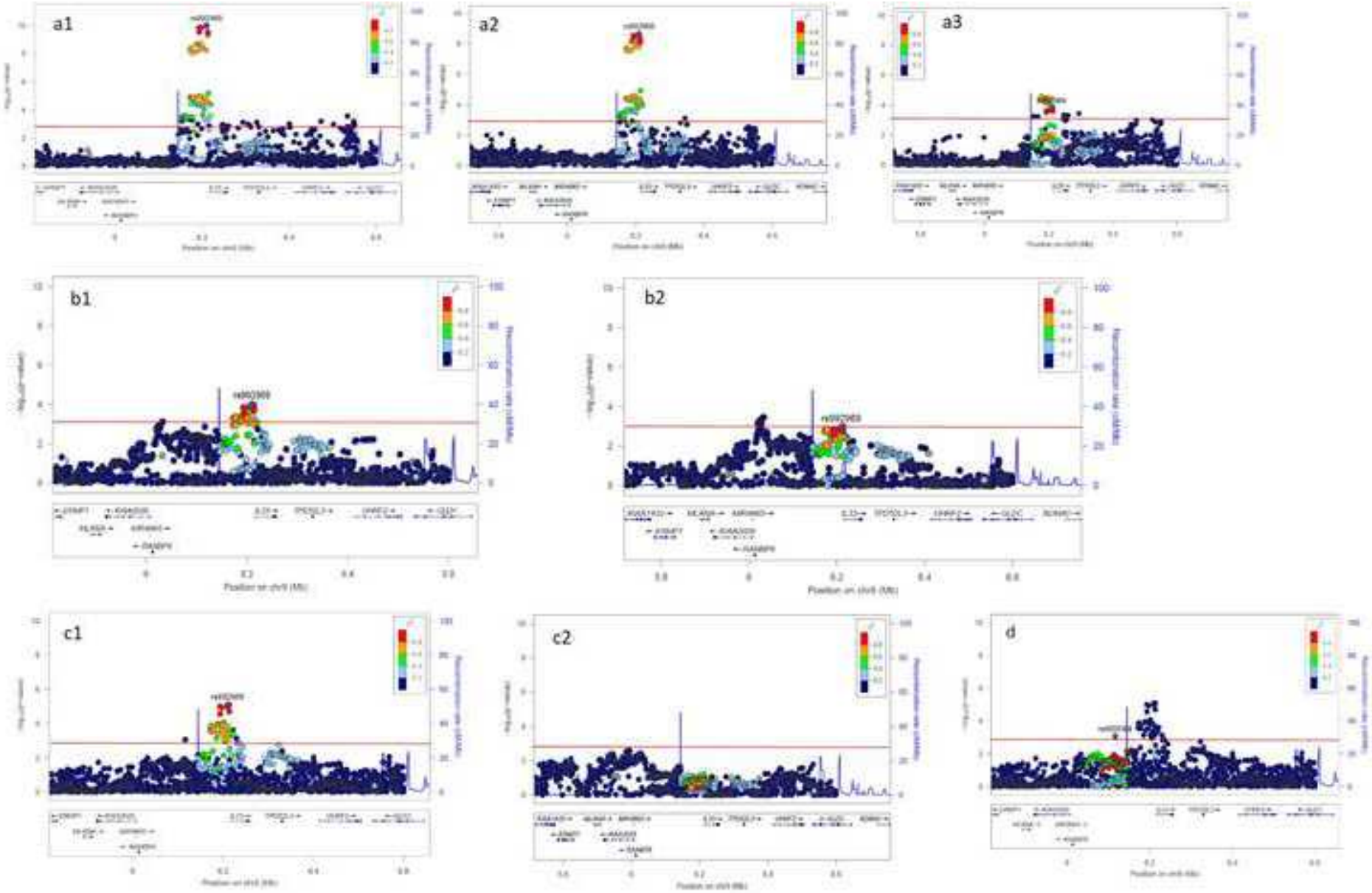


Figure No.5

Click here to access/download;Figure No.;Figure 5_eQTL.jpg

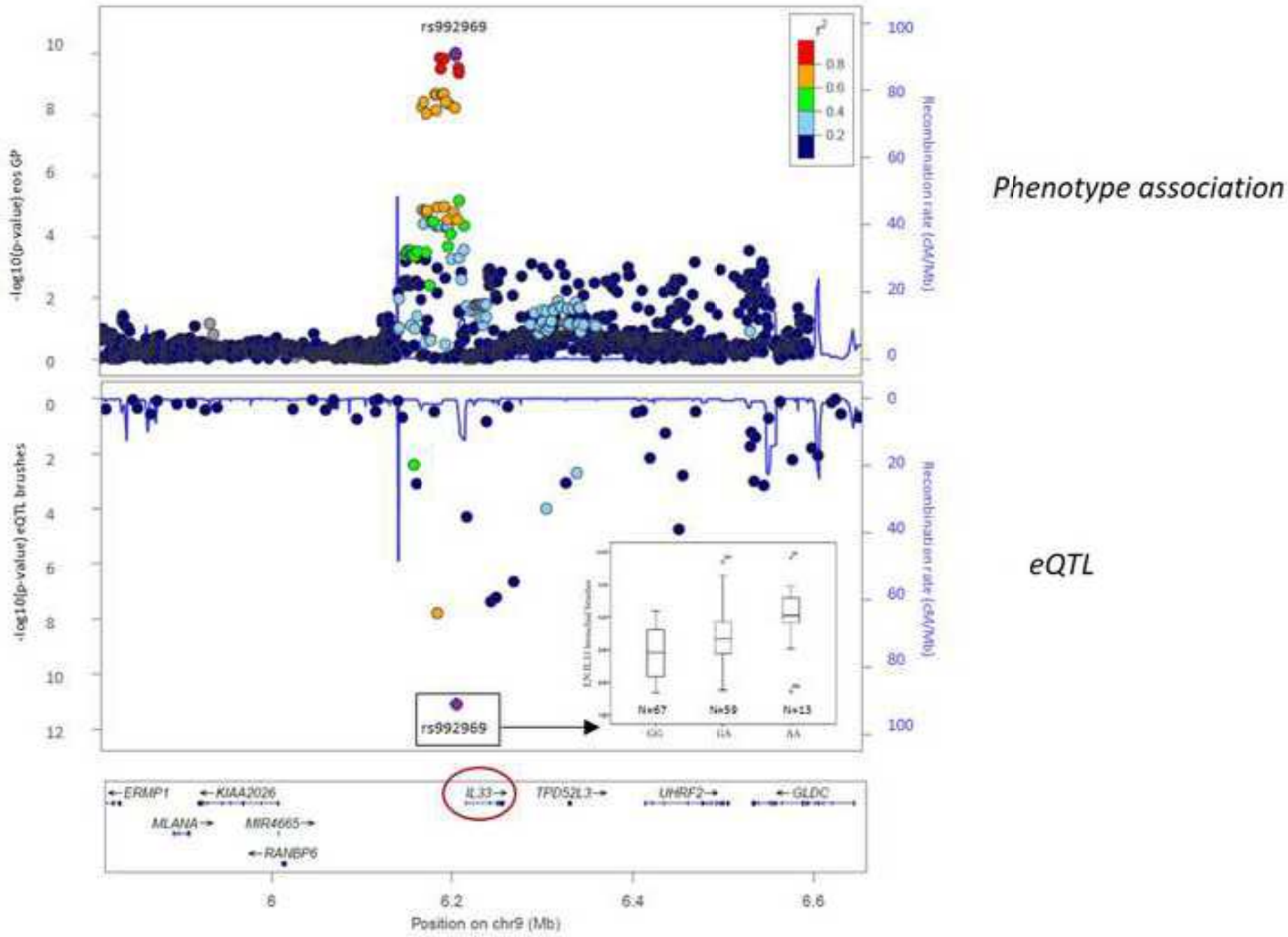
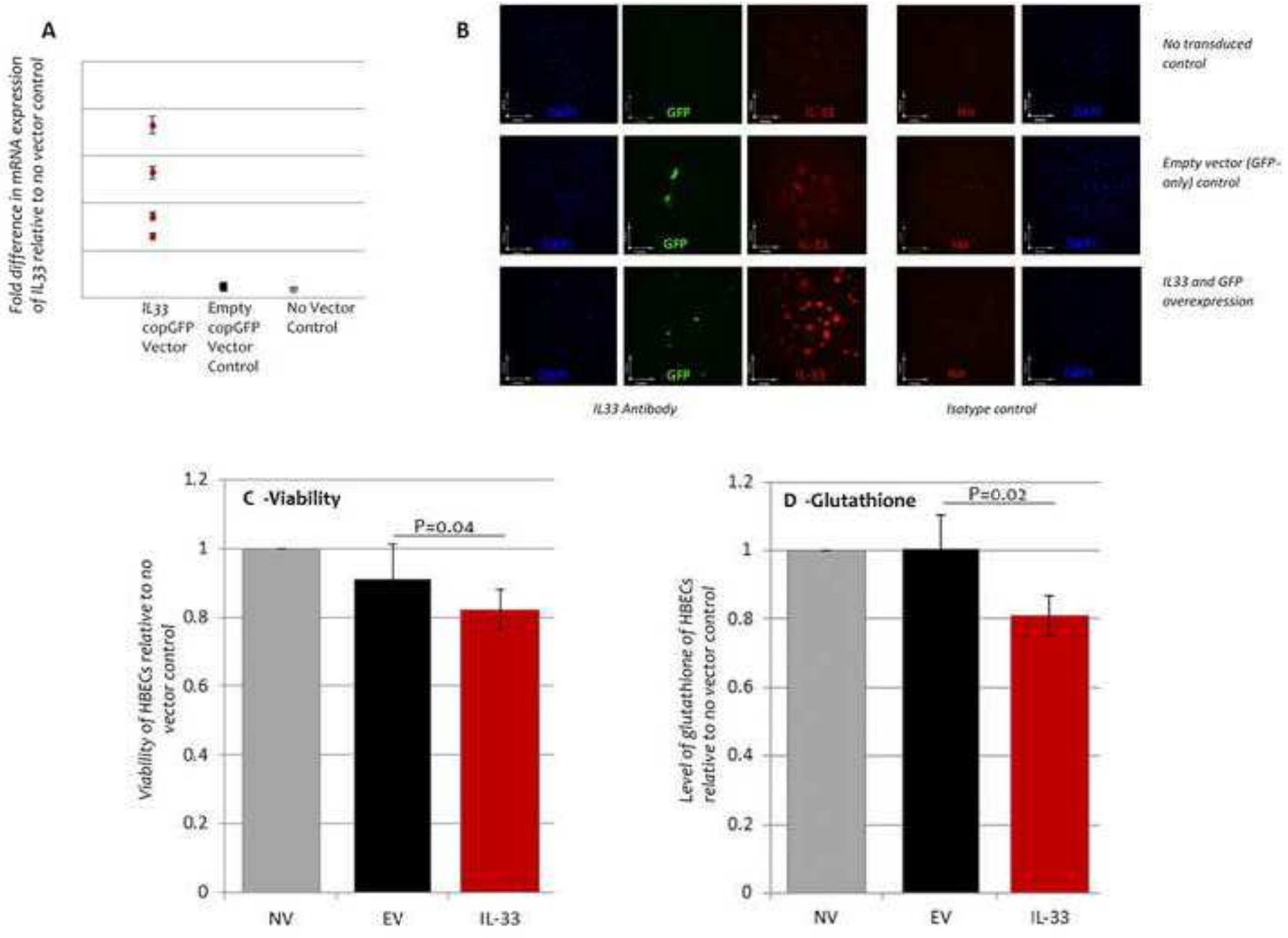


Figure No.6

Click here to access/download;Figure No.;Figure 6_FunctionalAssays.jpg



Supplemental material:

IL33 functional genetics-Ketelaar, Portelli, Dijk et al

Table of Contents

Supplemental Methods

M1- Cohort descriptions and details of genotype-phenotype analyses (see also supplemental table S1-S2):

Lifelines general population cohort (table S1)

Lifelines is a multi-disciplinary prospective population-based cohort study examining in a unique three-generation design the health and health-related behaviours of 167,729 persons living in the North of The Netherlands. It employs a broad range of investigative procedures in assessing the biomedical, socio-demographic, behavioural, physical and psychological factors which contribute to the health and disease of the general population, with a special focus on multi-morbidity and complex genetics. The cohort profile of the Lifelines study has been extensively described in Scholtens et al(1). Summarizing, the participants' baseline visit took place between December 2006 and December 2013. All general practitioners in the three northern provinces of the Netherlands were asked to invite their registered patients aged 25–49 years. All persons who consented to participate were asked to provide contact details to invite their family members (i.e., partner, parents and children), resulting in a three-generation study. Baseline data were collected from 167,729 participants, aged from 6 months to 93 years. Collected data include physical examinations, DNA, blood and urine samples, and comprehensive questionnaires on history of diseases, quality of life, lifestyle, individual socioeconomic status, work, psychosocial characteristics and medication use. Follow-up is planned for at least 30 years, with questionnaires administered every 1.5 years and a physical examination scheduled every 5 years. At current, a subset of the adult participants have both phenotypic and imputed genotype information available (n=13,395).

Participants of the Lifelines cohort were genotyped on the HumanCytoSNP-12 BeadChip (Illumina). Quality control before imputation was performed using ImputationTool2(2), excluding SNPs with a call-rate <95%, with a HWE-P value <0.001, MAF <0.01%. Samples were excluded in case of ambiguous sex (genetic mismatch with reported sex), of non-Caucasian origin (based on self-report, IBS and population stratification using EIGENSTRAT(3), and in case a pair of samples was discovered as first degree relatives using genetic cryptic relatedness, the sample with the best genotype quality was included only. Imputation was performed through Beagle 3.1.0 against the EUR panel from the 1000 genomes project (version March 2012)(4).

Klijs et al (2015)(5) concluded that the Lifelines adult study population is broadly representative for the adult population of the north of the Netherlands. The recruitment strategy had minor effect on the level of representativeness. These findings indicate that the risk of selection bias is low and that risk estimates in Lifelines can be generalized to the general population.

Genetics of Severe Asthma Phenotypes cohort (GASP) (table S2)

Asthmatic individuals from the GASP cohort (n=2,536) were used in the current study, this cohort has been recruited across UK hospitals as part of an Asthma UK initiative and is enriched for patients with British Thoracic Society Step 3 and above (~2,200 moderate-severe asthma, remainder mild asthma)(6). Asthma was defined as a doctor's diagnosis of asthma through the presence of symptoms and medical treatment, while age of onset of asthma was determined through patient records. Asthma related clinical phenotypes used in the current study focused to lung function (FEV_1 pre bronchodilator, FEV_1/FVC pre-bronchodilator), atopic status (positive skin prick test), Blood Eosinophil Count (x10^9/L) and Blood IgE levels (kU/L). Total peripheral blood eosinophil levels were calculated using a counting chamber while total Immunoglobulin E (IgE) levels were measured by ImmunoCAP™. Finally, atopy was defined as a positive response to a skin prick test (SPT) to any allergen from a panel of 4-24 allergens.

Participants in the GASP cohort were genotyped using two platforms, initially 744 subjects using the Affymetrix Axiom® UK BiLEVE array and 2172 subjects using the Affymetrix Axiom® UK Biobank array. In each genotyping batch samples were excluded: (i) if their genetically inferred gender did not match their reported gender; (ii) if they had outlying heterozygosity within the batch (outside either 2 or 3 standard deviations from the mean depending on batch); (iii) if they had a call rate <95% across genotyped variants; (iv) if cryptically related to another sample, 1 sample of the pair was removed; (v) if the sample shows significant deviation from European ancestry as determined by a plot of the first two principal components. The batches were merged and SNPs not available in both batches were excluded from the dataset. Following quality control 692,060 SNPs were available for 2,536 subjects. Imputation was then performed using IMPUTE 2.0 against the reference data set of the EUR

panel of the 1000 Genomes project (version March 2012)(4). Genetic studies were approved by the Medical Ethics Committee of each participating centre or via our multicentre ethics approval.

Dutch Asthma GWAS cohort (DAG) (table S2)

The DAG cohort has been extensively described previously(7). In summary, the DAG cohort consists of 469 trios ascertained through a proband with asthma, combined with an additional case-control study of 452 asthmatics and 511 controls. Of these, we selected 909 unrelated asthma patients who underwent the same, standardized, comprehensive evaluation for asthma at Beatrixoord Hospital, Haren, The Netherlands between 1962-2003. Asthma was defined as a doctor's diagnosis of asthma, asthma symptoms, and bronchial hyperresponsiveness (BHR). FEV_1 was measured using a water-sealed spirometer (Lode Spirograph type DL, Lode b.v., Groningen, The Netherlands). Total peripheral blood eosinophils were counted in a counting chamber and IgE levels were measured in serum by an enzyme-linked fluorescence assay (Mini Vidas, Biomerieux Inc., Marcy, France). In subjects older than 12 years intracutaneous tests with 16 common aeroallergens were performed. In children younger than 12 years, a skin prick test was performed with 10 allergens. Subjects with a positive response to one or more intracutaneous or skin prick tests (SPT) were considered to be atopic. Age of asthma onset was based on data from medical records and questionnaires, indicating the start of asthma symptoms. Participants in the DAG cohort were genotyped on two platforms, the Illumina 317 Chip and the Illumina 370 Duo Chip (Illumina, San Diego, CA). Quality control (QC) was performed per chip with exclusion of individuals with missing genotype call rate >0.01, related individuals (identity by descent (IBD) >0.125) and non-Caucasian subjects, as assessed by principal components analysis performed with EIGENSTRAT(3). SNPs were excluded with a missing genotype rate >0.01, a Hardy-Weinberg equilibrium P-value <10-7 and a MAF <0.01. Markers with Mendelian errors in phase I were excluded from analysis. Following quality control, the chips were merged and SNPs not available in both cohorts were excluded from the dataset. A total of 294,775 SNPs remained. Imputation was performed using IMPUTE 2.0 against the reference data set of the EUR panel of the 1000 Genomes project (version March 2012)(4). Genetic studies were approved by the Medical Ethics Committee of the University Medical Center Groningen and all participants provided written informed consent.

Next-Generation DNA Sequencing (NGS) cohort.

DNA from 200 severe asthma cases (BTS 4, 5) from GASP and 200 non-asthmatic, non-atopic, non-wheeze controls from the Nottingham Gedling cohort(8), were selected for resequencing. Subjects were matched for age and gender (Supplemental Table 3). Next-generation Illumina sequencing of the IL33 region (chr9:5924967-6267982) was outsourced to Source Bioscience (Nottingham, UK) and was carried out using the SureSelect enrichment approach. The chromosome 9 locus previously associated with asthma [GRCh37.p9] was the focus and 120 base pair paired-end long read oligonucleotides (baits) were designed using the SureSelect™ e-array design software. Bait tiling (X5) was used across the region, presenting with a capture size range of 500Kb to 1.5Mb. The initial target region was 343,016bp; using 7,751 baits achieved 65.28% coverage of this region. Samples were pooled for sequencing (3 pools for cases and 3 pools for controls). Next-generation sequencing was carried out on these six samples on two separate lanes, one for cases and the other for controls, using the Illumina HiSeq2000™ systems pipeline (San Diego, USA). Sequencing used a paired end design using 100bp reads. Resequencing the IL33 region identified 981 variants that were considered valid calls by SNver. Case-control association analyses revealed 12 SNPs significantly associated with severe asthma, of which two were within LD block E, the remaining 10 SNPs were rare single variants, of which 7 SNPs were novel (table S14). Due to the low frequency (MAF<0.1) these were not followed-up functionally.

Details of Genotype – Phenotype associations

For the genotype-phenotype association analyses SNPs were selected with a MAF≥0.01 located 400kb up- and downstream the *IL33* gene (Chr9: 5,815,786−6,657,983), encompassing all known asthma association signals (see also table S1). There were 1,970 SNPs present in Lifelines, and there were 3,025 and 2,780 SNPs available in the GASP and DAG cohorts, respectively, with 2,457 shared for meta-analysis of GASP/DAG. All Lifelines SNPs were present in the pool of SNPs of the DAG/GASP meta-analysis. Annotated SNP location and function was determined with the use of HaploReg v4.1(9). All genetic data are presented relative to assembly GRCh37/hg19. In Lifelines we performed genetic association within the genotyped subset of the general population cohort (n=13,395 with both genotype and phenotype information), with eosinophil counts, FEV1 (%pred) and FEV1/FVC.

This was followed by a sensitivity analysis for blood eosinophil counts, where we removed asthma patients (n=1,066; doctor's diagnosed asthma) and asthmatic+allergic subjects (n=6,227) and associated *IL33* SNPs with blood eosinophils within this non-asthmatic, non-allergic population. Herein, allergy was defined as having at least one self-reported allergy the questionnaire covering eczema, rhinitis, food allergy, dust allergy, animal allergy, pollen allergy, medication allergy, contact allergy, and insect bite allergy.

From this Lifelines general population cohort we subsequently took the asthma subpopulation (n=1,066, doctor's diagnosed asthma) and investigated genetic association with eosinophil counts, eosinophilic asthma (asthma and blood eosinophils >150 cells/µL, n=707), FEV1, FEV1/FVC and asthma with airway obstruction (asthma and FEV1<80% of predicted (n=258) or FEV1/FVC<70% (n=324)). Subsequently, we performed association analyses in a meta-analysis of GASP (n=2,536) and DAG (n=909) asthma patients studying atopy, blood eosinophils, total serum IgE, age of asthma onset, and lung function (FEV1, FEV1/FVC).

Associations of SNPs in the *IL33* region with FEV_1, FEV_1/FVC, blood eosinophils/neutrophils, total IgE levels, atopy and age of asthma onset were performed with PLINK v1.90b6.7(10) (Lifelines) or SNPtest v2.5β(11) (DAG/GASP) using an additive genetic model. Eosinophils, neutrophils and IgE levels were logarithmically transformed before analysis. Age of onset was analyzed as a continuous variable and a binary trait with cases being defined as having asthma onset <16 years of age (childhood onset asthma). FEV_1 and FEV_1/FVC analyses were corrected for age, gender and height. Eosinophils, neutrophils, IgE and atopy were corrected for age and gender and the age of asthma onset analysis were corrected for gender. DAG/GASP were meta-analysed, which was performed in METAL(12). Associations at an adjusted p-value<0.05 (FDR) were considered statistically significant.

Summary of Phenotye-Genotype association study results:

In Lifelines, we found an (FDR-) significant association of *IL33* SNPs with blood eosinophil counts in the general population, with eosinophilic asthma (vs healthy controls) and with asthma (vs healthy controls) as can be found in table 1 (main text) and supplemental tables S7-S9, S11-S13. In Lifelines, we did not find a significant association *IL33* SNPs with blood eosinophil counts within asthma patients, nor with eosinophilic asthma vs non-eosinophilic asthma as can be seen in supplemental tables S7-S9 (FDR>0.05). Neither we found an (FDR) significant association with blood neutrophil counts, FEV1 or FEV1/FVC in Lifelines general population or within the Lifelines asthma population (FDR>0.05).

In the asthma cohort DAG/GASP meta-analysis, *IL33* SNPs were significantly associated with blood eosinophil counts (table 1 main text), FEV1/FVC and age of asthma onset (supplemental table 15). Within DAG/GASP, no (FDR-)significant association of *IL33* SNPs with blood neutrophil counts, FEV1, total IgE levels, and atopy were found.

M2- Cohort descriptions and details QTL analyses

M2.1 Sample collection

Lung tissue and bronchial brushes

Lung tissue samples for mRNA expression analyses had been collected previously(13) from patients who had undergone lung resection for lung cancer (non-tumorous part, N=1,022), were lung transplantation donors (disapproved lung, N=89) or were lung transplantation recipients (operated lung, N=313). Samples had been collected at three centers, as part of a collaborative effort at the University Medical Center Groningen (UMCG, Groningen), the University of British Columbia (UBC, Vancouver) and the Laval University (Laval, Quebec). The detailed study design and collection procedure has been published before(13), and a summary of the patient characteristics from the included subjects can be found in supplemental table 4.

Bronchial epithelium from brushings (Cellebrity brush Boston Scientific, Massachusetts, USA) had been collected at the UMCG for mRNA expression analyses, deriving from N=139 healthy volunteers. Further patient characteristics are presented in supplemental table 4. From both cohorts, patients were excluded who lacked information on their pre-operative lung function, smoking status, comorbidities, drug use, age, and/or gender.

Bronchial epithelial cells cultured in vitro

Passage 2/3 human bronchial epithelial cells (n= 35) obtained from bronchial brushes and biopsies from asthma patients (referred to as AHBECs) as previously described(14) were cultured on PureCol Type-I Bovine collagen (Advanced BioMatrix, 5005-B) in fresh growth factor-supplemented medium (BEGM, Lonza) until 90% confluence. Protein and RNA lysates were collected as previously described(15) and IL33 levels compared in a genotype-stratified way.

M2.2 mRNA/protein expression assays

Lung tissue/bronchial brushes-mRNA

Expression levels of *IL33* mRNA in the lung tissue samples had been determined previously(13) as part of a general gene expression profile using a customized mRNA array (Affymetrix US Ltd., GEO platform GPL10379). The mRNA expression levels in the bronchial brushing samples had been measured on a commercially available array, the Human genome ST v1.0 array (Affymetrix US Ltd), as described before(16).

Bronchial epithelial cells-mRNA and protein

HBEC complimentary DNA (cDNA) was synthesised from 1µg RNA using Superscript II (Invitrogen, UK) and random hexamer primers according to the manufacturer's instructions. TaqMan® Quantitative PCR (qPCR) was then utilised to quantify mRNA levels of IL33 and was performed and analysed as previously described(15). IL33 protein in cell supernatants was measured using Luminex assays (supplied by R&D, product code LXSAHM) according to the manufacturer's recommendations using a custom Magnetic Luminex Screening Assay with a Human Premixed Multi-Analyte Kit (R&D systems). Each experimental supernatant was assayed in duplicate.

M2.3 Genotyping

Genotypes of SNPs in the *IL33* region had been determined in DNA from peripheral blood mononuclear cells (PBMCs) or oral swabs.

For the lung tissue cohort genotyping had been done on the Human 1M-Duo BeadChip array (Illumina Inc, San Diego, USA) which were imputed against the 1000G phase 1 reference panel (EUR)(4) using IMPUTE2(2) to increase the coverage of genotypic information.

Subjects from the bronchial brushing cohort had been genotyped on two platforms: the Human CytoSNP 12 and OmniExpress Exome genotyping arrays (both Illumina Inc, San Diego, USA)

For the AHBEC cohort DNA was extracted using the Qiagen QIAamp® DNA Mini and Blood Mini Kit according to the manufacturer's instructions. SNP Genotyping was then carried out using TaqMan® Pre-designed assays.

M2.4 Quality control genotype data

Genotype data was quality checked on array, subject and SNP level:

All genotyping arrays passed a call rate of >98%. Subjects that failed gender confirmation (PLINK)(10), and ethnic inference check (EIGENSTRAT)(3) were excluded. SNPs were excluded if they had a SNP call rate <90%, a minor

allele frequency (MAF) <1%, and deviated from Hardy-Weinberg equilibrium (HWE) P<1.0*E-6. After QC a total of N=1,111 subjects from the lung tissue cohort had reliable genotypes available (out of N=1424), and N=129 from the bronchial epithelium cohort (out of N=139 healthy subjects).

M2.5 Details QTL models

eQTL in lung tissue and bronchial brushes

We tested for expression quantitative trait loci (eQTL) in lung tissue (n=1,111) and bronchial brushes (n=139) using a linear regression model to investigate the association of SNPs and log-transformed *IL33* expression data. Specifically, we employed an additive genetic model with age, gender, smoking status and the PCs explaining >1% of expression variance as covariates using R statistics. Since only 2 independent genetic signals were tested by look-up, a p-value <0.05 was considered statistically significant.

Baseline and inducible eQTL and pQTL in primary bronchial epithelial cells

Passage 2/3 AHBECs (n=35) obtained from bronchial brushes and biopsies from asthma patients as previously described(14), were cultured on PureCol Type-I Bovine collagen (Advanced BioMatrix, 5005-B) in growth factor-supplemented medium (BEGM, Lonza). These were studied for baseline QTL and inducible QTL.

For the inducible QTL, cells were stimulated with either house dust mite (HDM) or rhinovirus (RV16, MOI=1). BEGM was changed to basal medium 24 hours prior to stimulation with HDM (50µg/ml) (Greer XPB70D3A25 (Lot: 23187)), or PBS as a vehicle control. For stimulation with RV16 (Public Health England), BEGM was changed to infection medium (BEGM-I), i.e. BEGM lacking Bovine Pituitary Extract, 24 hours prior to infection with RV-16 virus MOI of 1. Cells were infected for a period of 1 hour, following which they were washed three times with sterile PBS and fresh BEGM-I was added. Cells were then incubated for 24 hours. Protein and RNA lysates were collected as previously described(15). Cells were stratified based on the genotypes of the two genetic signals and expression compared using the non-parametric Kruskall-Wallis test. A p-value<0.05 was considered statistically significant.

M3- Functional cell work

Lentiviral overexpression in human bronchial epithelial cells

In order to investigate the functional consequences of sustained *IL33* in asthma, we stably overexpressed human full-length *IL33* (aa1-270) in primary human bronchial epithelial cells (HBECs) isolated from n=5 healthy individuals (Lonza, #CC-2540). This was done using a three plasmid lentiviral system (as described before(17)). This consisted of a pCMV_VSV-G envelope plasmid (CellBiolabs RV110, Addgene plasmid # 8454)(18), a packaging plasmid pCMV_8.91 (Addgene plasmid #2221)(19) and the actual lentiviral overexpression plasmid (pCDH-CMV-MCS-EF1-copGFP, System Biosciences, #CD511B-1) wherein the full length human *IL33* sequence was ligated at the multiple cloning site under a CMV-promotor. A copGFP reporter gene was used under the EF1-promotor in the same overexpression cassette enabling a check of successful transfection/transduction. The *IL33* sequence was commercially derived from OriGene (#SC100114) and comprised the common *IL33* sequence based on the human CEU/Hg37 reference genome (transcript variant 1, NM_033439). The 3 plasmids were first purified and transfected in the vector cell line HEK-293 in a ratio of 8:7:1 (overexpression: packaging:envelope plasmid, ratio based on weight) to create lentiviral particles. The following experimental groups of lentiviral particles were created: lentivirus with *IL33+copGFP* sequence ('IL33'), lentivirus with *copGFP* sequence only ('EV') and a lentivirus without any modifications ('NV'). Then the HBECs were transduced with the lentiviral particles in a ratio of lentivirus:transduction agent that equalized *copGFP* expression between the *IL33-copGFP* and *copGFP*-only constructs (empirically determined based on level of copGFP mRNA expression, 1:5 for the *IL33-copGFP* and 1:2 for the *copGFP* only construct, see figure E12. The transduction agent was Polybrene (Sigma H9268-10G, used at 2ug/mL). *copGFP* expression and *IL33* overexpression was verified on mRNA level (qPCR) and IL33 expression on protein was determined using immunofluorescence; see figures E13 and E14.

IL33/copGFP PCR

Taqman qPCR was used to quantify copGFP and IL33 expression in the overexpression work in human bronchial epithelial cells. Total RNA was isolated from the cells and cDNA was synthesised from 1µg RNA using Superscript II (Invitrogen, UK) and random hexamer primers according to the manufacturer's instructions as also described

before(15). Based on the sequence of the copGFP from the overexpression plasmid (SBI #CD511B-1) we developed a probe/primerset that had the same qPCR efficiency as the IL33 qPCR, with the following sequences to detect copGFP: probe (5'-3') 6FAM-CGGCTACGAGAACCCCTTCC-TAMRA; forward primer (5'-3') ATGGGCTACGGCTTCTAC; reverse primer (5'-3') CTCGTACTTCTCGATGCG.
IL33 (Hs04931857_m1) was assayed using a commercially derived PDAR (#4331182, Applied Bioscience) with ROX™ as passive reference dye. HPRT (PDAR #4310890E, Applied Bioscience) and 18S (PDAR # 4310893E Applied Bioscience) were used as housekeeping genes. PCR was run with a 2 minutes step of 50^{o}C, 10 minutes of 95^{o}C and then 40 cycles of 95^{o}C (15s)+60^{o}C (1min). MxPro software was used to analyse data.

IL33 immunofluorescence

Passage 2 human bronchial epithelial cells that were transduced with lentivirus containing the IL33 expression cassette and controls were seeded in four-well chamberslides (polystyrene, tissue-culture treated, non-coated) at 30,000 cells/well and cultured until confluence. Then medium was removed, cells washed twice with PBS and fixed in 4% formaldehyde for 30min at ambient temperature on a rocker. All preparation steps were performed light-protected to safe the endogenous GFP fluorescence. After a PBS wash (twice, 5min), cells were permeabilized for 30min in 0.15%TritonX(Sigma X100) in 1%BSA(Sigma #A-8412)/PBS, washed and blocked overnight at 4^{o}C using 10%NGS (Sigma #G6767) in PBS. Cells were washed twice with PBS and incubated with two different primary antibodies against IL33 overnight at 4C. 1) A polyclonal rabbit IgG anti-IL-33, ProteinTech, O22 cat12372-1-AP, stock 260ug/mL, used 1:100 in 10%NGS/PBS. 2) A monoclonal mouse IgM anti-IL-33 (clone Ag21430) ProteinTech, cat66235-1-Ig, stock 1360ug/mL, used 1:200 in 10%NGS/PBS. Applicable isotype controls (polyclonal rabbit IgG and monoclonal mouse IgM, Invitrogen #10500C and #14-4752-82) were used at the same concentration as the primary antibodies. Wells were washed three times in PBS (5min each) on a rocker and then Rhodamine TRITC-labeled secondary antibodies applied for 1h at 37^{o}C while shaking, in a humidified tray. Being: 1) goat-anti rabbit IgG (Jackson ImmunoResearch laboratories/Stratech 111-025-003, stock 1.5mg/mL) and 2) goat anti-mouse IgG (ProteinTech, #SA00007-1) both used 1:100 in 10%NGS/PBS. Wells were washed three times in PBS (5min each), air-dried and mounted using VectaShield Mount+DAPI (Vector Laboratories, H-1500), a coverslip applied, dried at ambient temperature and stored light-protected at 4^{o}C until visualization of IL33 and GFP using confocal microscopy (within 48h after mounting).

Functional read-outs of genetically modified HBECs:

We cultured the genetically modified HBECs submerged in bronchial epithelial cell growth medium (BEGM™, Lonza) and performed several functional read-outs:

Cell count, viability and ROS-glutathione assays

Genetically modified and control cells were seeded at 50,000 cells/well in a 6-wells plate format and cultured for 96h. Cells were then harvested by trypsinization, n=2 technical replicates pooled and resuspended in 500uL BEGM. Cell count was performed using a lysis+propidium iodide (PI) based assay (PI-Cassette™) according to the manufacturers' instruction (Application note No. 3007. Rev. 1.3, Chemometec) at the NucleoCounter® NC-3000™ system. Viability (fraction of viable cells) was determined using a Hoechst+PI based assay according to the manufacturers' instructions (application note No. 3023. Rev. 1.4, Chemometec). ROS-capturing capacity (level of free thiols [reduced glutathione] in cells) was determined using a VitaBright-48™ assay according to the manufacturers' instructions (Application note No. 3005. Rev 1.4) (20).

Metabolic activity assay

Genetically modified and control cells were seeded at 4000 cells/well in 96-wells plates and cultured for 96h in total, a subset harvested every 24h to determine metabolic activity over time. Metabolic activity was measured using a colorimetric MTT assay (Sigma, M5655), according to the manufacturer's instruction. Briefly; medium was removed and replaced by a warm (37^{o}C) MTT suspension (0.5mg/mL in BEGM) which was incubated protected from light for 4h at 37^{o}C. The MTT suspension was carefully removed and the formed crystals dissolved in 200uL isopropanol per well. Optical density was measured at 570nm wavelength, including a correction at 670nm.

Electric cell substrate impedance sensing (ECIS) array

To investigate cell-cell contact and barrier formation of bronchial epithelial cells exposed to sustained IL-33, we cultured the genetically modified and control cells on electrode-containing arrays. Electric cell substrate impedance sensing (ECIS) was performed as previously described in detail by our group(21). As high frequency (reflecting spreading of cells and formation of a monolayer) 32kHz was taken, whilst for the low frequency (reflecting cell-cell-contacts and barrier formation) 400Hz was taken as previously established for primary bronchial epithelial cells(21). Resistance values were normalized against the averaged first hour to correct for technical variation/fluctuations often present at the start. Longitudinally, the area under the curve (AUC) was calculated for timepoints 12-24-36-48-60-72h, as well as a cross-sectional comparison of the resistance was made at 24h and 48h to investigate the influence of IL33 on aspects of barrier formation of the bronchial epithelial cells.

Statistical analyses in vitro *cell work*

Treatment and genotype groups were compared using the non-parametric Kruskall Wallis-test. For the expression, cell count, vitality, viability, metabolic assay results and for the cross sectional analysis of the ECIS, followed by Wilcoxon post hoc statistics. The longitudinal AUCs of the ECIS were compared using a Z-test. A p-value <0.05 was considered statistically significant.

M4- ENCODE annotation of phenotype-associated genetic signals

ENCODE was consulted to annotate potential functional elements to the selected phenotype associated genetic signals, using the integrative data level available via https://www.encodeproject.org/data/annotations/. SNPs with LD >0.3 with the selected tagSNP were included in these analyses. Within the ENCODE setting, GWAS studying SNPs associated with asthma and lung phenotypes and their potential functionality were consulted, as well as SNPs were checked for functionality using RegulomeDB, HaploReg and the Chromatin databases (ChromHMM, Segway).Dataset was last accessed on the 9th August 2019. (9,22,23)

A) Supplemental Tables

Table S1 Lead genetic variants of genomewide association ($P<5x10^{-8}$) with asthma in GWAS/GWAS meta-analyses from 2007-2019

SNP	*Effect allele*	*OR*	*P-value*	*Population*	*Ref*	*First author(s)*	*Journal*	*Year*	*Position*	*Signal*
rs1888909	T	1.12	4.20E-34	Caucasian	(24)	Kristjansson RP	Nat Genet	2019	chr9:6197392	
rs7848215	T	1.16	5.29E-62	Caucasian	(25)	Johansson A	Hum Mol Genet	2019	chr9:6213468	
rs992969	A	1.25	1.4E-11	Caucasian	(26)	Pividori M	Lancet Respir Med	2019	chr9:6209697	
	A	1.18	1.1E-17	Multi-ancestry analysis	(27)	Demenais F	Nat Genet	2018		
rs144829310	T	1.18	8.3E-58	Caucasian	(25)	Johansson A	Hum Mol Genet	2019	chr9:6208030	
	T	1.21	2.3E-20	Caucasian	(6)	Shrine N, Portelli MA, John C	Lancet Respir Med	2019		
	T	1.09	1.2E-35	Caucasian	(28)	Ferreira MA	Nat Genet	2017		**A**
	T	1.17	1.3E-31	Caucasian	(29)	Pickrell JK	Nat Genet	2016		
rs72699186	T	1.26	2.0E-09	Caucasian	(30)	Ferreira MA	J Allergy Clin Immunol	2013	chr9:6175855	
rs928413	G	1.50	4.2E-13	Caucasian	(31)	Bonnelykke K	Nat Genet	2013	chr9:6213387	
rs1342326	C	1.20	3.5E-14	Caucasian	(32)	Ferreira MA	Lancet	2011	chr9:6190076	
	C	1.20	9.2E-10	Caucasian	(33)	Moffatt MF	N Engl J Med	2010		
rs2381416	C	1.18	1.7E-12	Multi-ancestry analysis	(34)	Torgerson DG	Nat Genet	2011	chr9:6193455	
rs2066362	T	1.21	1.39E-08	Caucasian	(33)	Moffatt MF	N Engl J Med	2010	chr9:6219176	
rs343478	G	1.06	4.5E-13	Caucasian	(25)	Johansson A	Hum Mol Genet	2019	chr9:6051399	**E**
	G	1.03	2.6E-10	Caucasian	(28)	Ferreira MA	Nat Genet	2017		

In this table an overview is given of the genetic variants associated with asthma discovered at genomewide significant in GWAS and GWAS meta-analyses form 2007-2019. The last column indicates how these variants related to the genetic signals defined in our manuscript. *OR=Odds ratio, Ref=literature reference.*

382 Table S2-Population characteristics of the Lifelines cohort(1)

Characteristics	*General Population (N=13,395)*	*Asthma Population (N=1,066)*	*Stats* (compared to rest of GP)	*Healthy Control (N=6,863)*	*Stats* (compared to rest of GP)
Age (y). mean (SD)	48.1 (11.4)	46.2 (10.9)	P<0.001 (MWU)	49.2 (11.9)	P<0.001 (MWU)
Gender (N. %male)	5,598 (41.8%)	428 (40.2%)	P=0.233 (Chisq)	3213 (46.8%)	P<0.001
Height (cm). mean (SD)	174.4 (9.2)	173.9 (9.6)	P=0.018 (MWU)	175.0 (9.3)	P=0.012 (MWU)
BMI (kg/m^2) mean (SD)	26.4 (4.3)	27.2 (5.0)	P<0.001 (MWU)	26.3 (4.1)	P=0.875
Ethnicity (N. %Caucasian)	11,615 (99.4%)	895 (98.8%)	P=0.018 (Chisq)	5,906 (99.6%)	P=0.351
FEV1 (L). mean (SD)	3.4 (0.8)	3.1 (0.8)	P<0.001 (MWU)	3.4 (0.8)	P<0.001 (MWU)
FEV1/FVC. mean (SD)	0.76 (0.07)	0.73 (0.09)	P<0.001 (MWU)	0.77 (0.07)	P=0.001 (MWU)
Blood eos (10^{E9}/L). median (IQR)	0.16 (0.10-0.23)	0.20 (0.13-0.30)	P<0.001 (MWU)	0.15 (0.10-0.22)	P<0.001 (MWU)
Blood neutro (10^{E9}/L). median (IQR)	3.18 (2.55-3.97)	3.33 (2.66-4.18)	P<0.001 (MWU)	3.13 (2.51-3.92)	P<0.001 (MWU)
Asthma- N (%)	1,066 (8.7%)	1,066 (100%)		-	
Eosinophilic Asthma- N (%)	707 (5.5%)	707 (68.6%)		-	
Low FEV1%pred Asthma- N (%)	258 (1.9%)	258 (24.2%)		-	
Low FEV1/FVC Asthma- N (%)	324 (2.5%)	324 (30.4%)		-	
No asthma/COPD/Allergy- N (%)	6,863 (51.2%)	-		6,863 (100%)	

383 *Chisq= chi squared test, GP=general population, IQR+ inter quantile range, MWU= Mann Whitney-U test, N = number of*
384 *subjects data field available for, %pred= percentage predicted, SD: Standard of Deviation, Stats= statistical comparison*
385 *subgroup to rest of general population*
386
387
388
389

390 Table S3-Population characteristics independent asthma cohorts(6,7)

Characteristics	*DAG (N=909)*	*N*	*GASP (N=2,536)*	*N*
Age (y). mean (SD)	34.78 (15.80)	909	47.83 (15.51)	2,285
Gender. Male (%)	46.9	909	36.0	2,534
Height (m). mean (SD)	1.68 (0.16)	905	1.65 (0.09)	1,692
FEV_1 (L). mean (SD)	2.81 (0.94)	899	2.26 (0.86)	2,039
FEV_1/FVC. mean (SD)	0.79 (0.10)	262	0.69 (0.14)	1,881
Blood Eosinophils (10-9/L). median (range)	0.23 (0.00-1.90)	769	0.31 (0.00-5.42)	1,018
Total IgE (kU/L). median (range)	378.41 (0.00 – 12400.00)	772	407.47 (1.00 – 5000.00)	1,374
Atopy (%)*	578 (85.4)	677	1,072 (68.5)	1,559
Age of asthma onset (y). mean (SD)	10.07 (10.58)	689	23.20 (17.95)	1,176
Childhood onset asthma# N(%)	520 (75.5)	689	578 (46.2)	1,284
%/N of asthma patients	100%	909	100%	2,536

391 *N = number of subjects data field available for, SD: Standard of Deviation*
392 **Atopy was based on at least one positive response to intracutaneous or skin prick tests (SPT) #Childhood onset asthma defined*
393 *as a diagnosis occurring before the age of 16y.*
394
395
396

Table S4- Population characteristics Next Generation Sequencing cohort(6,8)

Characteristics	*GASP Cohort (cases)*	*Gedling Cohort (controls)*
Age (y), mean (SD)	48 (14.88)	57 (12.64)
Gender, Male (%)	30.6	27.0
Height (m), mean (SD)	1.64 (0.08)	1.66 (0.06)
FEV_1 (L), mean (SD	2.17 (0.84)	2.77 (0.79)
Smoking pack/years	11.82 (20.25)	8.40 (18.61)
Never Smokers (%)	52.0	53.5

Demographics for the sub-cohorts taken from GASP (200 cases)and GEDLING (200 non-asthmatic, non-atopic controls) used in the next-generation sequencing of the chromosome 9 locus. *SD: Standard of Deviation*

Table S5-Population characteristics eQTL cohort lung tissue(13) and bronchial brushes(16)

Characteristics	*Lung tissue (n=1,111)*	*Bronchial brushes (n=139)*
Age (y). mean (SD)	58.5 (13.0)	40.0 (18.0)
Gender (N. %male)	54.4%	34.7%
FEV1 (L), mean (SD	2.70 (0.99)	2.76 (0.87)
FEV1/FVC. mean (SD)	0.71 (0.11)	-
Smoking status %current smoker	24.1%	66.7%
BMI. mean (SD)	-	24.1 (3.4)

Demographics for the lung tissue and bronchial brush cohorts used for the expression quantitative (eQTL) analyses. *SD: Standard of Deviation*

Table S6-Population characteristics of cultured bronchial epithelial cells- asthma cohort (AHBEC)

Characteristics	*AHBEC (N total=35)*	*N*
Age (y), mean (SD)	50 (13.47)	20
Gender, Male (%)	43.5	23
Height (m), mean (SD)	1.71 (0.10)	13
FEV_1 (L), mean (SD)	2.70 (0.95)	25
FEV_1/FVC , mean (SD)	0.69 (0.11)	19
Atopy, number (%)*	7 (58.3)	12

Demographics for the cultured primary bronchial epithelial cells from asthma patients (AHBEC) used for the expression quantitative (eQTL) analyses. *N = number of subjects data field available for, SD: Standard of Deviation*

**Atopy was defined as a positive response to a skin prick test. Data was not available for the full cohort of 51 individuals.*

423 Table S7-Five LD blocks (r^2>0.1) and association results with eosinophilic phenotypes in all cohorts

Tag SNP (genetic signal)	*Location*	*Gene context*	*Pheno risk allele (AF)*	*Alt allele*	*Associated feature(s)*	*Cohort*	*Effect size risk allele (B or OR)*	*SE*	*P.adj (FDR)*
A-rs992969	9:6209697	~6kb 5' of *IL33*	A (0.25)	G	*eos levels in GenPop,*	Lifelines	0.058 (B)	0.009	7.09E-08
					eos asthma vs HC	Lifelines	1.321 (OR)	0.062	4.73E-03
					eos asthma vs non-eos asthma	Lifelines	1.216 (OR)	0.109	0.556
					eos asthma vs non-eos asthma	DAG/GASP	1.078 (OR)	0.161	0.633
					eos levels in asthma subjects	Lifelines	0.042 (B)	0.032	0.714
					eos levels in asthma subjects	DAG/GASP	0.002 (B)	0.014	0.991
B-rs1342327	9:6189874	~25kb 5'of *IL33*	G (0.15)	C	*eos levels in GenPop,*	Lifelines,	0.035 (B)	0.011	0.027
					eos asthma vs HC	Lifelines	1.107 (OR)	0.075	0.587
					eos asthma vs non-eos asthma	Lifelines	1.081 (OR)	0.125	0.845
					eos asthma vs non-eos asthma	DAG/GASP	1.181 (OR)	0.421	0.942
					eos levels in asthma subjects	Lifelines	0.0118 (B)	0.037	0.895
					eos levels in asthma subjects	DAG/GASP	0.057 (B)	0.018	0.039
C-rs74438701	9:6282794	~25kb 3'of *IL33*	T (0.83)	C	*eos levels in GenPop*	Lifelines	0.035 (B)	0.011	0.041
					eos asthma vs HC	Lifelines	1.195 (OR)	0.085	0.219
					eos asthma vs non-eos asthma	Lifelines	1.293 (OR)	0.136	0.556
					eos asthma vs non-eos asthma	DAG/GASP	1.144 (OR)	0.309	0.763
					eos levels in asthma subjects	Lifelines	0.074 (B)	0.041	0.714
					eos levels in asthma subjects	DAG/GASP	0.012 (B)	0.018	0.991
D-rs2282162	9:6534466	intronic of *GLDC*	G (0.56)	A	*eos levels in GenPop*	Lifelines	0.029 (B)	0.008	0.011
					eos asthma vs HC	Lifelines	1.081 (B)	0.058	0.583
					eos asthma vs non-eos asthma	Lifelines	1.140 (OR)	0.100	0.586
					eos asthma vs non-eos asthma	DAG/GASP	1.032 (OR)	0.072	0.846
					eos levels in asthma subjects	Lifelines	0.073 (B)	0.030	0.714
					eos levels in asthma subjects	DAG/GASP	0.004 (B)	0.014	0.991
E-rs4008366	9:6116407	intergenic	T (0.69)	C	*eos levels in GenPop*	Lifelines	0.010 (B)	0.009	0.647

	eos asthma vs HC	Lifelines	1.264 (OR)	0.070	0.045
	eos asthma vs non-eos asthma	Lifelines	1.130 (OR)	0.116	0.691
	eos asthma vs non-eos asthma	DAG/GASP	1.007 (OR)	0.676	0.991
	eos levels in asthma subjects	Lifelines	0.003(B)	0.035	0.968
	eos levels in asthma subjects	DAG/GASP	0.0002 (B)	0.015	0.999

The table shows the results of the association analyses of all eosinophilic phenotypes in each cohort for the 5 LD blocks/signals (r^2>0.1). Eosinophilic asthma was defined as asthma with blood eosinophil count >150cells/uL. **Underlined**: the two genetic signals taken forward in functional assesment in this study. *AF=frequency (EUR 1000G); Alt allele=alternative allele; B=beta; eos=eosinophils/eosinophilic; FDR=false discovery rate value at alpha 0.05; GWAS= genomewide association study; GenPop =general population; HC=healthy control; kb= kilo basepairs; Lit.=literature; OR=odds ratio; P.adj= FDR adjusted p-value; Pheno risk allele=phenotype associated allele; Ref=Reference; SE=standard error.*

Table S8-Lifelines association results with eosinophilic asthma (>150cells/uL and >300cells/uL)

Tag SNP (genetic signal)	*Location*	*Gene context*	*Pheno risk allele (AF)*	*Alt allele*	*Associated feature(s)*	*Cohort*	*Effect size risk allele (OR)*	*SE*	*P.adj (FDR)*
A-rs992969	9:6209697	~6kb 5' of *IL33*	A (0.25)	G	*eos asthma* vs HC*	Lifelines	1.321	0.062	4.73E-03
					*eos asthma** vs HC*	Lifelines	1.330	0.097	8.58E-02
B-rs1342327	9:6189874	~25kb 5'of *IL33*	G (0.15)	C	*eos asthma* vs HC*	Lifelines	1.107	0.075	0.500
					*eos asthma** vs HC*	Lifelines	1.112	0.120	0.999
C-rs74438701	9:6282794	~25kb 3'of *IL33*	T (0.83)	C	*eos asthma* vs HC*	Lifelines	1.195	0.085	0.183
					*eos asthma** vs HC*	Lifelines	1.469	0.144	0.198
D-rs2282162	9:6534466	intronic of *GLDC*	G (0.56)	A	*eos asthma* vs HC*	Lifelines	1.081	0.058	0.495
					*eos asthma** vs HC*	Lifelines	1.304	0.193	0.127
E-rs4008366	9:6116407	intergenic	T (0.69)	C	*eos asthma* vs HC*	Lifelines	1.264	0.070	0.045
					*eos asthma** vs HC*	Lifelines	1.273	0.110	0.076

The table shows the results of the association between the 5 main LD blocks and eosinophilic asthma using two definitions (>150cells/uL(*n=707) and >300cells/uL(**n=260)) in Lifelines. **Underlined**: the two genetic signals taken forward in functional assesment in this study. *AF=frequency (EUR 1000G); Alt allele=alternative allele; B=beta; eos=eosinophils/eosinophilic; FDR=false discovery rate value at alpha 0.05; GWAS= genomewide association study; GenPop =general population; HC=healthy control; kb= kilo basepairs; Lit.=literature; OR=odds ratio; P.adj= FDR adjusted p-value; Pheno risk allele=phenotype associated allele; Ref=Reference; SE=standard error.*

443 Table S9-DAG/GASP association results with eosinophilic asthma (>150cells/uL and >300cells/uL)

Tag SNP (genetic signal)	*Location*	*Gene context*	*Pheno risk allele (AF)*	*Alt allele*	*Associated feature(s)*	*Cohort*	*Effect size risk allele (OR)*	*SE*	*P.adj (FDR)*
A-rs992969	9:62096977	~6kb 5' of *IL33*	A (0.25)	G	*eos asthma* vs non-eos asthma*	DAG/GASP	1.078	0.161	0.633
					*eos asthma** vs non-eos asthma*	DAG/GASP	1.049	0.096	0.680
B-rs1342327	9:61898744	~25kb 5'of *IL33*	G (0.15)	C	*eos asthma* vs non-eos asthma*	DAG/GASP	1.181	0.421	0.942
					*eos asthma** vs non-eos asthma*	DAG/GASP	1.329	0.585	0.893
C-rs74438701	9:62827944	~25kb 3'of *IL33*	T (0.83)	C	*eos asthma* vs non-eos asthma*	DAG/GASP	1.144	0.309	0.763
					*eos asthma** vs non-eos asthma*	DAG/GASP	1.117	0.293	0.789
D-rs2282162	9:65344666	intronic of *GLDC*	G (0.56)	A	*eos asthma* vs non-eos asthma*	DAG/GASP	1.032	0.072	0.846
					*eos asthma** vs non-eos asthma*	DAG/GASP	1.033	0.127	0.931
E-rs4008366	9:61164007	intergenic	T (0.69)	C	*eos asthma* vs non-eos asthma*	DAG/GASP	1.007	0.676	0.991
					*eos asthma** vs non-eos asthma*	DAG/GASP	1.059	0.583	0.802

444

445 The table shows the results of the association between the 5 main LD blocks and eosinophilic asthma using two
446 definitions (>150cells/uL(*n=1,002) and >300cells/uL(**n=493)) in DAG/GASP. **Underlined**: the two genetic
447 signals taken forward in functional assesment in this study. *AF=frequency (EUR 1000G); Alt allele=alternative*
448 *allele; B=beta; eos=eosinophils/eosinophilic; FDR=false discovery rate value at alpha 0.05; GWAS= genomewide*
449 *association study; GenPop =general population; HC=healthy control; kb= kilo basepairs; Lit.=literature;*
450 *OR=odds ratio; P.adj= FDR adjusted p-value; Pheno risk allele=phenotype associated allele; Ref=Reference;*
451 *SE=standard error.*

452
453

Table S10-The two genetic signals with their tagSNPs and proxySNPs used in functional follow-up

Signals	*Pheno risk allele (AF)*	*Alt allele*	*Proxy eQTL brushes*	*Pheno risk allele (AF)*	*Alt allele*	*R^2*	*Proxy eQTL HBECs*	*Pheno risk allele (AF)*	*Alt allele*	*R^2*
Signal A rs992969	*A (0.25)*	*G*	*N/A*	*N/A*	*N/A*		*rs2381416*	*C (0.26)*	*A*	*0.95*
Signal E rs4008366	*T (0.69)*	*C*	*rs693838*	*T (0.69)*	*C*	*1.0*	*rs442246*	*T (0.69)*	*G*	*1.0*

***Table**: Proxies* for the two functionally investigated signals in this manuscript. Proxies were used for the two tagSNPs in the eQTL datasets in case the original tagSNP of the LD block was not genotyped in the eQTL dataset. Criteria proxySNP selection: i) highest R^2 with tagSNP, but minimum R^2=0.5 ii); minimum AF=0.10 (EUR 1000G); iii) should itself still significantly associate with applicable phenotype of the genetic signal. AF= allele frequency (EUR 1000G); N/A: proxy not applicable, since tagSNP present; ND: Not determined, no proxy available at MAF≥0.10; Pheno risk allele: allele at risk for high blood eosinophils, asthma and/or eosinophilic asthma; Alt allele: alternative allele*

**proxies used in the lung tissue eQTL dataset have not been included in this table as no significant eQTLs were present in this dataset for the investigated signals, although all proxies had R^2>0.5 with the tagSNP of each selected signal A/E*

B) Supplemental Figures

Supplemental figure E1- Genetic region studied at *IL33* locus

A region of 400kb +/- IL33 was studied, being chr9: 5,815,786–6,657,983 (GRCh37/hg19):

Supplemental figure E2- tagSNPs of the five phenotype-associated LD blocks

E2-a **R^2 of five tagSNPs**

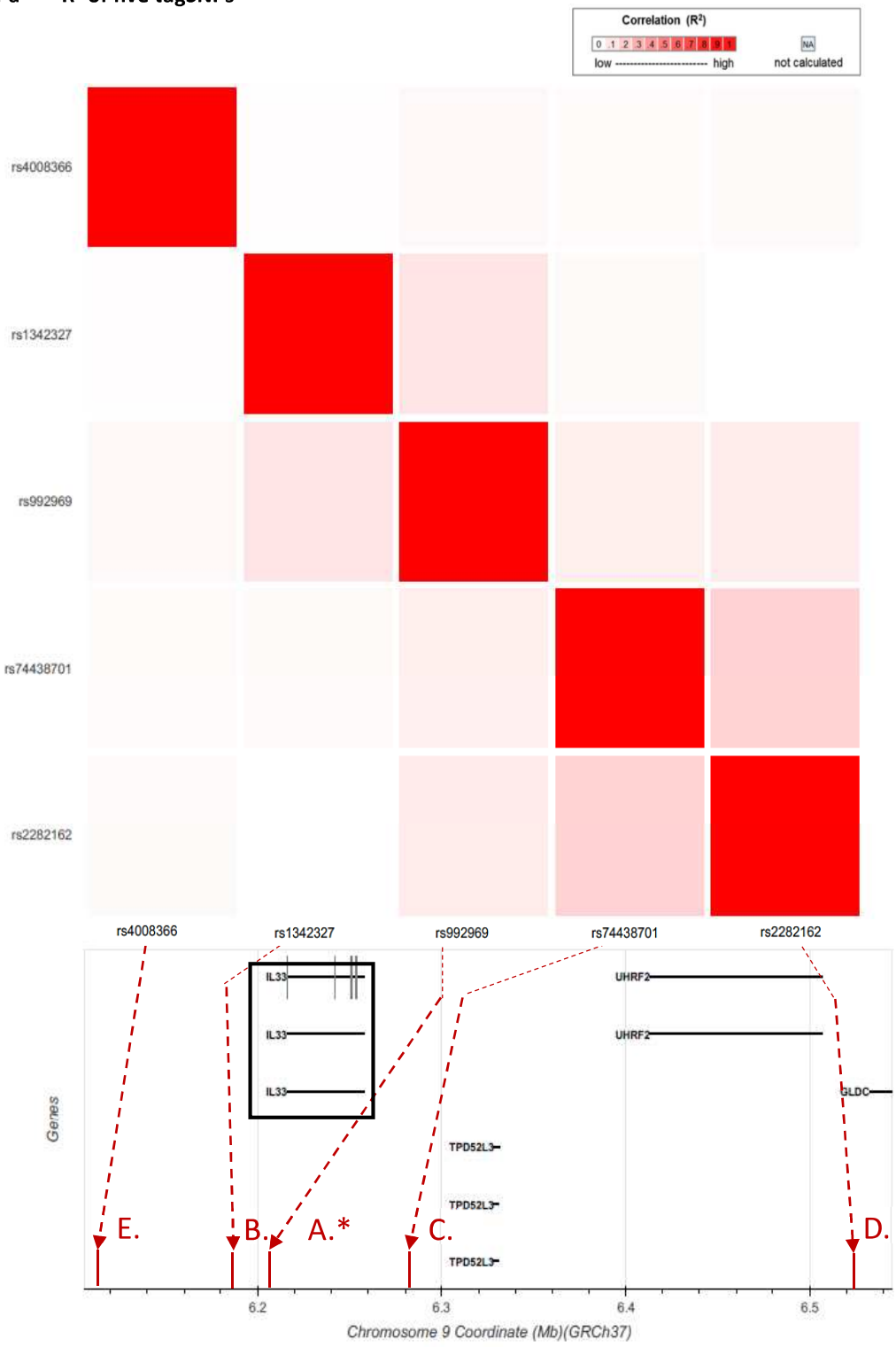


Figure-*The figure shows the LD pattern* ***(R²)*** *of the 5 tagSNPs representing LD blocks that were selected from the (in total 161) phenotype-associated SNPs, LD blocks of each signal defined using R^2>0.1. *Signals with known asthma-association from literature, see also table 1/S1.*

E2-b D’ of five tagSNPs

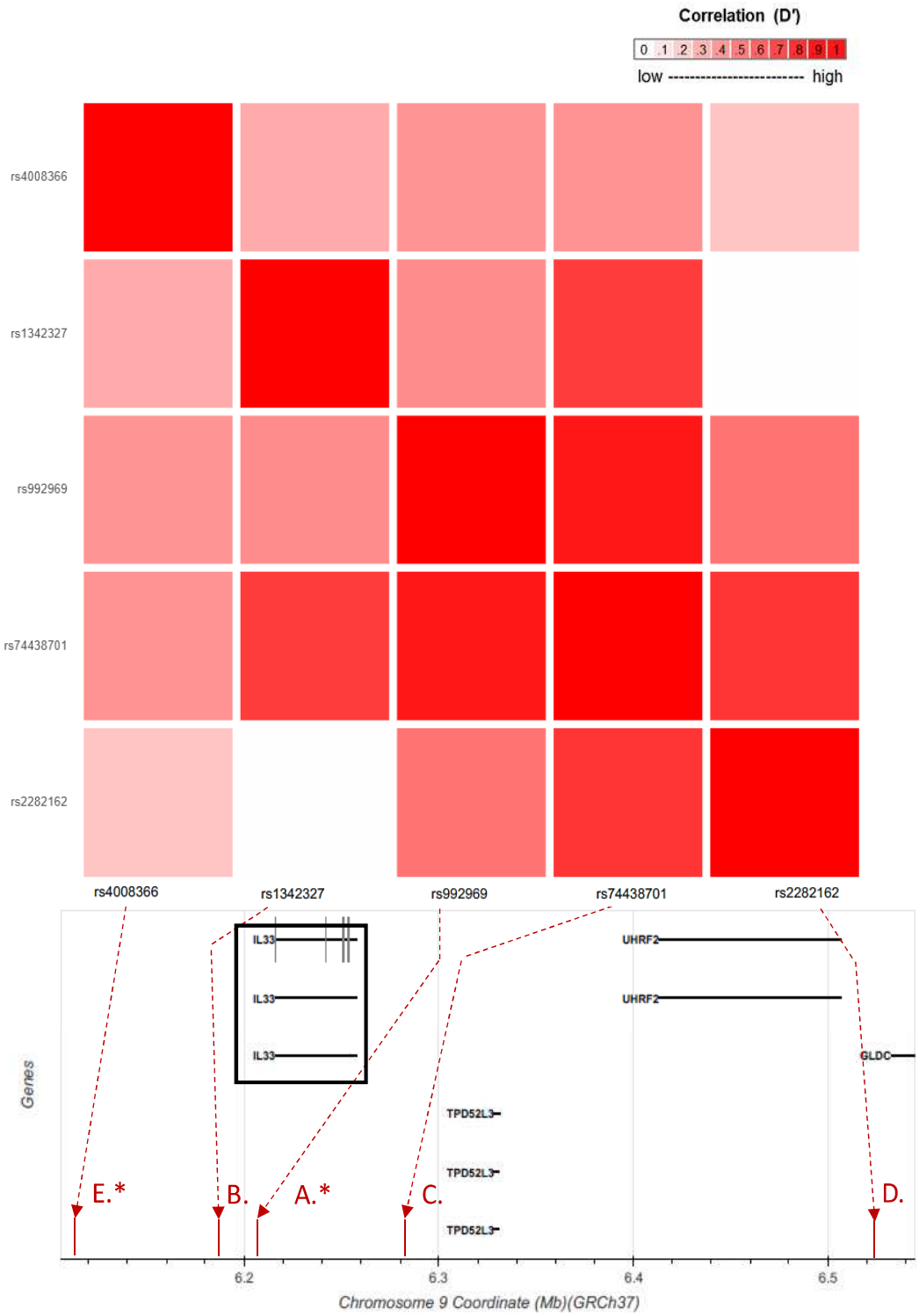


***Figure**-The figure shows the LD pattern **(D’)** of the 5 tagSNPs representing LD blocks that were selected from the (in total 161) phenotype-associated SNPs, LD blocks of each signal defined using R^2>0.1. it can be seen that the D’ is high as opposed to a low R^2 as shown in S2a. *Signals with known asthma-association from literature, see also table 1/S1. Image generated using the EUR population of the Phase I cohort of the 1000 genomes study via the online software tool LDlink 3.0, available at: https://analysistools.nci.nih.gov/LDlink/?tab=home.*

Supplemental figure E3- LD pattern within each of the five LD blocks (A-E)

E3-a LD Block A

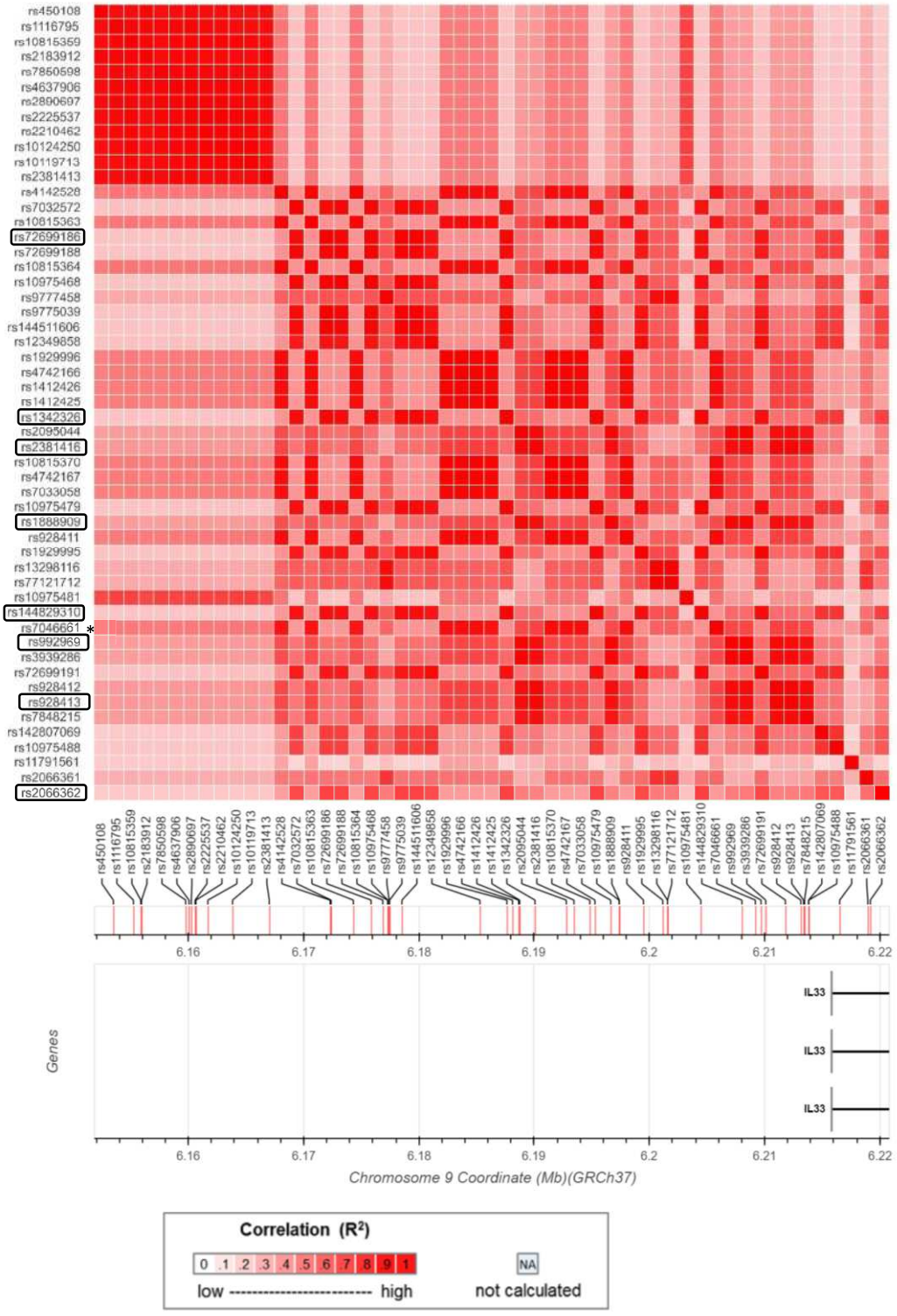


***Figure**: LD pattern (R^2) of the phenotype associated SNPs forming LD block A. Black squared SNP is also a GWAS-asthma SNP from literature. *TagSNP for LD block. Image generated using the EUR population of the Phase I cohort of the 1000 genomes study via the online software tool LDlink 3.0, available at: https://analysistools.nci.nih.gov/LDlink/?tab=home.*

E3-b LD Block B

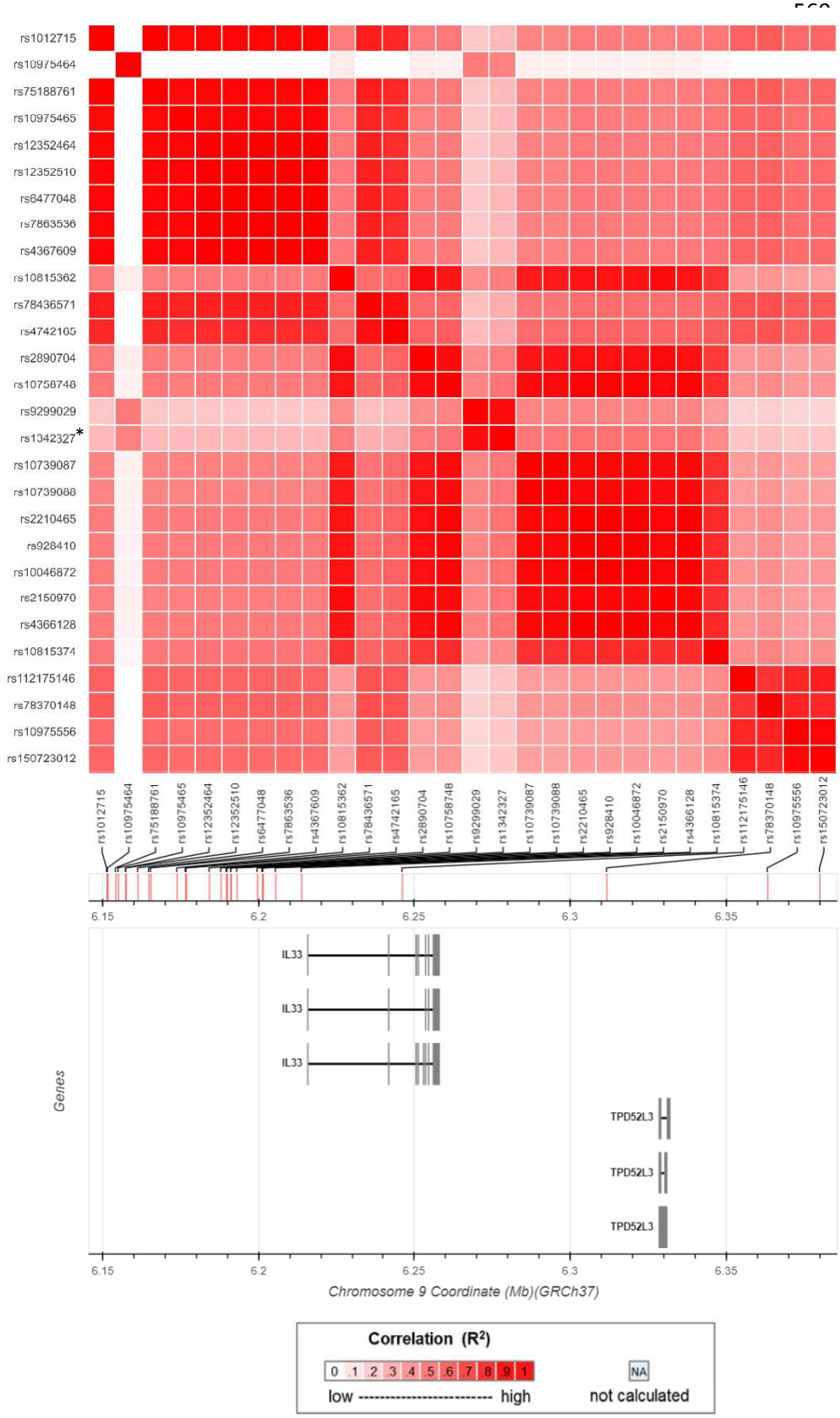


***Figure**: LD pattern (R^2) of the phenotype associated SNPs forming LD block B. *TagSNP for LD block. Image generated using the EUR population of the Phase I cohort of the 1000 genomes study via the online software tool LDlink 3.0, available at: https://analysistools.nci.nih.gov/LDlink/?tab=home.*

E3-c LD Block C

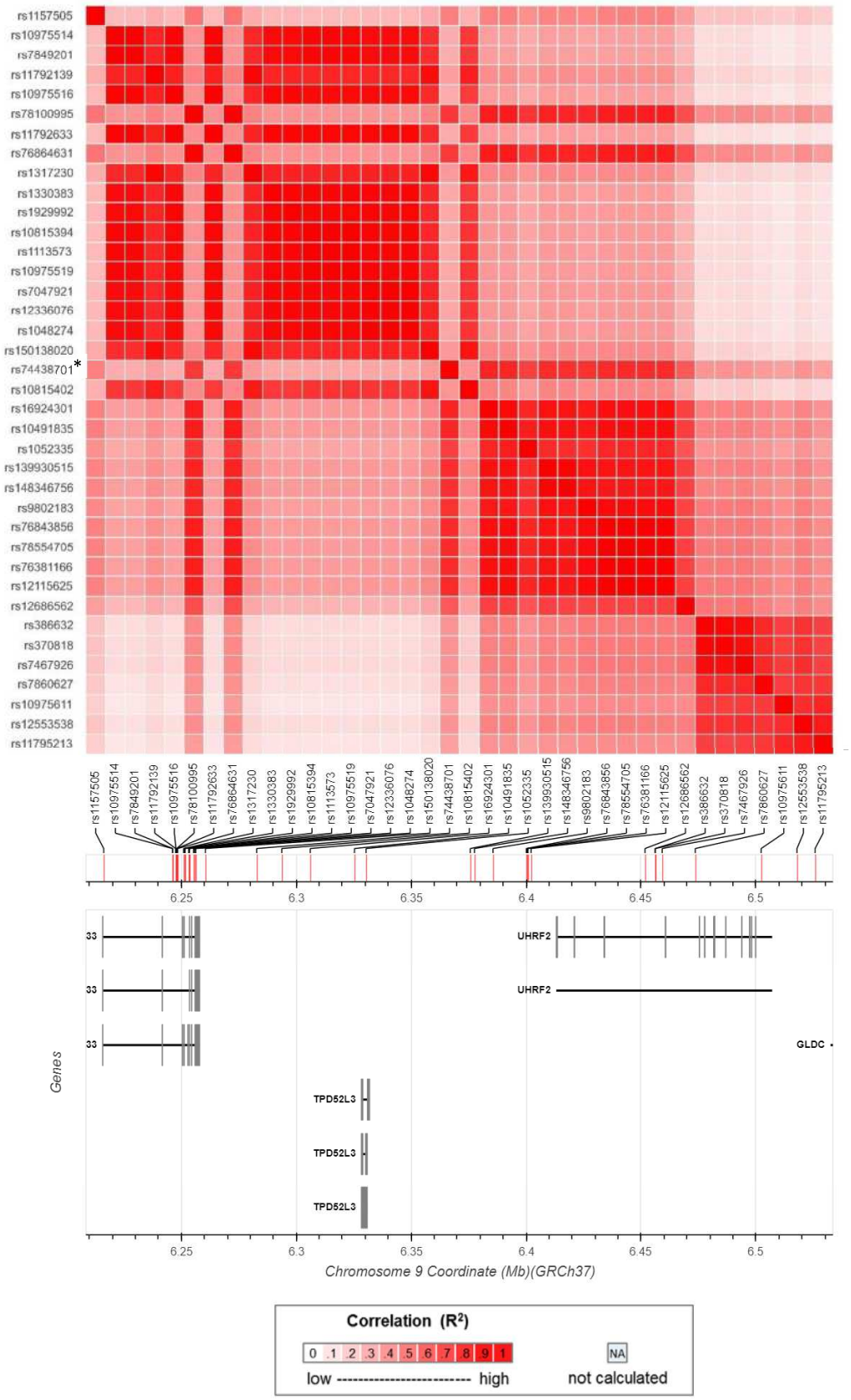


***Figure**: LD pattern (R^2) of the phenotype associated SNPs forming LD block C. *TagSNP for LD block. Image generated using the EUR population of the Phase I cohort of the 1000 genomes study via the online software tool LDlink 3.0, available at: https://analysistools.nci.nih.gov/LDlink/?tab=home.*

E3-d LD Block D

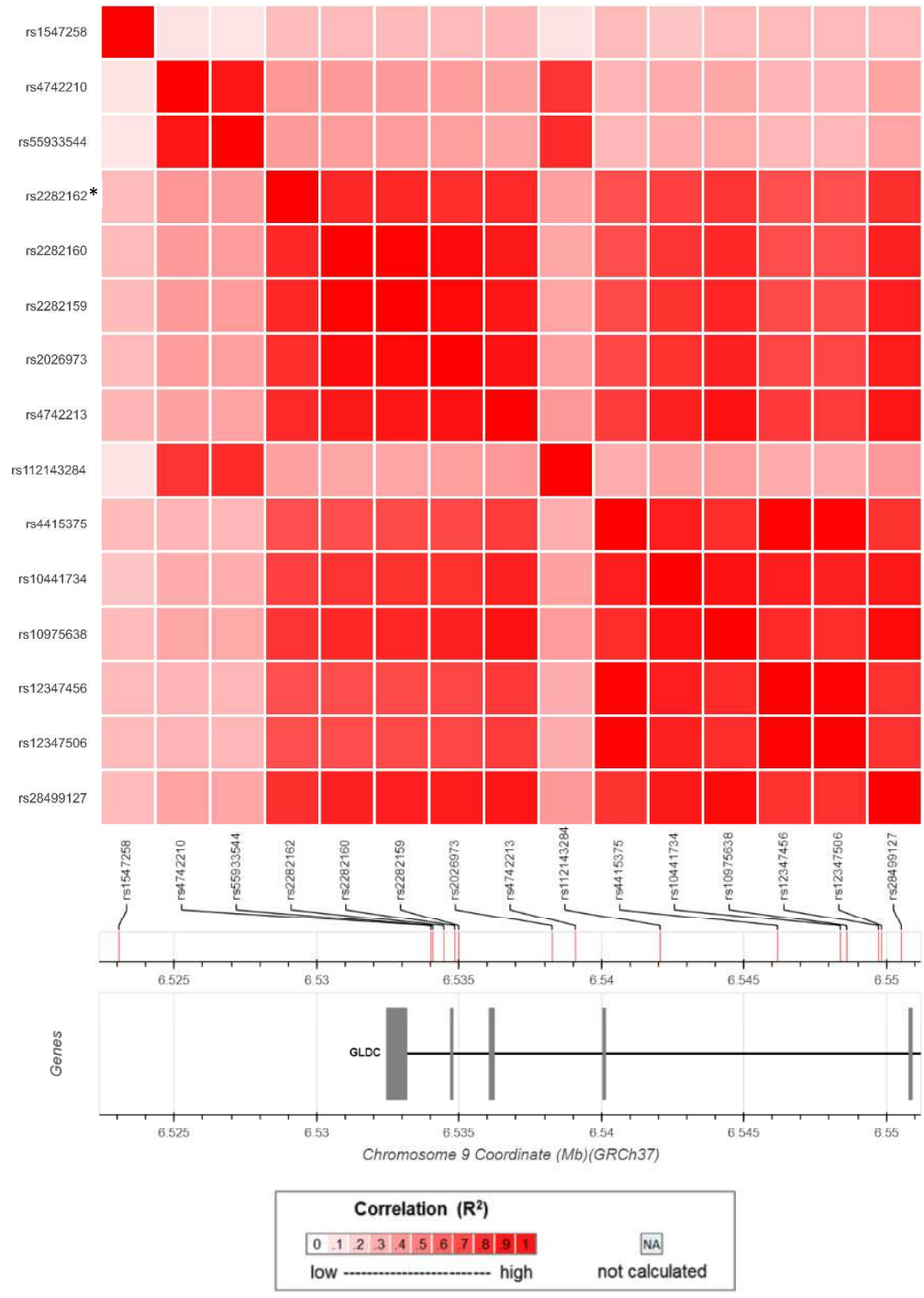


***Figure**: LD pattern (R^2) of the phenotype associated SNPs forming LD block D. *TagSNP for LD block. Image generated using the EUR population of the Phase I cohort of the 1000 genomes study via the online software tool LDlink 3.0, available at: https://analysistools.nci.nih.gov/LDlink/?tab=home.*

E3-e LD Block E

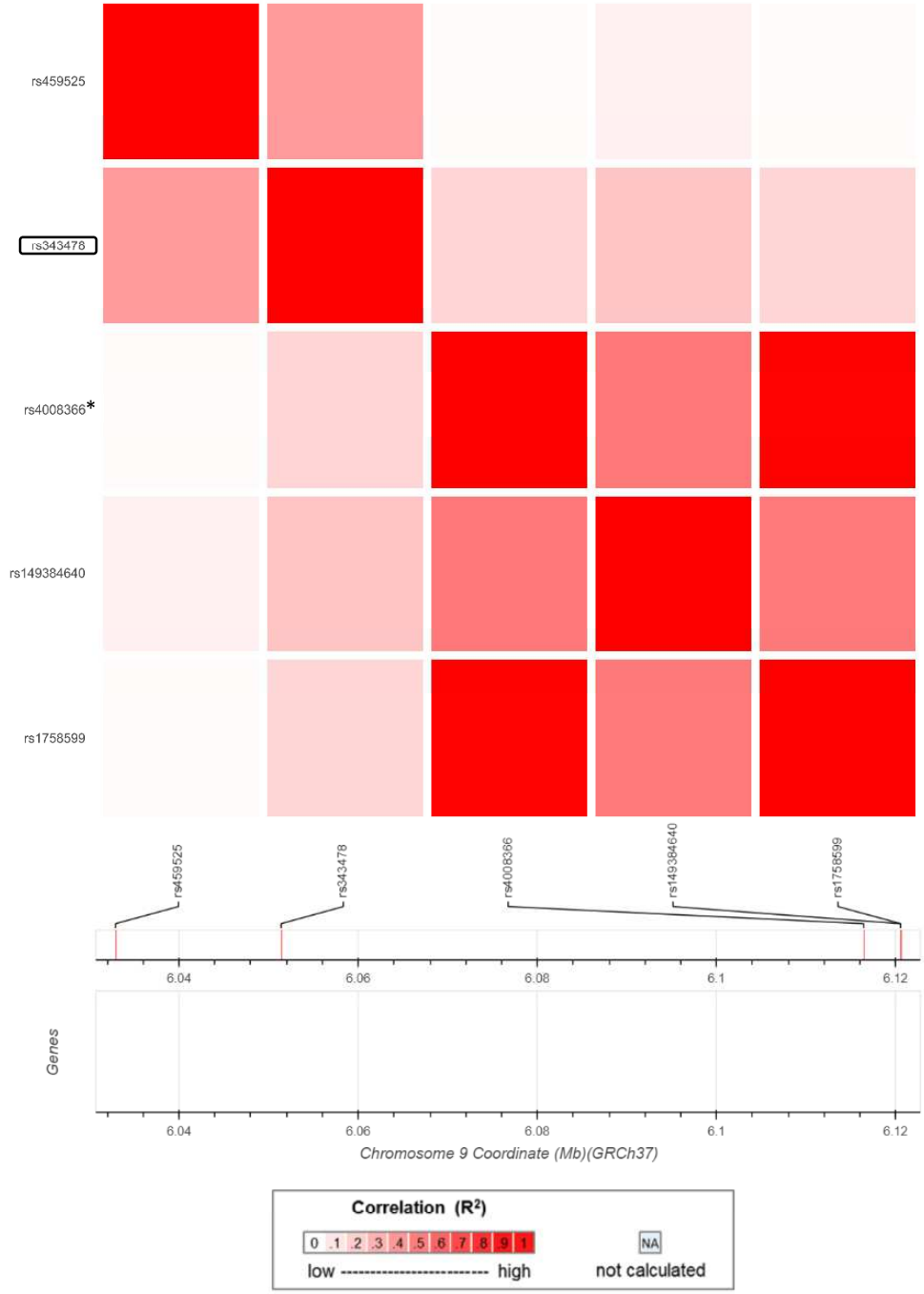
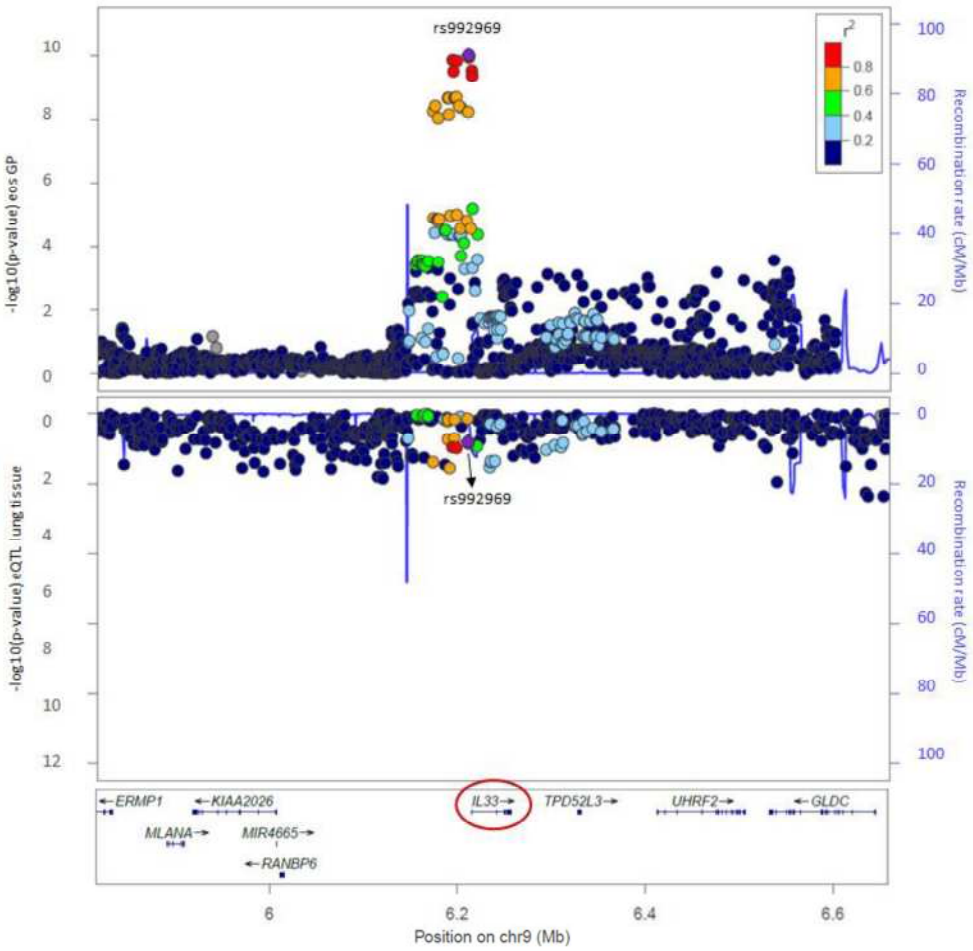


***Figure**: LD pattern (R^2) of the phenotype associated SNPs forming LD block E. Black squared SNP is a GWAS-asthma SNP from literature.*TagSNP for LD block. Image generated using the EUR population of the Phase I cohort of the 1000 genomes study via the online software tool LDlink 3.0, available at: https://analysistools.nci.nih.gov/LDlink/?tab=home.*

Supplemental figure E4- eQTL lung tissue in context of eosinophil associated signals

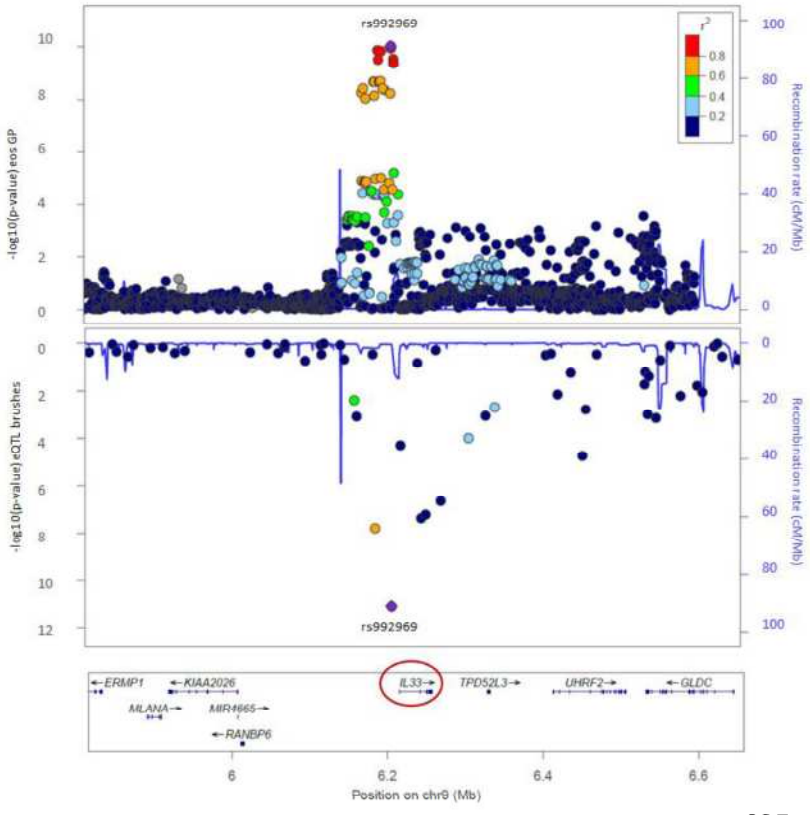


Upper panel: association of IL33 with blood eos GP Lifelines, n=13,395

Lower panel: IL33 eQTL lung tissue, n=1,1111

Plots generated using LocusZoom.(36)

Supplemental figure E5- eQTL bronchial brushes in context of eosinophil associated signals



Upper panel: association of IL33 with blood eos GP Lifelines, n=13,395

Lower panel: IL33 eQTL bronchial brushes n=139

Plots generated using LocusZoom.(36)

Supplemental figure E6- eQTL analyses in AHBECS: *IL33* mRNA levels stratified for *IL33* genotype of the phenotype associated signals.

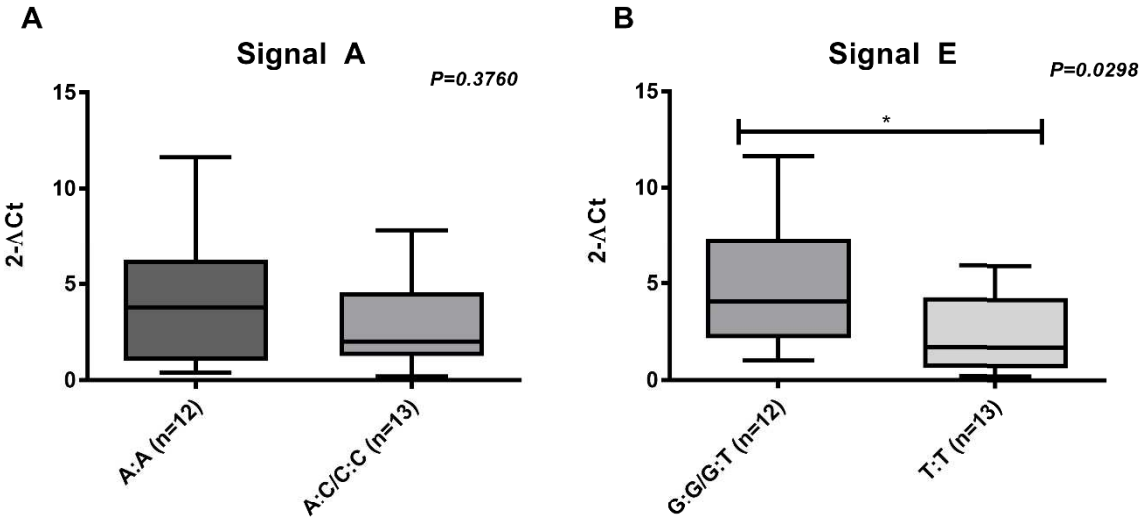


Figure E6- Cells were cultured *in vitro* and *IL33* mRNA levels were stratified based on the selected SNPs tagging distinct genetic signals of association at the *IL33* locus. Panel A represents Signal A tagged by rs2381416, and panel B represents Signal E tagged by rs442246. Statistics run was Mann-Whitney as appropriate for two group comparisons of non-normally distributed data. *P<0.05. Genotypes were grouped in a way to have at least an n=5 per group. Moreover, some variation in total number of subjects studied are present, as not each subject could contribute to all eQTL and pQTL experiments.

Supplemental figure E7- pQTL analyses in AHBECS: IL33 protein levels stratified for *IL33* genotype of the phenotype associated signals

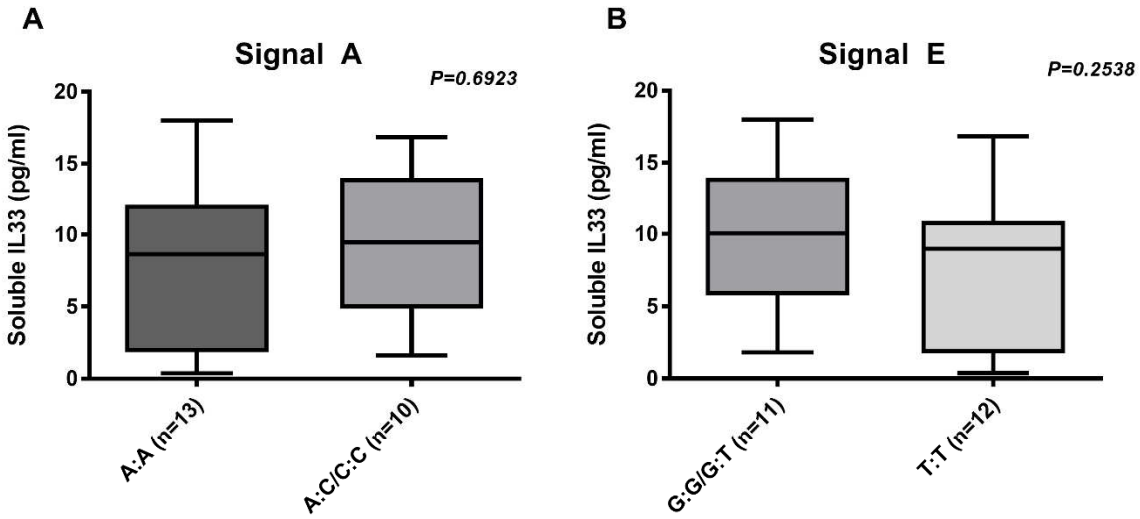
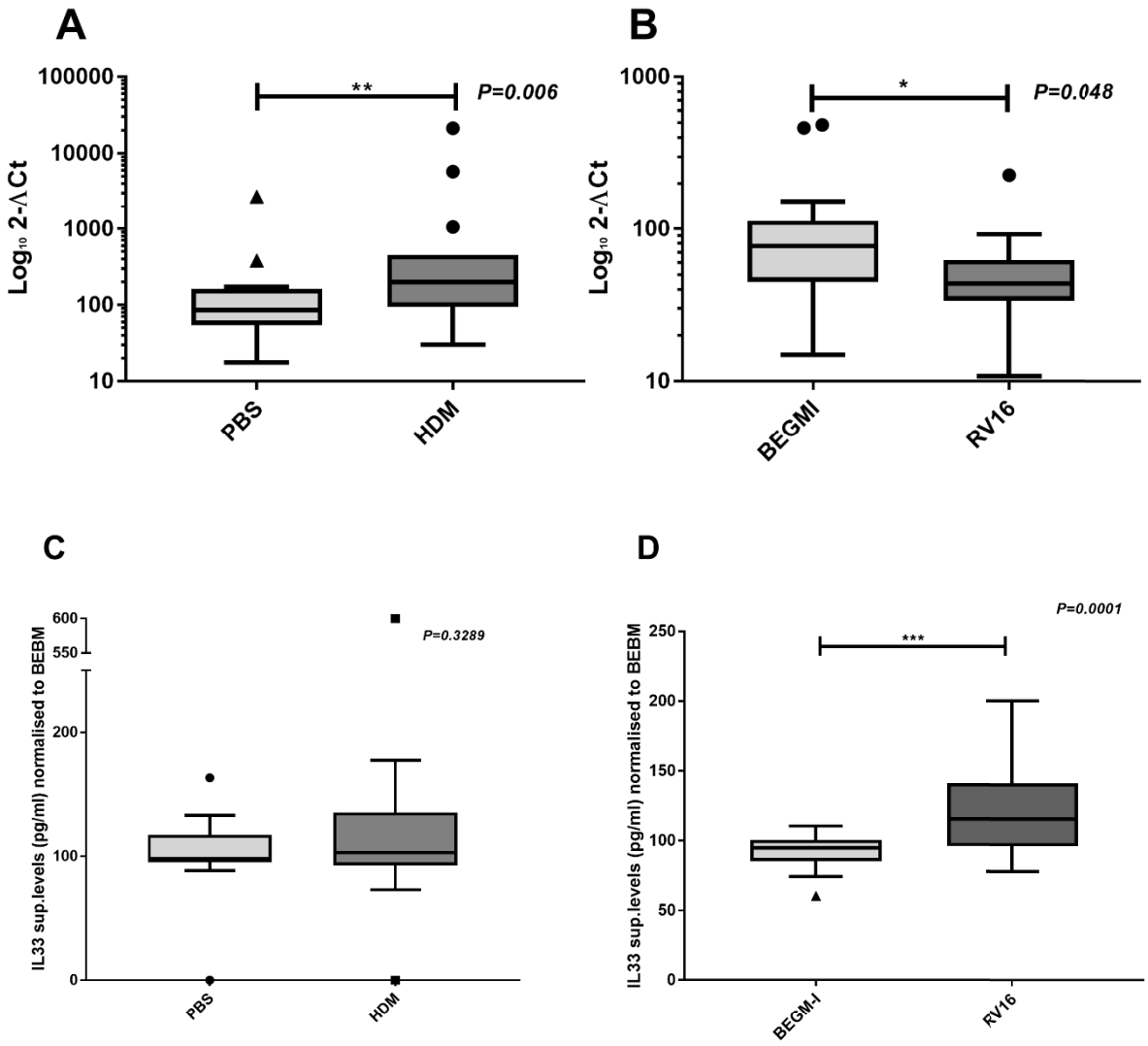


Figure E7: Cells were cultured in vitro and IL33 protein levels in cell supernatants (Luminex) were stratified based on the SNPs tagging distinct genetic signals of association at the *IL33* locus. Panel A represents Signal A tagged by rs2381416, and panel E represents Signal E tagged by rs442246. Statistics run was Mann-Whitney as appropriate for two group comparisons of non-normally distributed data. No statistically significant pQTLs were identified for these signals. Genotypes were grouped in a way to have at least an n=5 per group. Moreover, some variation in total number of subjects studied are present, as not each subject could contribute to all eQTL and pQTL experiments.

Supplemental figure E8- Effects of HDM and RV16 stimulation on IL33 mRNA and IL33 protein levels



Panel A/B: At the mRNA level, HDM stimulation (24h 50ug/mL) resulted in a 1.9-fold increase in *IL33* expression (Panel A, P=0.006, n=15), while stimulation with RV16 (MOI:1) for 24 hours reduced *IL33* mRNA levels 2.1-fold (Panel B, P=0.048, n=15). ***Panel C/D:*** protein level-Stimulation of cells with 50µg/ml HDM for 24 hours did not affect IL33 levels in the cellular supernatant (Panel C, P>0.05, n=18). RV-16 (MOI:1) stimulation for 24 hours however resulted in a 1.3-fold increase of IL33 protein in the cellular supernatant (Panel D, P=0.0001, n=18). Genotypes were grouped in a way to have at least an n=5 per group. Moreover, some variation in total number of subjects studied are present, as not each subject could contribute to all eQTL and pQTL experiments.

Supplemental figure E9- No effect of the phenotype associated signals on HDM and RV16 induced eQTL analyses

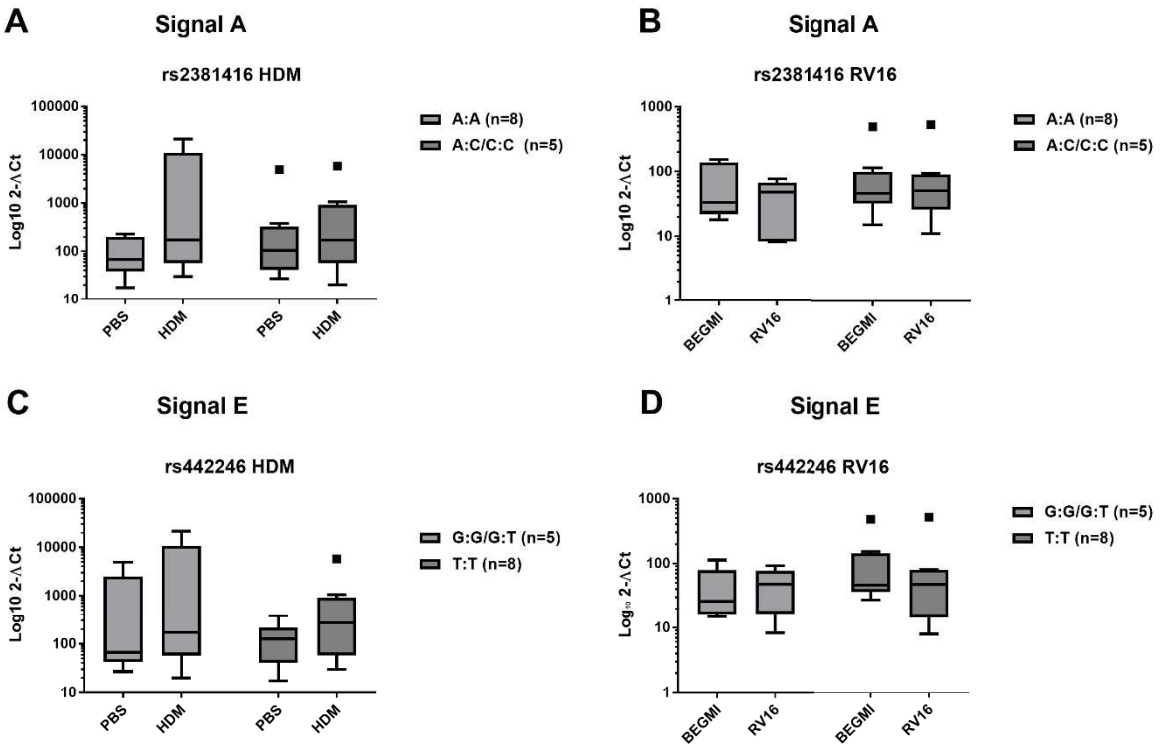


Figure E9: Complete analyses of levels of *IL33* mRNA in bronchial epithelial cells isolated from asthma patients and cultured *in vitro* under different asthma relevant micro-environments then stratified based on selected SNPs tagging genetic signals of association at the *IL33* locus. Each row represents the mRNA levels of a distinct genetic signal stimulated with HDM or RV16. Signal A (rs2381416): panel A (HDM)/ B (RV16). Signal E (rs442246): panel C (HDM)/ D (RV16). Statistics run were either Mann-Whitney or Kruskall Wallis as appropriate for two or three group comparisons. No statistically significant inducible eQTL were identified for these signals. Genotypes were grouped in a way to have at least an n=5 per group. Moreover, some variation in total number of subjects studied are present, as not each subject could contribute to all eQTL and pQTL experiments.

Supplemental figure E10- No effect of the phenotype associated signals on HDM and RV16 induced pQTL analyses

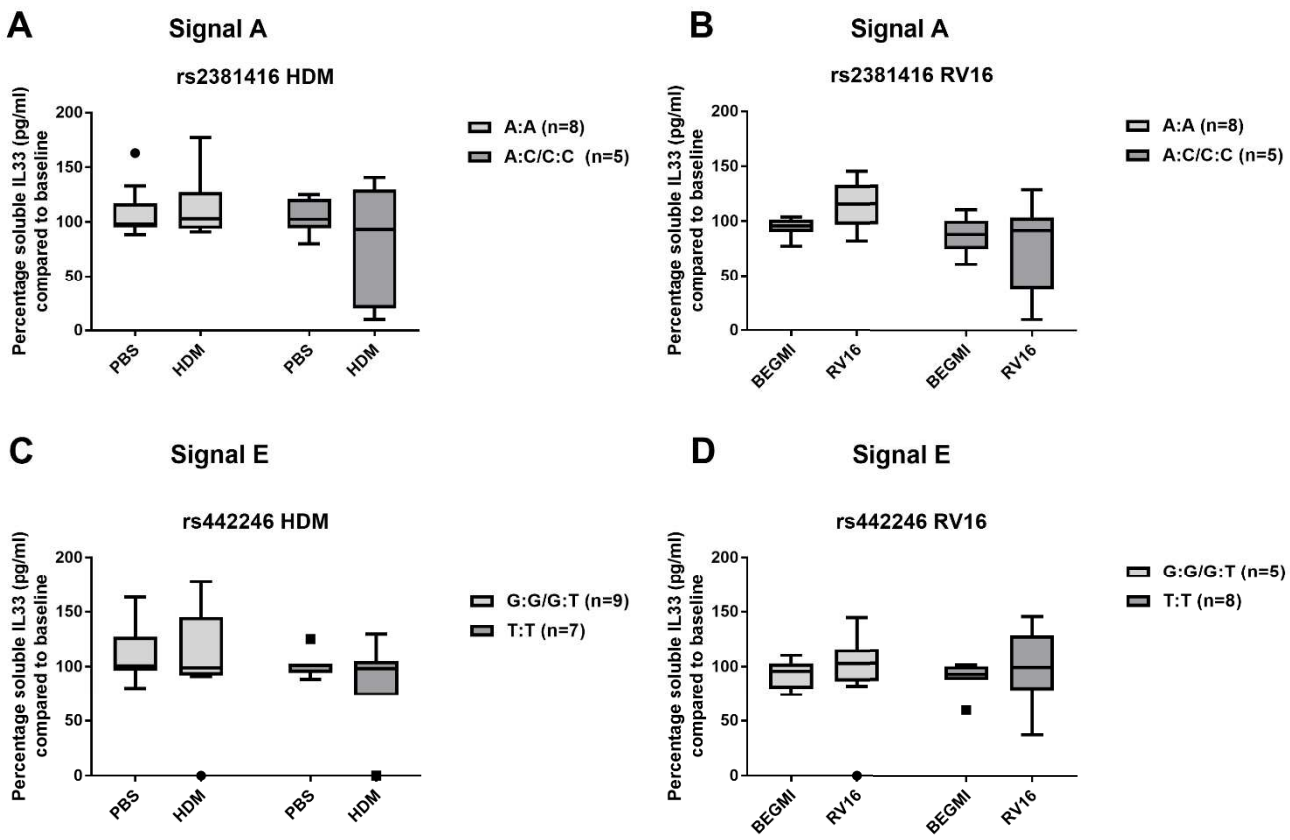


Figure E10: Complete analyses of levels of IL33 protein in supernatants of bronchial epithelial cells isolated from asthma patients and cultured *in vitro* under different asthma relevant micro-environments then stratified based on selected SNPs tagging genetic signals of association at the *IL33* locus. Each row represents the protein levels of a distinct genetic signal stimulated with HDM or RV16. Signal A (rs2381416): panel A (HDM)/ B (RV16). Signal E (rs442246): panel C (HDM) /D (RV16). Statistics run were either Mann-Whitney or Kruskall Wallis as appropriate for two or three group comparisons. No statistically significant inducible pQTL were identified for these signals. Genotypes were grouped in a way to have at least an n=5 per group. Moreover, small variation in total number of subjects studied are present, as not each subject could contribute to all eQTL and pQTL experiments.

Supplemental figure E11- Lentiviral overexpression method

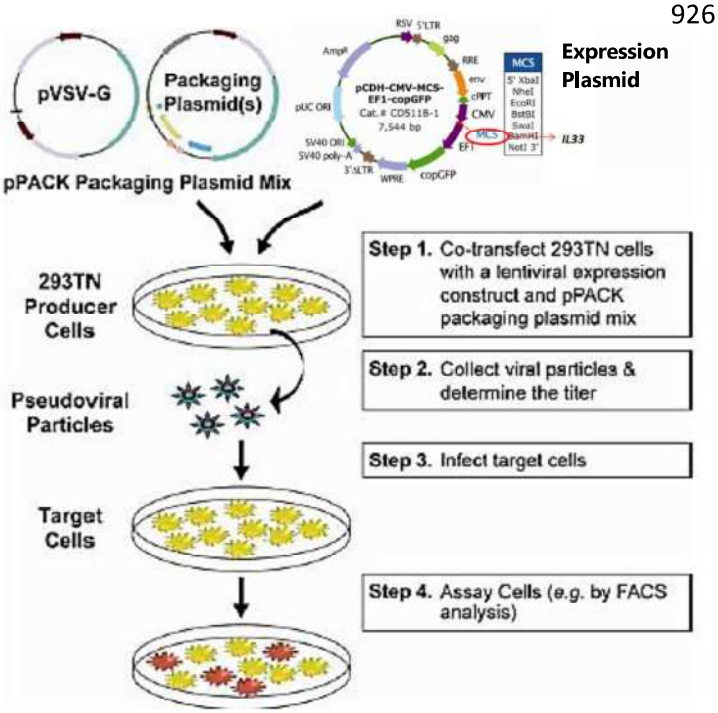


Figure E11- A three plasmid system was used for lentiviral overexpression of *IL33* in primary human bronchial epithelial cells, consisting of an envelope plasmid (pCMV_VSV-G CellBiolabs RV110, Addgene plasmid # 8454), a packaging plasmid (pCMV_8.91 (Addgene plasmid #2221)) and the plasmid containing the actual expression construct (human full-length *IL33* aa1-270, transcript variant 1, NM_033439) in pCDH-CMV-MCS-EF1-copGFP (System Biosciences, #CD511B-1). Controls were: transduction without plasmid (or NV), transduction with expression plasmid lacking *IL33* (empty vector control or EV). For details see supplemental methods above and Torr et al(17). *Figure adapted from SBI handbook 'pCDH cDNA Cloning and Expression Lentivectors CD- 500/800 series'* https://www.systembio.com/wp-content/uploads/Manual_pCDH_Vectors-1.pdf

Supplemental figure E12- GFP expression matched between *IL33* overexpression vector and empty vector

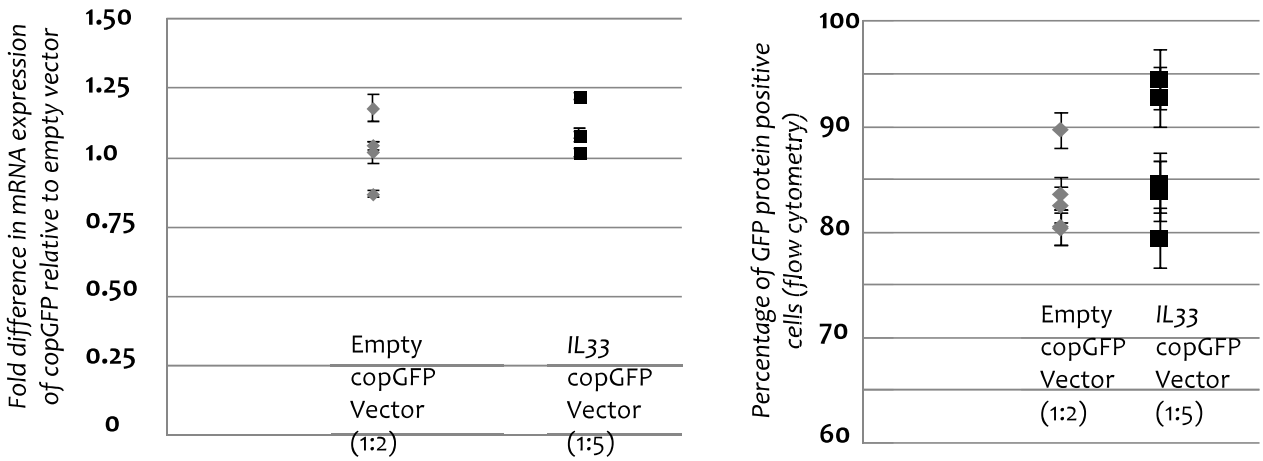


Figure E12- Using qPCR (left) and flow cytometry (right) (see methods), the copGFP expression was matched between the empty vector (GFP only, 'EV', grey squares) and the *IL33* overexpression vector ('IL33', black squares) on mRNA and protein level. This provided confidence that any differences were not due to different infection efficiency. These initial analyses identified infection ratio of lentiviral particles:polybrene of 1:2 for the EV and 1:5 for the IL33. Left (mRNA): data expressed as fold difference in copGFP mRNA levels compared to EV. Right: data expressed as percentage of copGFP positive cells as determined by flow cytometry. N=5 independent HBEC donors, data points represent mean +/-standard deviation for 2 technical replicates per donor.

Supplemental figure E13- *IL33* overexpression results in elevated IL33 mRNA in engineered cells

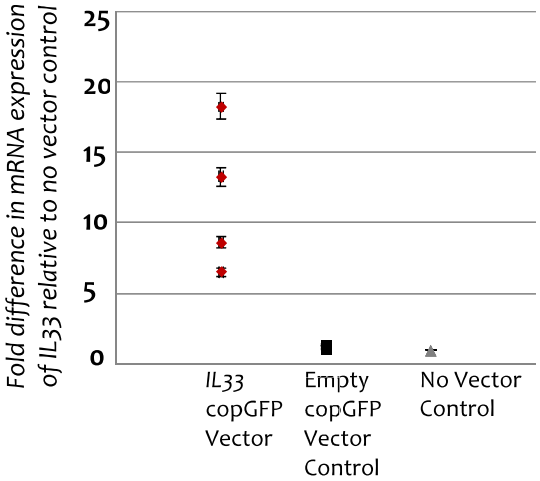
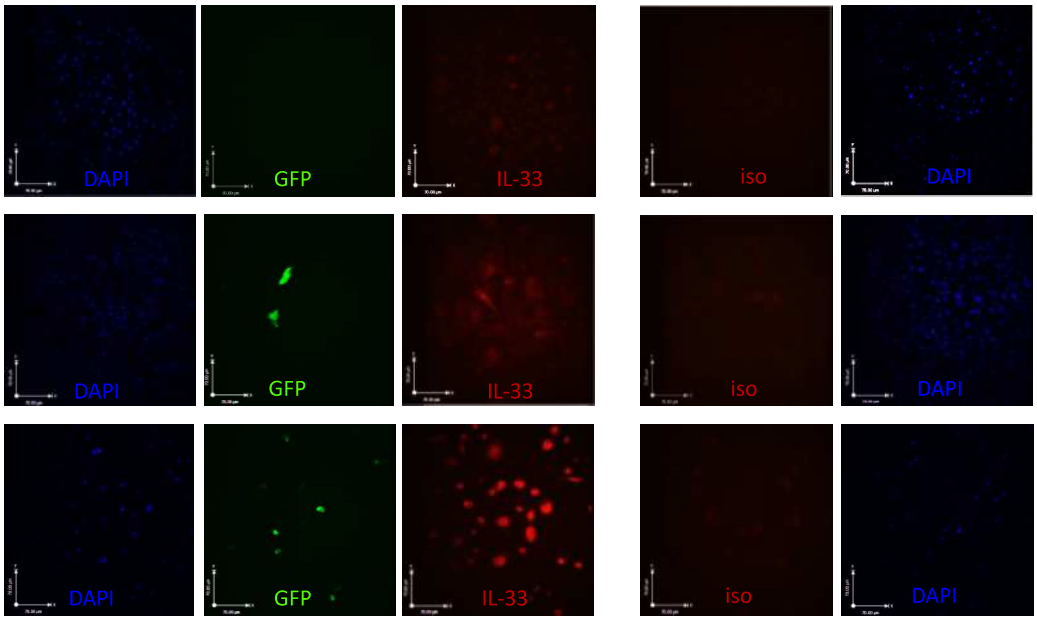


Figure E13- Using qPCR the overexpression of *IL33* was confirmed in the N=5 HBECs derived from healthy donors. Data expressed as fold difference in *IL33* mRNA levels compared to no vector control. N=5 HBEC donors, data points represent mean +/-standard deviation for 2 technical replicates per donor.

Supplemental figure E14- IL33 expression confirmed on protein level using 2 different antibodies

E14-a Polyclonal antibody against IL33 (ProteinTech)



E14-b Monoclonal antibody against IL33 (ProteinTech)

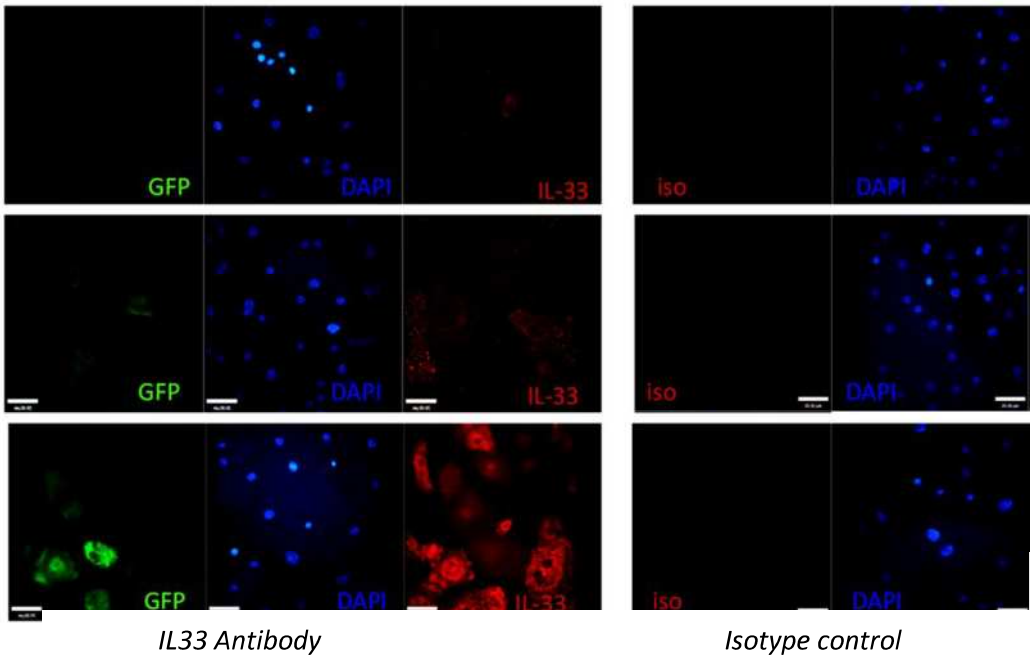


Figure E14- Using immunofluorescence the expression of IL33 on protein level was confirmed in HBECs transduced with lentivirus containing a human *IL33* expression cassette, whilst the exact cellular location was inconclusive considering the staining patterns of 2 different antibodies (red). Cells were processed for immunofluorescent staining at passage 2, two weeks after the lentiviral transduction when cells were considered virus-free.

Supplemental figure E15- Functional assays upon *IL33* overexpression in HBECs

E15-a No effect of *IL33* overexpression on cell number

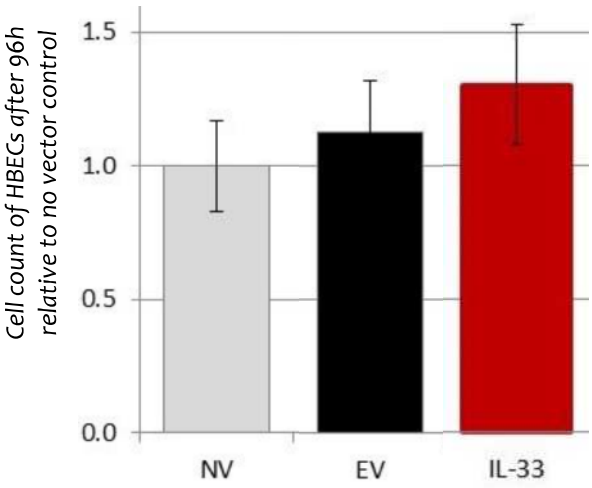


Figure E15a- HBECs overexpressing *IL33* (red) were cultured submerged in 6-wells plates and harvested+counted at 96h when they were in the log-phase of proliferation (empirically determined in pilot experiments) and compared against empty vector (EV) control and no vector (NV) control. Data represent n=3 HBEC donors, mean+ standard deviation, counts expressed relative to NV. Cross-sectional comparison of 3 conditions (MWU, p=0.21).

E15-b No effect of *IL33* overexpression on Metabolic activity (MTT)

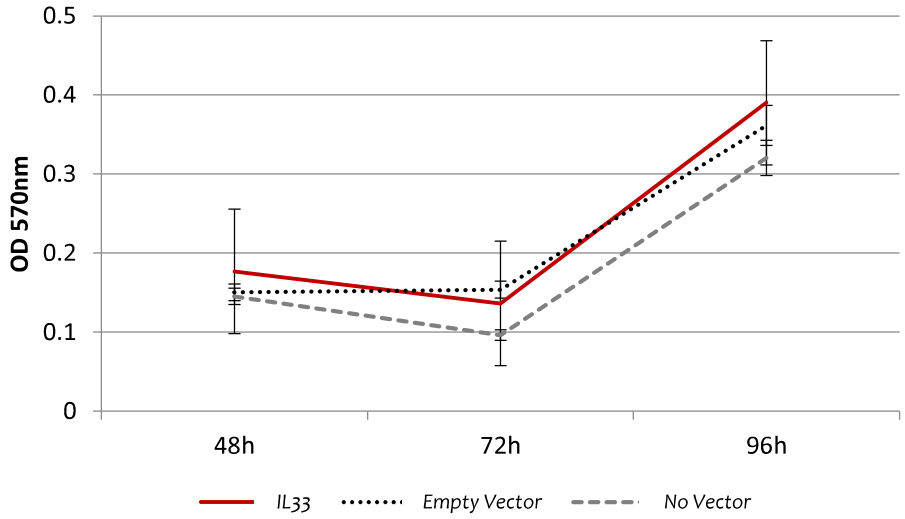
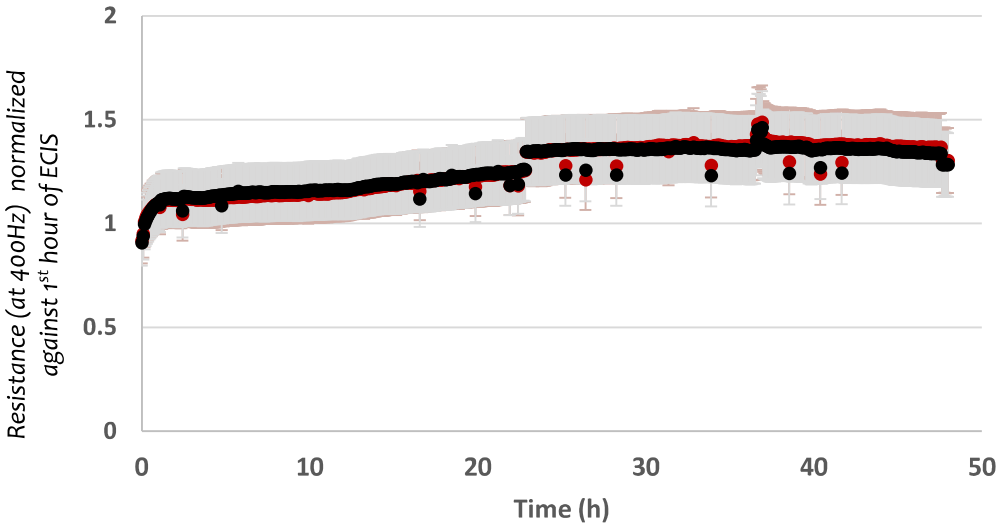


Figure E15b- Using an MTT assay (see methods) the metabolic activity of submerged cultured HBEC was determined upon sustained *IL33* overexpression. Cells were harvested and MTT activity measured every 24h between 48-96h, then cross-sectionally compared using Wilcoxon paired rank test (IL33 vs empty vector control at 48h p=0.65, at 72h p=0.72, at 96h p=0.57). Data represent n=5 HBEC donors per timepoint per condition; mean+/- standard deviation of OD values.

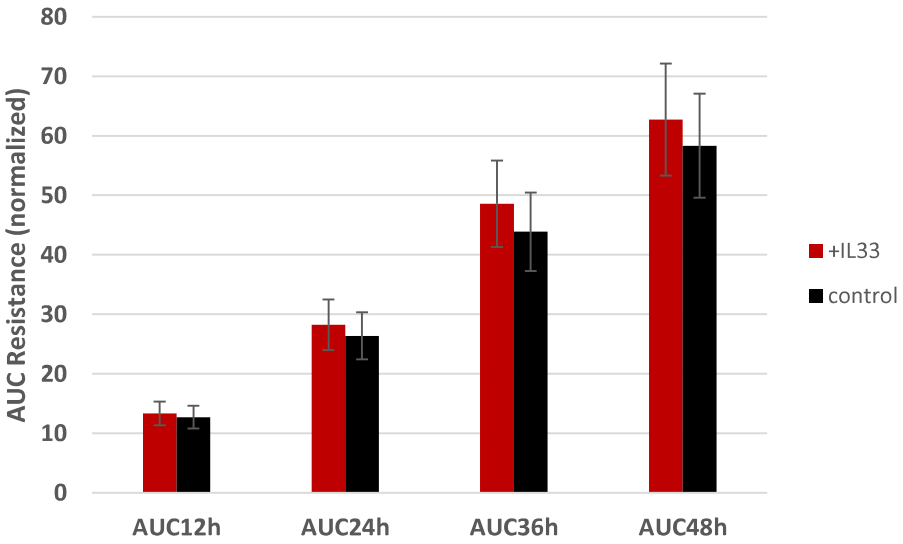
E15-c No effect of *IL33* expression on resistance (ECIS) over time

Figure E15c- HBECs overexpressing *IL33* were cultured to confluency on electrode-containing plates and resistance (at 400Hz) measured using ECIS every 10 minutes for 48h. Due to large donor variation in resistance values from the start onwards, data were normalized against the first hour to enable comparison of treatment groups. Mean +/- standard deviation of n=5 HBEC donors.



Upper panel: Longitudinal plot. Red dots with light red error bars: *IL33* overexpressing cells, black dots with grey error bars: empty vector (EV) controls.

Lower panel: Area under the curves (AUCs) were calculated, normalized against the average resistance of the first hour, then cross-sectionally compared every 12h. No difference in AUC was seen comparing *IL33* vs EV (12h; p=0.78, 24h; p=0.65, 36h; p=0.42, 48h; p=0.38, MWU). control=EV.



Supplemental material:

IL33 functional genetics-Ketelaar, Portelli, Dijk et al

Table of Contents

Supplemental Methods

M1- Cohort descriptions and details of genotype-phenotype analyses (see also supplemental table S1-S2):

Lifelines general population cohort (table S1)

Lifelines is a multi-disciplinary prospective population-based cohort study examining in a unique three-generation design the health and health-related behaviours of 167,729 persons living in the North of The Netherlands. It employs a broad range of investigative procedures in assessing the biomedical, socio-demographic, behavioural, physical and psychological factors which contribute to the health and disease of the general population, with a special focus on multi-morbidity and complex genetics. The cohort profile of the Lifelines study has been extensively described in Scholtens et al(1). Summarizing, the participants' baseline visit took place between December 2006 and December 2013. All general practitioners in the three northern provinces of the Netherlands were asked to invite their registered patients aged 25–49 years. All persons who consented to participate were asked to provide contact details to invite their family members (i.e., partner, parents and children), resulting in a three-generation study. Baseline data were collected from 167,729 participants, aged from 6 months to 93 years. Collected data include physical examinations, DNA, blood and urine samples, and comprehensive questionnaires on history of diseases, quality of life, lifestyle, individual socioeconomic status, work, psychosocial characteristics and medication use. Follow-up is planned for at least 30 years, with questionnaires administered every 1.5 years and a physical examination scheduled every 5 years. At current, a subset of the adult participants have both phenotypic and imputed genotype information available (n=13,395).

Participants of the Lifelines cohort were genotyped on the HumanCytoSNP-12 BeadChip (Illumina). Quality control before imputation was performed using ImputationTool2(2), excluding SNPs with a call-rate <95%, with a HWE-P value <0.001, MAF <0.01%. Samples were excluded in case of ambiguous sex (genetic mismatch with reported sex), of non-Caucasian origin (based on self-report, IBS and population stratification using EIGENSTRAT(3), and in case a pair of samples was discovered as first degree relatives using genetic cryptic relatedness, the sample with the best genotype quality was included only. Imputation was performed through Beagle 3.1.0 against the EUR panel from the 1000 genomes project (version March 2012)(4).

Klijs et al (2015)(5) concluded that the Lifelines adult study population is broadly representative for the adult population of the north of the Netherlands. The recruitment strategy had minor effect on the level of representativeness. These findings indicate that the risk of selection bias is low and that risk estimates in Lifelines can be generalized to the general population.

Genetics of Severe Asthma Phenotypes cohort (GASP) (table S2)

Asthmatic individuals from the GASP cohort (n=2,536) were used in the current study, this cohort has been recruited across UK hospitals as part of an Asthma UK initiative and is enriched for patients with British Thoracic Society Step 3 and above (~2,200 moderate-severe asthma, remainder mild asthma)(6). Asthma was defined as a doctor's diagnosis of asthma through the presence of symptoms and medical treatment, while age of onset of asthma was determined through patient records. Asthma related clinical phenotypes used in the current study focused to lung function (FEV_1 pre bronchodilator, FEV_1/FVC pre-bronchodilator), atopic status (positive skin prick test), Blood Eosinophil Count (x10^9/L) and Blood IgE levels (kU/L). Total peripheral blood eosinophil levels were calculated using a counting chamber while total Immunoglobulin E (IgE) levels were measured by ImmunoCAP™. Finally, atopy was defined as a positive response to a skin prick test (SPT) to any allergen from a panel of 4-24 allergens.

Participants in the GASP cohort were genotyped using two platforms, initially 744 subjects using the Affymetrix Axiom® UK BiLEVE array and 2172 subjects using the Affymetrix Axiom® UK Biobank array. In each genotyping batch samples were excluded: (i) if their genetically inferred gender did not match their reported gender; (ii) if they had outlying heterozygosity within the batch (outside either 2 or 3 standard deviations from the mean depending on batch); (iii) if they had a call rate <95% across genotyped variants; (iv) if cryptically related to another sample, 1 sample of the pair was removed; (v) if the sample shows significant deviation from European ancestry as determined by a plot of the first two principal components. The batches were merged and SNPs not available in both batches were excluded from the dataset. Following quality control 692,060 SNPs were available for 2,536 subjects. Imputation was then performed using IMPUTE 2.0 against the reference data set of the EUR

panel of the 1000 Genomes project (version March 2012)(4). Genetic studies were approved by the Medical Ethics Committee of each participating centre or via our multicentre ethics approval.

Dutch Asthma GWAS cohort (DAG) (table S2)

The DAG cohort has been extensively described previously(7). In summary, the DAG cohort consists of 469 trios ascertained through a proband with asthma, combined with an additional case-control study of 452 asthmatics and 511 controls. Of these, we selected 909 unrelated asthma patients who underwent the same, standardized, comprehensive evaluation for asthma at Beatrixoord Hospital, Haren, The Netherlands between 1962-2003. Asthma was defined as a doctor's diagnosis of asthma, asthma symptoms, and bronchial hyperresponsiveness (BHR). FEV_1 was measured using a water-sealed spirometer (Lode Spirograph type DL, Lode b.v., Groningen, The Netherlands). Total peripheral blood eosinophils were counted in a counting chamber and IgE levels were measured in serum by an enzyme-linked fluorescence assay (Mini Vidas, Biomerieux Inc., Marcy, France). In subjects older than 12 years intracutaneous tests with 16 common aeroallergens were performed. In children younger than 12 years, a skin prick test was performed with 10 allergens. Subjects with a positive response to one or more intracutaneous or skin prick tests (SPT) were considered to be atopic. Age of asthma onset was based on data from medical records and questionnaires, indicating the start of asthma symptoms. Participants in the DAG cohort were genotyped on two platforms, the Illumina 317 Chip and the Illumina 370 Duo Chip (Illumina, San Diego, CA). Quality control (QC) was performed per chip with exclusion of individuals with missing genotype call rate >0.01, related individuals (identity by descent (IBD) >0.125) and non-Caucasian subjects, as assessed by principal components analysis performed with EIGENSTRAT(3). SNPs were excluded with a missing genotype rate >0.01, a Hardy-Weinberg equilibrium P-value <10-7 and a MAF <0.01. Markers with Mendelian errors in phase I were excluded from analysis. Following quality control, the chips were merged and SNPs not available in both cohorts were excluded from the dataset. A total of 294,775 SNPs remained. Imputation was performed using IMPUTE 2.0 against the reference data set of the EUR panel of the 1000 Genomes project (version March 2012)(4). Genetic studies were approved by the Medical Ethics Committee of the University Medical Center Groningen and all participants provided written informed consent.

Next-Generation DNA Sequencing (NGS) cohort.

DNA from 200 severe asthma cases (BTS 4, 5) from GASP and 200 non-asthmatic, non-atopic, non-wheeze controls from the Nottingham Gedling cohort(8), were selected for resequencing. Subjects were matched for age and gender (Supplemental Table 3). Next-generation Illumina sequencing of the IL33 region (chr9:5924967-6267982) was outsourced to Source Bioscience (Nottingham, UK) and was carried out using the SureSelect enrichment approach. The chromosome 9 locus previously associated with asthma [GRCh37.p9] was the focus and 120 base pair paired-end long read oligonucleotides (baits) were designed using the SureSelect™ e-array design software. Bait tiling (X5) was used across the region, presenting with a capture size range of 500Kb to 1.5Mb. The initial target region was 343,016bp; using 7,751 baits achieved 65.28% coverage of this region. Samples were pooled for sequencing (3 pools for cases and 3 pools for controls). Next-generation sequencing was carried out on these six samples on two separate lanes, one for cases and the other for controls, using the Illumina HiSeq2000™ systems pipeline (San Diego, USA). Sequencing used a paired end design using 100bp reads. Resequencing the IL33 region identified 981 variants that were considered valid calls by SNver. Case-control association analyses revealed 12 SNPs significantly associated with severe asthma, of which two were within LD block E, the remaining 10 SNPs were rare single variants, of which 7 SNPs were novel (table S14). Due to the low frequency (MAF<0.1) these were not followed-up functionally.

Details of Genotype – Phenotype associations

For the genotype-phenotype association analyses SNPs were selected with a MAF≥0.01 located 400kb up- and downstream the *IL33* gene (Chr9: 5,815,786−6,657,983), encompassing all known asthma association signals (see also table S1). There were 1,970 SNPs present in Lifelines, and there were 3,025 and 2,780 SNPs available in the GASP and DAG cohorts, respectively, with 2,457 shared for meta-analysis of GASP/DAG. All Lifelines SNPs were present in the pool of SNPs of the DAG/GASP meta-analysis. Annotated SNP location and function was determined with the use of HaploReg v4.1(9). All genetic data are presented relative to assembly GRCh37/hg19. In Lifelines we performed genetic association within the genotyped subset of the general population cohort (n=13,395 with both genotype and phenotype information), with eosinophil counts, FEV1 (%pred) and FEV1/FVC.

This was followed by a sensitivity analysis for blood eosinophil counts, where we removed asthma patients (n=1,066; doctor's diagnosed asthma) and asthmatic+allergic subjects (n=6,227) and associated *IL33* SNPs with blood eosinophils within this non-asthmatic, non-allergic population. Herein, allergy was defined as having at least one self-reported allergy the questionnaire covering eczema, rhinitis, food allergy, dust allergy, animal allergy, pollen allergy, medication allergy, contact allergy, and insect bite allergy.

From this Lifelines general population cohort we subsequently took the asthma subpopulation (n=1,066, doctor's diagnosed asthma) and investigated genetic association with eosinophil counts, eosinophilic asthma (asthma and blood eosinophils >150 cells/µL, n=707), FEV1, FEV1/FVC and asthma with airway obstruction (asthma and FEV1<80% of predicted (n=258) or FEV1/FVC<70% (n=324)). Subsequently, we performed association analyses in a meta-analysis of GASP (n=2,536) and DAG (n=909) asthma patients studying atopy, blood eosinophils, total serum IgE, age of asthma onset, and lung function (FEV1, FEV1/FVC).

Associations of SNPs in the *IL33* region with FEV_1, FEV_1/FVC, blood eosinophils/neutrophils, total IgE levels, atopy and age of asthma onset were performed with PLINK v1.90b6.7(10) (Lifelines) or SNPtest v2.5β(11) (DAG/GASP) using an additive genetic model. Eosinophils, neutrophils and IgE levels were logarithmically transformed before analysis. Age of onset was analyzed as a continuous variable and a binary trait with cases being defined as having asthma onset <16 years of age (childhood onset asthma). FEV_1 and FEV_1/FVC analyses were corrected for age, gender and height. Eosinophils, neutrophils, IgE and atopy were corrected for age and gender and the age of asthma onset analysis were corrected for gender. DAG/GASP were meta-analysed, which was performed in METAL(12). Associations at an adjusted p-value<0.05 (FDR) were considered statistically significant.

Summary of Phenotye-Genotype association study results:

In Lifelines, we found an (FDR-) significant association of *IL33* SNPs with blood eosinophil counts in the general population, with eosinophilic asthma (vs healthy controls) and with asthma (vs healthy controls) as can be found in table 1 (main text) and supplemental tables S7-S9, S11-S13. In Lifelines, we did not find a significant association *IL33* SNPs with blood eosinophil counts within asthma patients, nor with eosinophilic asthma vs non-eosinophilic asthma as can be seen in supplemental tables S7-S9 (FDR>0.05). Neither we found an (FDR) significant association with blood neutrophil counts, FEV1 or FEV1/FVC in Lifelines general population or within the Lifelines asthma population (FDR>0.05).

In the asthma cohort DAG/GASP meta-analysis, *IL33* SNPs were significantly associated with blood eosinophil counts (table 1 main text), FEV1/FVC and age of asthma onset (supplemental table 15). Within DAG/GASP, no (FDR-)significant association of *IL33* SNPs with blood neutrophil counts, FEV1, total IgE levels, and atopy were found.

M2- Cohort descriptions and details QTL analyses

M2.1 Sample collection

Lung tissue and bronchial brushes

Lung tissue samples for mRNA expression analyses had been collected previously(13) from patients who had undergone lung resection for lung cancer (non-tumorous part, N=1,022), were lung transplantation donors (disapproved lung, N=89) or were lung transplantation recipients (operated lung, N=313). Samples had been collected at three centers, as part of a collaborative effort at the University Medical Center Groningen (UMCG, Groningen), the University of British Columbia (UBC, Vancouver) and the Laval University (Laval, Quebec). The detailed study design and collection procedure has been published before(13), and a summary of the patient characteristics from the included subjects can be found in supplemental table 4.

Bronchial epithelium from brushings (Cellebrity brush Boston Scientific, Massachusetts, USA) had been collected at the UMCG for mRNA expression analyses, deriving from N=139 healthy volunteers. Further patient characteristics are presented in supplemental table 4. From both cohorts, patients were excluded who lacked information on their pre-operative lung function, smoking status, comorbidities, drug use, age, and/or gender.

Bronchial epithelial cells cultured in vitro

Passage 2/3 human bronchial epithelial cells (n= 35) obtained from bronchial brushes and biopsies from asthma patients (referred to as AHBECs) as previously described(14) were cultured on PureCol Type-I Bovine collagen (Advanced BioMatrix, 5005-B) in fresh growth factor-supplemented medium (BEGM, Lonza) until 90% confluence. Protein and RNA lysates were collected as previously described(15) and IL33 levels compared in a genotype-stratified way.

M2.2 mRNA/protein expression assays

Lung tissue/bronchial brushes-mRNA

Expression levels of *IL33* mRNA in the lung tissue samples had been determined previously(13) as part of a general gene expression profile using a customized mRNA array (Affymetrix US Ltd., GEO platform GPL10379). The mRNA expression levels in the bronchial brushing samples had been measured on a commercially available array, the Human genome ST v1.0 array (Affymetrix US Ltd), as described before(16).

Bronchial epithelial cells-mRNA and protein

HBEC complimentary DNA (cDNA) was synthesised from 1µg RNA using Superscript II (Invitrogen, UK) and random hexamer primers according to the manufacturer's instructions. TaqMan® Quantitative PCR (qPCR) was then utilised to quantify mRNA levels of IL33 and was performed and analysed as previously described(15). IL33 protein in cell supernatants was measured using Luminex assays (supplied by R&D, product code LXSAHM) according to the manufacturer's recommendations using a custom Magnetic Luminex Screening Assay with a Human Premixed Multi-Analyte Kit (R&D systems). Each experimental supernatant was assayed in duplicate.

M2.3 Genotyping

Genotypes of SNPs in the *IL33* region had been determined in DNA from peripheral blood mononuclear cells (PBMCs) or oral swabs.

For the lung tissue cohort genotyping had been done on the Human 1M-Duo BeadChip array (Illumina Inc, San Diego, USA) which were imputed against the 1000G phase 1 reference panel (EUR)(4) using IMPUTE2(2) to increase the coverage of genotypic information.

Subjects from the bronchial brushing cohort had been genotyped on two platforms: the Human CytoSNP 12 and OmniExpress Exome genotyping arrays (both Illumina Inc, San Diego, USA)

For the AHBEC cohort DNA was extracted using the Qiagen QIAamp® DNA Mini and Blood Mini Kit according to the manufacturer's instructions. SNP Genotyping was then carried out using TaqMan® Pre-designed assays.

M2.4 Quality control genotype data

Genotype data was quality checked on array, subject and SNP level:

All genotyping arrays passed a call rate of >98%. Subjects that failed gender confirmation (PLINK)(10), and ethnic inference check (EIGENSTRAT)(3) were excluded. SNPs were excluded if they had a SNP call rate <90%, a minor

allele frequency (MAF) <1%, and deviated from Hardy-Weinberg equilibrium (HWE) P<1.0*E-6. After QC a total of N=1,111 subjects from the lung tissue cohort had reliable genotypes available (out of N=1424), and N=129 from the bronchial epithelium cohort (out of N=139 healthy subjects).

M2.5 Details QTL models

eQTL in lung tissue and bronchial brushes

We tested for expression quantitative trait loci (eQTL) in lung tissue (n=1,111) and bronchial brushes (n=139) using a linear regression model to investigate the association of SNPs and log-transformed *IL33* expression data. Specifically, we employed an additive genetic model with age, gender, smoking status and the PCs explaining >1% of expression variance as covariates using R statistics. Since only 2 independent genetic signals were tested by look-up, a p-value <0.05 was considered statistically significant.

Baseline and inducible eQTL and pQTL in primary bronchial epithelial cells

Passage 2/3 AHBECs (n=35) obtained from bronchial brushes and biopsies from asthma patients as previously described(14), were cultured on PureCol Type-I Bovine collagen (Advanced BioMatrix, 5005-B) in growth factor-supplemented medium (BEGM, Lonza). These were studied for baseline QTL and inducible QTL.

For the inducible QTL, cells were stimulated with either house dust mite (HDM) or rhinovirus (RV16, MOI=1). BEGM was changed to basal medium 24 hours prior to stimulation with HDM (50µg/ml) (Greer XPB70D3A25 (Lot: 23187)), or PBS as a vehicle control. For stimulation with RV16 (Public Health England), BEGM was changed to infection medium (BEGM-I), i.e. BEGM lacking Bovine Pituitary Extract, 24 hours prior to infection with RV-16 virus MOI of 1. Cells were infected for a period of 1 hour, following which they were washed three times with sterile PBS and fresh BEGM-I was added. Cells were then incubated for 24 hours. Protein and RNA lysates were collected as previously described(15). Cells were stratified based on the genotypes of the two genetic signals and expression compared using the non-parametric Kruskall-Wallis test. A p-value<0.05 was considered statistically significant.

M3- Functional cell work

Lentiviral overexpression in human bronchial epithelial cells

In order to investigate the functional consequences of sustained *IL33* in asthma, we stably overexpressed human full-length *IL33* (aa1-270) in primary human bronchial epithelial cells (HBECs) isolated from n=5 healthy individuals (Lonza, #CC-2540). This was done using a three plasmid lentiviral system (as described before(17)). This consisted of a pCMV_VSV-G envelope plasmid (CellBiolabs RV110, Addgene plasmid # 8454)(18), a packaging plasmid pCMV_8.91 (Addgene plasmid #2221)(19) and the actual lentiviral overexpression plasmid (pCDH-CMV-MCS-EF1-copGFP, System Biosciences, #CD511B-1) wherein the full length human *IL33* sequence was ligated at the multiple cloning site under a CMV-promotor. A copGFP reporter gene was used under the EF1-promotor in the same overexpression cassette enabling a check of successful transfection/transduction. The *IL33* sequence was commercially derived from OriGene (#SC100114) and comprised the common *IL33* sequence based on the human CEU/Hg37 reference genome (transcript variant 1, NM_033439). The 3 plasmids were first purified and transfected in the vector cell line HEK-293 in a ratio of 8:7:1 (overexpression: packaging:envelope plasmid, ratio based on weight) to create lentiviral particles. The following experimental groups of lentiviral particles were created: lentivirus with *IL33+copGFP* sequence ('IL33'), lentivirus with *copGFP* sequence only ('EV') and a lentivirus without any modifications ('NV'). Then the HBECs were transduced with the lentiviral particles in a ratio of lentivirus:transduction agent that equalized *copGFP* expression between the *IL33-copGFP* and *copGFP*-only constructs (empirically determined based on level of copGFP mRNA expression, 1:5 for the *IL33-copGFP* and 1:2 for the *copGFP* only construct, see figure E12. The transduction agent was Polybrene (Sigma H9268-10G, used at 2ug/mL). *copGFP* expression and *IL33* overexpression was verified on mRNA level (qPCR) and IL33 expression on protein was determined using immunofluorescence; see figures E13 and E14.

IL33/copGFP PCR

Taqman qPCR was used to quantify copGFP and IL33 expression in the overexpression work in human bronchial epithelial cells. Total RNA was isolated from the cells and cDNA was synthesised from 1µg RNA using Superscript II (Invitrogen, UK) and random hexamer primers according to the manufacturer's instructions as also described

before(15). Based on the sequence of the copGFP from the overexpression plasmid (SBI #CD511B-1) we developed a probe/primerset that had the same qPCR efficiency as the IL33 qPCR, with the following sequences to detect copGFP: probe (5'-3') 6FAM-CGGCTACGAGAACCCCTTCC-TAMRA; forward primer (5'-3') ATGGGCTACGGCTTCTAC; reverse primer (5'-3') CTCGTACTTCTCGATGCG.
IL33 (Hs04931857_m1) was assayed using a commercially derived PDAR (#4331182, Applied Bioscience) with ROX™ as passive reference dye. HPRT (PDAR #4310890E, Applied Bioscience) and 18S (PDAR # 4310893E Applied Bioscience) were used as housekeeping genes. PCR was run with a 2 minutes step of 50^{o}C, 10 minutes of 95^{o}C and then 40 cycles of 95^{o}C (15s)+60^{o}C (1min). MxPro software was used to analyse data.

IL33 immunofluorescence

Passage 2 human bronchial epithelial cells that were transduced with lentivirus containing the IL33 expression cassette and controls were seeded in four-well chamberslides (polystyrene, tissue-culture treated, non-coated) at 30,000 cells/well and cultured until confluence. Then medium was removed, cells washed twice with PBS and fixed in 4% formaldehyde for 30min at ambient temperature on a rocker. All preparation steps were performed light-protected to safe the endogenous GFP fluorescence. After a PBS wash (twice, 5min), cells were permeabilized for 30min in 0.15%TritonX(Sigma X100) in 1%BSA(Sigma #A-8412)/PBS, washed and blocked overnight at 4^{o}C using 10%NGS (Sigma #G6767) in PBS. Cells were washed twice with PBS and incubated with two different primary antibodies against IL33 overnight at 4C. 1) A polyclonal rabbit IgG anti-IL-33, ProteinTech, O22 cat12372-1-AP, stock 260ug/mL, used 1:100 in 10%NGS/PBS. 2) A monoclonal mouse IgM anti-IL-33 (clone Ag21430) ProteinTech, cat66235-1-Ig, stock 1360ug/mL, used 1:200 in 10%NGS/PBS. Applicable isotype controls (polyclonal rabbit IgG and monoclonal mouse IgM, Invitrogen #10500C and #14-4752-82) were used at the same concentration as the primary antibodies. Wells were washed three times in PBS (5min each) on a rocker and then Rhodamine TRITC-labeled secondary antibodies applied for 1h at 37^{o}C while shaking, in a humidified tray. Being: 1) goat-anti rabbit IgG (Jackson ImmunoResearch laboratories/Stratech 111-025-003, stock 1.5mg/mL) and 2) goat anti-mouse IgG (ProteinTech, #SA00007-1) both used 1:100 in 10%NGS/PBS. Wells were washed three times in PBS (5min each), air-dried and mounted using VectaShield Mount+DAPI (Vector Laboratories, H-1500), a coverslip applied, dried at ambient temperature and stored light-protected at 4^{o}C until visualization of IL33 and GFP using confocal microscopy (within 48h after mounting).

Functional read-outs of genetically modified HBECs:

We cultured the genetically modified HBECs submerged in bronchial epithelial cell growth medium (BEGM™, Lonza) and performed several functional read-outs:

Cell count, viability and ROS-glutathione assays

Genetically modified and control cells were seeded at 50,000 cells/well in a 6-wells plate format and cultured for 96h. Cells were then harvested by trypsinization, n=2 technical replicates pooled and resuspended in 500uL BEGM. Cell count was performed using a lysis+propidium iodide (PI) based assay (PI-Cassette™) according to the manufacturers' instruction (Application note No. 3007. Rev. 1.3, Chemometec) at the NucleoCounter® NC-3000™ system. Viability (fraction of viable cells) was determined using a Hoechst+PI based assay according to the manufacturers' instructions (application note No. 3023. Rev. 1.4, Chemometec). ROS-capturing capacity (level of free thiols [reduced glutathione] in cells) was determined using a VitaBright-48™ assay according to the manufacturers' instructions (Application note No. 3005. Rev 1.4) (20).

Metabolic activity assay

Genetically modified and control cells were seeded at 4000 cells/well in 96-wells plates and cultured for 96h in total, a subset harvested every 24h to determine metabolic activity over time. Metabolic activity was measured using a colorimetric MTT assay (Sigma, M5655), according to the manufacturer's instruction. Briefly; medium was removed and replaced by a warm (37^{o}C) MTT suspension (0.5mg/mL in BEGM) which was incubated protected from light for 4h at 37^{o}C. The MTT suspension was carefully removed and the formed crystals dissolved in 200uL isopropanol per well. Optical density was measured at 570nm wavelength, including a correction at 670nm.

Electric cell substrate impedance sensing (ECIS) array

To investigate cell-cell contact and barrier formation of bronchial epithelial cells exposed to sustained IL-33, we cultured the genetically modified and control cells on electrode-containing arrays. Electric cell substrate impedance sensing (ECIS) was performed as previously described in detail by our group(21). As high frequency (reflecting spreading of cells and formation of a monolayer) 32kHz was taken, whilst for the low frequency (reflecting cell-cell-contacts and barrier formation) 400Hz was taken as previously established for primary bronchial epithelial cells(21). Resistance values were normalized against the averaged first hour to correct for technical variation/fluctuations often present at the start. Longitudinally, the area under the curve (AUC) was calculated for timepoints 12-24-36-48-60-72h, as well as a cross-sectional comparison of the resistance was made at 24h and 48h to investigate the influence of IL33 on aspects of barrier formation of the bronchial epithelial cells.

Statistical analyses in vitro *cell work*

Treatment and genotype groups were compared using the non-parametric Kruskall Wallis-test. For the expression, cell count, vitality, viability, metabolic assay results and for the cross sectional analysis of the ECIS, followed by Wilcoxon post hoc statistics. The longitudinal AUCs of the ECIS were compared using a Z-test. A p-value <0.05 was considered statistically significant.

M4- ENCODE annotation of phenotype-associated genetic signals

ENCODE was consulted to annotate potential functional elements to the selected phenotype associated genetic signals, using the integrative data level available via https://www.encodeproject.org/data/annotations/. SNPs with LD >0.3 with the selected tagSNP were included in these analyses. Within the ENCODE setting, GWAS studying SNPs associated with asthma and lung phenotypes and their potential functionality were consulted, as well as SNPs were checked for functionality using RegulomeDB, HaploReg and the Chromatin databases (ChromHMM, Segway).Dataset was last accessed on the 9th August 2019. (9,22,23)

A) Supplemental Tables

Table S1 Lead genetic variants of genomewide association ($P<5x10^{-8}$) with asthma in GWAS/GWAS meta-analyses from 2007-2019

SNP	*Effect allele*	*OR*	*P-value*	*Population*	*Ref*	*First author(s)*	*Journal*	*Year*	*Position*	*Signal*
rs1888909	T	1.12	4.20E-34	Caucasian	(24)	Kristjansson RP	Nat Genet	2019	chr9:6197392	
rs7848215	T	1.16	5.29E-62	Caucasian	(25)	Johansson A	Hum Mol Genet	2019	chr9:6213468	
rs992969	A	1.25	1.4E-11	Caucasian	(26)	Pividori M	Lancet Respir Med	2019	chr9:6209697	
	A	1.18	1.1E-17	Multi-ancestry analysis	(27)	Demenais F	Nat Genet	2018		
rs144829310	T	1.18	8.3E-58	Caucasian	(25)	Johansson A	Hum Mol Genet	2019	chr9:6208030	
	T	1.21	2.3E-20	Caucasian	(6)	Shrine N, Portelli MA, John C	Lancet Respir Med	2019		
	T	1.09	1.2E-35	Caucasian	(28)	Ferreira MA	Nat Genet	2017		**A**
	T	1.17	1.3E-31	Caucasian	(29)	Pickrell JK	Nat Genet	2016		
rs72699186	T	1.26	2.0E-09	Caucasian	(30)	Ferreira MA	J Allergy Clin Immunol	2013	chr9:6175855	
rs928413	G	1.50	4.2E-13	Caucasian	(31)	Bonnelykke K	Nat Genet	2013	chr9:6213387	
rs1342326	C	1.20	3.5E-14	Caucasian	(32)	Ferreira MA	Lancet	2011	chr9:6190076	
	C	1.20	9.2E-10	Caucasian	(33)	Moffatt MF	N Engl J Med	2010		
rs2381416	C	1.18	1.7E-12	Multi-ancestry analysis	(34)	Torgerson DG	Nat Genet	2011	chr9:6193455	
rs2066362	T	1.21	1.39E-08	Caucasian	(33)	Moffatt MF	N Engl J Med	2010	chr9:6219176	
rs343478	G	1.06	4.5E-13	Caucasian	(25)	Johansson A	Hum Mol Genet	2019	chr9:6051399	**E**
	G	1.03	2.6E-10	Caucasian	(28)	Ferreira MA	Nat Genet	2017		

In this table an overview is given of the genetic variants associated with asthma discovered at genomewide significant in GWAS and GWAS meta-analyses form 2007-2019. The last column indicates how these variants related to the genetic signals defined in our manuscript. *OR=Odds ratio, Ref=literature reference.*

382 Table S2-Population characteristics of the Lifelines cohort(1)

Characteristics	*General Population (N=13,395)*	*Asthma Population (N=1,066)*	*Stats* (compared to rest of GP)	*Healthy Control (N=6,863)*	*Stats* (compared to rest of GP)
Age (y). mean (SD)	48.1 (11.4)	46.2 (10.9)	P<0.001 (MWU)	49.2 (11.9)	P<0.001 (MWU)
Gender (N. %male)	5,598 (41.8%)	428 (40.2%)	P=0.233 (Chisq)	3213 (46.8%)	P<0.001
Height (cm). mean (SD)	174.4 (9.2)	173.9 (9.6)	P=0.018 (MWU)	175.0 (9.3)	P=0.012 (MWU)
BMI (kg/m^2) mean (SD)	26.4 (4.3)	27.2 (5.0)	P<0.001 (MWU)	26.3 (4.1)	P=0.875
Ethnicity (N. %Caucasian)	11,615 (99.4%)	895 (98.8%)	P=0.018 (Chisq)	5,906 (99.6%)	P=0.351
FEV1 (L). mean (SD)	3.4 (0.8)	3.1 (0.8)	P<0.001 (MWU)	3.4 (0.8)	P<0.001 (MWU)
FEV1/FVC. mean (SD)	0.76 (0.07)	0.73 (0.09)	P<0.001 (MWU)	0.77 (0.07)	P=0.001 (MWU)
Blood eos (10^{E9}/L). median (IQR)	0.16 (0.10-0.23)	0.20 (0.13-0.30)	P<0.001 (MWU)	0.15 (0.10-0.22)	P<0.001 (MWU)
Blood neutro (10^{E9}/L). median (IQR)	3.18 (2.55-3.97)	3.33 (2.66-4.18)	P<0.001 (MWU)	3.13 (2.51-3.92)	P<0.001 (MWU)
Asthma- N (%)	1,066 (8.7%)	1,066 (100%)		-	
Eosinophilic Asthma- N (%)	707 (5.5%)	707 (68.6%)		-	
Low FEV1%pred Asthma- N (%)	258 (1.9%)	258 (24.2%)		-	
Low FEV1/FVC Asthma- N (%)	324 (2.5%)	324 (30.4%)		-	
No asthma/COPD/Allergy- N (%)	6,863 (51.2%)	-		6,863 (100%)	

383 *Chisq= chi squared test, GP=general population, IQR+ inter quantile range, MWU= Mann Whitney-U test, N = number of*
384 *subjects data field available for, %pred= percentage predicted, SD: Standard of Deviation, Stats= statistical comparison*
385 *subgroup to rest of general population*
386
387
388
389

390 Table S3-Population characteristics independent asthma cohorts(6,7)

Characteristics	*DAG (N=909)*	*N*	*GASP (N=2,536)*	*N*
Age (y). mean (SD)	34.78 (15.80)	909	47.83 (15.51)	2,285
Gender. Male (%)	46.9	909	36.0	2,534
Height (m). mean (SD)	1.68 (0.16)	905	1.65 (0.09)	1,692
FEV_1 (L). mean (SD)	2.81 (0.94)	899	2.26 (0.86)	2,039
FEV_1/FVC. mean (SD)	0.79 (0.10)	262	0.69 (0.14)	1,881
Blood Eosinophils (10-9/L). median (range)	0.23 (0.00-1.90)	769	0.31 (0.00-5.42)	1,018
Total IgE (kU/L). median (range)	378.41 (0.00 – 12400.00)	772	407.47 (1.00 – 5000.00)	1,374
Atopy (%)*	578 (85.4)	677	1,072 (68.5)	1,559
Age of asthma onset (y). mean (SD)	10.07 (10.58)	689	23.20 (17.95)	1,176
Childhood onset asthma# N(%)	520 (75.5)	689	578 (46.2)	1,284
%/N of asthma patients	100%	909	100%	2,536

391 *N = number of subjects data field available for, SD: Standard of Deviation*
392 **Atopy was based on at least one positive response to intracutaneous or skin prick tests (SPT) #Childhood onset asthma defined*
393 *as a diagnosis occurring before the age of 16y.*
394
395
396

Table S4- Population characteristics Next Generation Sequencing cohort(6,8)

Characteristics	*GASP Cohort (cases)*	*Gedling Cohort (controls)*
Age (y), mean (SD)	48 (14.88)	57 (12.64)
Gender, Male (%)	30.6	27.0
Height (m), mean (SD)	1.64 (0.08)	1.66 (0.06)
FEV_1 (L), mean (SD	2.17 (0.84)	2.77 (0.79)
Smoking pack/years	11.82 (20.25)	8.40 (18.61)
Never Smokers (%)	52.0	53.5

Demographics for the sub-cohorts taken from GASP (200 cases)and GEDLING (200 non-asthmatic, non-atopic controls) used in the next-generation sequencing of the chromosome 9 locus. *SD: Standard of Deviation*

Table S5-Population characteristics eQTL cohort lung tissue(13) and bronchial brushes(16)

Characteristics	*Lung tissue (n=1,111)*	*Bronchial brushes (n=139)*
Age (y). mean (SD)	58.5 (13.0)	40.0 (18.0)
Gender (N. %male)	54.4%	34.7%
FEV1 (L), mean (SD	2.70 (0.99)	2.76 (0.87)
FEV1/FVC. mean (SD)	0.71 (0.11)	-
Smoking status %current smoker	24.1%	66.7%
BMI. mean (SD)	-	24.1 (3.4)

Demographics for the lung tissue and bronchial brush cohorts used for the expression quantitative (eQTL) analyses. *SD: Standard of Deviation*

Table S6-Population characteristics of cultured bronchial epithelial cells- asthma cohort (AHBEC)

Characteristics	*AHBEC (N total=35)*	*N*
Age (y), mean (SD)	50 (13.47)	20
Gender, Male (%)	43.5	23
Height (m), mean (SD)	1.71 (0.10)	13
FEV_1 (L), mean (SD)	2.70 (0.95)	25
FEV_1/FVC , mean (SD)	0.69 (0.11)	19
Atopy, number (%)*	7 (58.3)	12

Demographics for the cultured primary bronchial epithelial cells from asthma patients (AHBEC) used for the expression quantitative (eQTL) analyses. *N = number of subjects data field available for, SD: Standard of Deviation*

**Atopy was defined as a positive response to a skin prick test. Data was not available for the full cohort of 51 individuals.*

423 Table S7-Five LD blocks (r^2>0.1) and association results with eosinophilic phenotypes in all cohorts

Tag SNP (genetic signal)	*Location*	*Gene context*	*Pheno risk allele (AF)*	*Alt allele*	*Associated feature(s)*	*Cohort*	*Effect size risk allele (B or OR)*	*SE*	*P.adj (FDR)*
A-rs992969	9:6209697	~6kb 5' of *IL33*	A (0.25)	G	*eos levels in GenPop,*	Lifelines	0.058 (B)	0.009	7.09E-08
					eos asthma vs HC	Lifelines	1.321 (OR)	0.062	4.73E-03
					eos asthma vs non-eos asthma	Lifelines	1.216 (OR)	0.109	0.556
					eos asthma vs non-eos asthma	DAG/GASP	1.078 (OR)	0.161	0.633
					eos levels in asthma subjects	Lifelines	0.042 (B)	0.032	0.714
					eos levels in asthma subjects	DAG/GASP	0.002 (B)	0.014	0.991
B-rs1342327	9:6189874	~25kb 5'of *IL33*	G (0.15)	C	*eos levels in GenPop,*	Lifelines,	0.035 (B)	0.011	0.027
					eos asthma vs HC	Lifelines	1.107 (OR)	0.075	0.587
					eos asthma vs non-eos asthma	Lifelines	1.081 (OR)	0.125	0.845
					eos asthma vs non-eos asthma	DAG/GASP	1.181 (OR)	0.421	0.942
					eos levels in asthma subjects	Lifelines	0.0118 (B)	0.037	0.895
					eos levels in asthma subjects	DAG/GASP	0.057 (B)	0.018	0.039
C-rs74438701	9:6282794	~25kb 3'of *IL33*	T (0.83)	C	*eos levels in GenPop*	Lifelines	0.035 (B)	0.011	0.041
					eos asthma vs HC	Lifelines	1.195 (OR)	0.085	0.219
					eos asthma vs non-eos asthma	Lifelines	1.293 (OR)	0.136	0.556
					eos asthma vs non-eos asthma	DAG/GASP	1.144 (OR)	0.309	0.763
					eos levels in asthma subjects	Lifelines	0.074 (B)	0.041	0.714
					eos levels in asthma subjects	DAG/GASP	0.012 (B)	0.018	0.991
D-rs2282162	9:6534466	intronic of *GLDC*	G (0.56)	A	*eos levels in GenPop*	Lifelines	0.029 (B)	0.008	0.011
					eos asthma vs HC	Lifelines	1.081 (B)	0.058	0.583
					eos asthma vs non-eos asthma	Lifelines	1.140 (OR)	0.100	0.586
					eos asthma vs non-eos asthma	DAG/GASP	1.032 (OR)	0.072	0.846
					eos levels in asthma subjects	Lifelines	0.073 (B)	0.030	0.714
					eos levels in asthma subjects	DAG/GASP	0.004 (B)	0.014	0.991
E-rs4008366	9:6116407	intergenic	T (0.69)	C	*eos levels in GenPop*	Lifelines	0.010 (B)	0.009	0.647

	eos asthma vs HC	Lifelines	1.264 (OR)	0.070	0.045
	eos asthma vs non-eos asthma	Lifelines	1.130 (OR)	0.116	0.691
	eos asthma vs non-eos asthma	DAG/GASP	1.007 (OR)	0.676	0.991
	eos levels in asthma subjects	Lifelines	0.003(B)	0.035	0.968
	eos levels in asthma subjects	DAG/GASP	0.0002 (B)	0.015	0.999

The table shows the results of the association analyses of all eosinophilic phenotypes in each cohort for the 5 LD blocks/signals (r^2>0.1). Eosinophilic asthma was defined as asthma with blood eosinophil count >150cells/uL. **Underlined**: the two genetic signals taken forward in functional assesment in this study. *AF=frequency (EUR 1000G); Alt allele=alternative allele; B=beta; eos=eosinophils/eosinophilic; FDR=false discovery rate value at alpha 0.05; GWAS= genomewide association study; GenPop =general population; HC=healthy control; kb= kilo basepairs; Lit.=literature; OR=odds ratio; P.adj= FDR adjusted p-value; Pheno risk allele=phenotype associated allele; Ref=Reference; SE=standard error.*

Table S8-Lifelines association results with eosinophilic asthma (>150cells/uL and >300cells/uL)

Tag SNP (genetic signal)	*Location*	*Gene context*	*Pheno risk allele (AF)*	*Alt allele*	*Associated feature(s)*	*Cohort*	*Effect size risk allele (OR)*	*SE*	*P.adj (FDR)*
A-rs992969	9:6209697	~6kb 5' of *IL33*	A (0.25)	G	*eos asthma* vs HC*	Lifelines	1.321	0.062	4.73E-03
					*eos asthma** vs HC*	Lifelines	1.330	0.097	8.58E-02
B-rs1342327	9:6189874	~25kb 5'of *IL33*	G (0.15)	C	*eos asthma* vs HC*	Lifelines	1.107	0.075	0.500
					*eos asthma** vs HC*	Lifelines	1.112	0.120	0.999
C-rs74438701	9:6282794	~25kb 3'of *IL33*	T (0.83)	C	*eos asthma* vs HC*	Lifelines	1.195	0.085	0.183
					*eos asthma** vs HC*	Lifelines	1.469	0.144	0.198
D-rs2282162	9:6534466	intronic of *GLDC*	G (0.56)	A	*eos asthma* vs HC*	Lifelines	1.081	0.058	0.495
					*eos asthma** vs HC*	Lifelines	1.304	0.193	0.127
E-rs4008366	9:6116407	intergenic	T (0.69)	C	*eos asthma* vs HC*	Lifelines	1.264	0.070	0.045
					*eos asthma** vs HC*	Lifelines	1.273	0.110	0.076

The table shows the results of the association between the 5 main LD blocks and eosinophilic asthma using two definitions (>150cells/uL(*n=707) and >300cells/uL(**n=260)) in Lifelines. **Underlined**: the two genetic signals taken forward in functional assesment in this study. *AF=frequency (EUR 1000G); Alt allele=alternative allele; B=beta; eos=eosinophils/eosinophilic; FDR=false discovery rate value at alpha 0.05; GWAS= genomewide association study; GenPop =general population; HC=healthy control; kb= kilo basepairs; Lit.=literature; OR=odds ratio; P.adj= FDR adjusted p-value; Pheno risk allele=phenotype associated allele; Ref=Reference; SE=standard error.*

443 Table S9-DAG/GASP association results with eosinophilic asthma (>150cells/uL and >300cells/uL)

Tag SNP (genetic signal)	*Location*	*Gene context*	*Pheno risk allele (AF)*	*Alt allele*	*Associated feature(s)*	*Cohort*	*Effect size risk allele (OR)*	*SE*	*P.adj (FDR)*
A-rs992969	9:62096977	~6kb 5' of *IL33*	A (0.25)	G	*eos asthma* vs non-eos asthma*	DAG/GASP	1.078	0.161	0.633
					*eos asthma** vs non-eos asthma*	DAG/GASP	1.049	0.096	0.680
B-rs1342327	9:61898744	~25kb 5'of *IL33*	G (0.15)	C	*eos asthma* vs non-eos asthma*	DAG/GASP	1.181	0.421	0.942
					*eos asthma** vs non-eos asthma*	DAG/GASP	1.329	0.585	0.893
C-rs74438701	9:62827944	~25kb 3'of *IL33*	T (0.83)	C	*eos asthma* vs non-eos asthma*	DAG/GASP	1.144	0.309	0.763
					*eos asthma** vs non-eos asthma*	DAG/GASP	1.117	0.293	0.789
D-rs2282162	9:65344666	intronic of *GLDC*	G (0.56)	A	*eos asthma* vs non-eos asthma*	DAG/GASP	1.032	0.072	0.846
					*eos asthma** vs non-eos asthma*	DAG/GASP	1.033	0.127	0.931
E-rs4008366	9:61164007	intergenic	T (0.69)	C	*eos asthma* vs non-eos asthma*	DAG/GASP	1.007	0.676	0.991
					*eos asthma** vs non-eos asthma*	DAG/GASP	1.059	0.583	0.802

444

445 The table shows the results of the association between the 5 main LD blocks and eosinophilic asthma using two
446 definitions (>150cells/uL(*n=1,002) and >300cells/uL(**n=493)) in DAG/GASP. **Underlined**: the two genetic
447 signals taken forward in functional assesment in this study. *AF=frequency (EUR 1000G); Alt allele=alternative*
448 *allele; B=beta; eos=eosinophils/eosinophilic; FDR=false discovery rate value at alpha 0.05; GWAS= genomewide*
449 *association study; GenPop =general population; HC=healthy control; kb= kilo basepairs; Lit.=literature;*
450 *OR=odds ratio; P.adj= FDR adjusted p-value; Pheno risk allele=phenotype associated allele; Ref=Reference;*
451 *SE=standard error.*

452
453

Table S10-The two genetic signals with their tagSNPs and proxySNPs used in functional follow-up

Signals	*Pheno risk allele (AF)*	*Alt allele*	*Proxy eQTL brushes*	*Pheno risk allele (AF)*	*Alt allele*	*R^2*	*Proxy eQTL HBECs*	*Pheno risk allele (AF)*	*Alt allele*	*R^2*
Signal A rs992969	*A (0.25)*	*G*	*N/A*	*N/A*	*N/A*		*rs2381416*	*C (0.26)*	*A*	*0.95*
Signal E rs4008366	*T (0.69)*	*C*	*rs693838*	*T (0.69)*	*C*	*1.0*	*rs442246*	*T (0.69)*	*G*	*1.0*

***Table**: Proxies* for the two functionally investigated signals in this manuscript. Proxies were used for the two tagSNPs in the eQTL datasets in case the original tagSNP of the LD block was not genotyped in the eQTL dataset. Criteria proxySNP selection: i) highest R^2 with tagSNP, but minimum R^2=0.5 ii); minimum AF=0.10 (EUR 1000G); iii) should itself still significantly associate with applicable phenotype of the genetic signal. AF= allele frequency (EUR 1000G); N/A: proxy not applicable, since tagSNP present; ND: Not determined, no proxy available at MAF≥0.10; Pheno risk allele: allele at risk for high blood eosinophils, asthma and/or eosinophilic asthma; Alt allele: alternative allele*

**proxies used in the lung tissue eQTL dataset have not been included in this table as no significant eQTLs were present in this dataset for the investigated signals, although all proxies had R^2>0.5 with the tagSNP of each selected signal A/E*

B) Supplemental Figures

Supplemental figure E1- Genetic region studied at *IL33* locus

A region of 400kb +/- IL33 was studied, being chr9: 5,815,786–6,657,983 (GRCh37/hg19):

Supplemental figure E2- tagSNPs of the five phenotype-associated LD blocks

E2-a **R^2 of five tagSNPs**

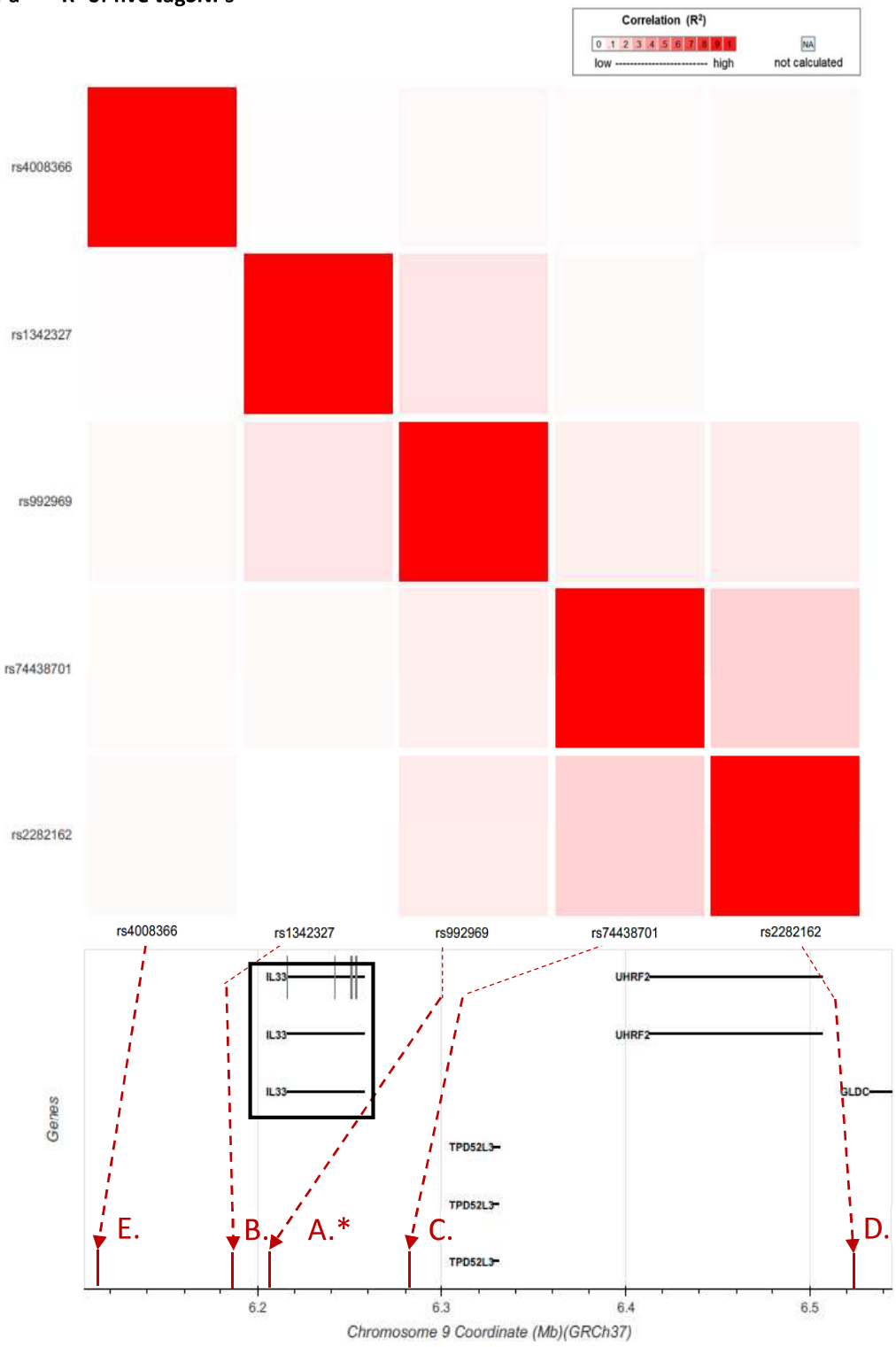


Figure-*The figure shows the LD pattern* ***(R²)*** *of the 5 tagSNPs representing LD blocks that were selected from the (in total 161) phenotype-associated SNPs, LD blocks of each signal defined using R^2>0.1. *Signals with known asthma-association from literature, see also table 1/S1.*

E2-b D’ of five tagSNPs

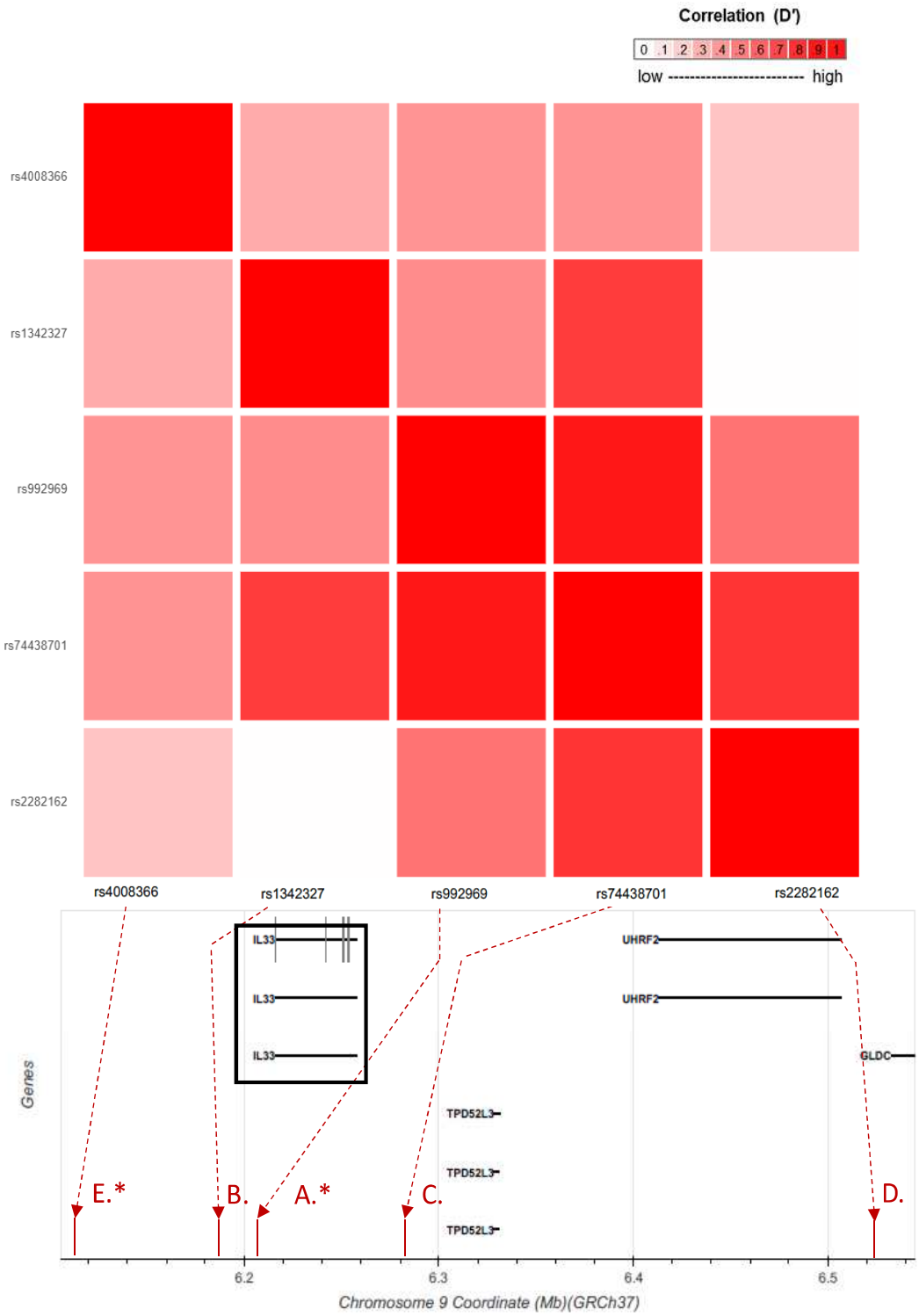


***Figure**-The figure shows the LD pattern **(D’)** of the 5 tagSNPs representing LD blocks that were selected from the (in total 161) phenotype-associated SNPs, LD blocks of each signal defined using R^2>0.1. it can be seen that the D’ is high as opposed to a low R^2 as shown in S2a. *Signals with known asthma-association from literature, see also table 1/S1. Image generated using the EUR population of the Phase I cohort of the 1000 genomes study via the online software tool LDlink 3.0, available at: https://analysistools.nci.nih.gov/LDlink/?tab=home.*

Supplemental figure E3- LD pattern within each of the five LD blocks (A-E)

E3-a LD Block A

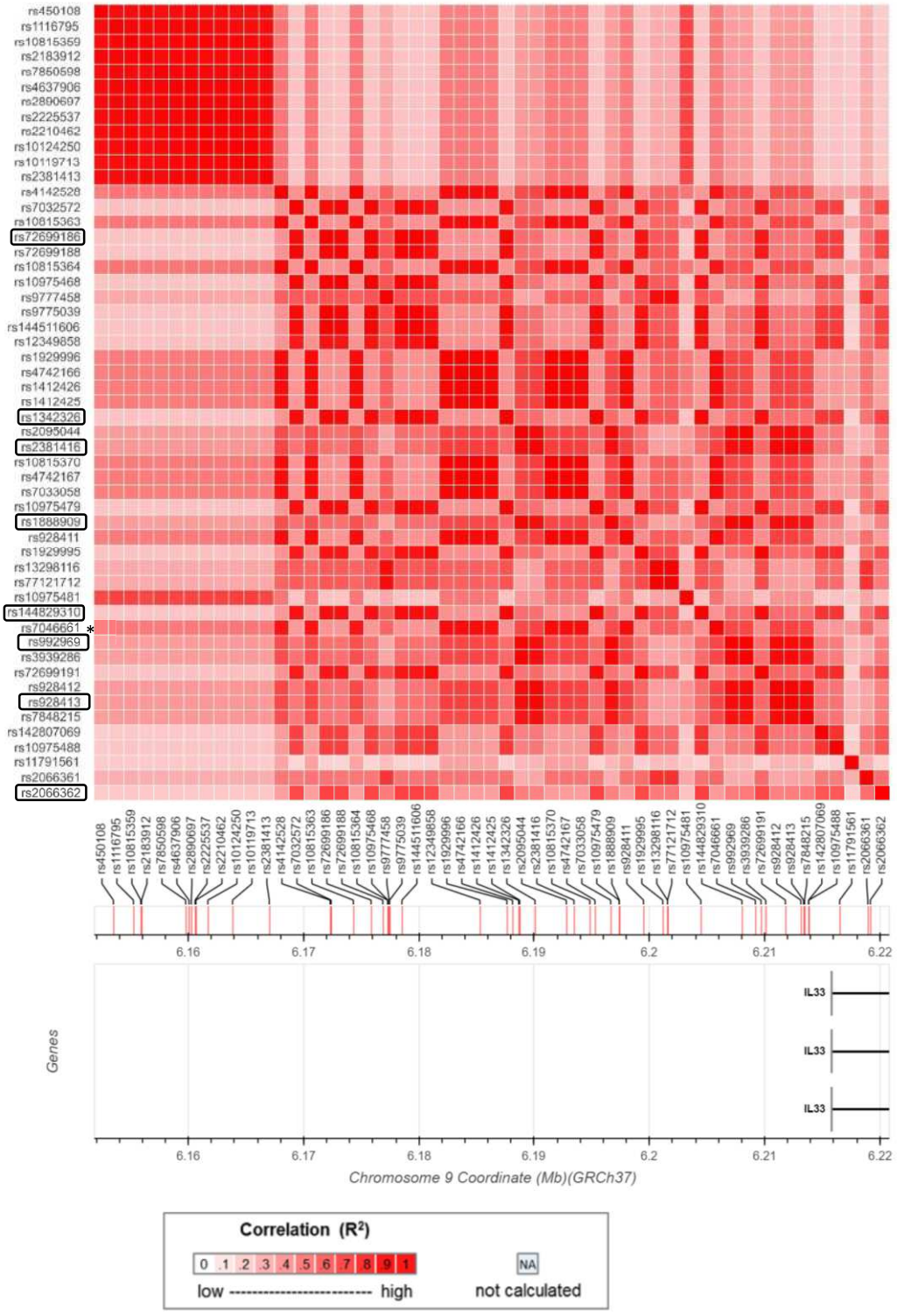


***Figure**: LD pattern (R^2) of the phenotype associated SNPs forming LD block A. Black squared SNP is also a GWAS-asthma SNP from literature. *TagSNP for LD block. Image generated using the EUR population of the Phase I cohort of the 1000 genomes study via the online software tool LDlink 3.0, available at: https://analysistools.nci.nih.gov/LDlink/?tab=home.*

E3-b LD Block B

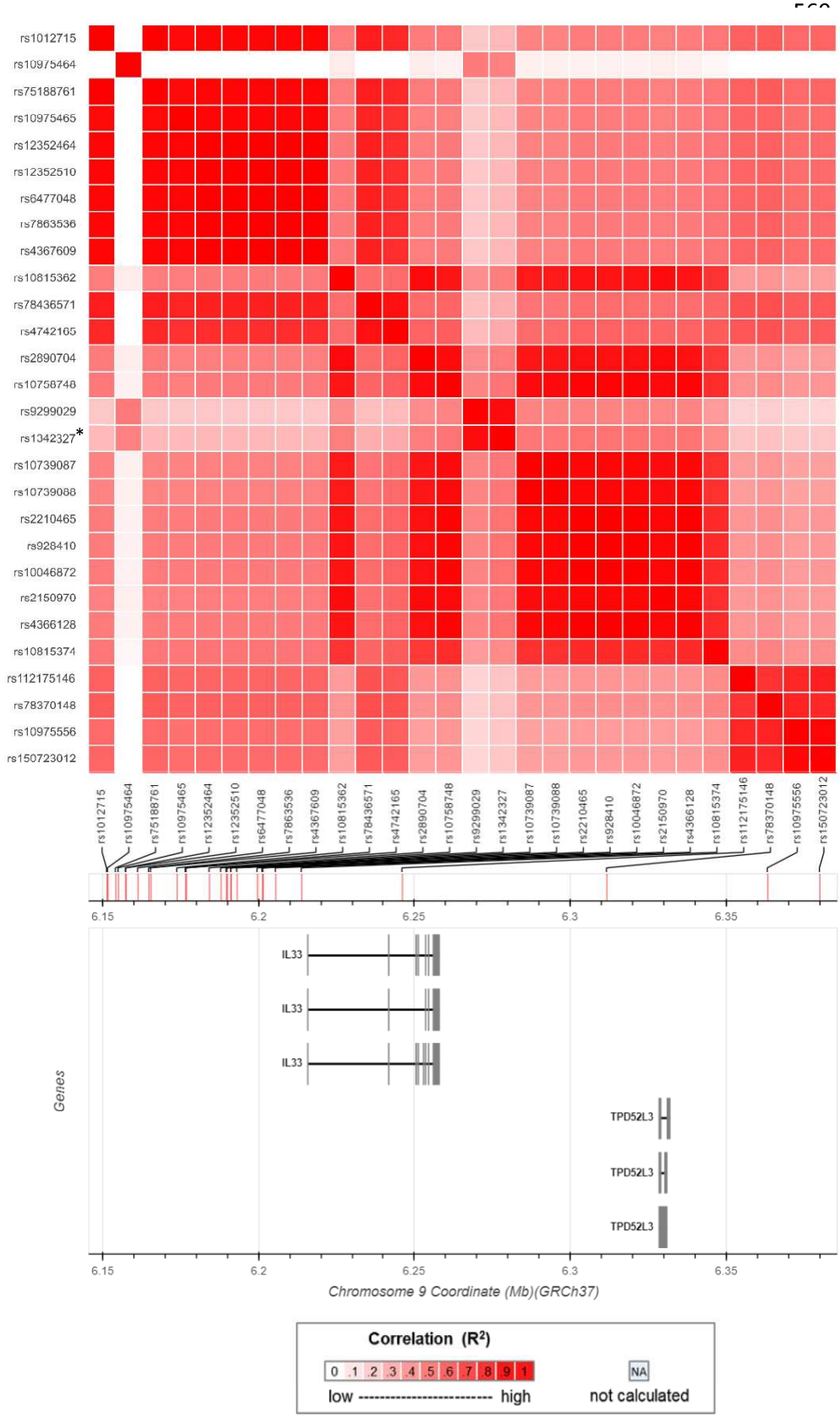


***Figure**: LD pattern (R^2) of the phenotype associated SNPs forming LD block B. *TagSNP for LD block. Image generated using the EUR population of the Phase I cohort of the 1000 genomes study via the online software tool LDlink 3.0, available at: https://analysistools.nci.nih.gov/LDlink/?tab=home.*

E3-c LD Block C

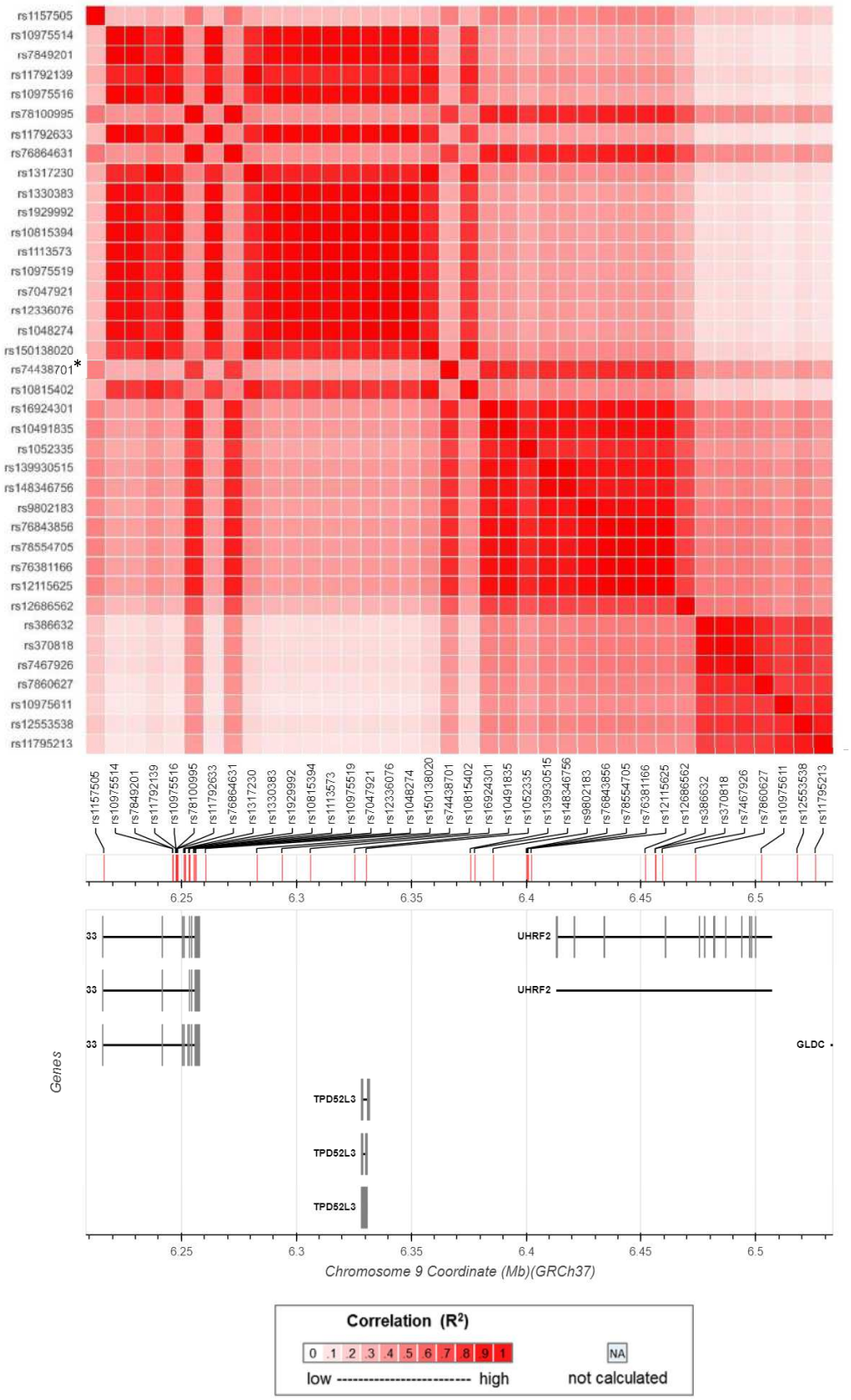


***Figure**: LD pattern (R^2) of the phenotype associated SNPs forming LD block C. *TagSNP for LD block. Image generated using the EUR population of the Phase I cohort of the 1000 genomes study via the online software tool LDlink 3.0, available at: https://analysistools.nci.nih.gov/LDlink/?tab=home.*

E3-d LD Block D

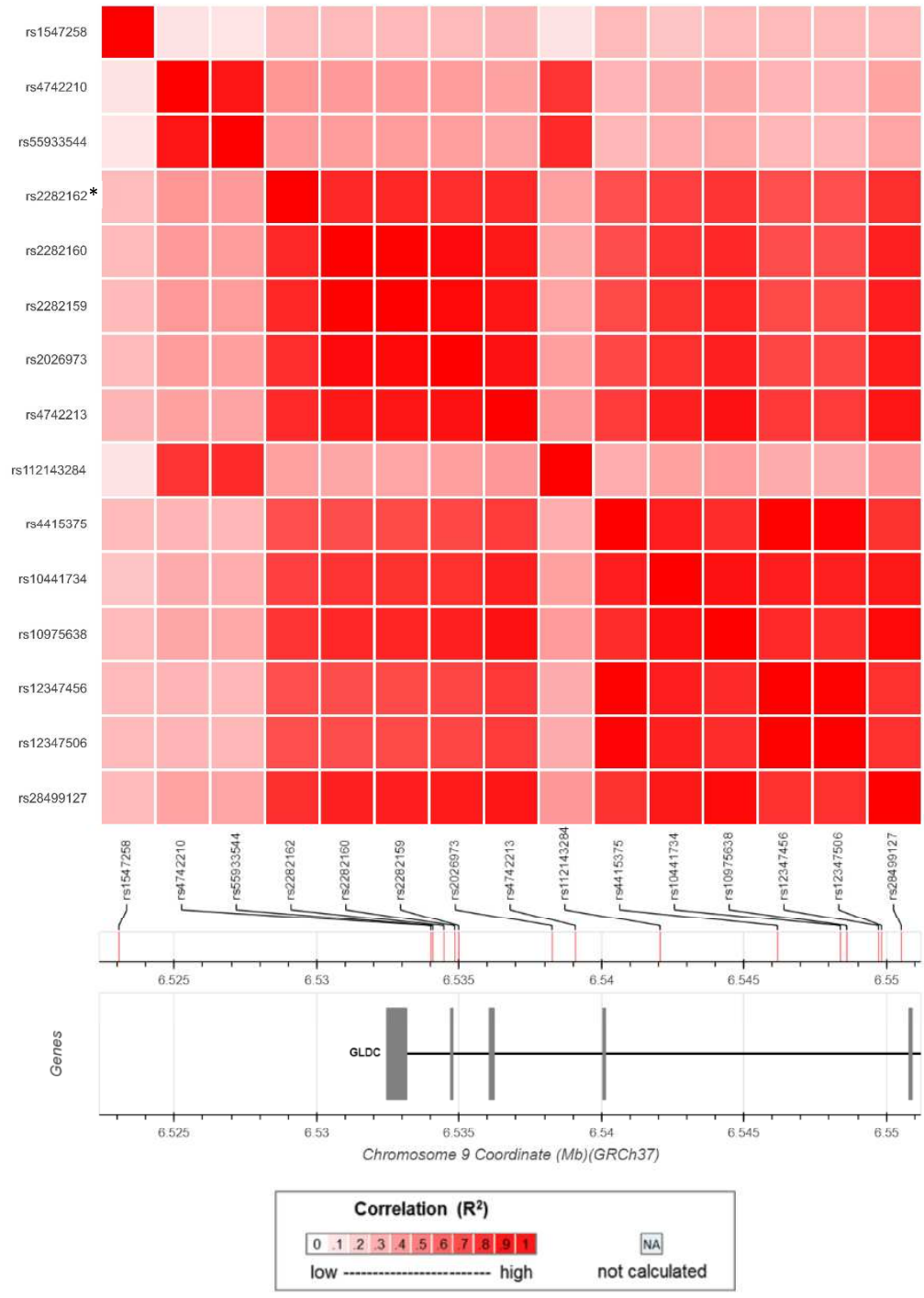


***Figure**: LD pattern (R^2) of the phenotype associated SNPs forming LD block D. *TagSNP for LD block. Image generated using the EUR population of the Phase I cohort of the 1000 genomes study via the online software tool LDlink 3.0, available at: https://analysistools.nci.nih.gov/LDlink/?tab=home.*

E3-e LD Block E

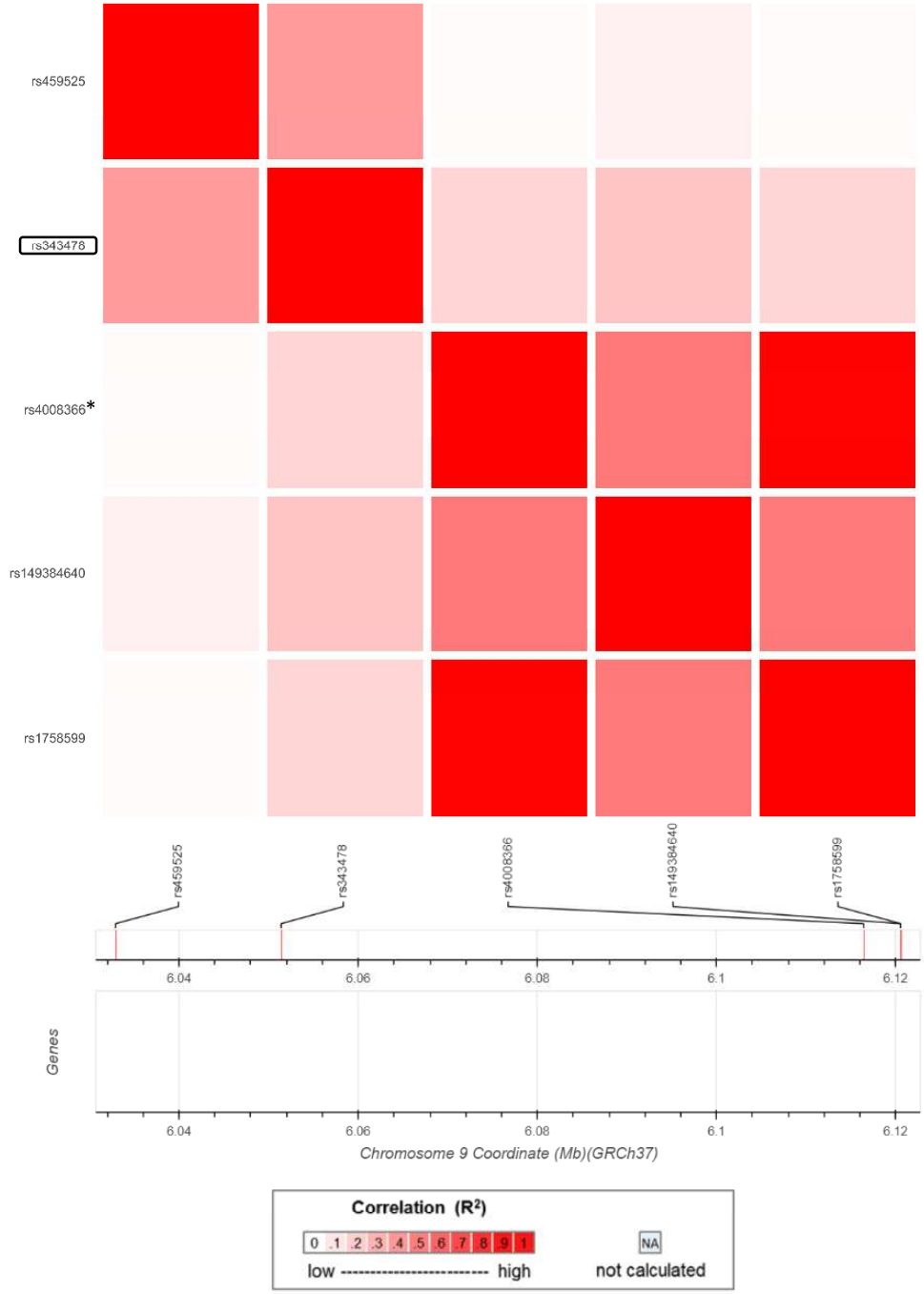
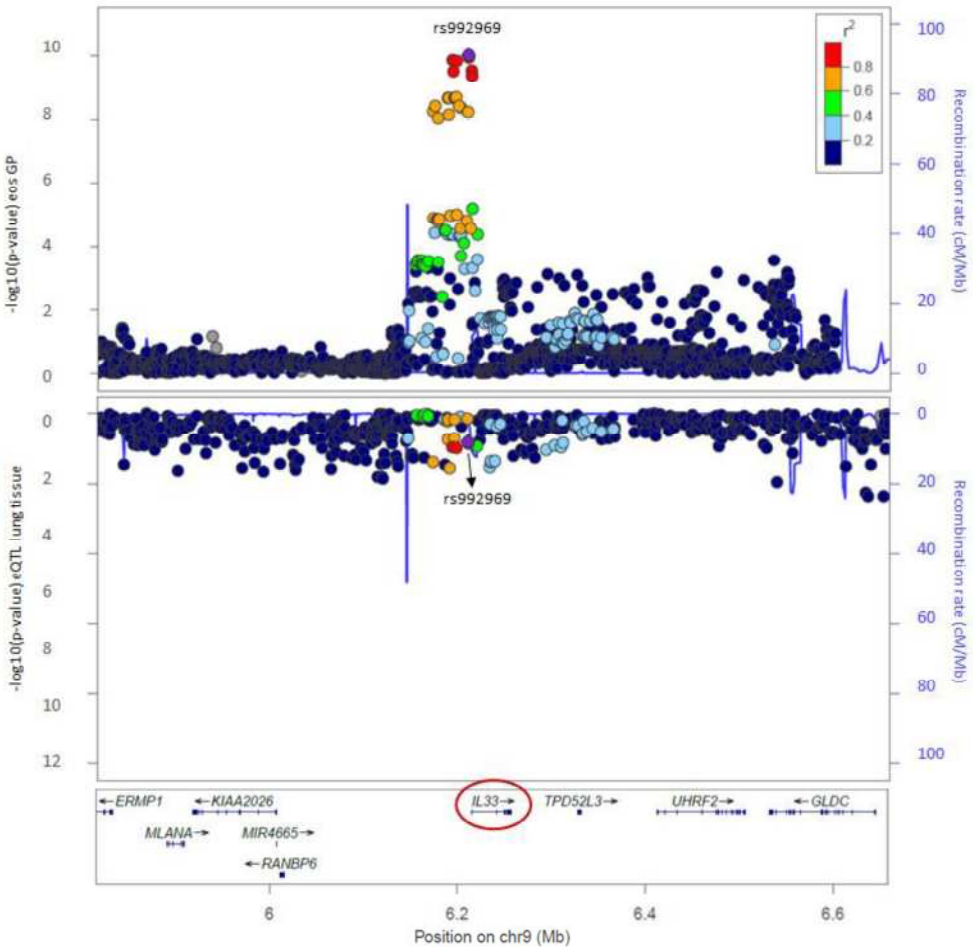


***Figure**: LD pattern (R^2) of the phenotype associated SNPs forming LD block E. Black squared SNP is a GWAS-asthma SNP from literature.*TagSNP for LD block. Image generated using the EUR population of the Phase I cohort of the 1000 genomes study via the online software tool LDlink 3.0, available at: https://analysistools.nci.nih.gov/LDlink/?tab=home.*

Supplemental figure E4- eQTL lung tissue in context of eosinophil associated signals

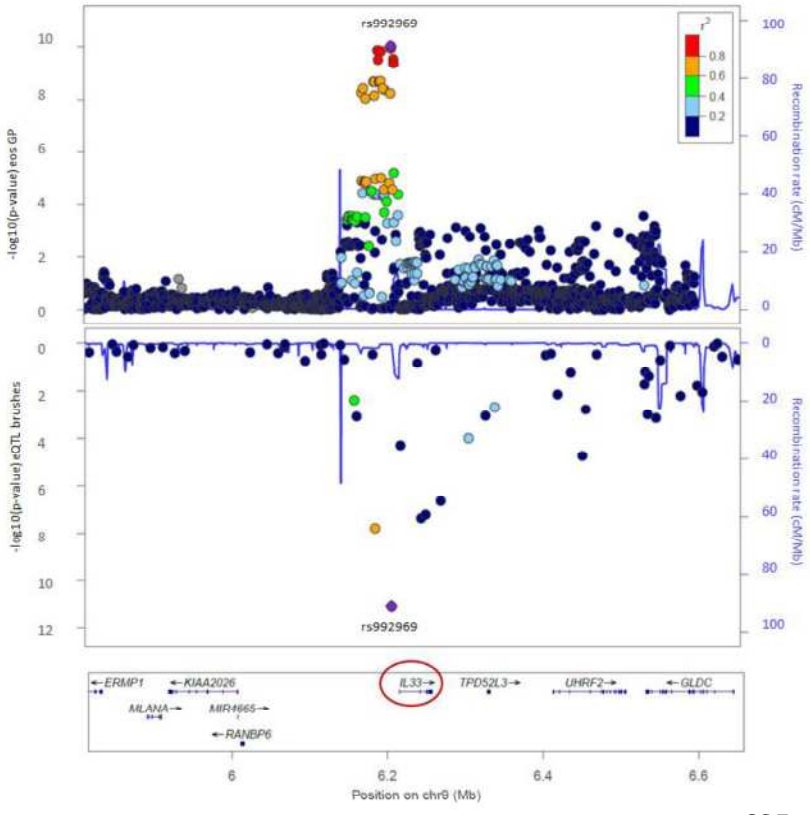


Upper panel: association of IL33 with blood eos GP Lifelines, n=13,395

Lower panel: IL33 eQTL lung tissue, n=1,1111

Plots generated using LocusZoom.(36)

Supplemental figure E5- eQTL bronchial brushes in context of eosinophil associated signals



Upper panel: association of IL33 with blood eos GP Lifelines, n=13,395

Lower panel: IL33 eQTL bronchial brushes n=139

Plots generated using LocusZoom.(36)

Supplemental figure E6- eQTL analyses in AHBECS: *IL33* mRNA levels stratified for *IL33* genotype of the phenotype associated signals.

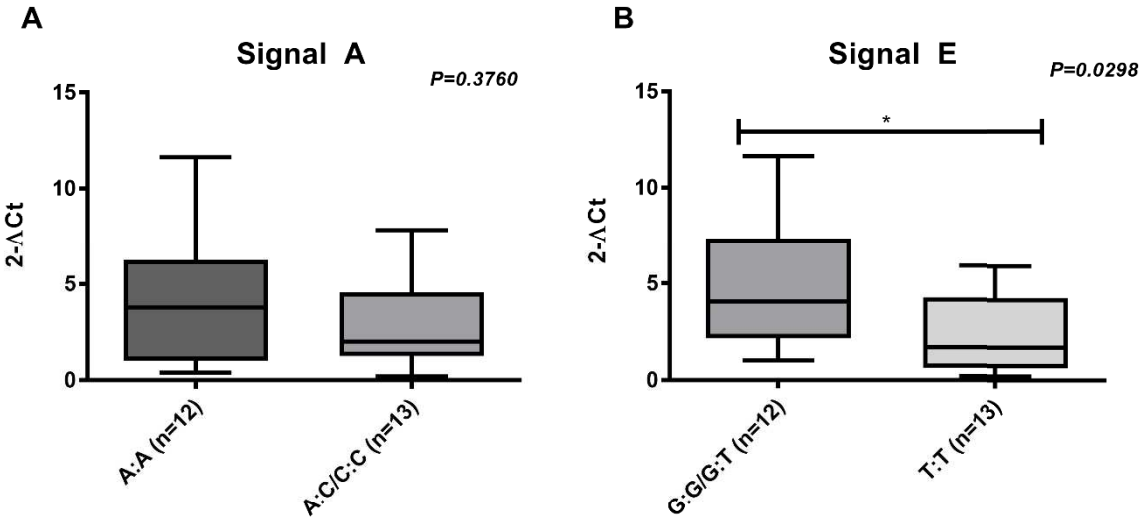


Figure E6- Cells were cultured *in vitro* and *IL33* mRNA levels were stratified based on the selected SNPs tagging distinct genetic signals of association at the *IL33* locus. Panel A represents Signal A tagged by rs2381416, and panel B represents Signal E tagged by rs442246. Statistics run was Mann-Whitney as appropriate for two group comparisons of non-normally distributed data. *P<0.05. Genotypes were grouped in a way to have at least an n=5 per group. Moreover, some variation in total number of subjects studied are present, as not each subject could contribute to all eQTL and pQTL experiments.

Supplemental figure E7- pQTL analyses in AHBECS: IL33 protein levels stratified for *IL33* genotype of the phenotype associated signals

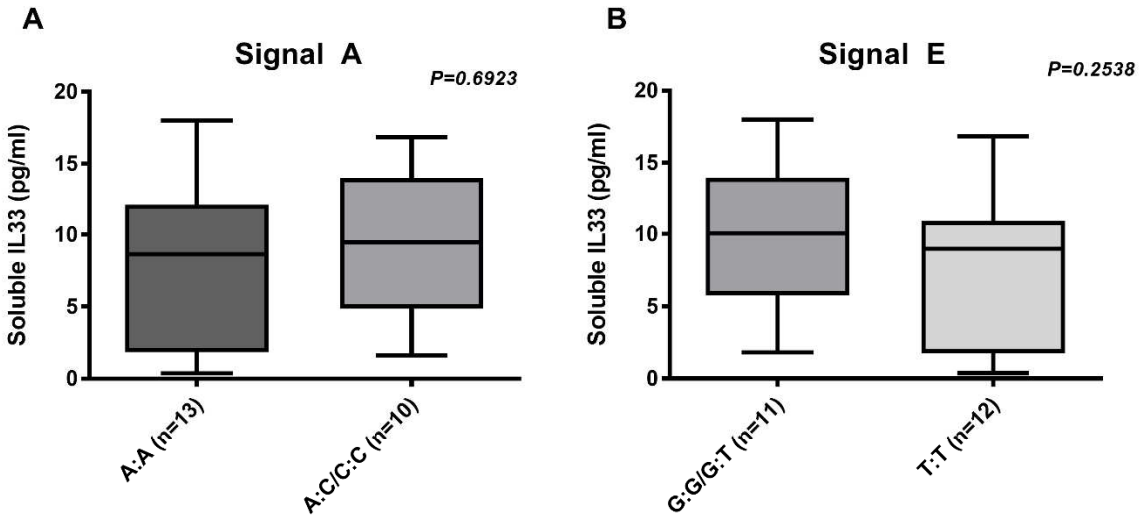
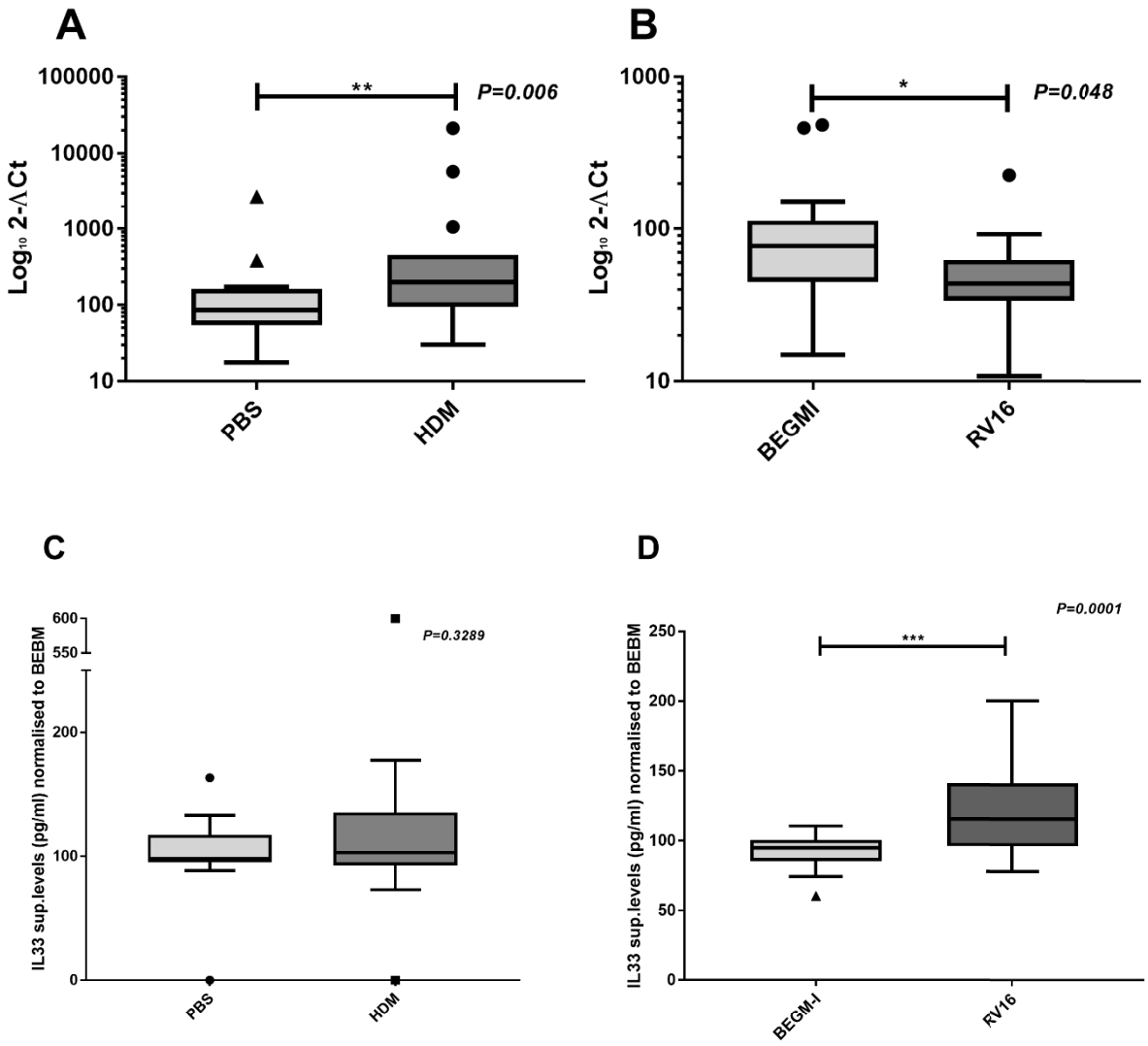


Figure E7: Cells were cultured in vitro and IL33 protein levels in cell supernatants (Luminex) were stratified based on the SNPs tagging distinct genetic signals of association at the *IL33* locus. Panel A represents Signal A tagged by rs2381416, and panel E represents Signal E tagged by rs442246. Statistics run was Mann-Whitney as appropriate for two group comparisons of non-normally distributed data. No statistically significant pQTLs were identified for these signals. Genotypes were grouped in a way to have at least an n=5 per group. Moreover, some variation in total number of subjects studied are present, as not each subject could contribute to all eQTL and pQTL experiments.

Supplemental figure E8- Effects of HDM and RV16 stimulation on IL33 mRNA and IL33 protein levels



Panel A/B: At the mRNA level, HDM stimulation (24h 50ug/mL) resulted in a 1.9-fold increase in *IL33* expression (Panel A, P=0.006, n=15), while stimulation with RV16 (MOI:1) for 24 hours reduced *IL33* mRNA levels 2.1-fold (Panel B, P=0.048, n=15). ***Panel C/D:*** protein level-Stimulation of cells with 50µg/ml HDM for 24 hours did not affect IL33 levels in the cellular supernatant (Panel C, P>0.05, n=18). RV-16 (MOI:1) stimulation for 24 hours however resulted in a 1.3-fold increase of IL33 protein in the cellular supernatant (Panel D, P=0.0001, n=18). Genotypes were grouped in a way to have at least an n=5 per group. Moreover, some variation in total number of subjects studied are present, as not each subject could contribute to all eQTL and pQTL experiments.

Supplemental figure E9- No effect of the phenotype associated signals on HDM and RV16 induced eQTL analyses

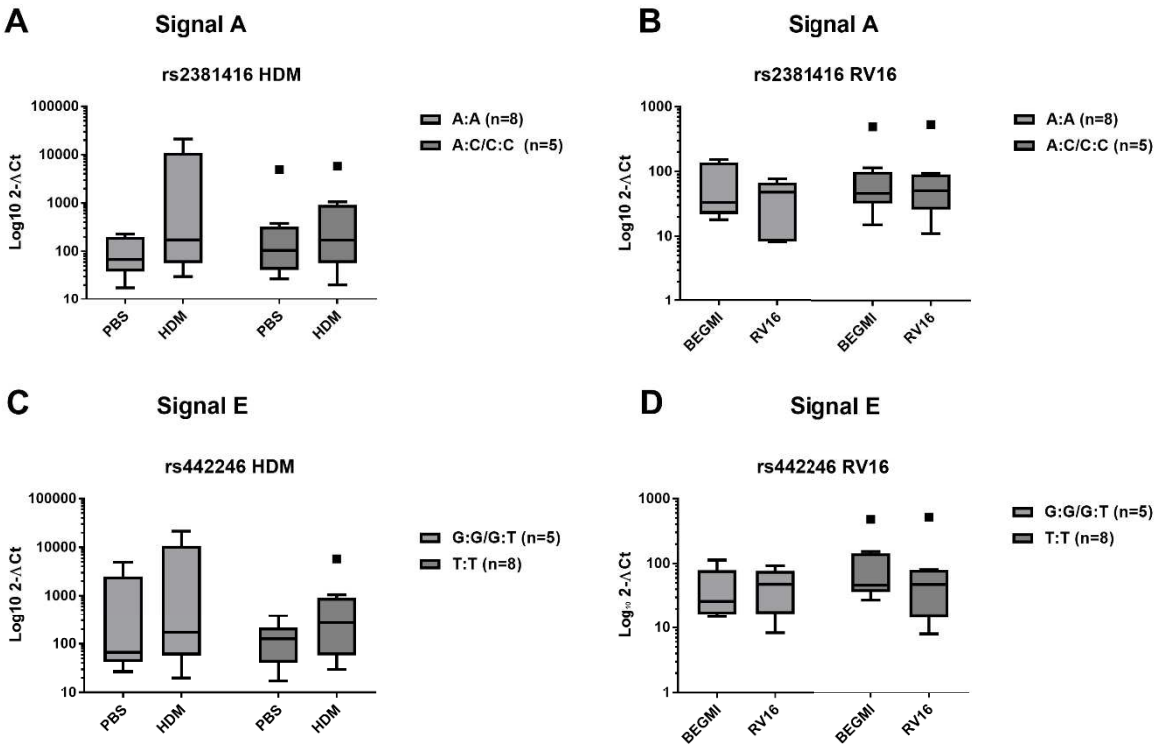


Figure E9: Complete analyses of levels of *IL33* mRNA in bronchial epithelial cells isolated from asthma patients and cultured *in vitro* under different asthma relevant micro-environments then stratified based on selected SNPs tagging genetic signals of association at the *IL33* locus. Each row represents the mRNA levels of a distinct genetic signal stimulated with HDM or RV16. Signal A (rs2381416): panel A (HDM)/ B (RV16). Signal E (rs442246): panel C (HDM)/ D (RV16). Statistics run were either Mann-Whitney or Kruskall Wallis as appropriate for two or three group comparisons. No statistically significant inducible eQTL were identified for these signals. Genotypes were grouped in a way to have at least an n=5 per group. Moreover, some variation in total number of subjects studied are present, as not each subject could contribute to all eQTL and pQTL experiments.

Supplemental figure E10- No effect of the phenotype associated signals on HDM and RV16 induced pQTL analyses

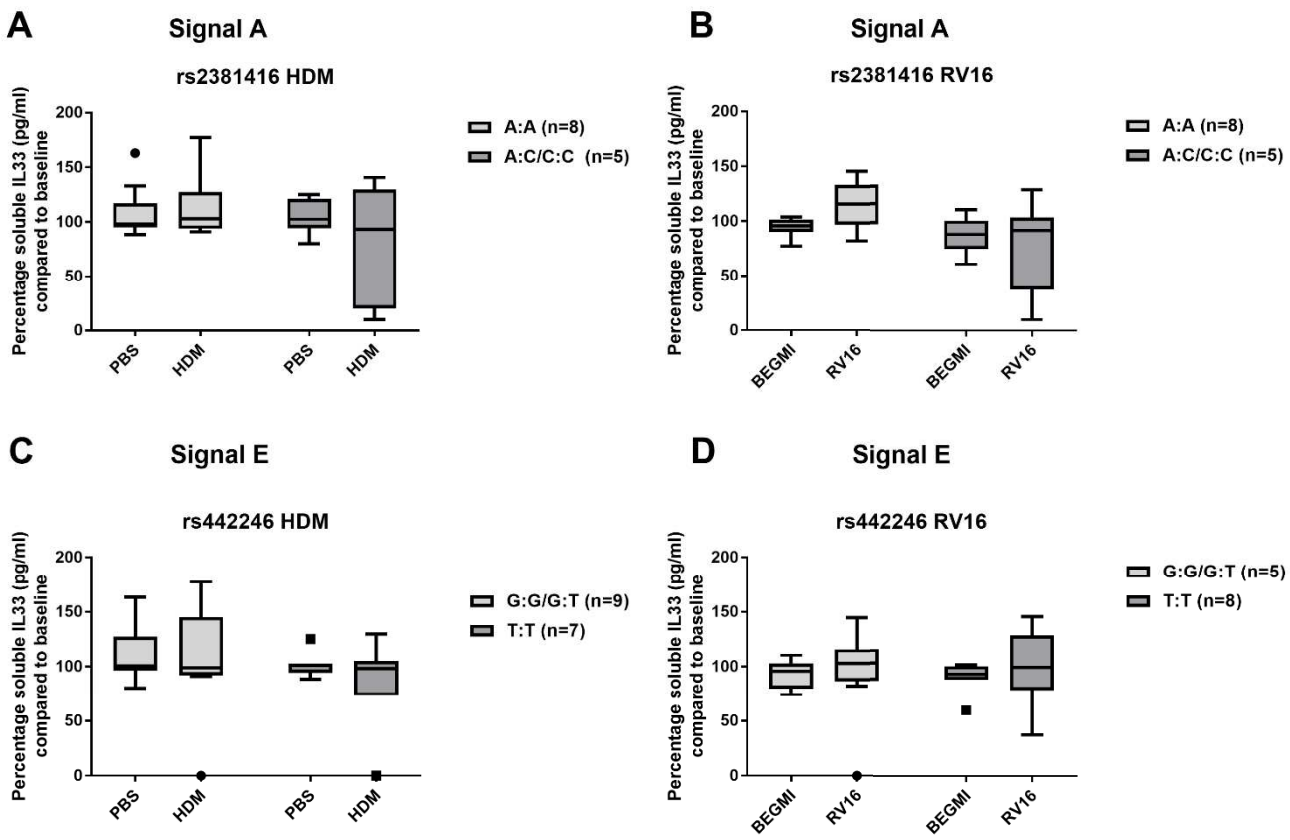


Figure E10: Complete analyses of levels of IL33 protein in supernatants of bronchial epithelial cells isolated from asthma patients and cultured *in vitro* under different asthma relevant micro-environments then stratified based on selected SNPs tagging genetic signals of association at the *IL33* locus. Each row represents the protein levels of a distinct genetic signal stimulated with HDM or RV16. Signal A (rs2381416): panel A (HDM)/ B (RV16). Signal E (rs442246): panel C (HDM) /D (RV16). Statistics run were either Mann-Whitney or Kruskall Wallis as appropriate for two or three group comparisons. No statistically significant inducible pQTL were identified for these signals. Genotypes were grouped in a way to have at least an n=5 per group. Moreover, small variation in total number of subjects studied are present, as not each subject could contribute to all eQTL and pQTL experiments.

Supplemental figure E11- Lentiviral overexpression method

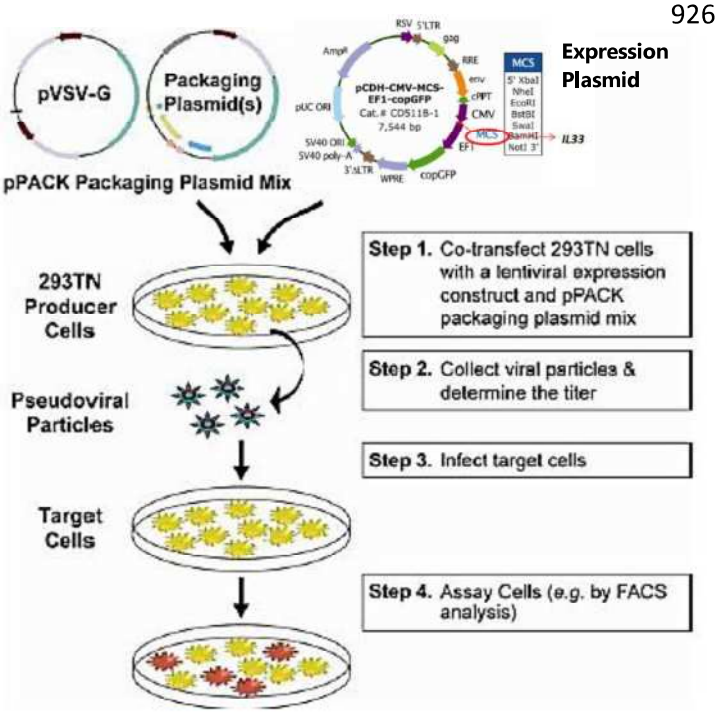


Figure E11- A three plasmid system was used for lentiviral overexpression of *IL33* in primary human bronchial epithelial cells, consisting of an envelope plasmid (pCMV_VSV-G CellBiolabs RV110, Addgene plasmid # 8454), a packaging plasmid (pCMV_8.91 (Addgene plasmid #2221)) and the plasmid containing the actual expression construct (human full-length *IL33* aa1-270, transcript variant 1, NM_033439) in pCDH-CMV-MCS-EF1-copGFP (System Biosciences, #CD511B-1). Controls were: transduction without plasmid (or NV), transduction with expression plasmid lacking *IL33* (empty vector control or EV). For details see supplemental methods above and Torr et al(17). *Figure adapted from SBI handbook 'pCDH cDNA Cloning and Expression Lentivectors CD- 500/800 series'* https://www.systembio.com/wp-content/uploads/Manual_pCDH_Vectors-1.pdf

Supplemental figure E12- GFP expression matched between *IL33* overexpression vector and empty vector

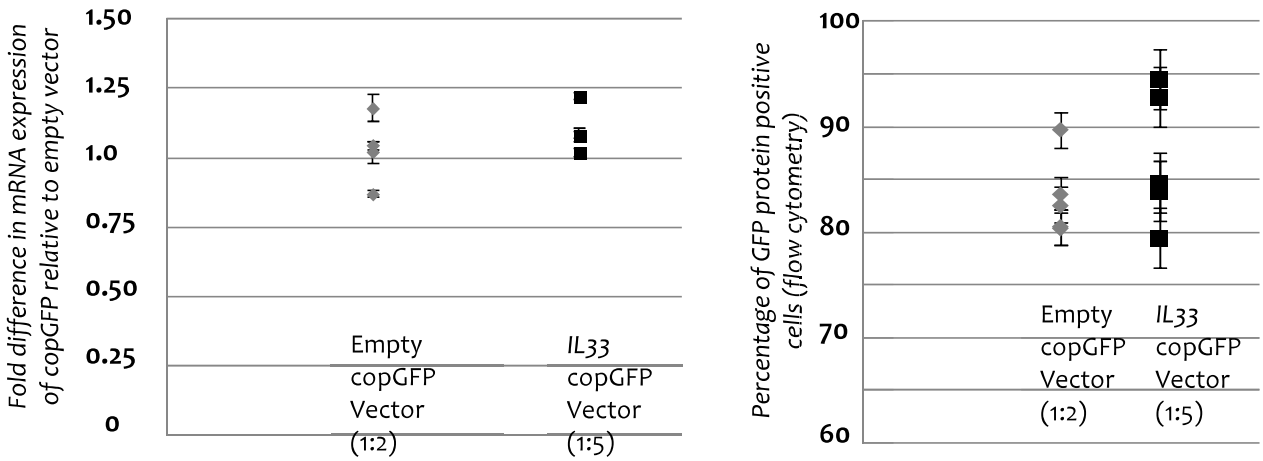


Figure E12- Using qPCR (left) and flow cytometry (right) (see methods), the copGFP expression was matched between the empty vector (GFP only, 'EV', grey squares) and the *IL33* overexpression vector ('IL33', black squares) on mRNA and protein level. This provided confidence that any differences were not due to different infection efficiency. These initial analyses identified infection ratio of lentiviral particles:polybrene of 1:2 for the EV and 1:5 for the IL33. Left (mRNA): data expressed as fold difference in copGFP mRNA levels compared to EV. Right: data expressed as percentage of copGFP positive cells as determined by flow cytometry. N=5 independent HBEC donors, data points represent mean +/-standard deviation for 2 technical replicates per donor.

Supplemental figure E13- *IL33* overexpression results in elevated IL33 mRNA in engineered cells

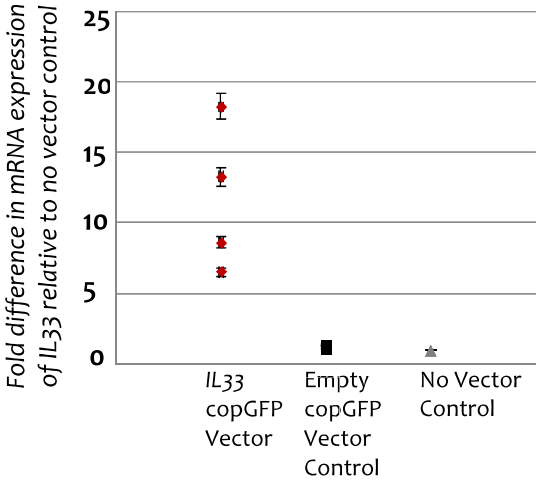
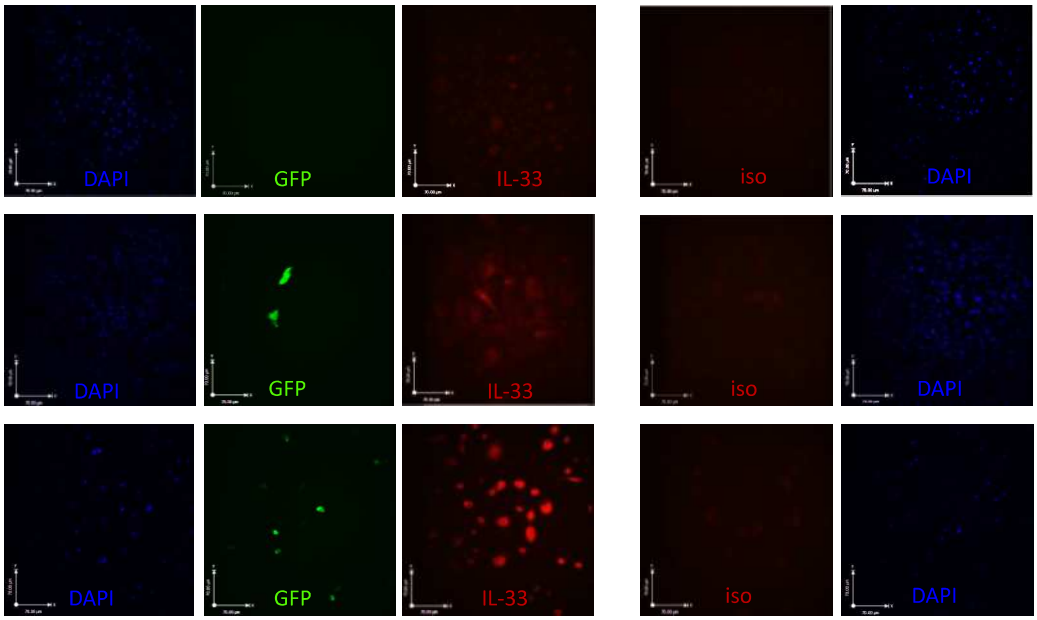


Figure E13- Using qPCR the overexpression of *IL33* was confirmed in the N=5 HBECs derived from healthy donors. Data expressed as fold difference in *IL33* mRNA levels compared to no vector control. N=5 HBEC donors, data points represent mean +/-standard deviation for 2 technical replicates per donor.

Supplemental figure E14- IL33 expression confirmed on protein level using 2 different antibodies

E14-a Polyclonal antibody against IL33 (ProteinTech)



E14-b Monoclonal antibody against IL33 (ProteinTech)

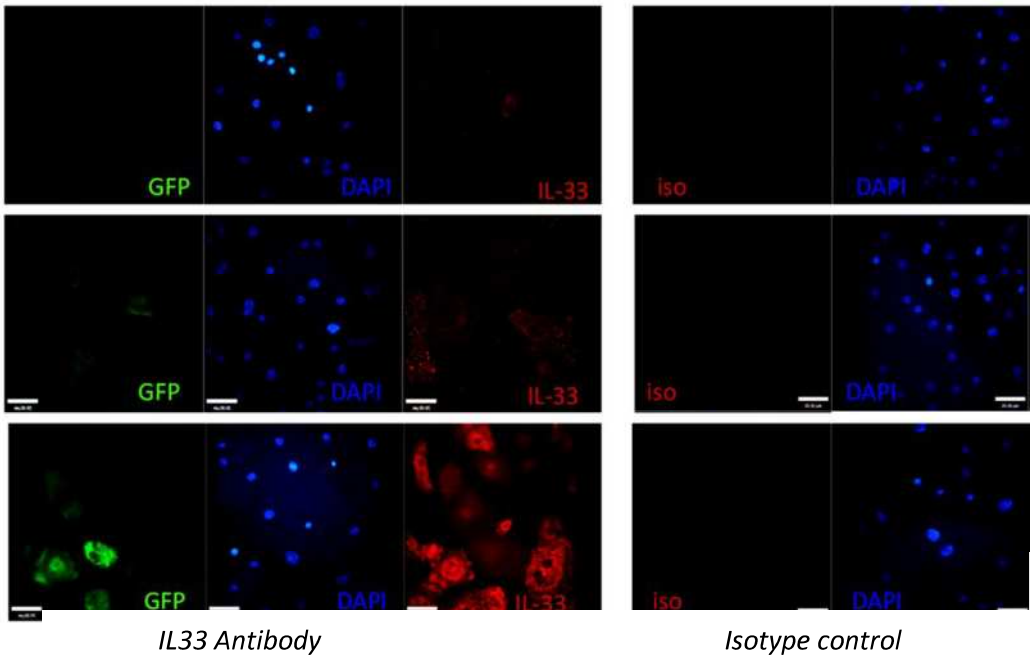


Figure E14- Using immunofluorescence the expression of IL33 on protein level was confirmed in HBECs transduced with lentivirus containing a human *IL33* expression cassette, whilst the exact cellular location was inconclusive considering the staining patterns of 2 different antibodies (red). Cells were processed for immunofluorescent staining at passage 2, two weeks after the lentiviral transduction when cells were considered virus-free.

Supplemental figure E15- Functional assays upon *IL33* overexpression in HBECs

E15-a No effect of *IL33* overexpression on cell number

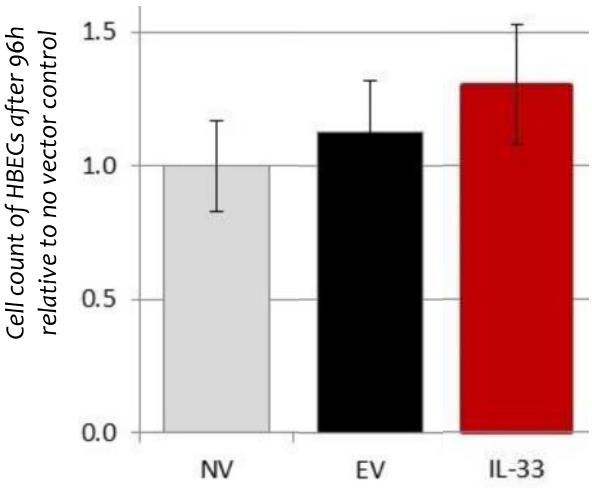


Figure E15a- HBECs overexpressing *IL33* (red) were cultured submerged in 6-wells plates and harvested+counted at 96h when they were in the log-phase of proliferation (empirically determined in pilot experiments) and compared against empty vector (EV) control and no vector (NV) control. Data represent n=3 HBEC donors, mean+ standard deviation, counts expressed relative to NV. Cross-sectional comparison of 3 conditions (MWU, p=0.21).

E15-b No effect of *IL33* overexpression on Metabolic activity (MTT)

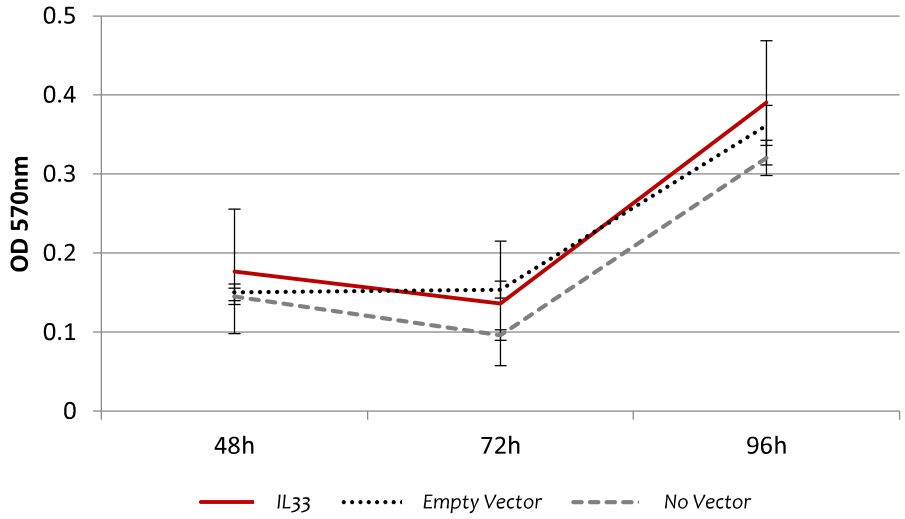
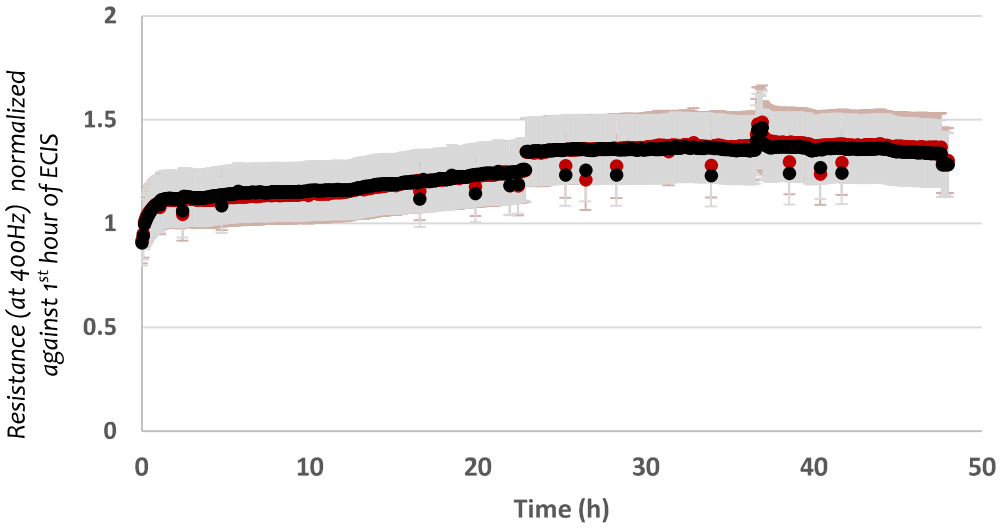


Figure E15b- Using an MTT assay (see methods) the metabolic activity of submerged cultured HBEC was determined upon sustained *IL33* overexpression. Cells were harvested and MTT activity measured every 24h between 48-96h, then cross-sectionally compared using Wilcoxon paired rank test (IL33 vs empty vector control at 48h p=0.65, at 72h p=0.72, at 96h p=0.57). Data represent n=5 HBEC donors per timepoint per condition; mean+/- standard deviation of OD values.

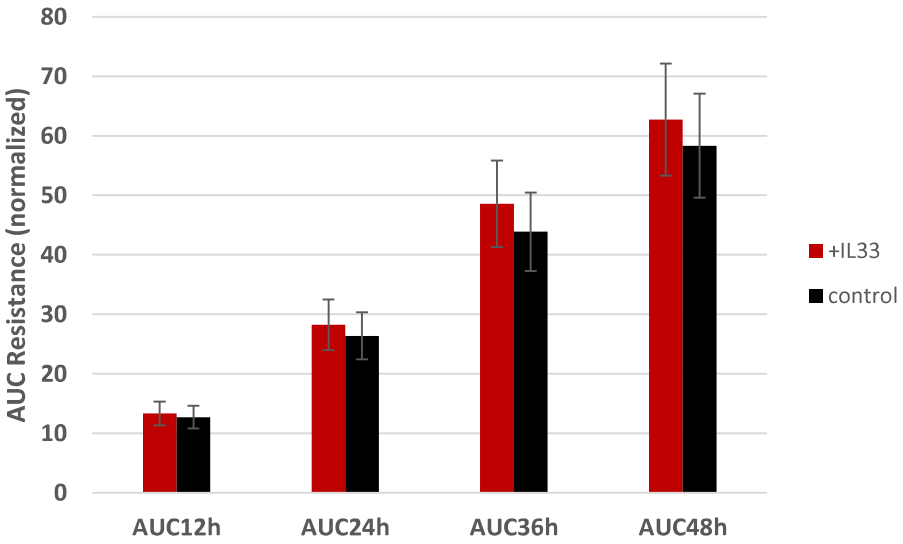
E15-c No effect of *IL33* expression on resistance (ECIS) over time

Figure E15c- HBECs overexpressing *IL33* were cultured to confluency on electrode-containing plates and resistance (at 400Hz) measured using ECIS every 10 minutes for 48h. Due to large donor variation in resistance values from the start onwards, data were normalized against the first hour to enable comparison of treatment groups. Mean +/- standard deviation of n=5 HBEC donors.

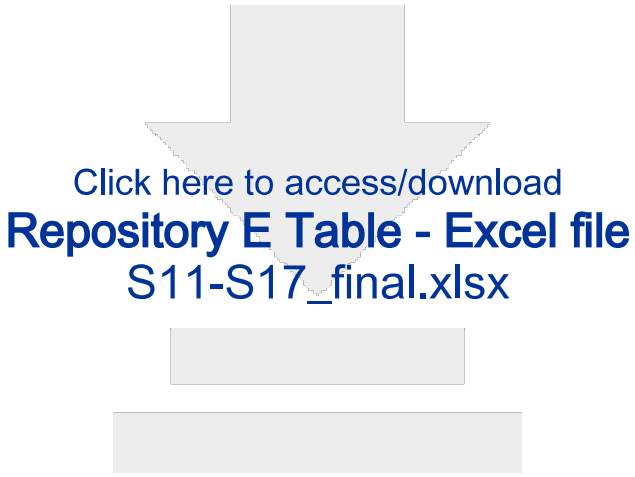


Upper panel: Longitudinal plot. Red dots with light red error bars: *IL33* overexpressing cells, black dots with grey error bars: empty vector (EV) controls.

Lower panel: Area under the curves (AUCs) were calculated, normalized against the average resistance of the first hour, then cross-sectionally compared every 12h. No difference in AUC was seen comparing *IL33* vs EV (12h; p=0.78, 24h; p=0.65, 36h; p=0.42, 48h; p=0.38, MWU). control=EV.



Repository E Table - Excel file

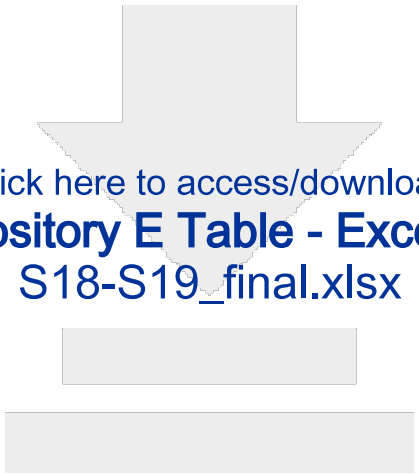


Click here to access/download

Repository E Table - Excel file

S11-S17_final.xlsx

Repository E Table - Excel file



Click here to access/download

Repository E Table - Excel file

S18-S19_final.xlsx