

In the format provided by the authors and unedited.

Multiancestry association study identifies new asthma risk loci that colocalize with immune-cell enhancer marks

Lists of the authors and affiliations appear in the end of the text.

Multi-ancestry association study identifies new asthma risk loci that co-localize with immune cell enhancer marks

Florence Demenais, Patricia Margaritte-Jeannin, Kathleen C Barnes, William OC Cookson *et al.**

*A full list of co-authors names appears in the main paper

Supplementary Information

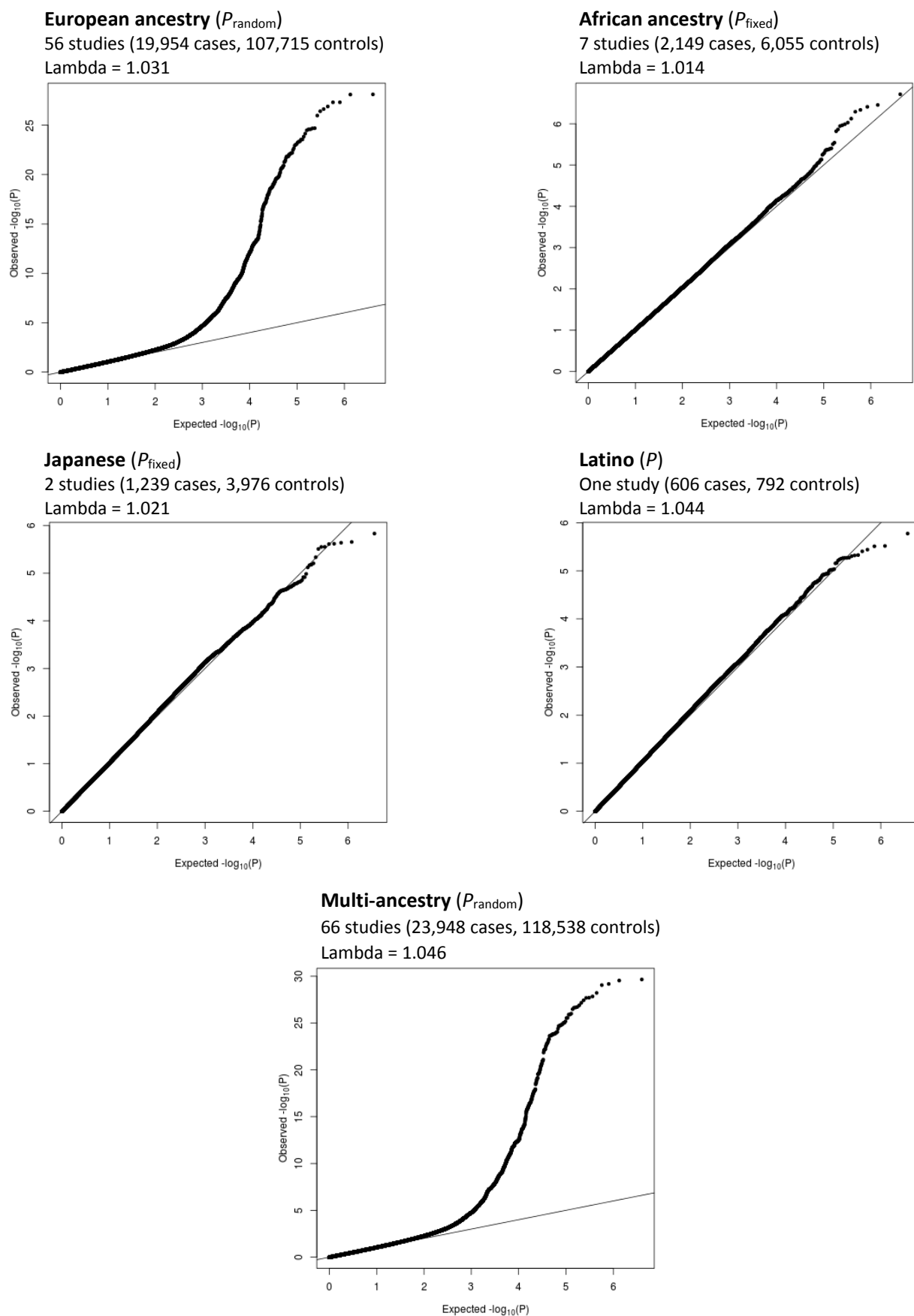
Contents

Supplementary Figures	3
Supplementary Figure 1. Quantile-quantile (QQ) plots for the ancestry-specific and multi-ancestry meta-analyses of asthma	4
Supplementary Figure 2. Manhattan plots of asthma for each meta-analysis in African-ancestry, Japanese and Latino populations	5
Supplementary Figure 3. Regional plots for the 18 asthma loci reaching genome-wide significance	6
Supplementary Figure 4. Forest Plots of the lead SNPs at the 18 genome-wide significant asthma loci	24
Supplementary Figure 5. Quantile-quantile (QQ) plot for the meta-analysis of pediatric asthma	31
Supplementary Figure 6. Manhattan plot for the meta-analysis of pediatric asthma	32
Supplementary Figure 7. Regional plots for the 17q12-21 locus in multi-ancestry and pediatric subgroup meta-analyses	33
Supplementary Tables	34
Supplementary Table 1. Description of TAGC studies included in the meta-analysis	35
Supplementary Table 2. Information on genotyping methods, imputation, and statistical analysis by study	35
Supplementary Table 3. Effect sizes of the lead SNPs in ancestry-specific and multi-ancestry meta-analyses	36
Supplementary Table 4. Genome-wide significant SNPs in the European-ancestry meta-analysis	37
Supplementary Table 5. Genetic loci associated with asthma in the African-ancestry meta-analysis	38
Supplementary Table 6. Genetic loci associated with asthma in the Japanese meta-analysis	39
Supplementary Table 7. Genetic loci associated with asthma in Latinos	39
Supplementary Table 8. Genome-wide significant SNPs in the multi-ancestry meta-analysis	40
Supplementary Table 9. Tests of association of the lead SNPs with asthma with and without Genomic Control (GC) correction	41

Supplementary Table 10. Genetic loci associated with pediatric asthma	42
Supplementary Table 11. Association of 17q12-21 SNPs with asthma in multi-ancestry and pediatric meta-analyses	43
Supplementary Table 12. Association between cis-gene transcripts in blood and lung and 17q12-21 SNPs.....	44
Supplementary Table 13. Asthma loci reported in the GWAS catalog that did not reach genome-wide significance in the TAGC meta-analyses	45
Supplementary Table 14. Association analyses of the 1q21.3 and 22q12.3 SNPs in two sets of TAGC datasets.....	47
Supplementary Table 15. Approximate conditional analysis of asthma-associated loci	48
Supplementary Table 16. Associations between SNPs at the nine novel asthma loci and cis-gene expression in blood and lung	49
Supplementary Table 17. Overlap between TAGC asthma association signals ($P_{\text{random}} < 10^{-4}$) and GWAS signals with specific diseases/traits in the GWAS catalog	50
Supplementary Table 18. Overlap between TAGC asthma signals ($P_{\text{random}} < 10^{-3}$) and GWAS signals with diseases/traits in the GWAS catalog	51
Supplementary Table 19. Enrichment of asthma risk loci in promoter and enhancer marks by cell type.....	51
Supplementary Table 20. Enrichment of TAGC asthma-associated loci in DNase I hypersensitive sites (DHSs) by cell type.....	52
Supplementary Table 21. Results of GRAIL analysis for the asthma risk loci.....	53
 Supplementary Note	 54
Study Descriptions	55
Funding and Acknowledgements	80
Contributing groups for TAGC and by study	87
 Supplementary References	 91

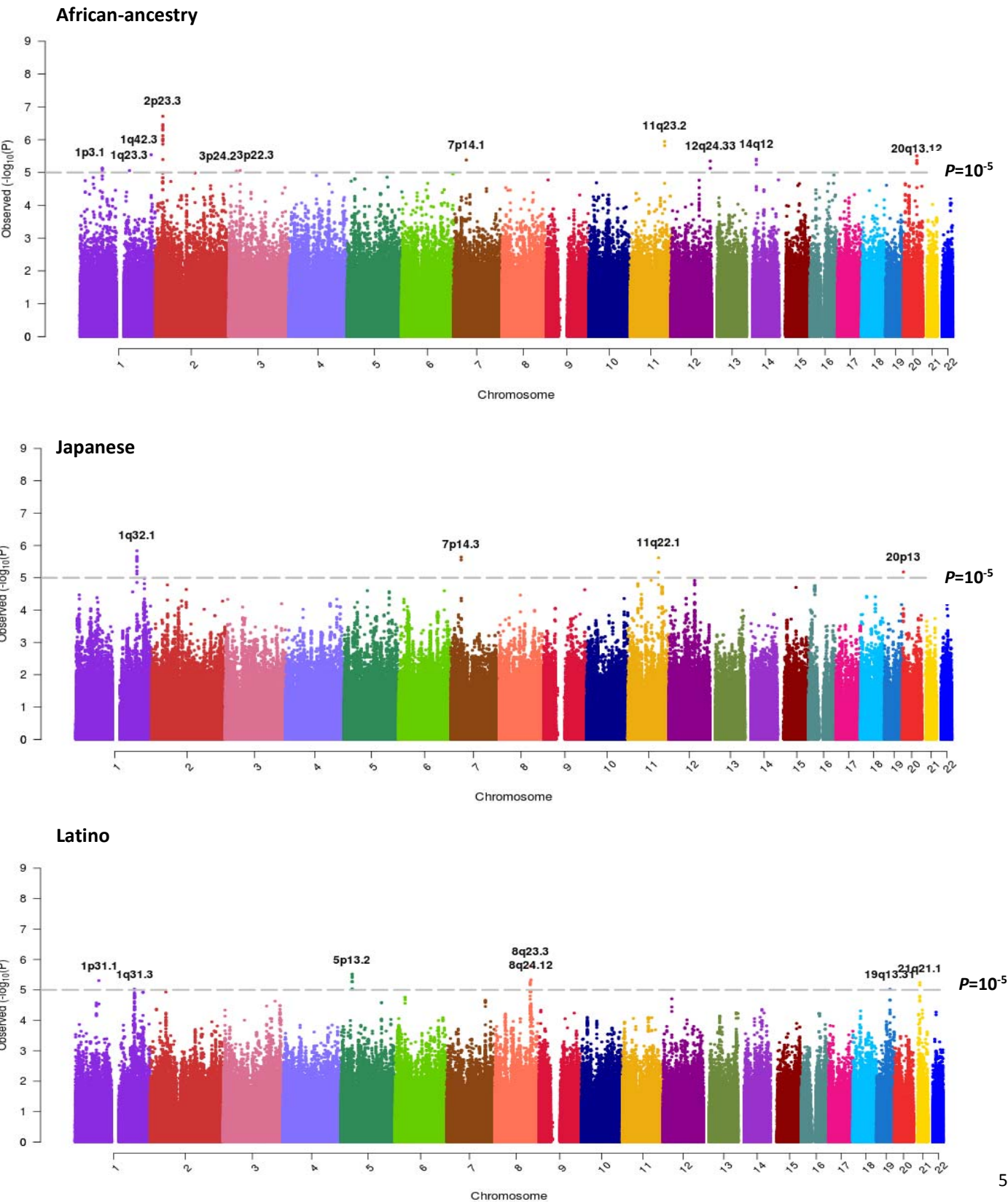
Supplementary Figures

Supplementary Figure 1. Quantile-quantile (QQ) plots for the ancestry-specific and multi-ancestry meta-analyses of asthma. QQ plots of P_{random} or P_{fixed} values for all SNPs passing QC in a least two-thirds of studies.



Supplementary Figure 2. Manhattan plots for each meta-analysis in African-ancestry, Japanese and Latino populations. These analyses were carried out under a fixed-effects model.

The x axis represents chromosomal location and the y axis represents $-\log_{10}(P \text{ value})$; the dashed horizontal line denotes $P = 10^{-5}$. Each locus is annotated by its cytogenetic band location.



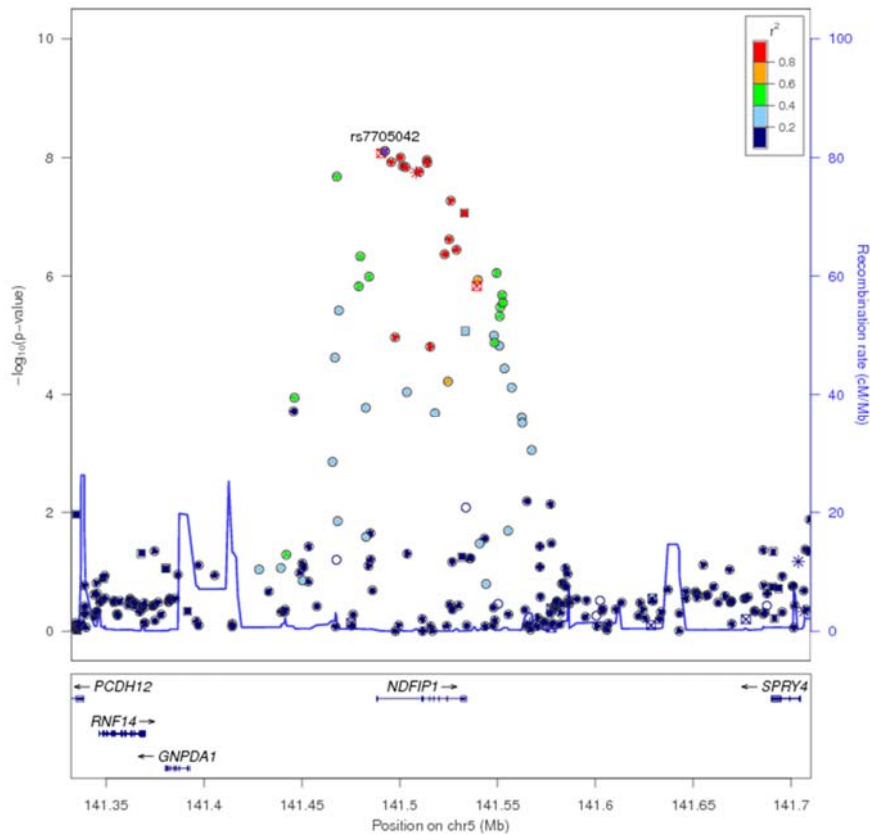
Supplementary Figure 3. Regional plots for the 18 asthma loci reaching genome-wide significance: a. New asthma loci; b. Known asthma loci

The x axis represents physical location in megabase (build 37) and the Y axis represents $-\log_{10}(P \text{ value})$ for tests of association between individual SNPs and asthma in either the multi-ancestry or European ancestry meta-analysis (random-effects model) using Locus Zoom (see URLs). The rs ID is shown for the lead SNP in the region (the lead SNP is the variant with the strongest evidence for association in the European-ancestry or multi-ancestry meta-analysis in the region, as shown in Table 1). For remaining SNPs the color indicates the r^2 measure of LD with the lead SNP.

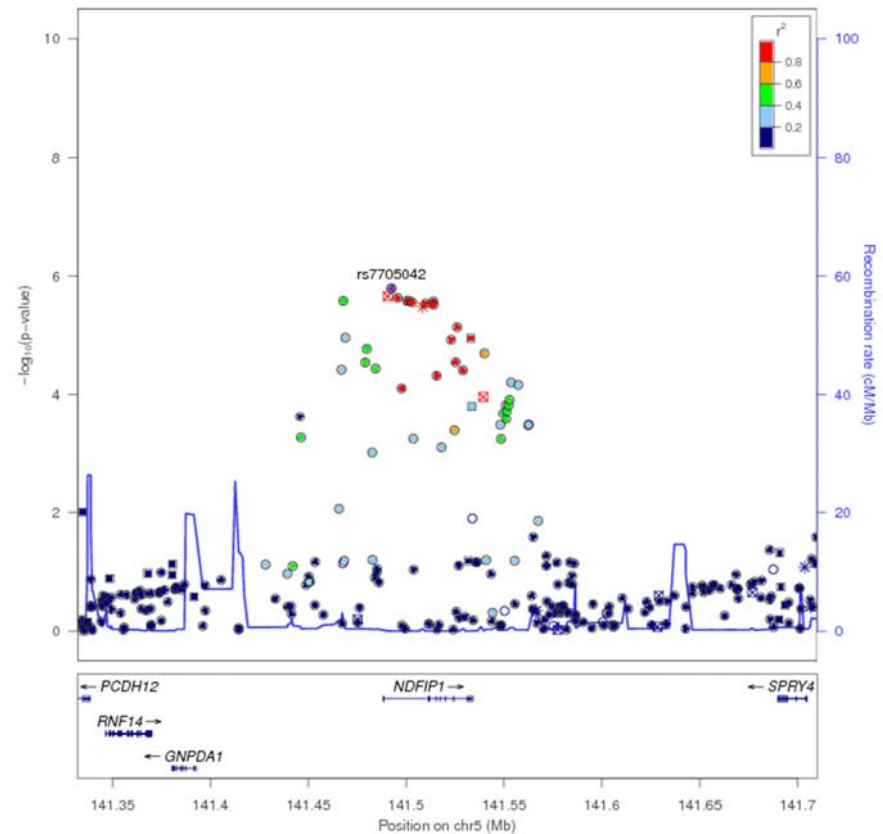
a. New asthma loci

5q31.3 - rs7705042

Multi-ancestry

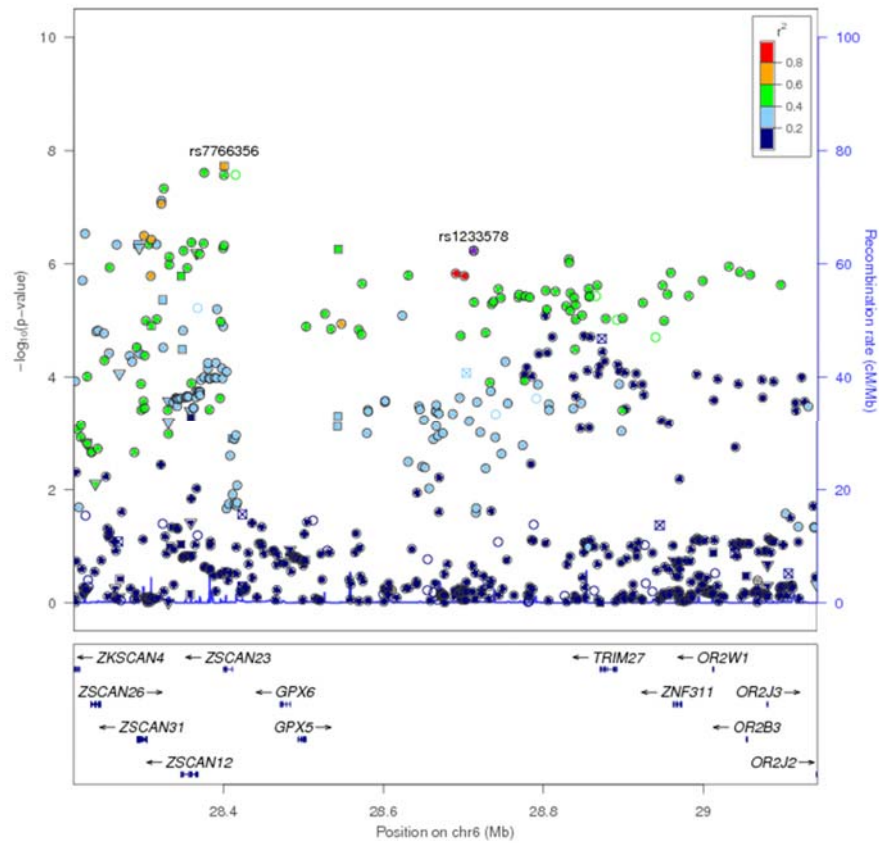


European-ancestry

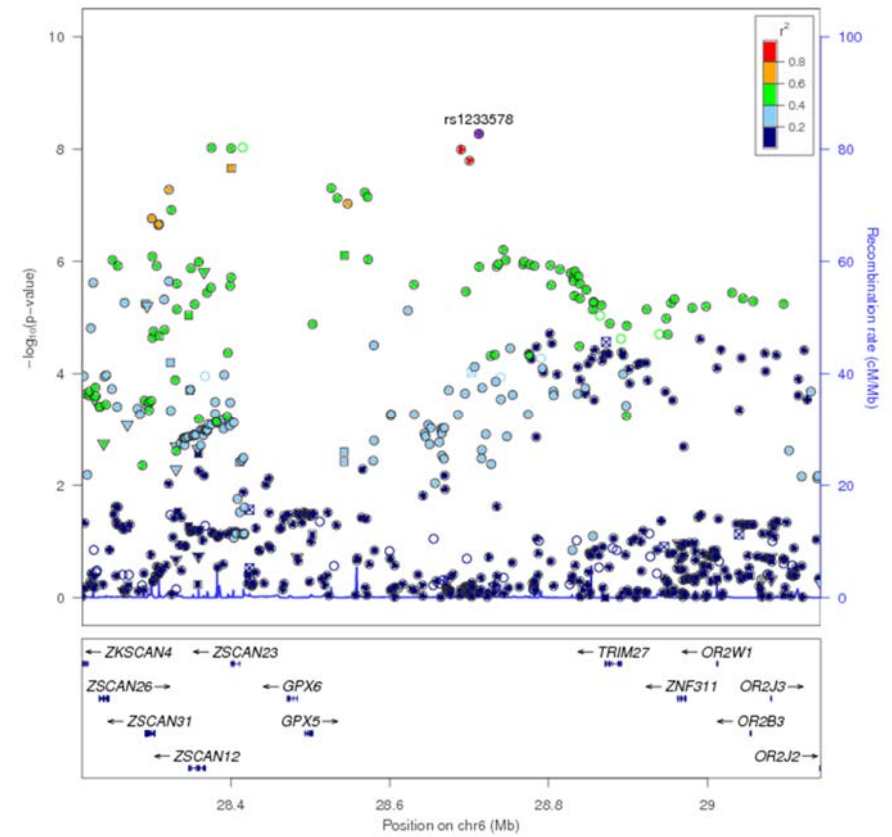


6p22.1 - rs1233578

Multi-ancestry

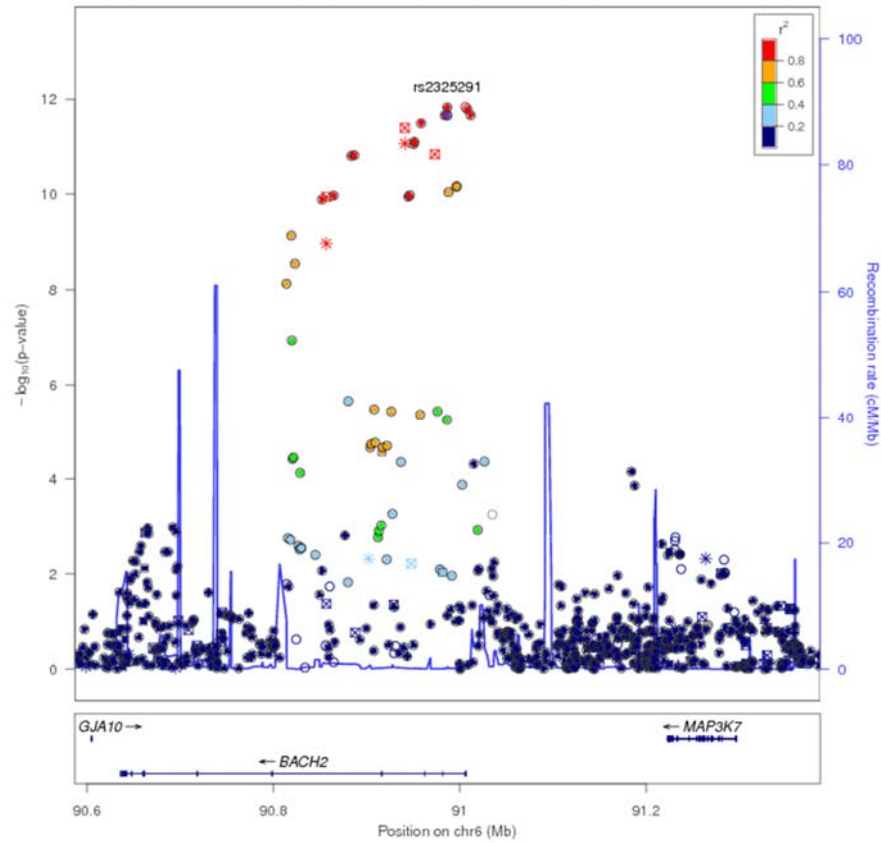


European-ancestry

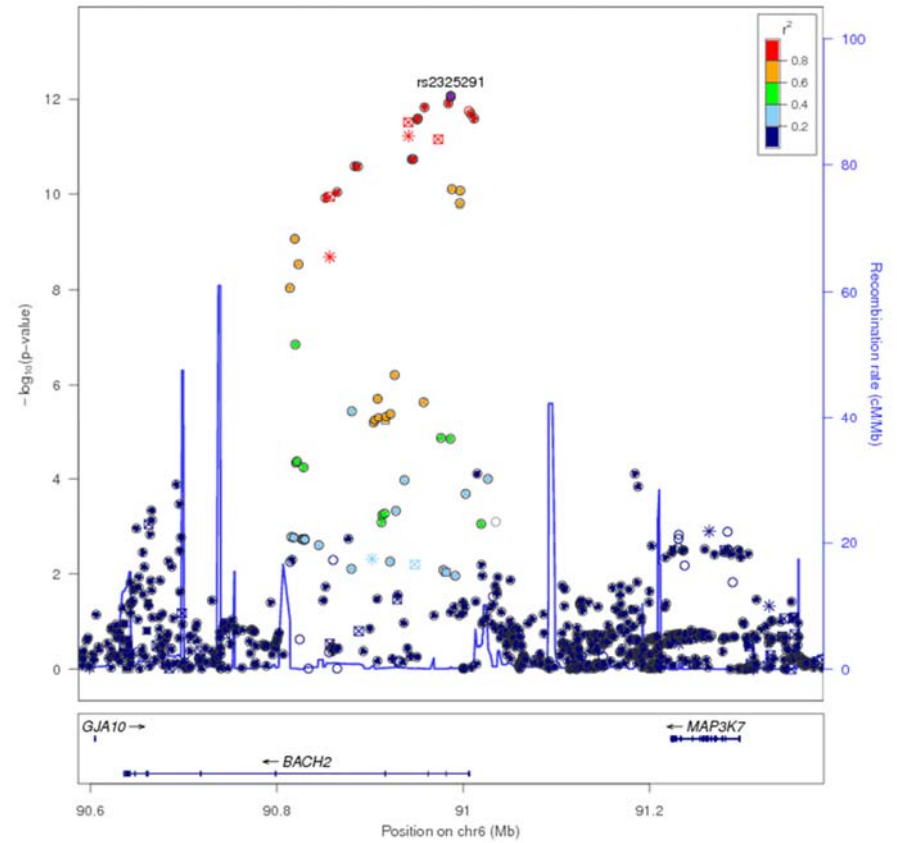


6q15 - rs2325291

Multi-ancestry

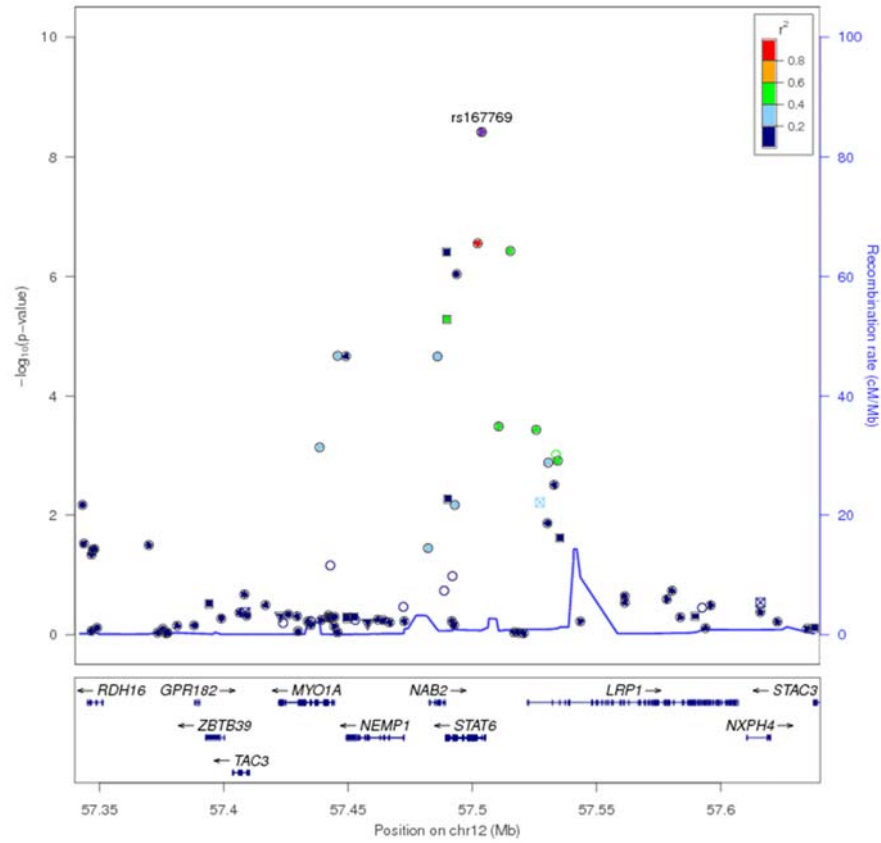


European-ancestry

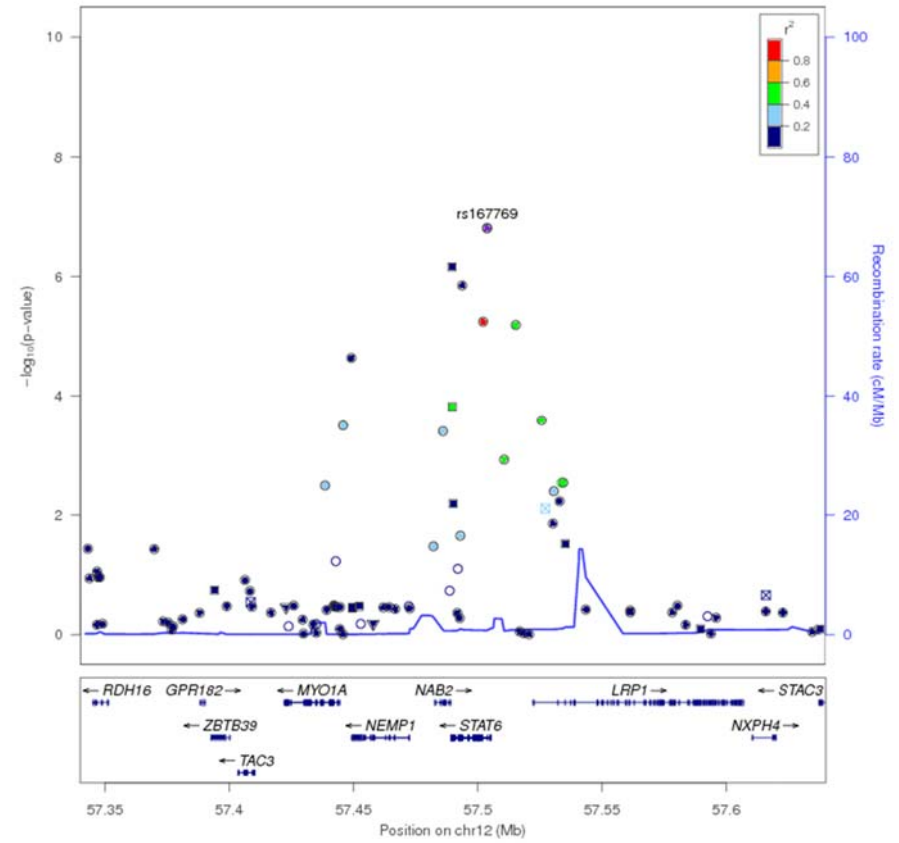


12q13.3 - rs167769

Multi-ancestry

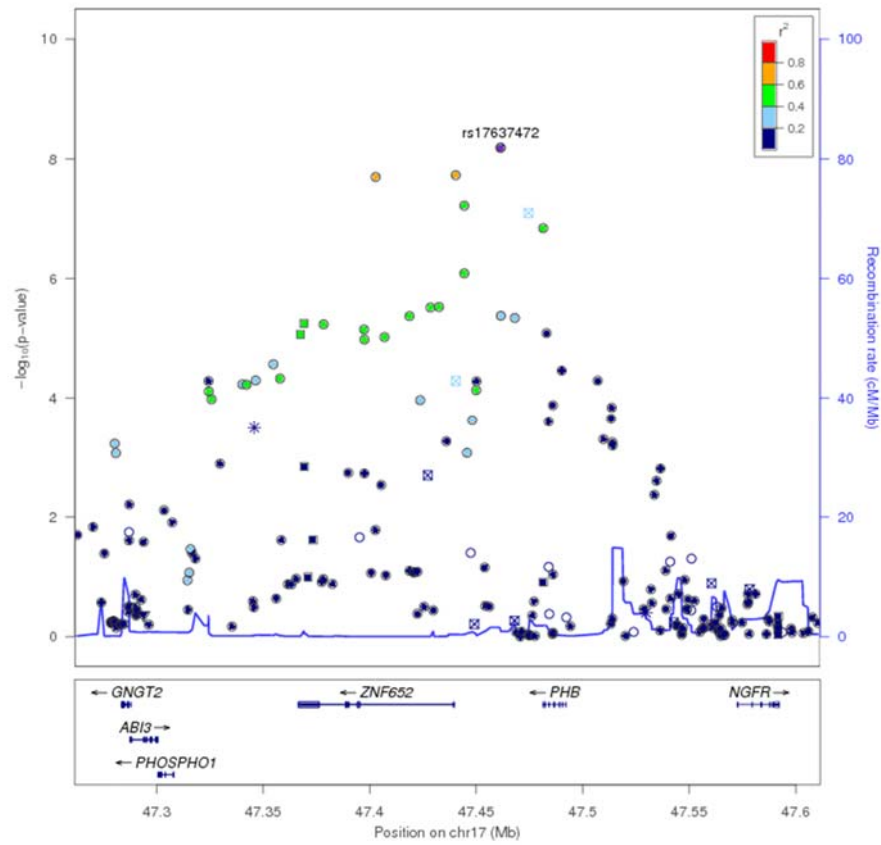


European-ancestry

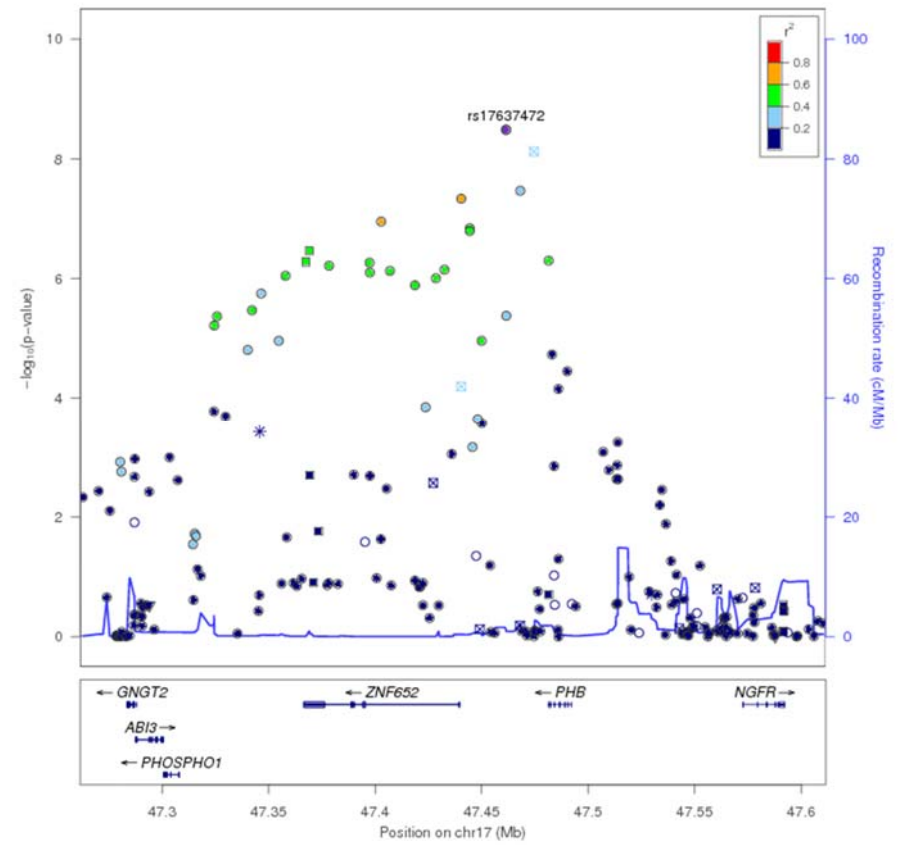


17q21.33 - rs17637472

Multi-ancestry

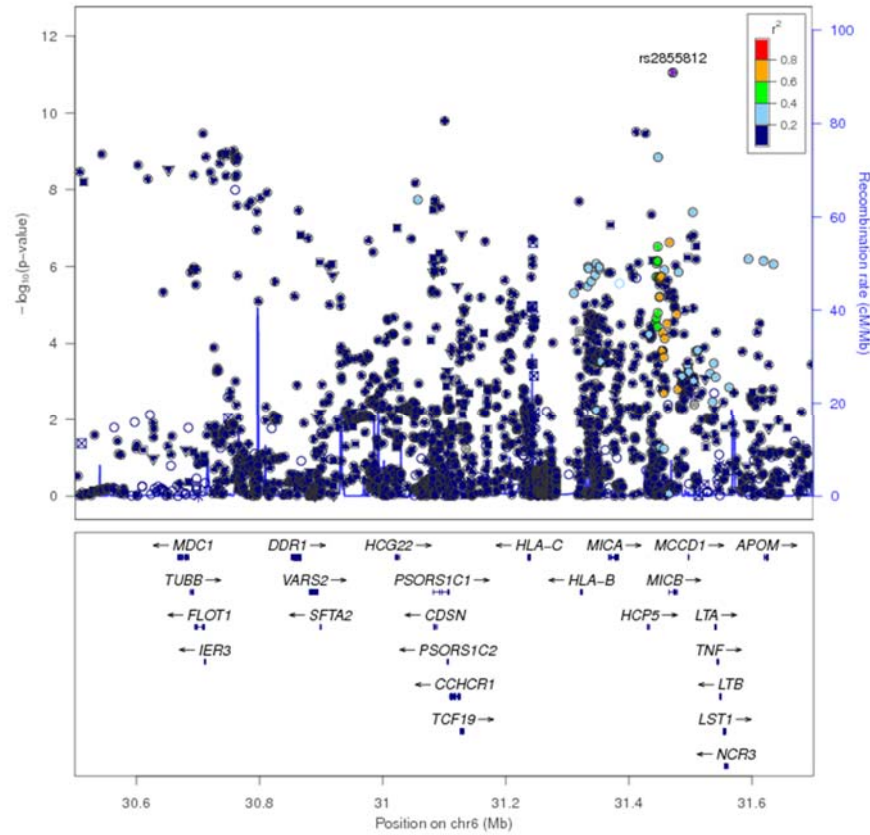


European-ancestry

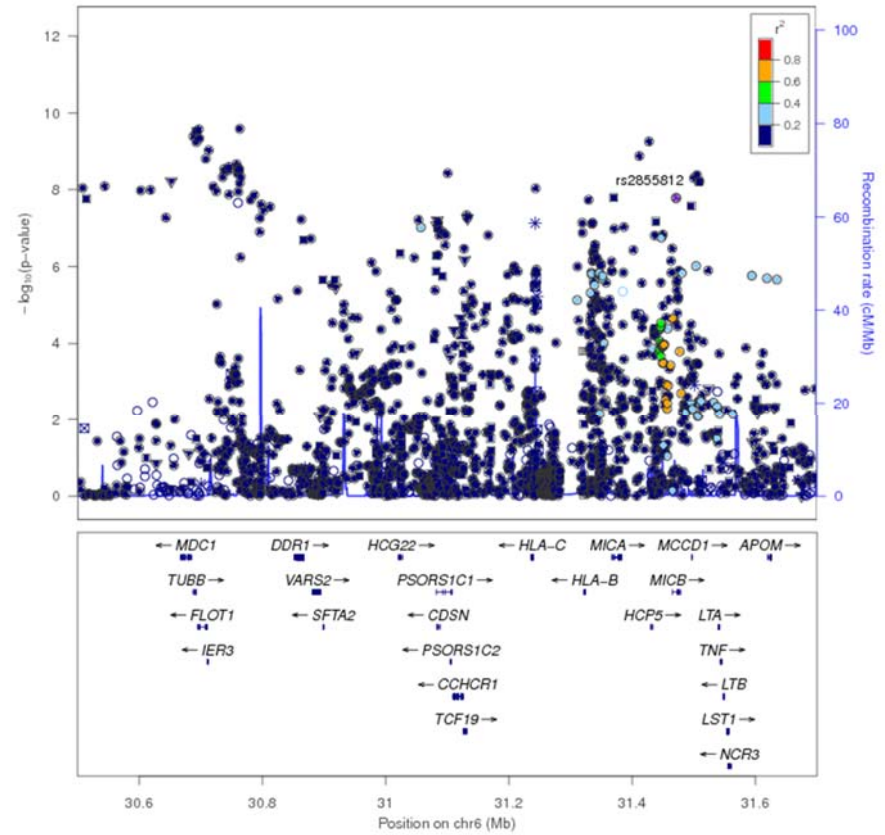


6p21.33 - rs2855812

Multi-ancestry

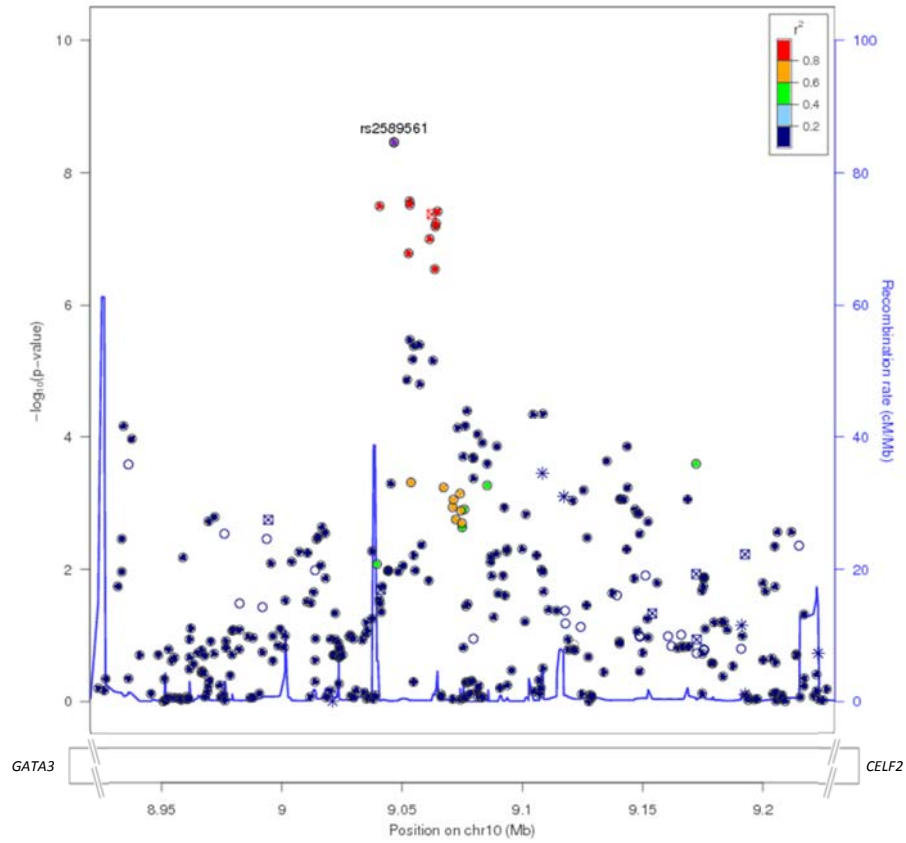


European-ancestry

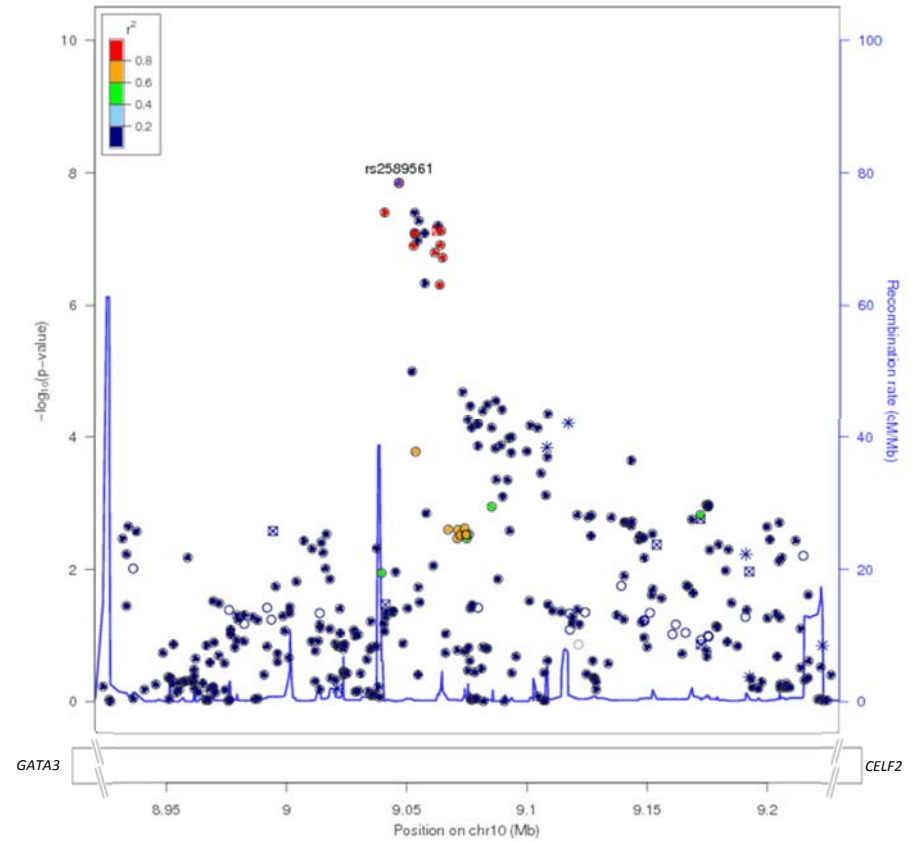


10p14 - rs2589561

Multi-ancestry

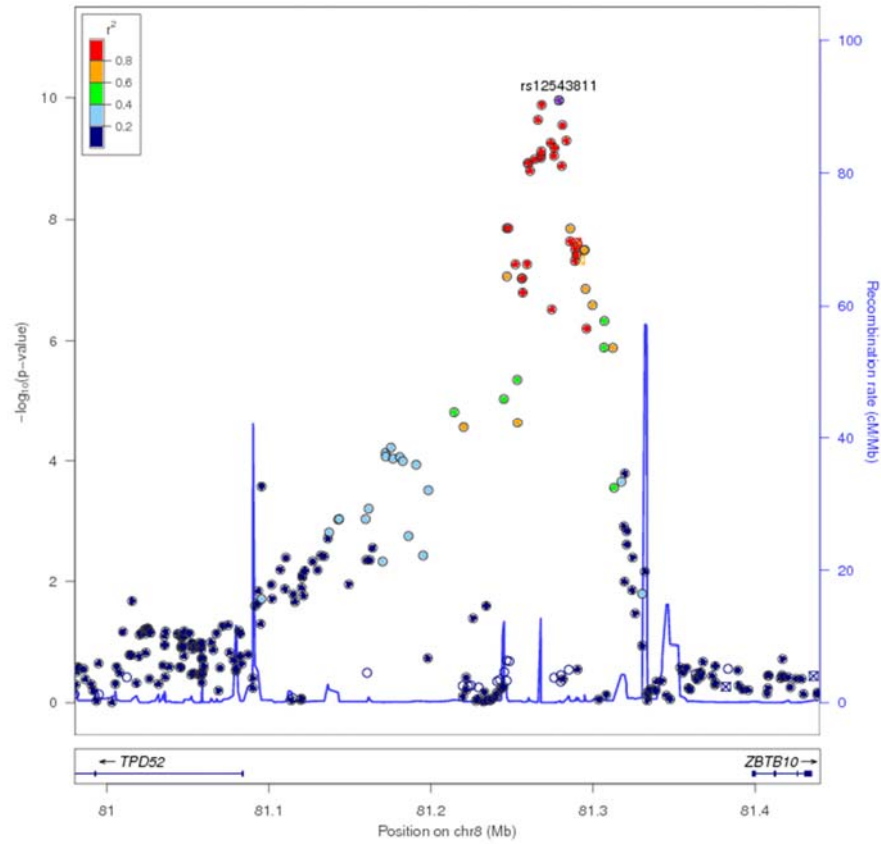


European-ancestry

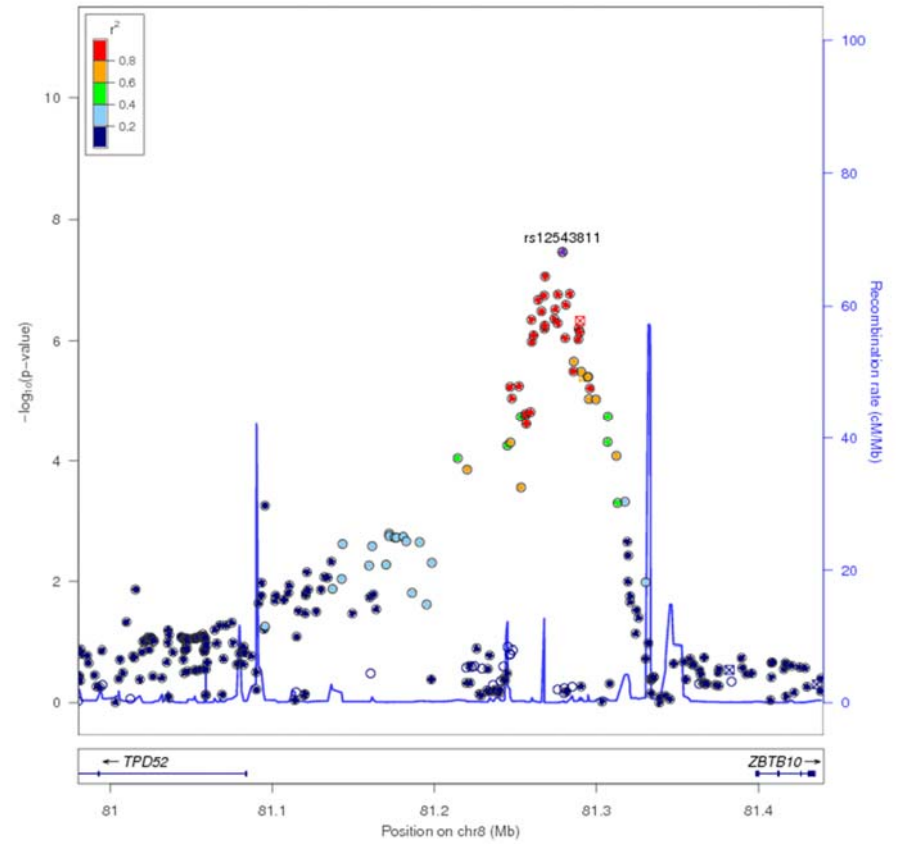


8q21.13 - rs12543811

Multi-ancestry

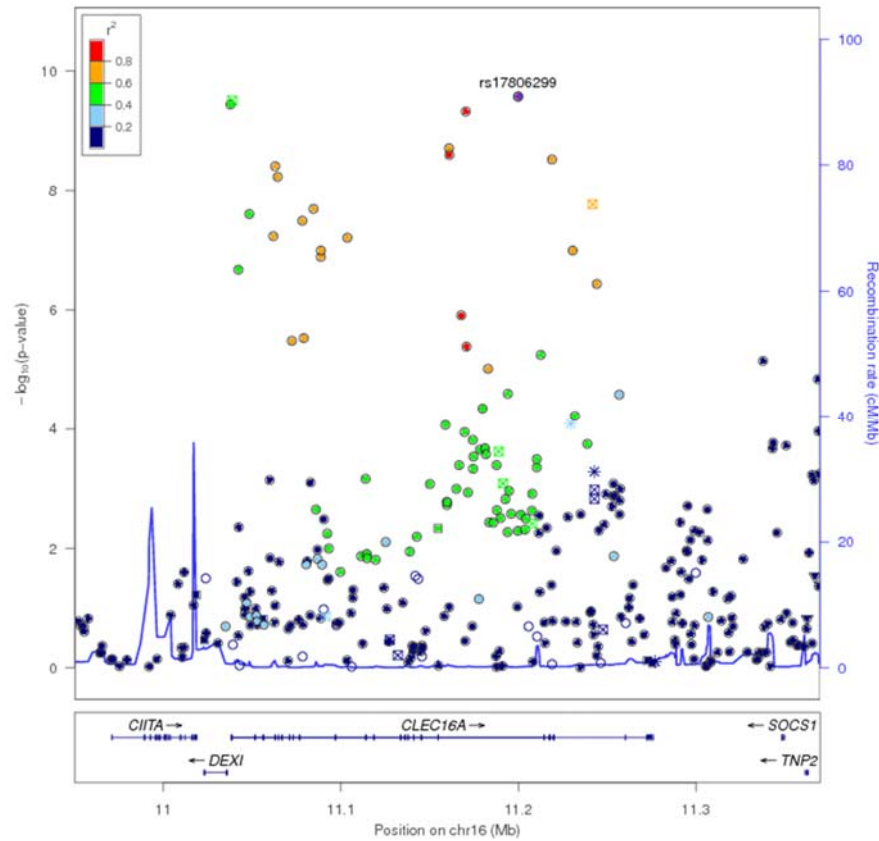


European-ancestry

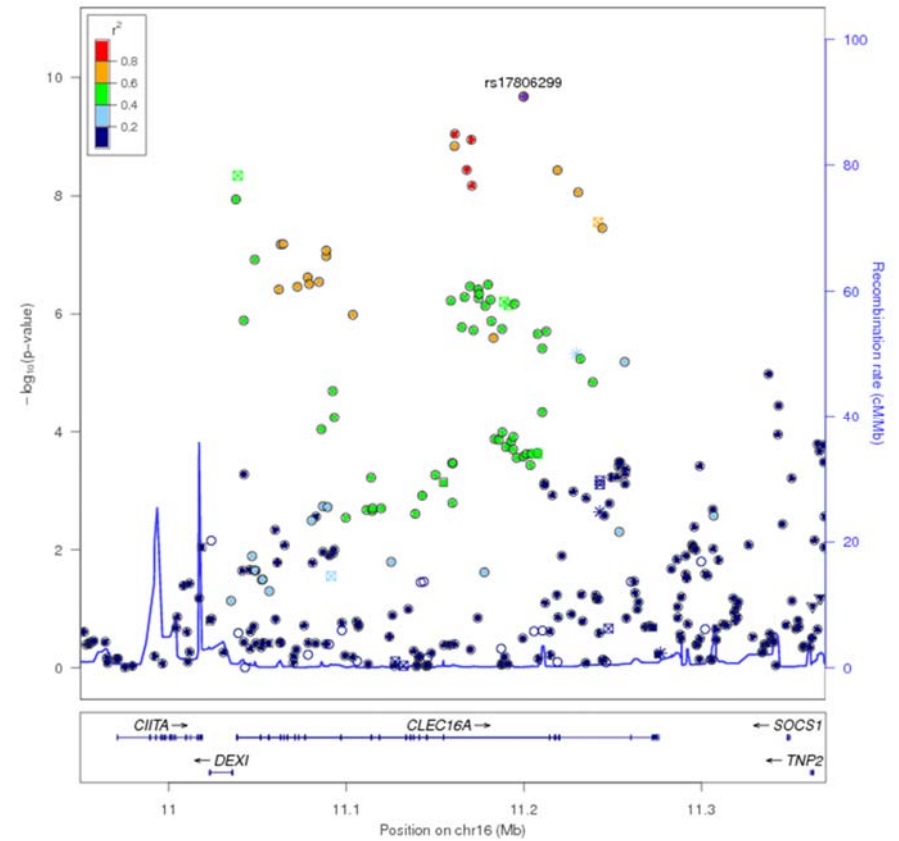


16p13.13 - rs17806299

Multi-ancestry

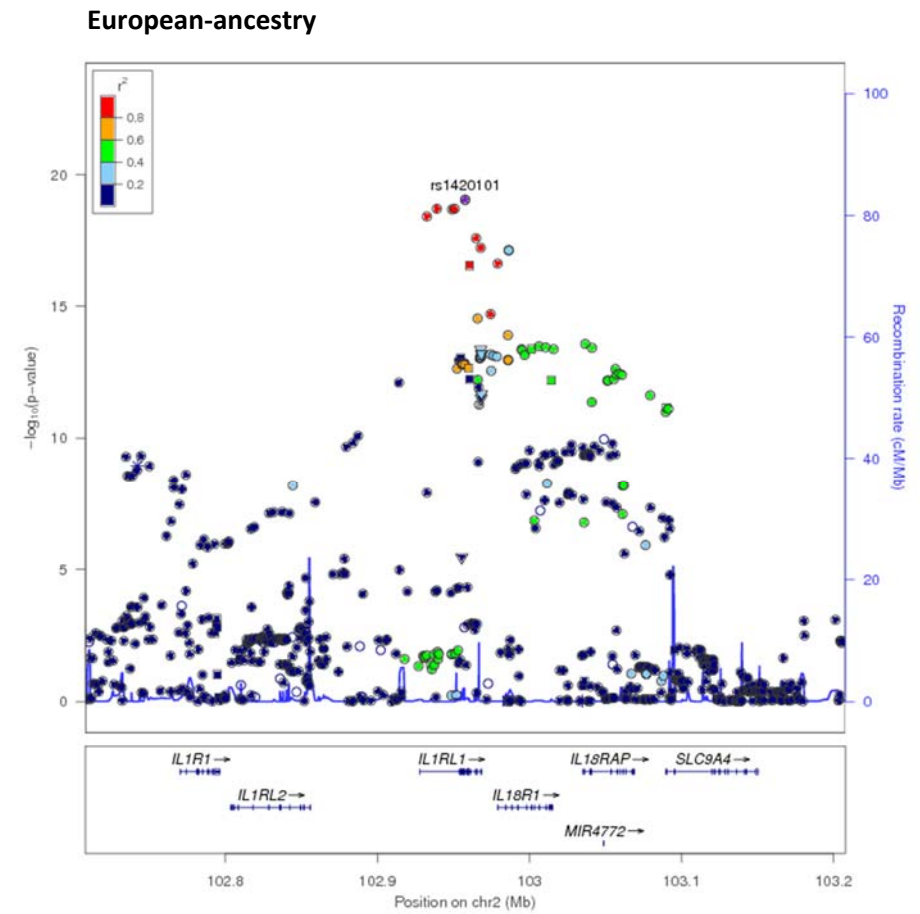
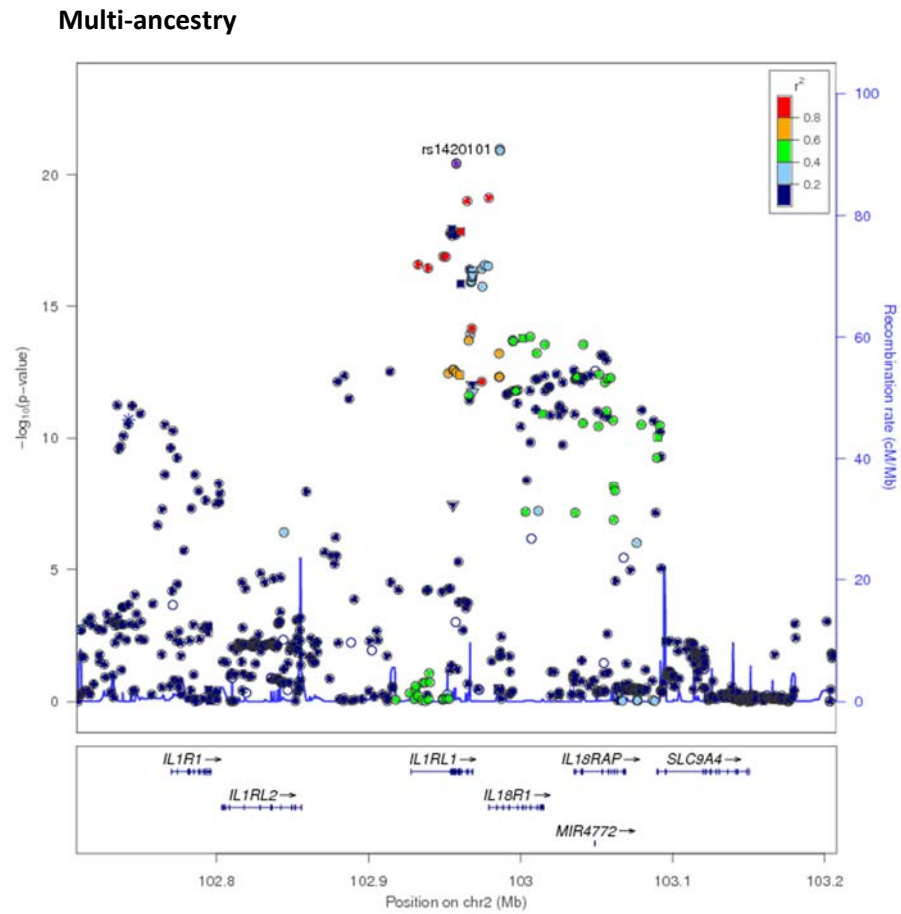


European-ancestry



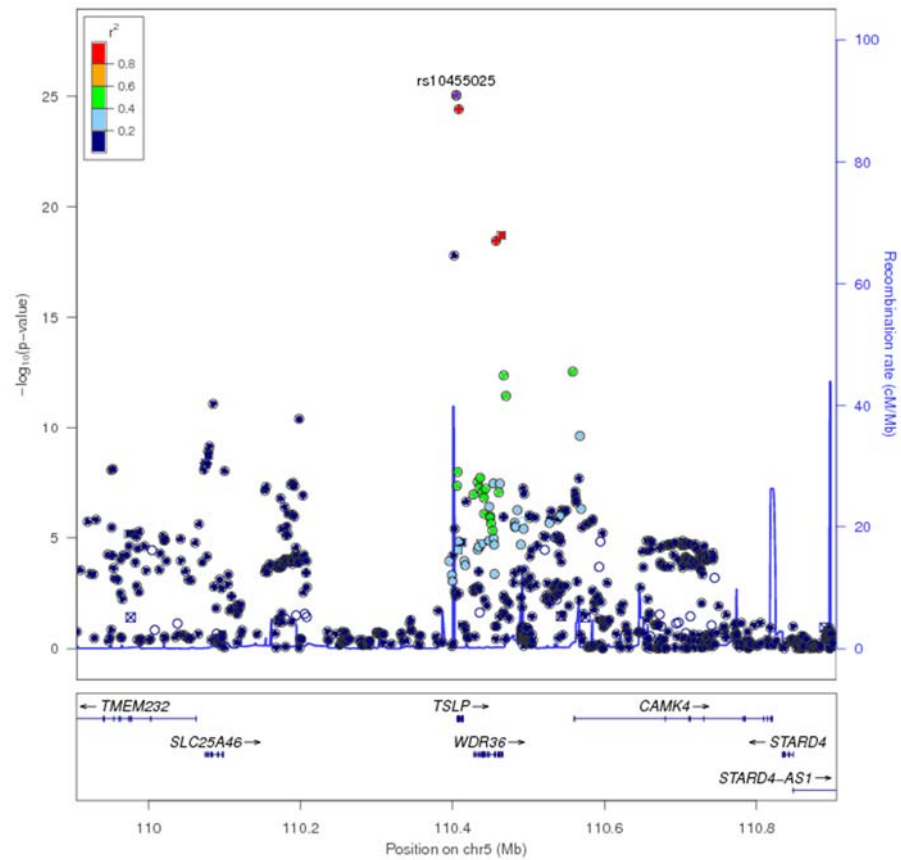
b. Known asthma loci

2q12 – rs1420101

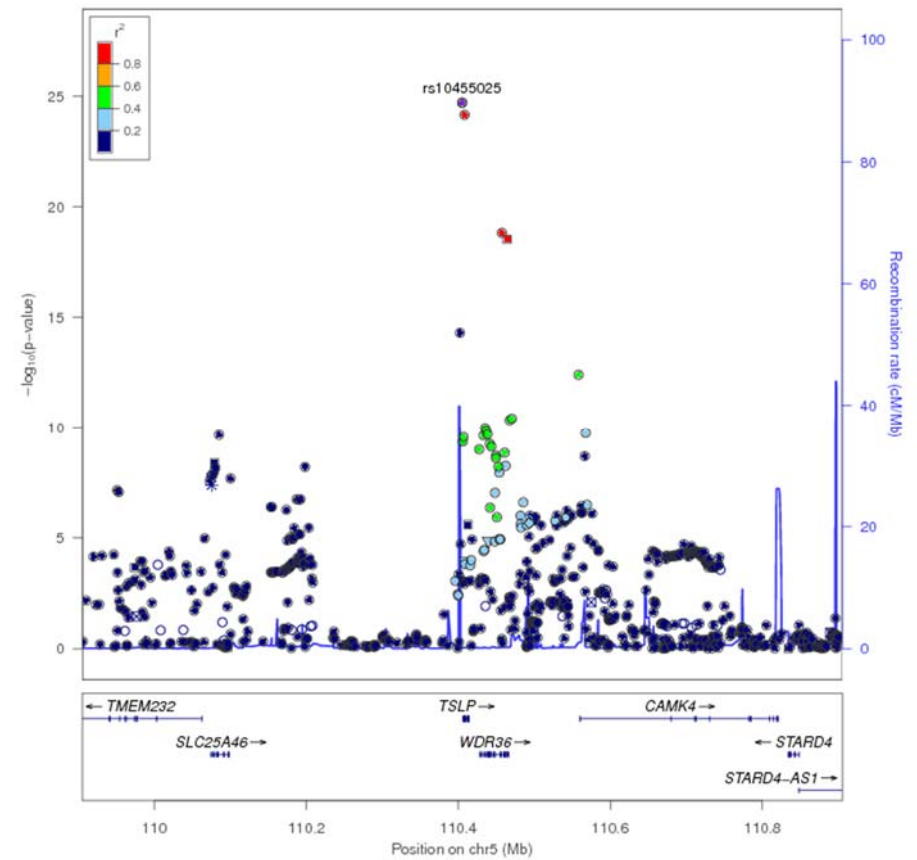


5q22.1 – rs10455025

Multi-ancestry

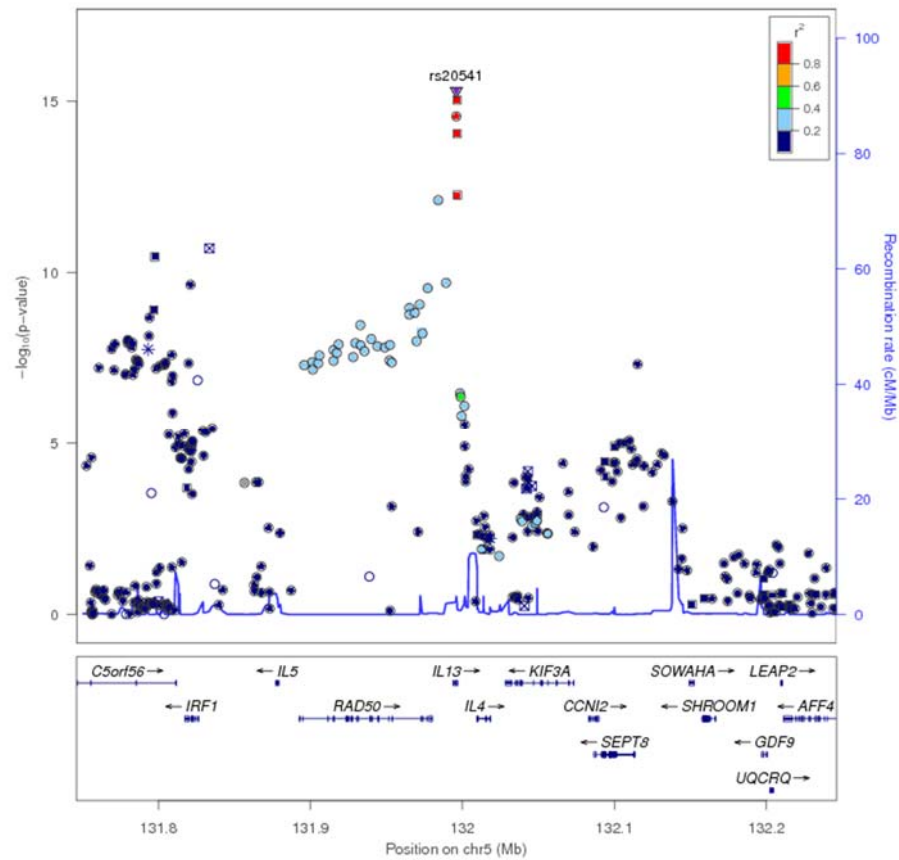


European-ancestry

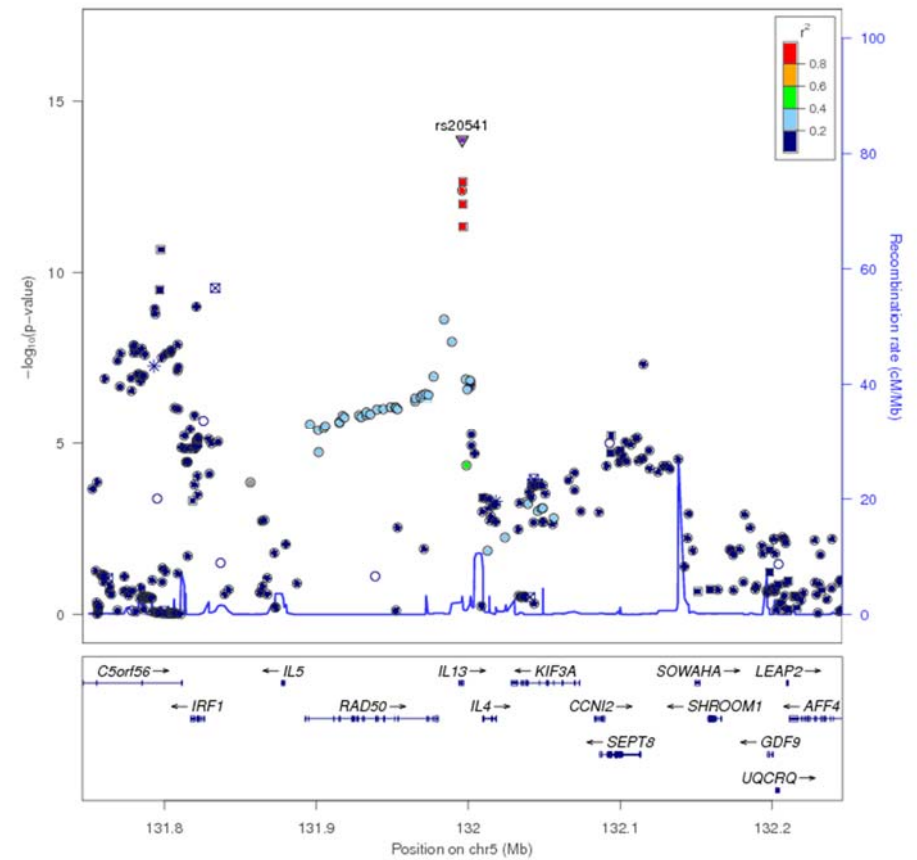


5q31 – rs20541

Multi-ancestry

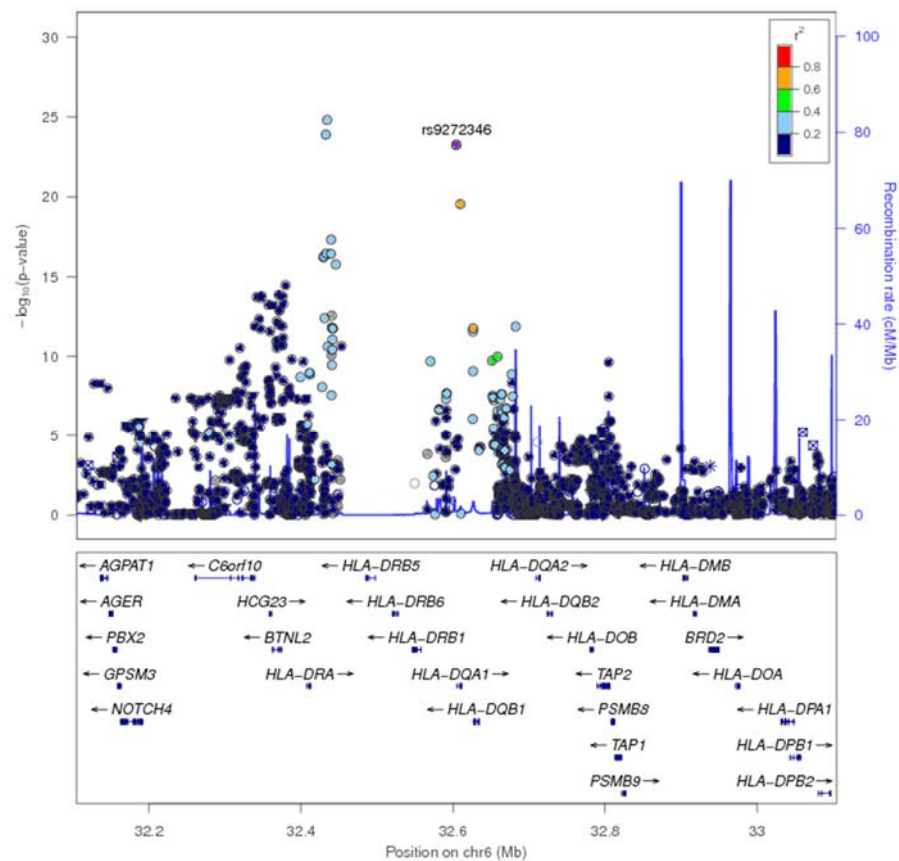


European-ancestry

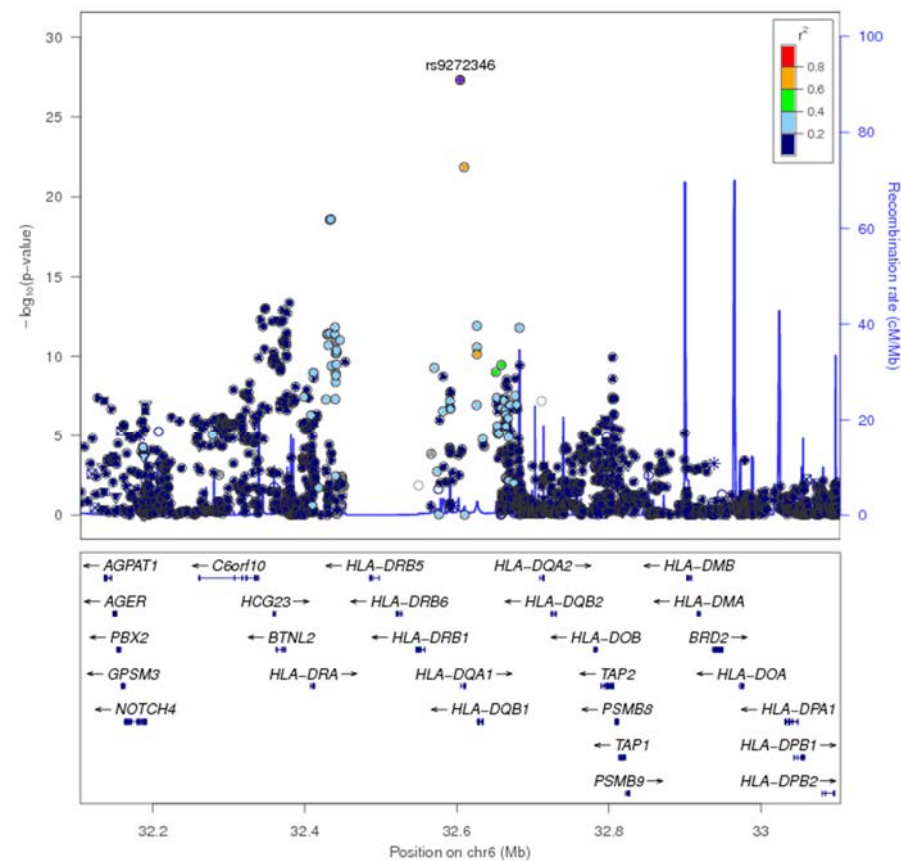


6p21.32 – rs9272346

Multi-ancestry

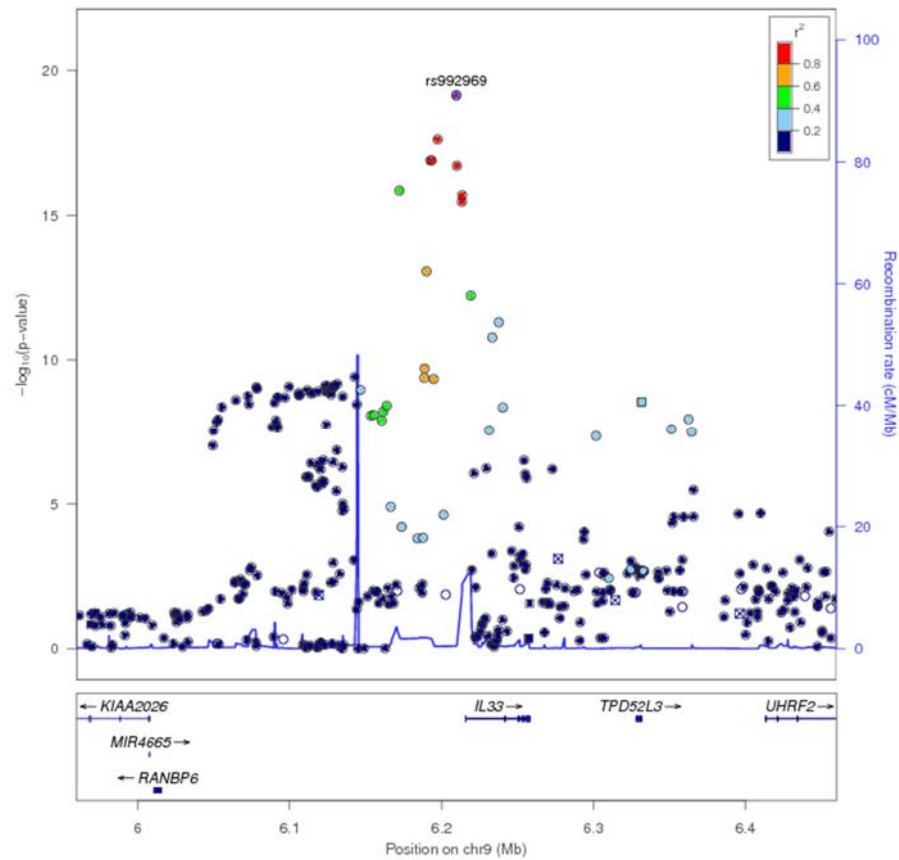


European-ancestry

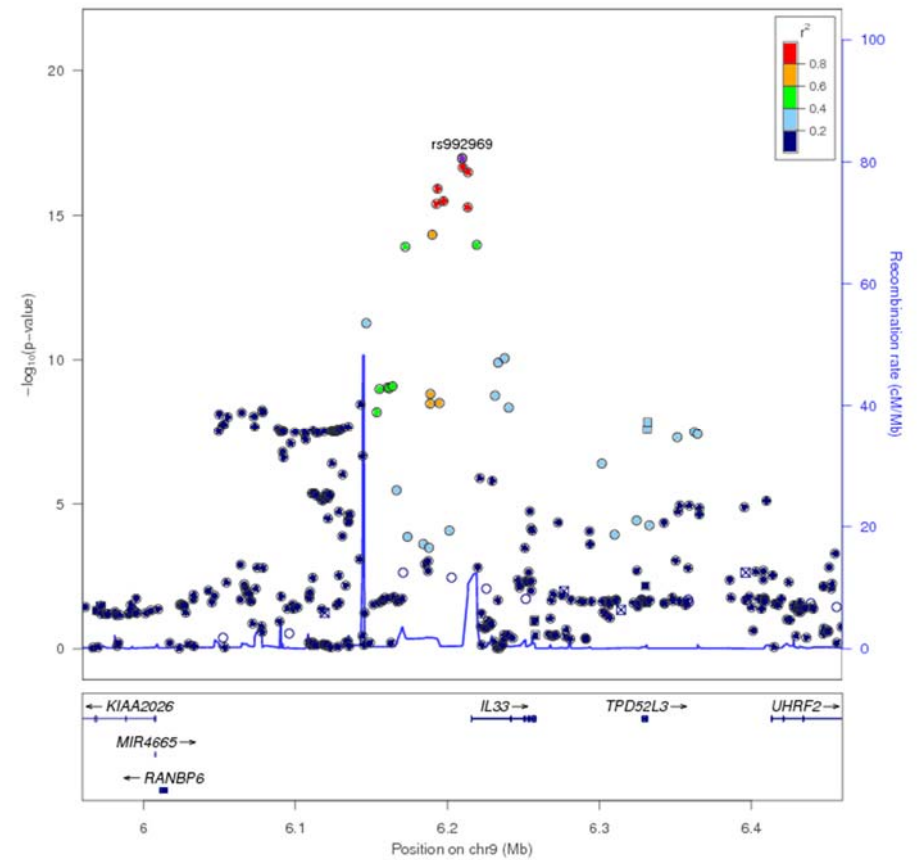


9p24.1 – rs992969

Multi-ancestry

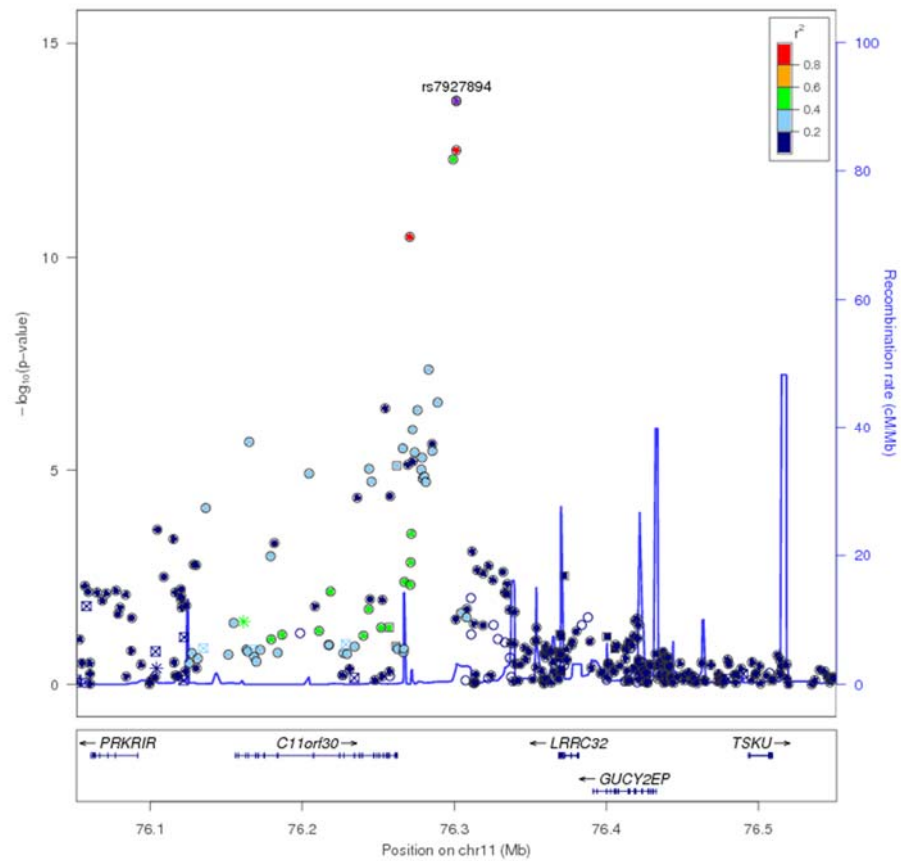


European-ancestry

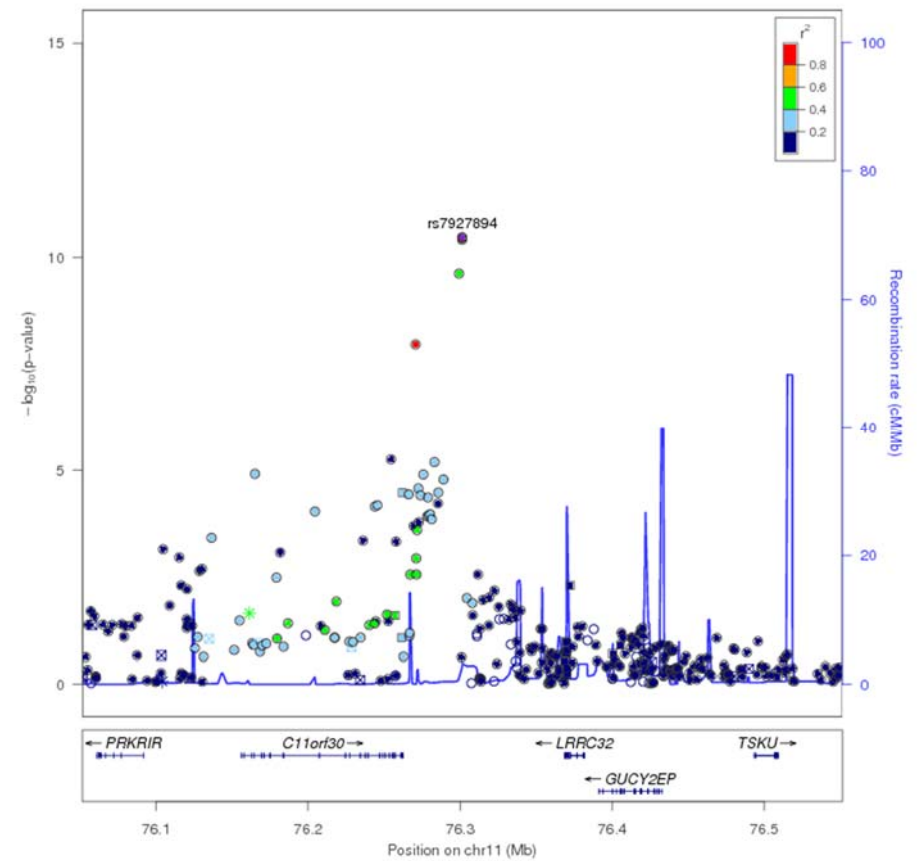


11q13.5 – rs7927894

Multi-ancestry

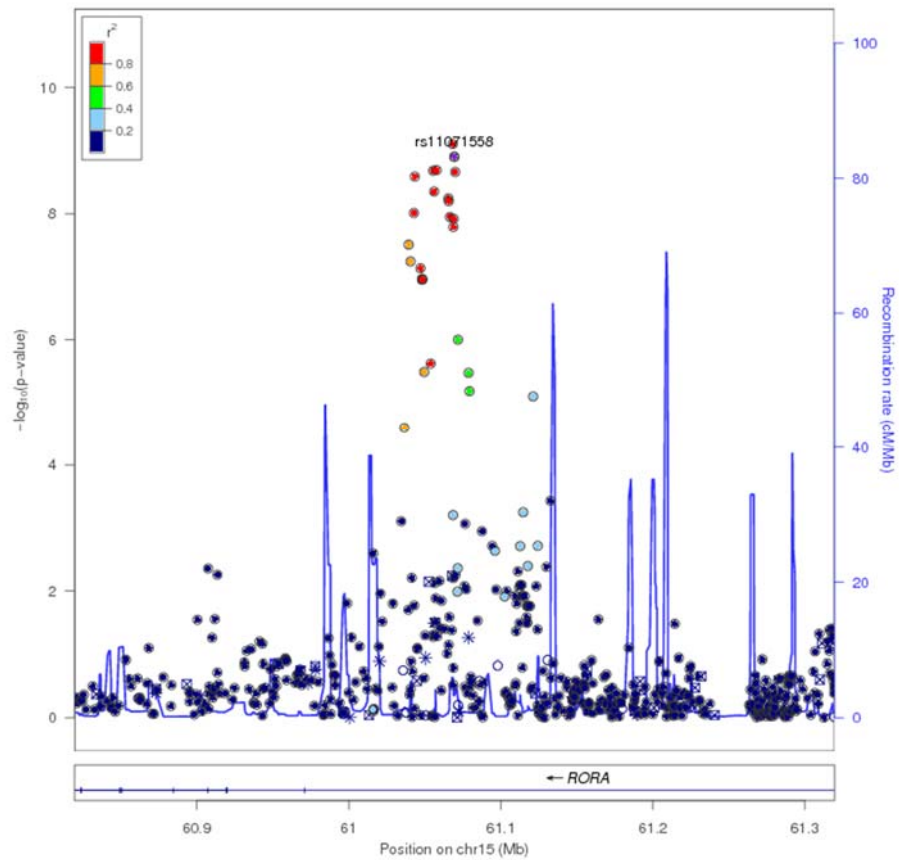


European-ancestry

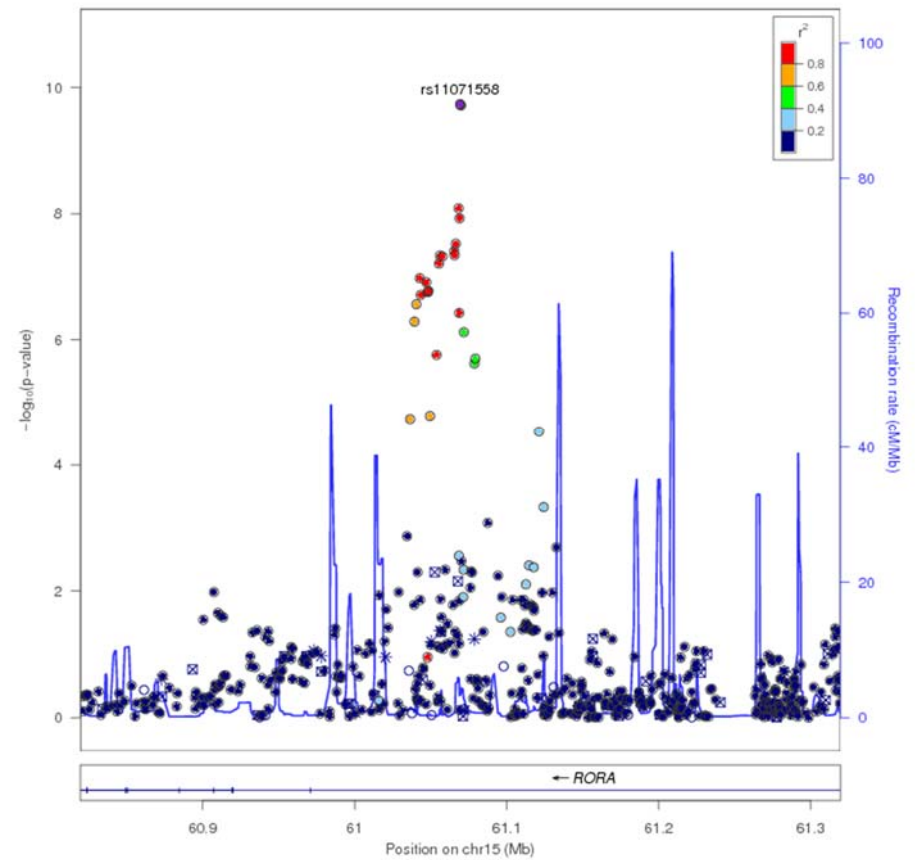


15q22.2 – rs11071558

Multi-ancestry

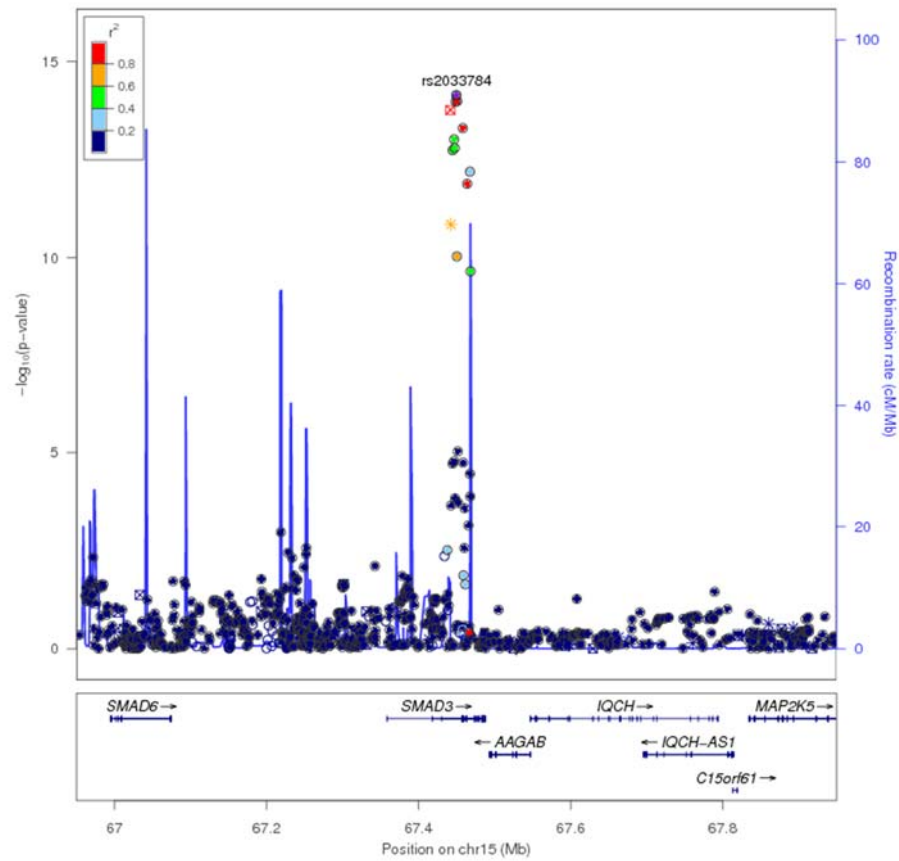


European-ancestry

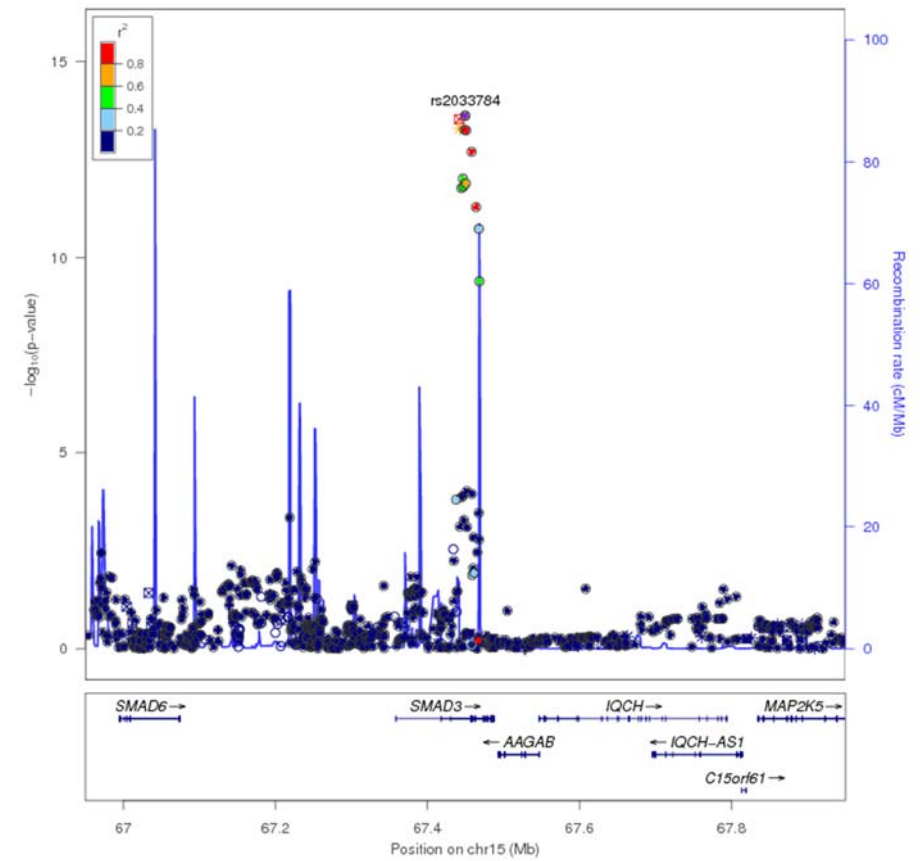


15q22.33 – rs2033784

Multi-ancestry

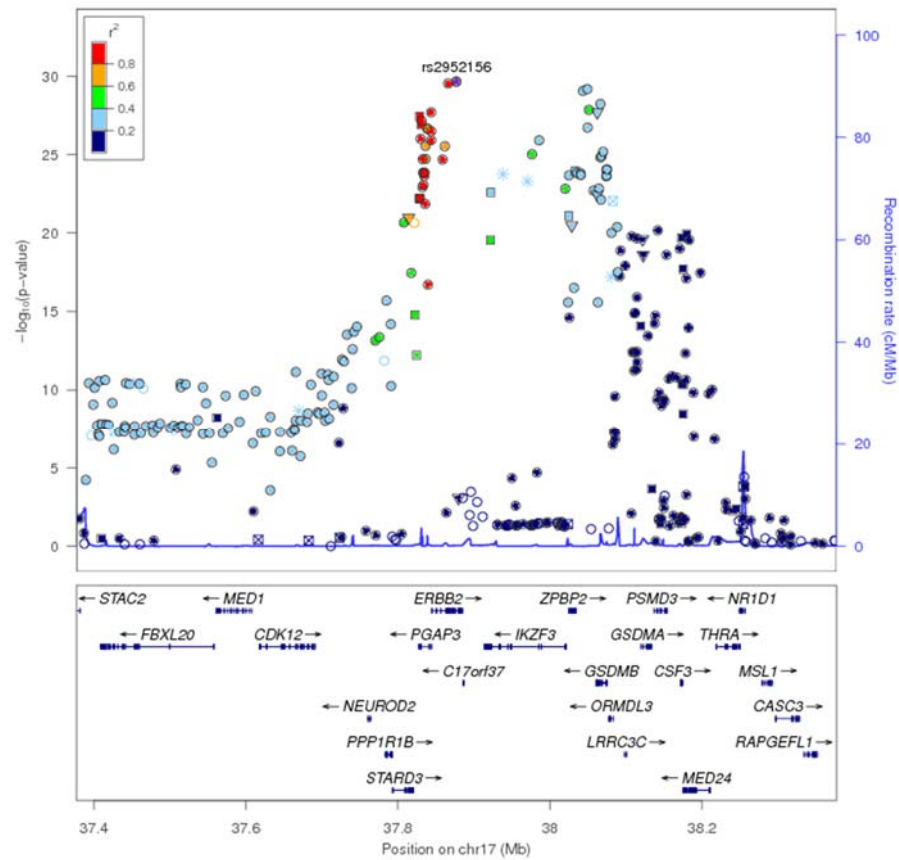


European-ancestry

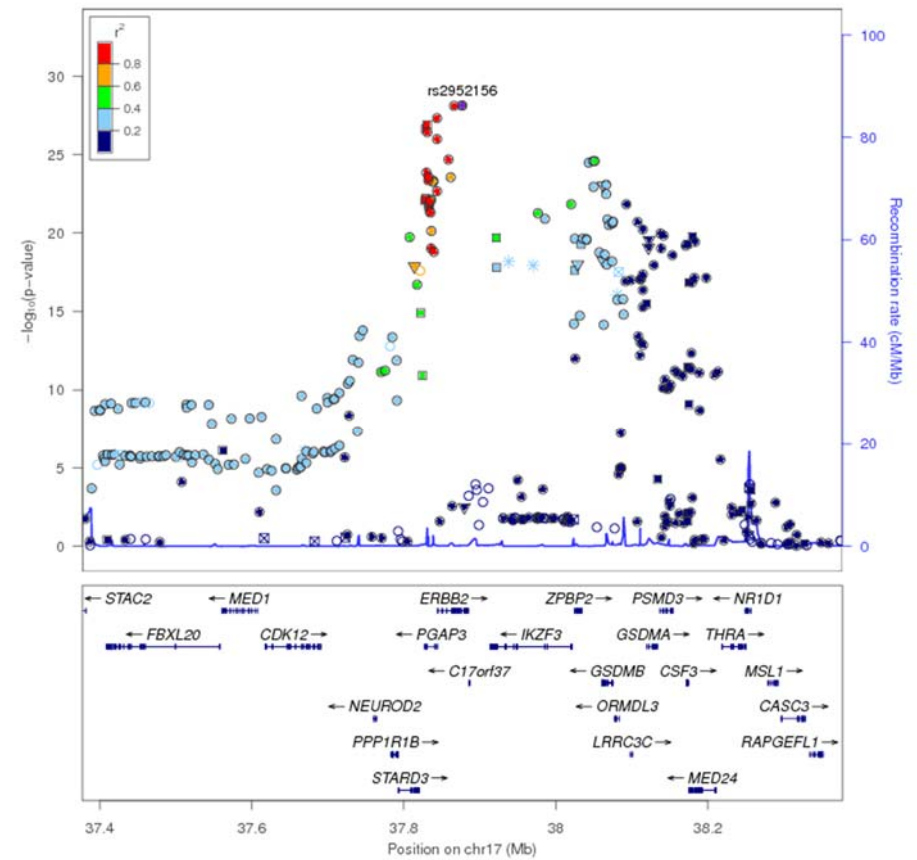


17q12-21 – rs2952156

Multi-ancestry



European-ancestry

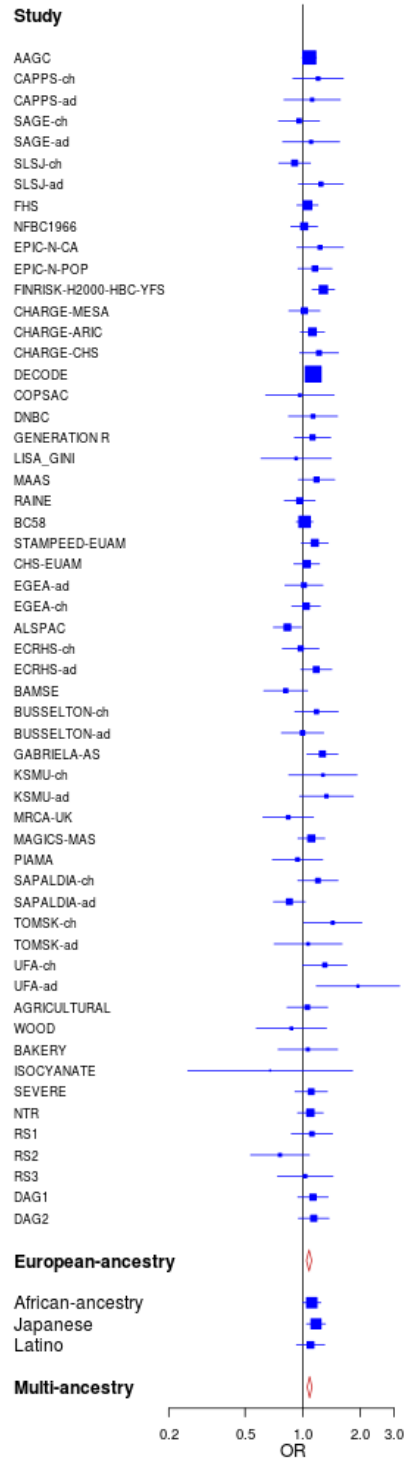


Supplementary Figure 4. Forest Plots of the lead SNPs at the 18 genome-wide significant asthma loci

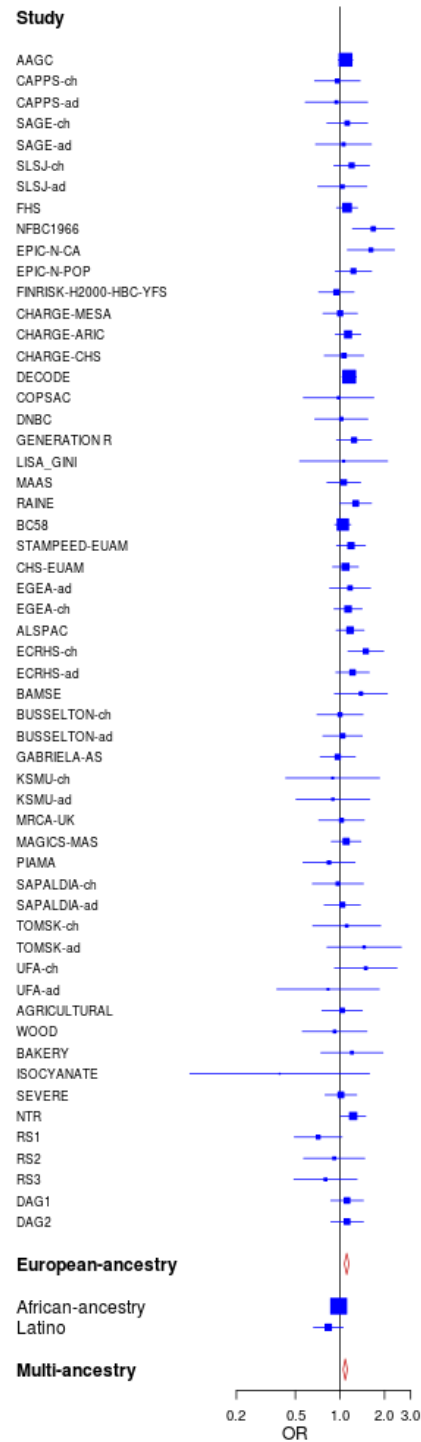
(nine loci harbouring new signals for asthma followed by nine known asthma loci, as shown in Table 1)

Odds-ratio (OR) and 95% Confidence Intervals (CI) are plotted by study and by population. European ancestry and multi-ancestry random-effects meta-analysis results are plotted as a diamond.

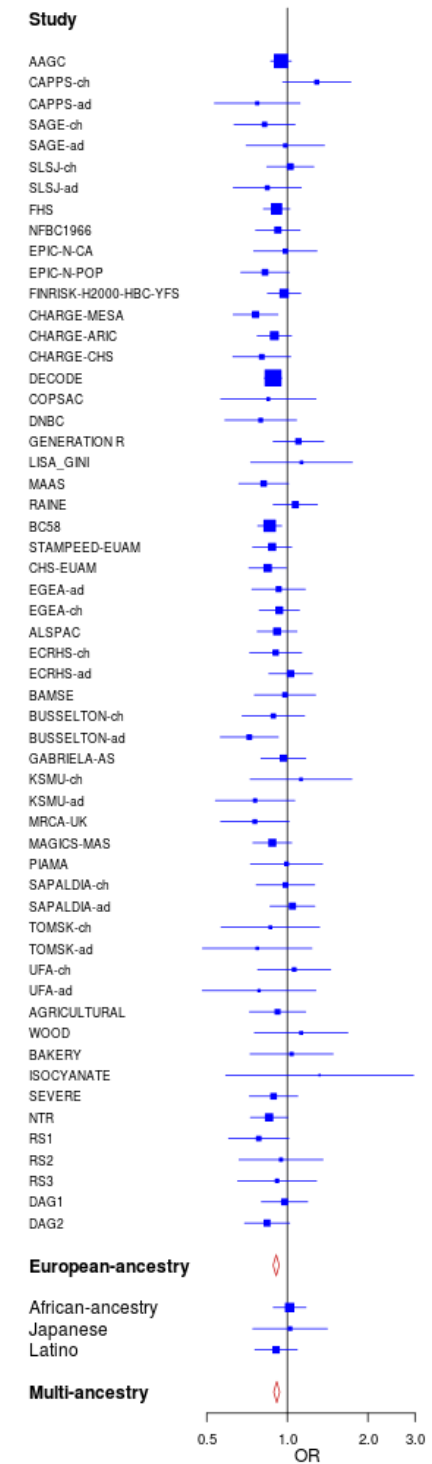
5q31.3, rs7705042



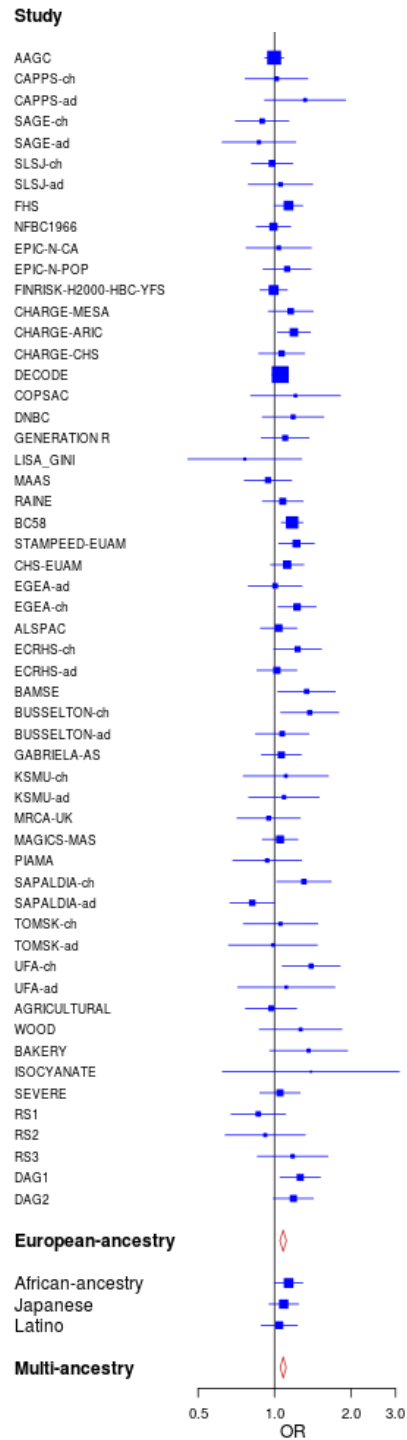
6p22.1, rs1233578



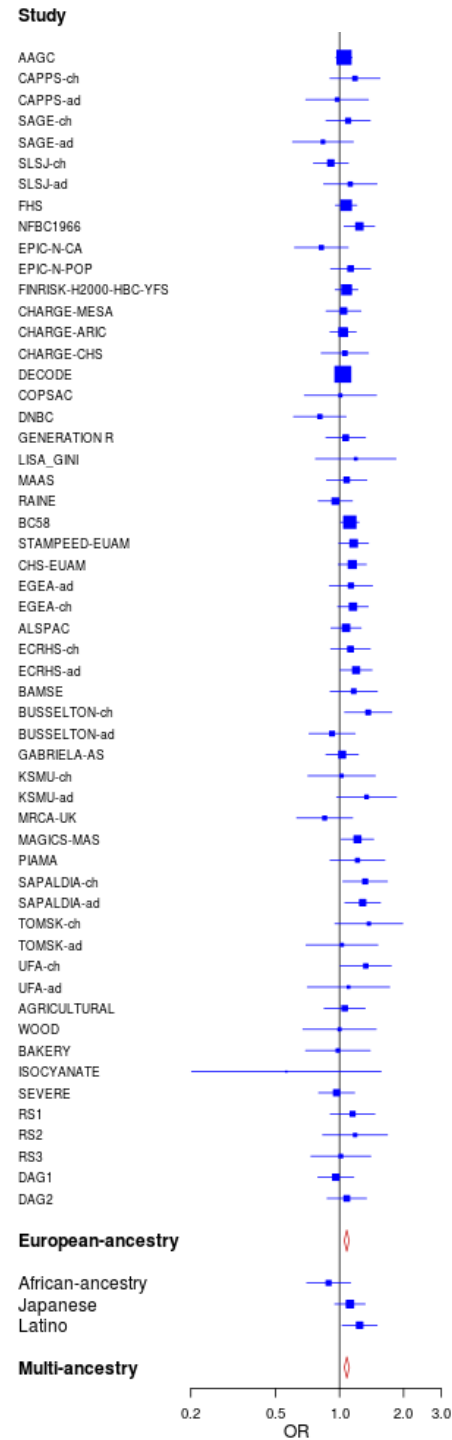
6q15, rs2325291



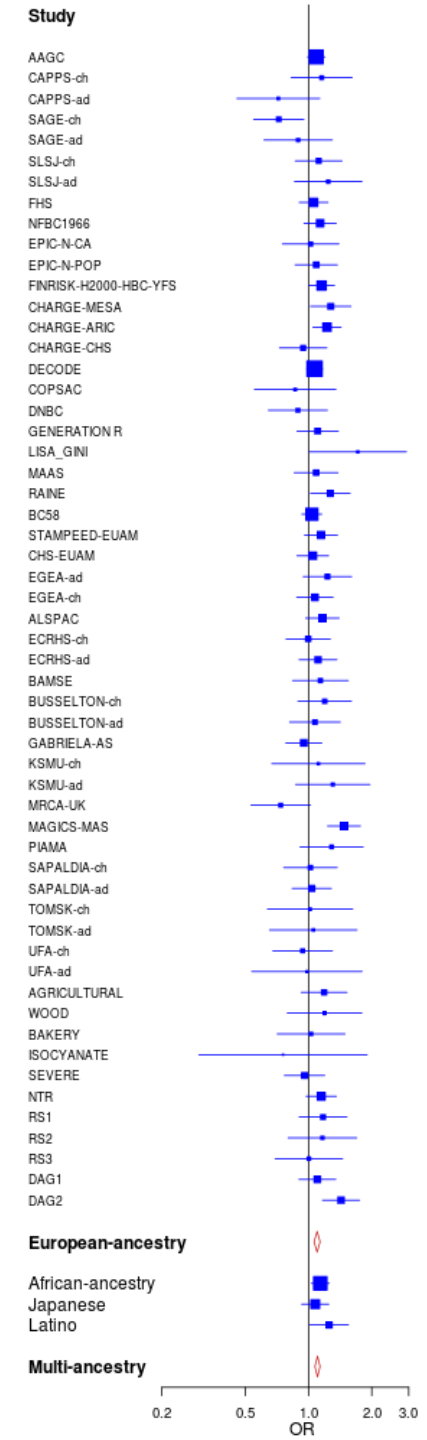
12q13.3, rs167769



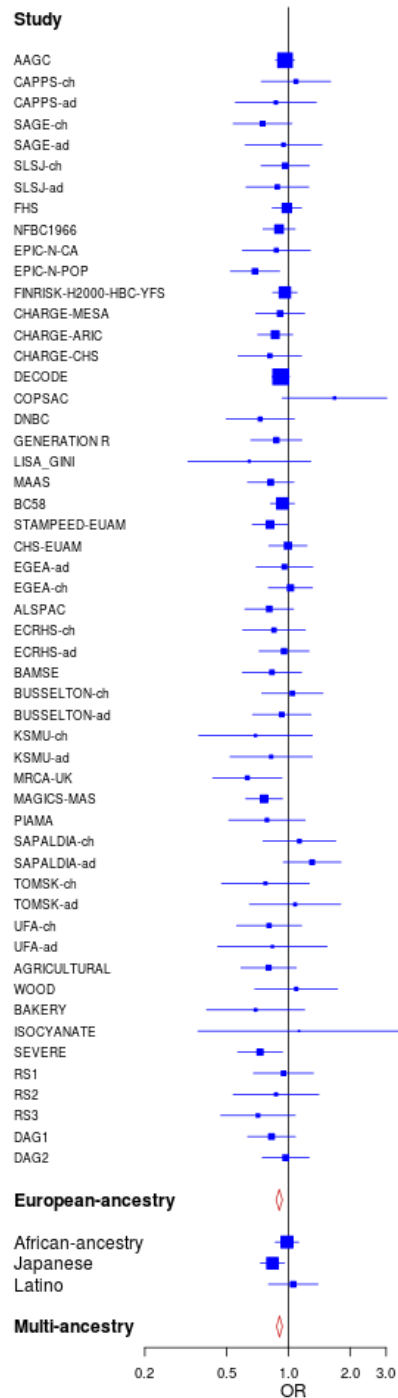
17q21.33, rs17637472



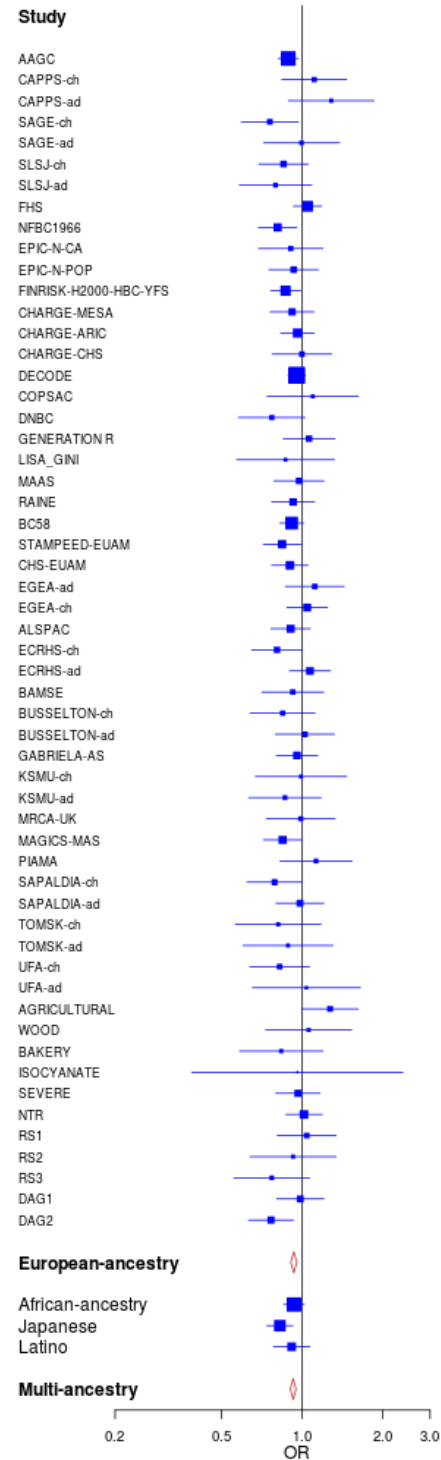
6p21.33, rs2855812



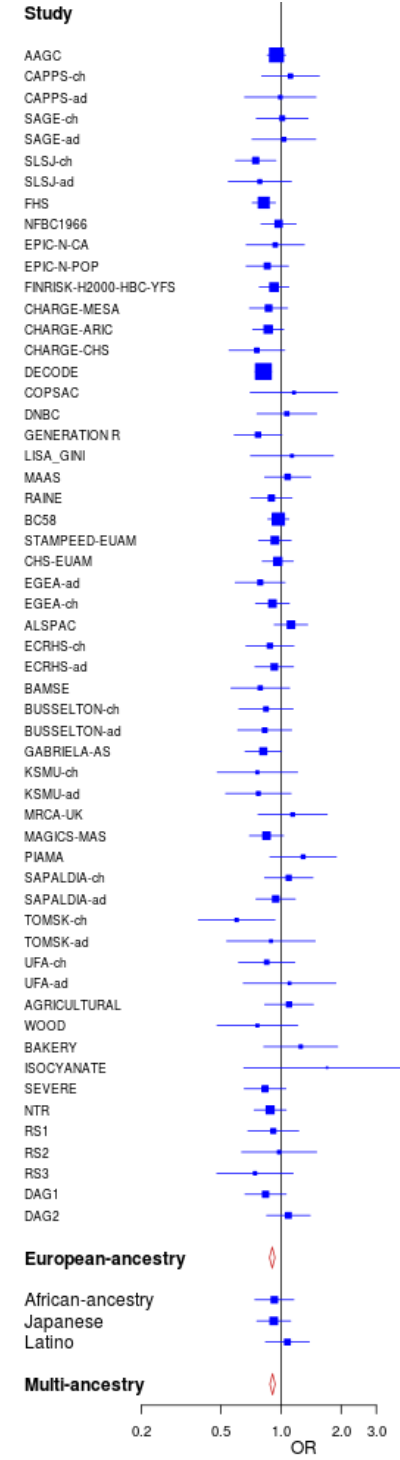
10p14, rs2589561



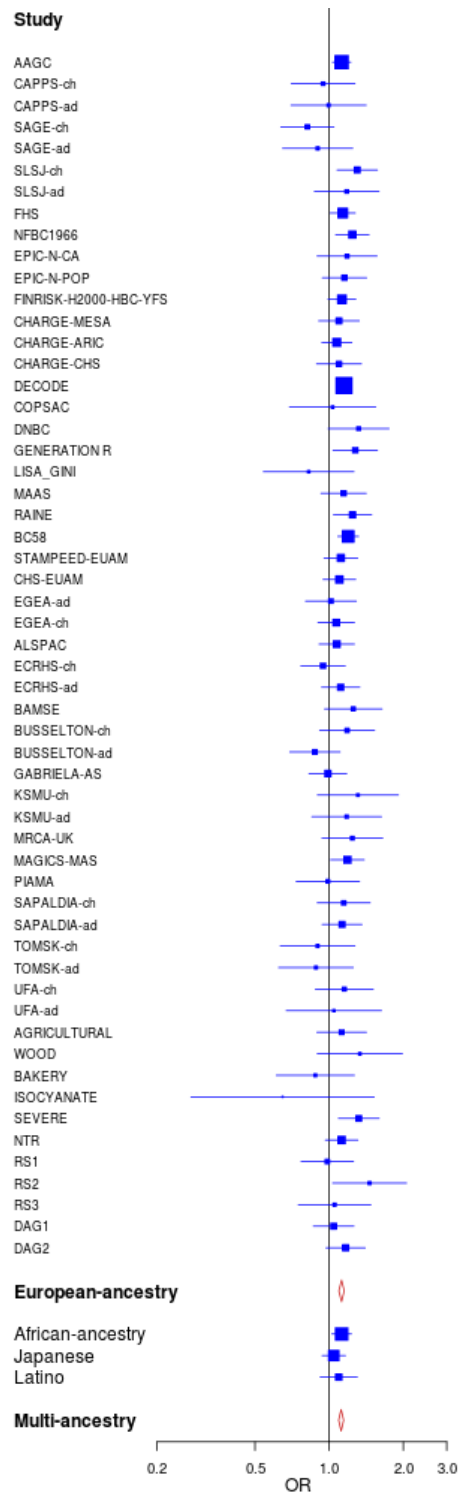
8q21.13, rs12543811



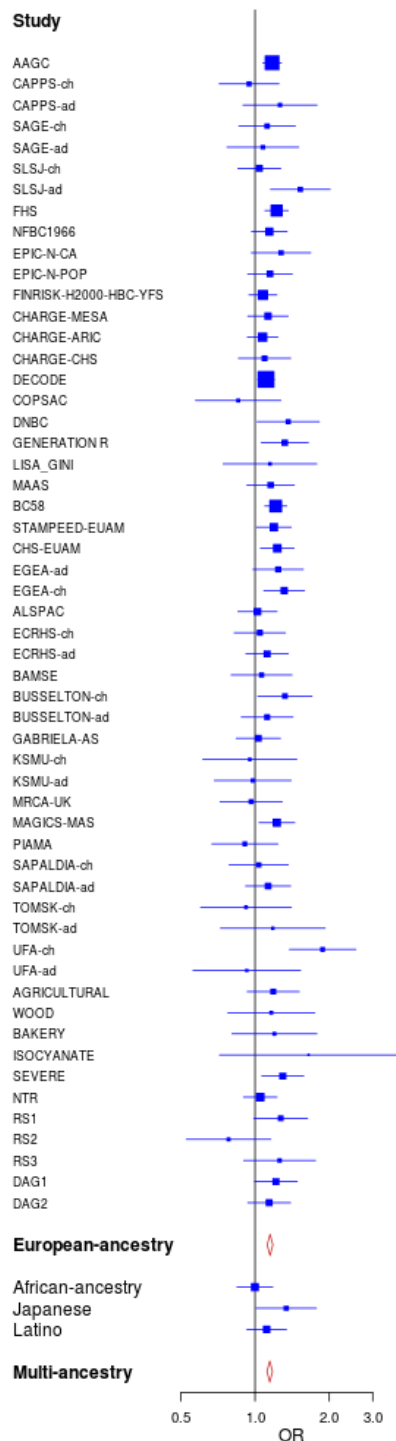
16p13.13, rs17806299



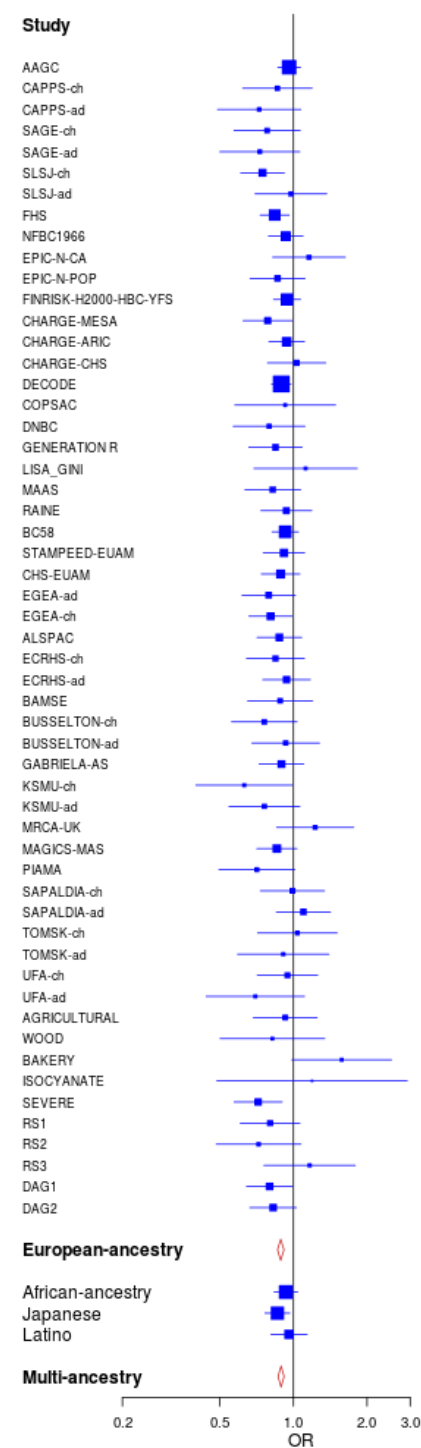
2q12, rs1420101



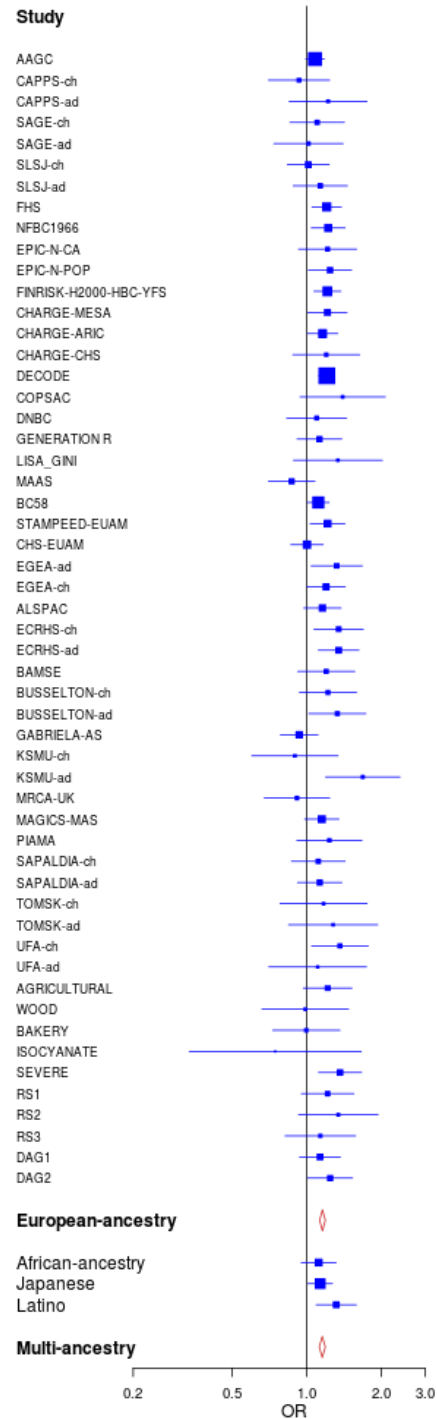
5q22.1, rs10455025



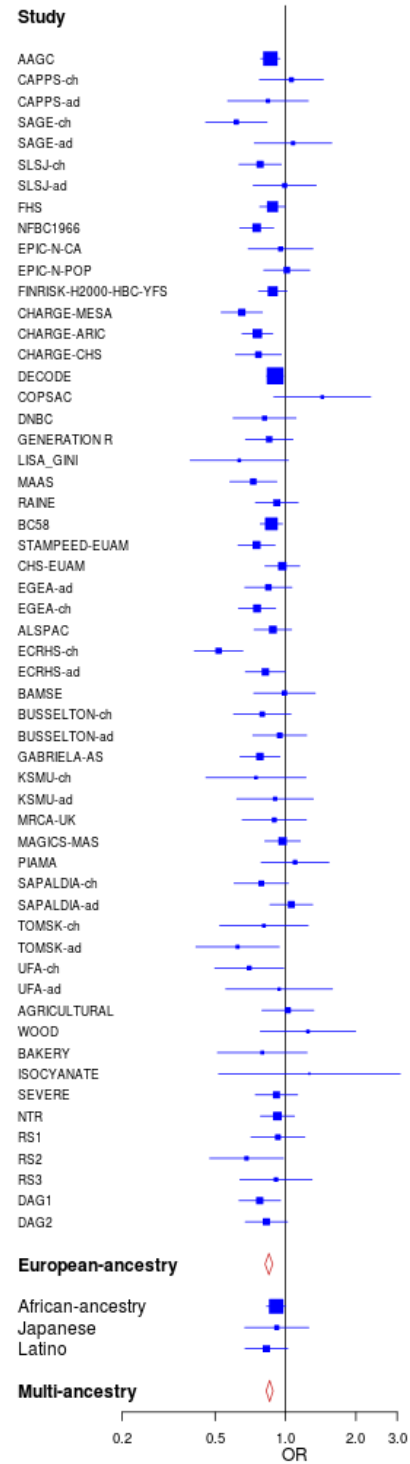
5q31, rs20541



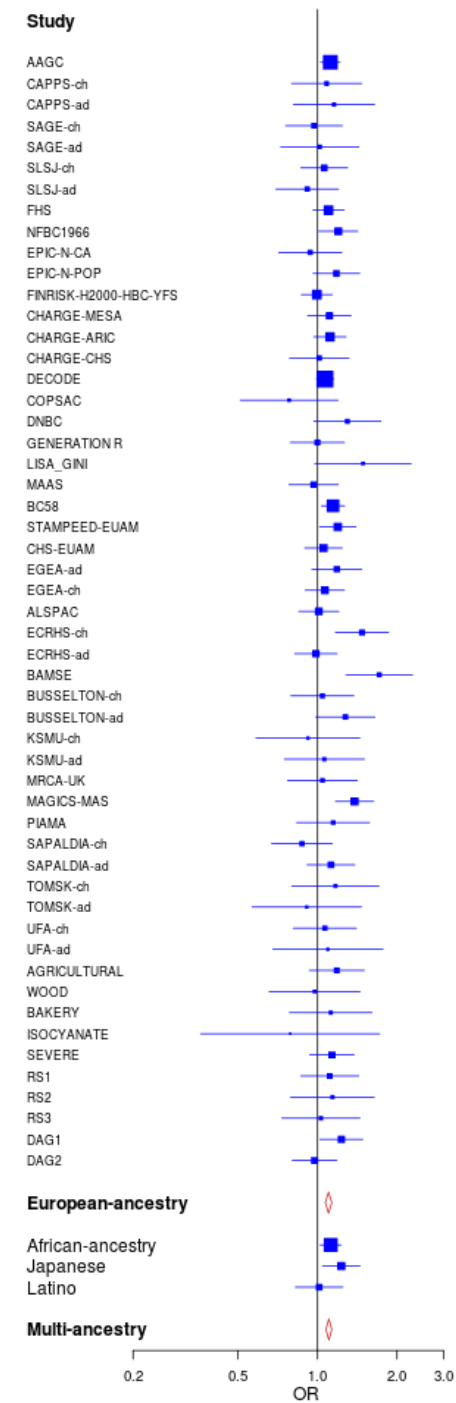
6p21.32, rs9272346



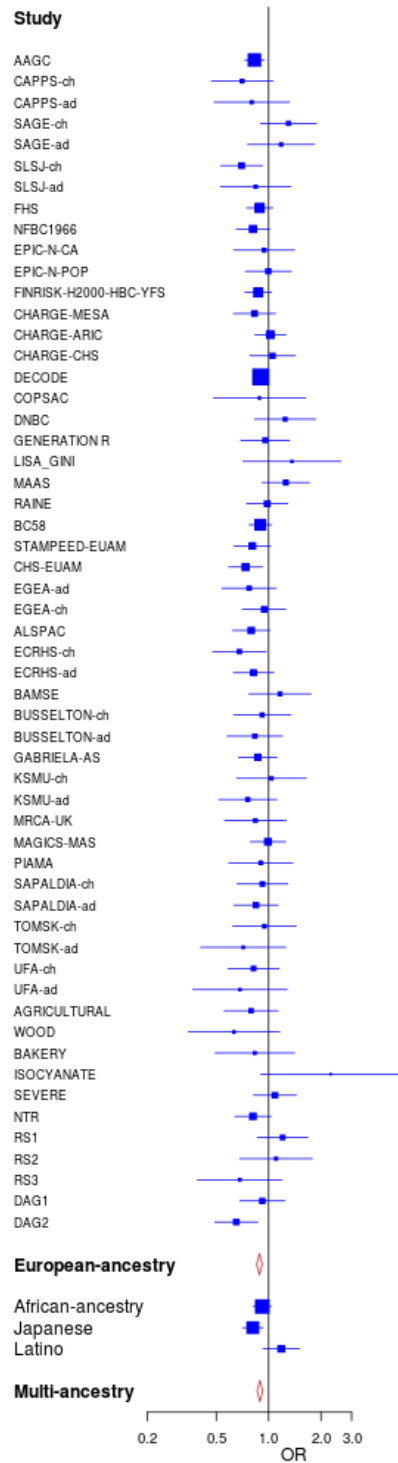
9p24.1, rs992969



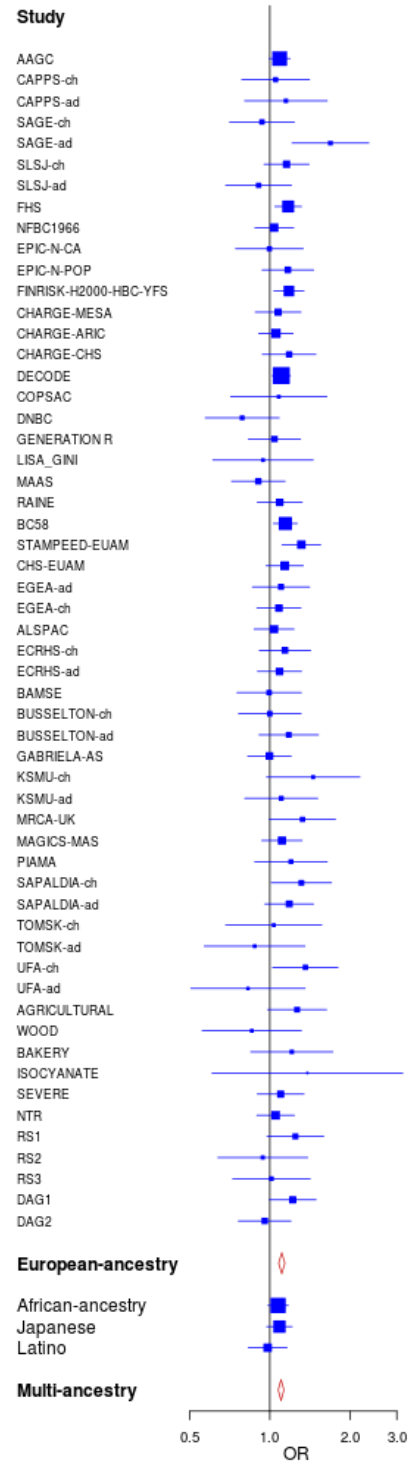
11q13.5, rs7927894



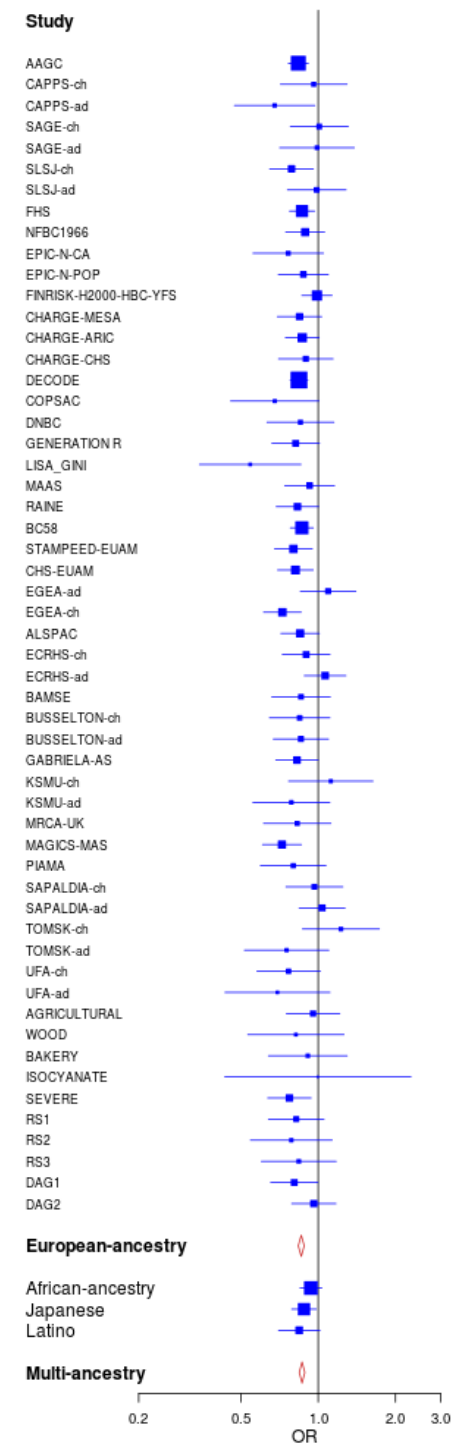
15q22.2, rs11071558



15q22.33, rs2033784

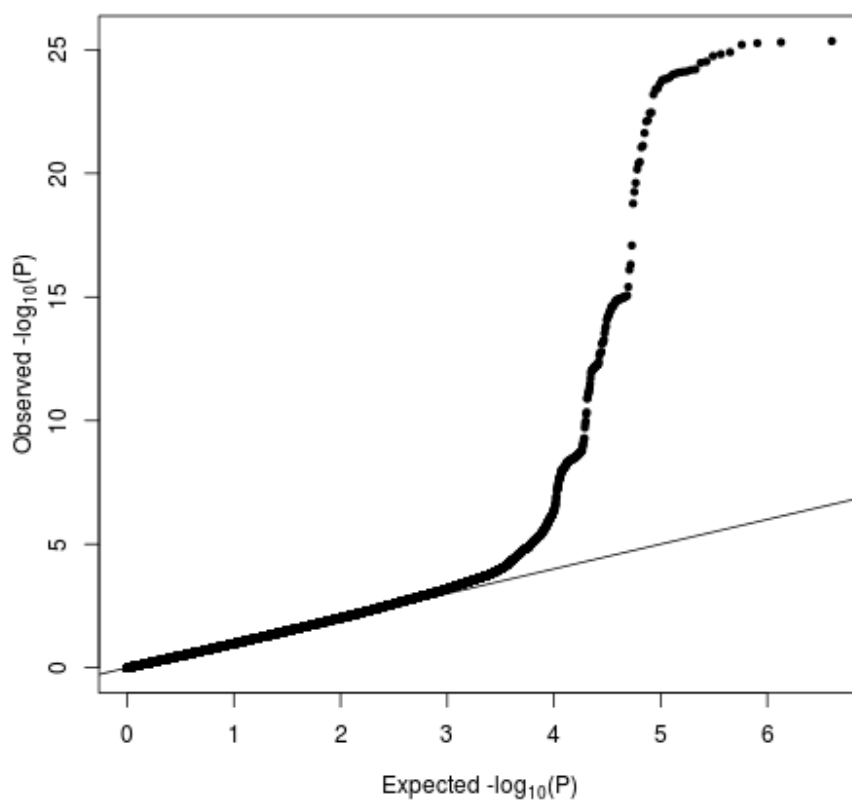


17q12-21, rs2952156



Supplementary Figure 5. Quantile-quantile (QQ) plot for the meta-analysis of pediatric asthma
QQ plots of P_{random} values for all SNPs passing QC in at least two-thirds of the studies
(27 studies; 8,976 cases, 18,399 controls)

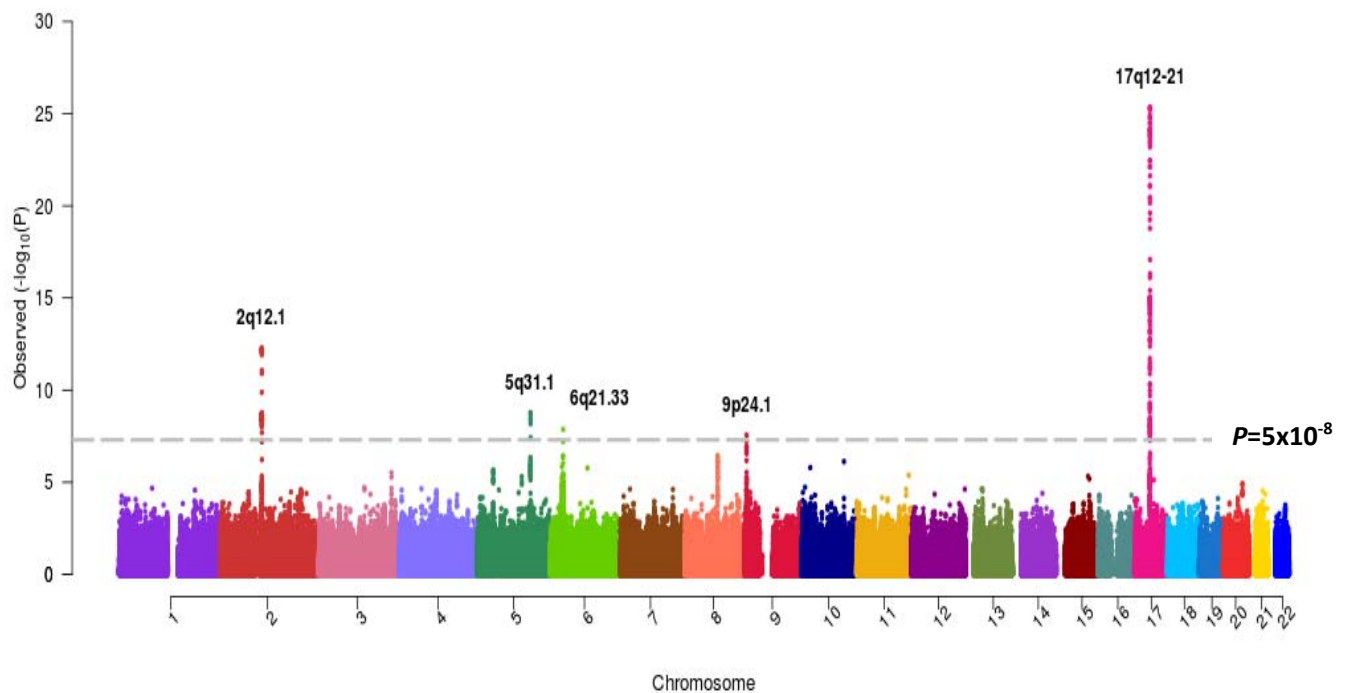
Lambda = 0.951



Supplementary Figure 6. Manhattan plot for the meta-analysis of pediatric asthma

(random-effects model)

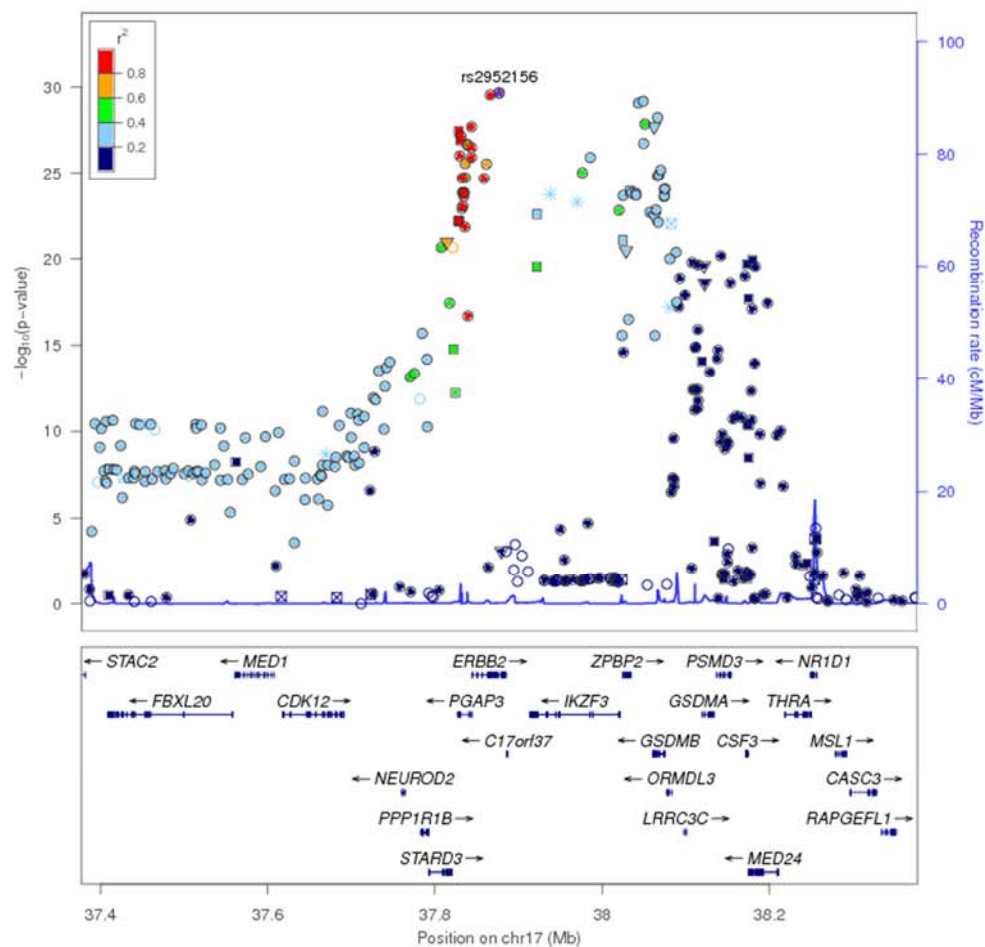
The x axis represents chromosomal location and the y axis represents $-\log_{10}(P \text{ value})$; the dashed horizontal line denotes $P = 5 \times 10^{-8}$. Each locus is annotated by its cytogenetic band location.



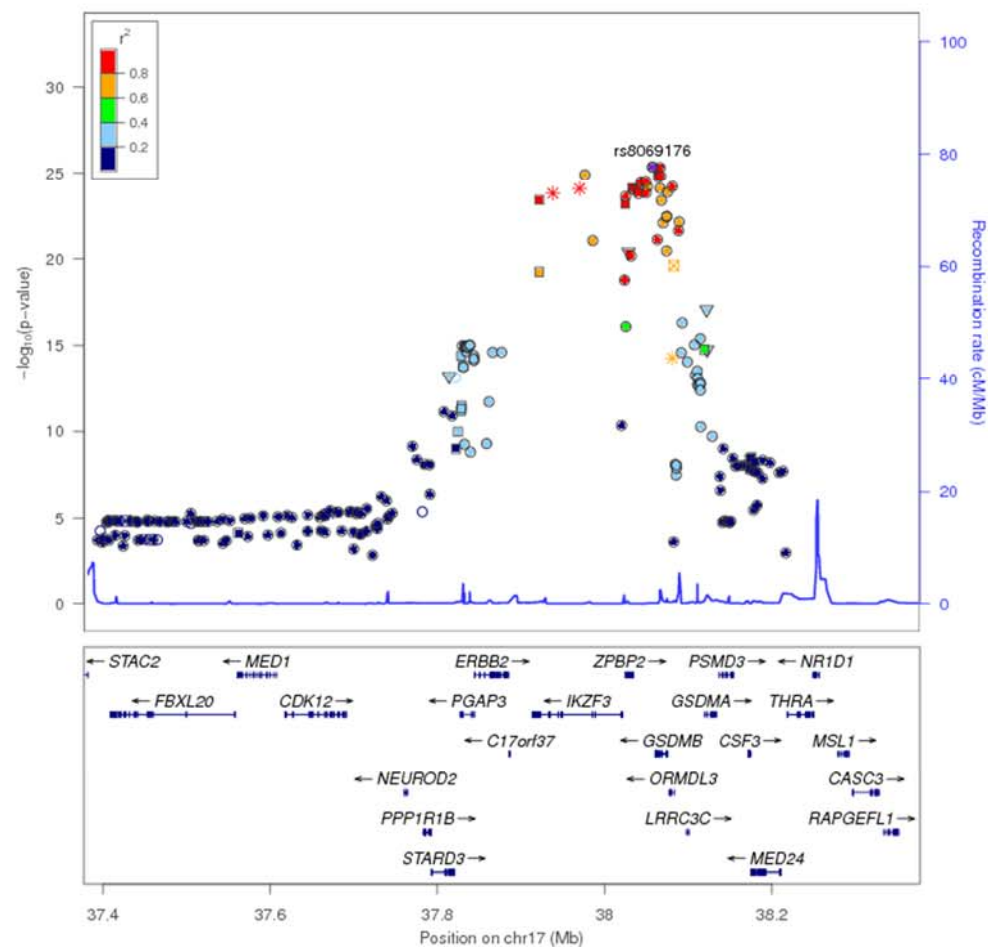
Supplementary Figure 7. Regional plots for the 17q12-21 locus in multi-ancestry and pediatric subgroup meta-analyses

The x axis represents physical location in megabase (build 37) and the Y axis represents $-\log_{10}$ (P value) for tests of association in either the multi-ancestry or pediatric subgroup meta-analysis (random-effects model) using LocusZoom (see URLs). The rs ID is shown for the lead SNP in the region (purple diamond). For remaining SNPs, the color indicates the r^2 measure of LD with the lead SNP.

Multi-ancestry meta-analysis



Pediatric subgroup meta-analysis



Supplementary Tables

Supplementary Tables 1 and 2 can be found in the Supplementary excel files

Supplementary Table 1. Description of TAGC studies included in the meta-analysis

Supplementary Table 2. Information on genotyping methods, imputation, and statistical analysis by study

Supplementary Table 3. Effect sizes of the lead SNPs in ancestry-specific and multi-ancestry meta-analyses at the 18 significant loci ($P_{\text{random}} \leq 5 \times 10^{-8}$)

Region ^a	SNP	Position Build 37	Allele (R/E) ^b	EUROPEAN ANCESTRY (56 studies)					AFRICAN ANCESTRY (7 studies)				JAPANESE (2 studies)				LATINOS (1 study)			MULTTI-ANCESTRY (66 studies)					P ⁱ _{ethnic}	
				EAF ^c	OR (95% CI) ^d	P _{random} ^e	P _{fixed} ^f	P _{het} ^g	EAF ^c	OR (95% CI) ^d	P _{fixed} ^f	P _{het} ^g	EAF ^c	OR (95% CI) ^d	P _{fixed} ^f	P _{het} ^g	EAF ^c	OR (95% CI) ^d	P ^h	OR (95% CI) ^d	P _{random} ^e	P _{fixed} ^f	P _{het} ^g			
New asthma susceptibility loci																										
5q31.3	rs7705042	141,492,419	C/A	0.63	1.08 (1.05-1.11)	1.6x10 ⁻⁶	8.5x10 ⁻¹⁰	0.07	0.77	1.12 (1.0-1.24)	0.04	0.39	0.63	1.17 (1.05-1.31)	5.5x10 ⁻³	0.98	0.67	1.10 (0.93-1.30)	0.27	1.09 (1.06-1.12)	7.9x10 ⁻⁹	3.1x10 ⁻¹²	0.11	0.55		
6p22.1	rs1233578	28,712,247	A/G	0.13	1.11 (1.07-1.15)	5.3x10 ⁻⁹	5.3x10 ⁻⁹	0.82	0.41	0.98 (0.90-1.07)	0.65	0.73	NA		NA		0.15	0.83 (0.66-1.05)	0.12	1.09 (1.05-1.12)	5.9x10 ⁻⁷	5.9x10 ⁻⁷	0.56	0.003		
6q15	rs2325291	90,986,686	G/A	0.33	0.91 (0.89-0.93)	8.6x10 ⁻¹³	8.6x10 ⁻¹³	0.78	0.12	1.02 (0.89-1.17)	0.78	0.49	0.03	1.02 (0.74-1.41)	0.88	0.69	0.30	0.91 (0.76-1.09)	0.30	0.91 (0.89-0.94)	2.2x10 ⁻¹²	2.2x10 ⁻¹²	0.80	0.39		
12q13.3	rs167769	57,503,775	C/T	0.40	1.08 (1.05-1.11)	1.6x10 ⁻⁷	5.5x10 ⁻⁹	0.19	0.14	1.14 (1.0-1.29)	0.05	0.71	0.21	1.09 (0.95-1.24)	0.22	0.18	0.37	1.04 (0.89-1.23)	0.60	1.08 (1.05-1.11)	3.9x10 ⁻⁹	4.6x10 ⁻¹⁰	0.31	0.85		
17q21.33	rs17637472	47,461,433	G/A	0.39	1.08 (1.05-1.11)	3.3x10 ⁻⁹	3.3x10 ⁻⁹	0.56	0.08	0.91 (0.77-1.09)	0.32	0.13	0.13	1.12 (0.95-1.32)	0.18	0.58	0.22	1.24 (1.03-1.50)	0.02	1.08 (1.05-1.11)	6.6x10 ⁻⁹	1.1x10 ⁻⁹	0.35	0.12		
New signals at loci previously associated with asthma in ancestry-specific populations																										
6p21.33	rs2855812	31,472,720	G/T	0.23	1.10 (1.06-1.13)	1.7x10 ⁻⁸	2.8x10 ⁻¹⁰	0.23	0.25	1.14 (1.03-1.26)	0.01	0.82	0.15	1.08 (0.93-1.25)	0.34	0.89	0.15	1.25 (1.01-1.55)	0.04	1.10 (1.07-1.13)	8.9x10 ⁻¹²	1.9x10 ⁻¹²	0.39	0.58		
10p14	rs2589561	9,046,645	A/G	0.82	0.90 (0.87-0.94)	1.4x10 ⁻⁸	1.4x10 ⁻⁸	0.78	0.86	0.99 (0.87-1.12)	0.83	0.74	0.81	0.84 (0.73-0.96)	9.8x10 ⁻³	0.60	0.89	1.06 (0.80-1.39)	0.69	0.91(0.88-0.94)	3.5x10 ⁻⁹	3.5x10 ⁻⁹	0.82	0.25		
Asthma signals previously reported for asthma plus hay fever																										
8q21.13	rs12543811	81,278,885	G/A	0.66	0.93 (0.91-0.95)	3.4x10 ⁻⁸	3.0x10 ⁻⁸	0.47	0.59	0.93 (0.85-1.02)	0.12	0.72	0.38	0.82 (0.74-0.92)	7.6x10 ⁻⁴	0.71	0.52	0.91 (0.78-1.07)	0.25	0.92 (0.9-0.95)	1.1x10 ⁻¹⁰	1.1x10 ⁻¹⁰	0.54	0.24		
16p13.13	rs17806299	11,199,980	G/A	0.20	0.90 (0.88-0.93)	2.1x10 ⁻¹⁰	2.1x10 ⁻¹⁰	0.51	0.06	0.94 (0.78-1.13)	0.50	0.20	0.09	0.92 (0.76-1.12)	0.40	0.78	0.11	1.08 (0.84-1.39)	0.56	0.91 (0.88-0.94)	2.7x10 ⁻¹⁰	2.7x10 ⁻¹⁰	0.49	0.58		
Knwon asthma loci																										
2q12	rs1420101	102,957,716	C/T	0.37	1.12 (1.10-1.15)	9.1x10 ⁻²⁰	9.1x10 ⁻²⁰	0.63	0.33	1.12 (1.03-1.23)	0.01	0.59	0.42	1.04 (0.94-1.16)	0.44	0.05	0.27	1.09 (0.92-1.30)	0.31	1.12 (1.09-1.15)	3.9x10 ⁻²¹	3.9x10 ⁻²¹	0.61	0.64		
5q22.1	rs10455025	110,404,999	A/C	0.34	1.15 (1.12-1.18)	2.0x10 ⁻²⁵	2.0x10 ⁻²⁵	0.53	0.09	1.0 (0.84-1.18)	0.99	0.55	0.04	1.34 (1.01-1.77)	0.04	0.85	0.25	1.11 (0.93-1.34)	0.25	1.15 (1.12-1.18)	9.4x10 ⁻²⁶	9.4x10 ⁻²⁶	0.57	0.27		
5q31	rs20541	131,995,964	A/G	0.79	0.89 (0.86-0.91)	1.4x10 ⁻¹⁴	1.4x10 ⁻¹⁴	0.73	0.82	0.93 (0.83-1.04)	0.21	0.43	0.69	0.86 (0.77-0.97)	0.01	0.58	0.56	0.96 (0.81-1.14)	0.63	0.89 (0.87-0.92)	5.0x10 ⁻¹⁶	5.0x10 ⁻¹⁶	0.77	0.62		
6p21.32	rs9272346	32,604,372	G/A	0.56	1.16 (1.13-1.19)	4.8x10 ⁻²⁸	2.4x10 ⁻²⁸	0.46	0.51	1.11 (1.02-1.21)	0.02	0.004	0.53	1.13 (1.01-1.27)	0.03	0.90	0.69	1.32 (1.09-1.59)	0.004	1.16 (1.12-1.19)	5.7x10 ⁻²⁴	8.2x10 ⁻³²	0.14	0.43		
9p24.1	rs992969	6,209,697	A/G	0.75	0.85 (0.82-0.88)	1.1x10 ⁻¹⁷	4.3x10 ⁻²⁹	0.008	0.70	0.91 (0.83-1.0)	0.05	0.60	0.97	0.92 (0.67-1.26)	0.59	0.66	0.83	0.83 (0.67-1.02)	0.08	0.86 (0.83-0.88)	7.2x10 ⁻²⁰	3.3x10 ⁻³⁰	0.02	0.57		
11q13.5	rs7927894	76,301,316	C/T	0.37	1.10 (1.07-1.13)	3.5x10 ⁻¹¹	3.2x10 ⁻¹²	0.38	0.35	1.12 (1.02-1.23)	0.01	0.86	0.13	1.23 (1.05-1.45)	0.01	0.88	0.24	1.01 (0.82-1.25)	0.89	1.10 (1.08-1.13)	2.2x10 ⁻¹⁴	2.2x10 ⁻¹⁴	0.56	0.47		
15q22.2	rs11071558	61,069,421	A/G	0.14	0.89 (0.85-0.92)	1.9x10 ⁻¹⁰	8.3x10 ⁻¹¹	0.44	0.32	0.92 (0.83-1.01)	0.08	0.21	0.23	0.81 (0.71-0.93)	2.6x10 ⁻³	0.07	0.12	1.18 (0.93-1.50)	0.18	0.89 (0.86-0.92)	1.3x10 ⁻⁹	3.1x10 ⁻¹²	0.19	0.06		
15q22.33	rs2033784	67,449,660	A/G	0.30	1.11 (1.08-1.14)	2.5x10 ⁻¹⁴	2.5x10 ⁻¹⁴	0.75	0.44	1.07 (0.98-1.17)	0.11	0.51	0.40	1.09 (0.97-1.21)	0.14	0.24	0.36	0.98 (0.83-1.16)	0.83	1.10 (1.08-1.13)	7.4x10 ⁻¹⁵	7.4x10 ⁻¹⁵	0.76	0.48		
17q12-21	rs2952156	37,876,835	A/G	0.70	0.86 (0.84-0.88)	7.6x10 ⁻²⁹	7.6x10 ⁻²⁹	0.55	0.60	0.94 (0.85-1.03)	0.17	0.36	0.51	0.88 (0.79-0.98)	0.02	0.37	0.57	0.84 (0.7-1.02)	0.07	0.87 (0.84-0.89)	2.2x10 ⁻³⁰	2.2x10 ⁻³⁰	0.52	0.35		
Each locus is represented by the variant with the strongest evidence for association in the European-ancestry or multi-ancestry meta-analysis (as shown in Table 1).																										
^a Cytogenetic band																										
^b R=reference allele /E=effect allele. ^c EAF = Effect allele frequency																										
^d Odds-ratios (ORs) and 95% Confidence Intervals (CI) were computed for the effect allele under a random-effects model (European ancestry and multi-ancestry meta-analyses) or fixed-effects model (African ancestry and Japanese meta-analysis) or in the Latino GWAS																										
^e P _{random} for test of association between SNP and asthma under a random-effects model																										
^f P _{fixed} for test of association between SNP and asthma under a fixed-effects model																										
^g P _{het} for test of heterogeneity across studies with the use of Cochran's Q test																										
^h P for test of association between SNP and asthma in the Latino GWAS																										
ⁱ P _{ethnic} for test of heterogeneity between ancestry-specific effect sizes																										
NA= the SNP did not pass QC in the Japanese studies																										

Supplementary Table 4 can be found in the **Supplementary excel files**

Supplementary Table 4. Genome-wide significant SNPs in the European-ancestry meta-analysis.

Supplementary Table 5. Genetic loci associated with asthma in the African-ancestry meta-analysis

Results of this analysis (2,149 cases, 6,055 controls) are shown for the lead SNP at each locus having $P_{\text{fixed}} < 10^{-5}$

Region ^a	Lead SNP	Position build 37	Nearby genes ^b	Allele (R/E) ^c	EAf ^d	OR (95%CI) ^e	P_{fixed} ^f	P_{het} ^g
1p3.1	rs10157802	74,110,649	<i>NEGR1, LRRIQ3</i>	G/A	0.82	0.78 (0.69-0.87)	7.3×10^{-6}	0.66
1q23.3	rs1348135	163,776,298	<i>NUF2, PBX1</i>	C/T	0.05	0.57 (0.44-0.73)	8.8×10^{-6}	0.28
1q42.3	rs10924970	235,362,960	<i>ARID4B, RBM34, GGPS1</i>	T/C	0.61	1.25 (1.14-1.37)	2.9×10^{-6}	0.58
2p23.3	rs4268898	24,490,413	<i>ITSN2, C2orf84, NCOA1</i>	C/T	0.56	0.79 (0.73-0.87)	1.9×10^{-7}	0.48
3p24.2	rs12634582	24,282,432	<i>THRB, NR1D2, RARB</i>	T/C	0.38	1.26 (1.14-1.40)	9.0×10^{-6}	0.88
3p22.3	rs9870718	36,239,406	<i>ARPP21, STAC</i>	C/T	0.36	0.81 (0.74-0.89)	8.7×10^{-6}	0.75
7p14.1	rs10233459	42,040,986	<i>GLI3, INHBA, C7orf25</i>	A/G	0.05	1.57 (1.30-1.91)	4.21×10^{-6}	0.51
11q23.2	rs4938096	114,235,730	<i>NNMT, C11orf71</i>	C/T	0.21	1.30 (1.17-1.44)	1.1×10^{-6}	0.31
12q24.33	rs10773588	129,511,610	<i>GLT1D1, TMEM132D</i>	G/T	0.06	0.59 (0.47-0.74)	4.5×10^{-6}	0.29
14q12	rs10141207	32,691,785	<i>ARHGAP5, AKAP6</i>	G/A	0.35	1.24 (1.13-1.36)	3.9×10^{-6}	0.40
20q13.12	rs16989837	43,859,099	<i>SLP1, SEMG2</i>	C/T	0.10	1.39 (1.21-1.60)	3.1×10^{-6}	0.13

^aCytogenetic band

^bGenes coding for proteins; the gene where eventually the lead SNP lies is first indicated followed by the previous gene and next gene

^cR=reference allele /E=effect allele; ^dEAf = Effect allele frequency

^eOdds-ratios (ORs) and 95% Confidence Intervals (CI) were computed for the effect allele under a fixed-effects model

^f P_{fixed} for test of association between SNP and asthma under a fixed-effects model

^g P_{het} for test of heterogeneity across studies with the use of Cochran's Q test

Supplementary Table 6. Genetic loci associated with asthma in the Japanese meta-analysis

Results of this analysis (1,239 cases, 3,976 controls) are shown for the lead SNP at each locus having $P_{\text{fixed}} < 10^{-5}$

Region ^a	Lead SNP	Position build 37	Nearby genes ^b	Allele (R/E) ^c	EAf ^d	OR (95%CI) ^e	P_{fixed} ^f	P_{het} ^g
1q32.1	rs1122396	201,138,428	<i>TMEM9,IGFN1</i>	G/A	0.51	0.76 (0.68-0.85)	1.5×10^{-6}	0.83
7p14.3	rs10951405	34,043,348	<i>BMPER,BBS9,NPSR1</i>	C/T	0.89	1.44 (1.24-1.68)	2.8×10^{-6}	0.28
11q22.1	rs3758697	100,709,894	<i>ARHGAP42,CNTN5,TMEM133</i>	G/A	0.30	1.33 (1.18-1.50)	2.4×10^{-6}	0.20
20p13	rs6084352	3,348,407	<i>C20orf194,SLC4A11,ATRN</i>	A/G	0.30	1.31 (1.16-1.47)	6.6×10^{-6}	0.77

^aCytogenetic band

^bGenes coding for proteins; the gene where eventually the lead SNP lies is first indicated followed by the previous gene and next gene

^cR=reference allele /E=effect allele. ^dEAf = Effect allele frequency

^eOdds-ratios (ORs) and 95% Confidence Intervals (CI) were computed for the effect allele under a fixed-effects model

^f P_{fixed} for test of association between SNP and asthma under a fixed-effects model

^g P_{het} for test of heterogeneity across studies using the Cochran's Q test

Supplementary Table 7. Genetic loci associated with asthma in Latinos

Results of this analysis (606 cases, 792 controls) are shown for the lead SNP at each locus having $P < 10^{-5}$

Region ^a	Lead SNP	Position build 37	Nearby genes ^b	Allele (R/E) ^c	EAf ^d	OR (95%CI) ^e	P^f
1p31.1	rs2352521	79,256,105	<i>IFI44,ELTD1</i>	C/T	0.72	1.66 (1.33-2.06)	4.9×10^{-6}
1q31.3	rs6694672	196,945,789	<i>CFHR2,CFHR5</i>	T/G	0.05	2.13 (1.52-2.97)	9.4×10^{-6}
5p13.2	rs12521260	37,881,414	<i>GDNF,EGFLAM</i>	G/T	0.28	1.52 (1.27-1.81)	3.0×10^{-6}
8q23.3	rs12542922	117,333,797	<i>TRPS1,EIF3H</i>	A/G	0.28	0.61 (0.49-0.74)	1.7×10^{-6}
8q24.12	rs2073617	119,964,283	<i>TNFRSF11B,SAMD12,COLEC10</i>	G/A	0.37	1.46 (1.24-1.71)	4.7×10^{-6}
19q13.31	rs4802207	44,600,377	<i>ZNF224,ZNF284,ZNF225</i>	C/T	0.59	0.68 (0.57-0.81)	9.5×10^{-6}
21q21.1	rs2825968	21,415,693	<i>TMPRSS15,NCAM2</i>	T/C	0.51	0.69 (0.59-0.81)	5.8×10^{-6}

^aCytogenetic band

^bGenes coding for proteins; the gene where eventually the lead SNP lies is first indicated followed by the previous gene and next gene

^cR=reference allele /E=effect allele. ^dEAf = Effect allele frequency

^eOdds-ratios (ORs) and 95% Confidence Intervals (CI) were computed for the effect allele

^f P for test of association between SNP and asthma

Supplementary Table 8 can be found in the Supplementary excel files

Supplementary Table 8. Genome-wide significant SNPs in the multi-ancestry meta-analysis.

Supplementary Table 9. Tests of association of the 18 lead SNPs with asthma with and without Genomic Control (GC) correction

Region ^a	Lead SNP	Position ^b	Nearby genes ^c	Allele (R/F) ^d	European-ancestry analysis			Multi-ancestry analysis	
					EAf ^e	<i>P</i> _{random} without GC correction ^f	<i>P</i> _{random} with GC correction ^g	<i>P</i> _{random} without GC correction ^f	<i>P</i> _{random} with GC correction ^g
New asthma susceptibility loci									
5q31.3	rs7705042	141,492,419	<i>NDFIP1, GNDPA1, SPRY4</i>	C/A	0.63	1.6 × 10 ⁻⁶	2.3 × 10 ⁻⁶	7.9 × 10 ⁻⁹	1.7 × 10 ⁻⁸
6p22.1	rs1233578	28,712,247	<i>GPX5, TRIM27</i>	A/G	0.13	5.3 × 10 ⁻⁹	9.0 × 10 ⁻⁹	5.9 × 10 ⁻⁷	1.0 × 10 ⁻⁶
6q15	rs2325291	90,986,686	<i>BACH2, GJA10, MAP3K7</i>	G/A	0.33	8.6 × 10 ⁻¹³	1.9 × 10 ⁻¹²	2.2 × 10 ⁻¹²	6.6 × 10 ⁻¹²
12q13.3	rs167769	57,503,775	<i>STAT6, NAB2, LRP1</i>	C/T	0.40	1.6 × 10 ⁻⁷	2.4 × 10 ⁻⁷	3.9 × 10 ⁻⁹	8.5 × 10 ⁻⁹
17q21.33	rs17637472	47,461,433	<i>ZNF652, PHB</i>	G/A	0.39	3.3 × 10 ⁻⁹	5.6 × 10 ⁻⁹	6.6 × 10 ⁻⁹	1.4 × 10 ⁻⁸
New signals at loci previously associated with asthma in ancestry specific populations									
6p21.33	rs2855812	31,472,720	<i>MICB, HCP5, MCCD1</i>	G/T	0.23	1.7 × 10 ⁻⁸	2.7 × 10 ⁻⁸	8.9 × 10 ⁻¹²	2.5 × 10 ⁻¹¹
10p14	rs2589561	9,046,645	<i>GATA3, CELF2</i>	A/G	0.82	1.4 × 10 ⁻⁸	2.4 × 10 ⁻⁸	3.5 × 10 ⁻⁹	7.7 × 10 ⁻⁹
Asthma signals previously reported for asthma plus hay fever									
8q21.13	rs12543811	81,278,885	<i>TPD52, ZBTB10</i>	G/A	0.66	3.4 × 10 ⁻⁸	5.5 × 10 ⁻⁸	1.1 × 10 ⁻¹⁰	2.8 × 10 ⁻¹⁰
16p13.13	rs17806299	11,199,980	<i>CLEC16A, DEXI, SOCS1</i>	G/A	0.20	2.1 × 10 ⁻¹⁰	3.9 × 10 ⁻¹⁰	2.7 × 10 ⁻¹⁰	6.5 × 10 ⁻¹⁰
Known asthma loci									
2q12	rs1420101	102,957,716	<i>IL1RL1, IL1RL2, IL18R1</i>	C/T	0.37	9.1 × 10 ⁻²⁰	3.2 × 10 ⁻¹⁹	3.9 × 10 ⁻²¹	2.8 × 10 ⁻²⁰
5q22.1	rs10455025	110,404,999	<i>SLC25A46, TSLP</i>	A/C	0.34	2.0 × 10 ⁻²⁵	1.1 × 10 ⁻²⁴	9.4 × 10 ⁻²⁶	1.1 × 10 ⁻²⁴
5q31	rs20541	131,995,964	<i>IL13, RAD50, IL4</i>	A/G	0.79	1.4 × 10 ⁻¹⁴	3.4 × 10 ⁻¹⁴	5.0 × 10 ⁻¹⁶	2.2 × 10 ⁻¹⁵
6p21.32	rs9272346	32,604,372	<i>HLA-DRB1, HLA-DQA1</i>	G/A	0.56	4.8 × 10 ⁻²⁸	3.0 × 10 ⁻²⁷	5.7 × 10 ⁻²⁴	5.5 × 10 ⁻²³
9p24.1	rs992969	6,209,697	<i>RANBP6, IL33</i>	A/G	0.75	1.1 × 10 ⁻¹⁷	3.3 × 10 ⁻¹⁷	7.2 × 10 ⁻²⁰	4.6 × 10 ⁻¹⁹
11q13.5	rs7927894	76,301,316	<i>C11orf30, LRRC32</i>	C/T	0.37	3.5 × 10 ⁻¹¹	6.9 × 10 ⁻¹¹	2.2 × 10 ⁻¹⁴	8.2 × 10 ⁻¹⁴
15q22.2	rs11071558	61,069,421	<i>RORA, NARG2, VPS13C</i>	A/G	0.14	1.9 × 10 ⁻¹⁰	3.5 × 10 ⁻¹⁰	1.3 × 10 ⁻⁹	2.9 × 10 ⁻⁹
15q22.33	rs2033784	67,449,660	<i>SMAD3, SMAD6, AAGAB</i>	A/G	0.30	2.5 × 10 ⁻¹⁴	6.0 × 10 ⁻¹⁴	7.4 × 10 ⁻¹⁵	2.9 × 10 ⁻¹⁴
17q12-21	rs2952156	37,876,835	<i>ERBB2, PGAP3, C17orf37</i>	A/G	0.70	7.6 × 10 ⁻²⁹	5.0 × 10 ⁻²⁸	2.2 × 10 ⁻³⁰	4.1 × 10 ⁻²⁹

Each asthma risk locus is represented by the variant with the strongest evidence for association in the European-ancestry or multi-ancestry analysis (as shown in Table 1)

^aCytogenetic band

^bSNP Position, build 37.

^cGenes coding for proteins; the gene where eventually the SNP lies is first indicated followed by the previous gene and next gene.

^dR=reference allele /E=effect allele.

^eEAf=Effect allele frequency

^f*P*-value for test of association of SNP with asthma (random-effects model) without GC correction.

^g*P*-value for test of association of SNP with asthma (random-effects model) with GC correction, the lambda parameter is 1.031 in the European-ancestry analysis and 1.046 in the multi-ancestry analysis

Supplementary Table 10. Genetic loci associated with pediatric asthma

Results of this analysis (8,976 cases, 18,399 controls) are shown for the lead SNP at genome-wide significant loci ($P_{\text{random}} < 5 \times 10^{-8}$)

Region ^a	Lead SNP	Position build 37	Nearby Genes ^b	Alleles (R/E) ^c	OR (95%CI) ^d	P_{random} ^e	P_{fixed} ^f	P_{het} ^g
2q12.1	rs4988958	102,968,285	<i>IL1RL1, IL1RL2, IL18R1</i>	T/C	0.85 (0.82-0.89)	5.4×10^{-13}	5.4×10^{-13}	0.78
5q31.1	rs1295685	131,996,445	<i>IL13, RAD50, IL4</i>	A/G	0.86 (0.82-0.9)	1.7×10^{-9}	1.7×10^{-9}	0.89
6p21.33	rs2596464	31,412,961	<i>MICA, MICB</i>	T/C	1.12 (1.08-1.17)	1.4×10^{-8}	1.4×10^{-8}	0.47
9p24.1	rs12551256	6,231,239	<i>IL33, RANBP6, TPD52L3</i>	A/G	0.89 (0.86-0.93)	2.8×10^{-8}	2.8×10^{-8}	0.51
17q12-21	rs8069176	38,057,197	<i>ZPBP2, GSDMB</i>	G/A	0.80 (0.77-0.83)	4.4×10^{-26}	4.4×10^{-26}	0.77

^aCytogenetic band

^bGenes coding for proteins; the gene where eventually the lead SNP lies is first indicated followed by the previous gene and next gene

^cR=reference allele /E=effect allele.

^dOdds-ratios (ORs) and 95% Confidence Intervals (CI) were computed for the effect allele under a random-effects model

^e P_{random} for test of association between SNP and asthma under a random-effects model

^f P_{fixed} for test of association between SNP and asthma under a fixed-effects model

^g P_{het} for test of heterogeneity across studies with the use of Cochran's Q test

Supplementary Table 11 can be found in the **Supplementary excel files**

Supplementary Table 11. Association of 17q12-21 SNPs with asthma in multi-ancestry and pediatric meta-analyses

Supplementary Table 12. Association between cis-gene transcripts in blood and lung and 17q12-21 SNPs

a. Association between gene transcripts and lead asthma SNPs (rs2952156 in multi-ancestry meta-analysis and rs8069176 in pediatric meta-analysis)

Asthma SNP	Position build 37	Nearby genes ^a	Source database ^b	Tissue ^c	Gene transcript	$P_{\text{transcript}}^d$	Gene transcript	$P_{\text{transcript}}^d$	Gene transcript	$P_{\text{transcript}}^d$	Gene transcript	$P_{\text{transcript}}^d$	PMID for eQTL database
rs2952156	37,876,835	<i>ERBB2</i>	Blood_eQTL	Blood	<i>ORMDL3</i>	5.2×10^{-109}	<i>GSDMB</i>	3.7×10^{-90}	<i>PPP1R1B</i>	1.8×10^{-8}			24013639
		<i>PGAP3</i>	GTEx	Blood	<i>GSDMB</i>	1.1×10^{-9}	<i>ORMDL3</i>	3.7×10^{-6}					25954001
		<i>C17orf37</i>	MuTHER	LCLs	<i>ORMDL3</i>	3.3×10^{-16}	<i>ZBP2</i>	3.3×10^{-11}	<i>GSDMB</i>	9.9×10^{-7}			22941192
			MRCA	LCLs	<i>ORMDL3</i>	1.1×10^{-13}	<i>GSDMB</i>	9.3×10^{-8}					23345460
			MRCE	LCLs	<i>GSDMB</i>	7.7×10^{-17}	<i>ORMDL3</i>	5.6×10^{-15}					23345460
			GTEx	Lung	<i>PGAP3</i>	1.0×10^{-9}	<i>ORMDL3</i>	1.5×10^{-7}					25954001
			Hao, 2012	Lung	<i>CDK12</i>	2.0×10^{-20}	<i>GSDMA</i>	1.3×10^{-13}	<i>PGAP3</i>	1.1×10^{-8}	<i>GSDMB</i>	3.9×10^{-8}	23209423
rs8069176	38,057,197	<i>ZBP2</i>	Blood_eQTL	Blood	<i>ORMDL3</i>	9.8×10^{-198}	<i>GSDMB</i>	9.8×10^{-198}					24013639
		<i>GSDMB</i>	GTEx	Blood	<i>ORMDL3</i>	1.2×10^{-26}	<i>GSDMB</i>	3.1×10^{-22}					25954001
			MuTHER	LCLs	<i>ORMDL3</i>	2.6×10^{-29}	<i>ZBP2</i>	5.5×10^{-14}	<i>GSDMB</i>	1.9×10^{-12}			22941192
			MRCA	LCLs	<i>ORMDL3</i>	1.0×10^{-24}	<i>GSDMB</i>	4.3×10^{-15}					23345460
			MRCE	LCLs	<i>GSDMB</i>	1.0×10^{-33}	<i>ORMDL3</i>	2.7×10^{-29}					23345460
			GTEx	Lung	<i>ORMDL3</i>	3.6×10^{-9}	<i>GSDMA</i>	1.6×10^{-8}					25954001
			Hao, 2012	Lung	<i>GSDMA</i>	2.2×10^{-31}	<i>GSDMB</i>	3.8×10^{-12}	<i>CDK12</i>	3.0×10^{-7}	<i>ORMDL3</i>	1.6×10^{-5}	23209423

^aGenes coding for proteins; the gene where eventually the SNP lies is first indicated, followed by the previous and next genes.

^beQTL databases are described in Online Methods

^cBlood is whole blood; LCL = lymphoblastoid cell lines; Lung is whole lung tissue

^d $P_{\text{transcript}}$ for test of association of SNP with gene transcript

b. Association analysis of gene transcripts with lead asthma SNPs (rs2952156, rs8069176) and peak SNPs for gene transcripts

This analysis used the summary statistics from GTEx eQTL-database, available in both blood (n=338) and lung (n=278)

To assess whether the asthma-associated SNP accounted for a substantial proportion of the association of the peak SNP with the transcript, we performed approximate conditional analysis, using the GCTA software (PMID 23756893), by conditioning the peak SNP for a particular gene transcript on the asthma-associated SNP (we used the ECRHS dataset as the reference sample to estimate LD).

As can be seen in this table (results in bold), rs8069176 accounts for a large part of the association of the peak SNP (rs9303281) with *ORMDL3* transcript in blood while rs2952156 accounts for a large part of the association of the peak SNP (rs2941505) with *PGAP3* transcript in lung.

Tissue	Gene transcript	Association of lead asthma SNP with transcript			Association of peak SNP with transcript				
		Asthma SNP	Position	$P_{\text{transcript}}^a$	Peak SNP ^b	Position	r^2 ^c	$P_{\text{transcript}}^d$	P_{adj}^e
Blood	<i>ORMDL3</i>	rs2952156	37,876,835	3.7×10^{-6}	rs9303281	38,074,046	0.39	6.3×10^{-30}	7.6×10^{-23}
		rs8069176	38,057,197	1.2×10^{-26}	rs9303281	38,074,046	0.79	6.3×10^{-30}	2.2×10^{-4}
	<i>GSDMB</i>	rs2952156	37,876,835	1.1×10^{-9}	rs8067378	38,051,348	0.46	1.0×10^{-29}	1.0×10^{-18}
		rs8069176	38,057,197	3.1×10^{-22}	rs8067378	38,051,348	0.81	1.0×10^{-29}	8.6×10^{-8}
Lung	<i>PGAP3</i>	rs2952156	37,876,835	1.0×10^{-9}	rs2941505	37,832,704	0.84	5.6×10^{-12}	1.5×10^{-3}
		rs8069176	38,057,197	1.4×10^{-3}	rs2941505	37,832,704	0.35	5.6×10^{-12}	2.6×10^{-9}
	<i>ORMDL3</i>	rs2952156	37,876,835	1.5×10^{-7}	rs3816470	37,985,801	0.45	6.1×10^{-15}	2.2×10^{-8}
		rs8069176	38,057,197	3.6×10^{-9}	rs3816470	37,985,801	0.74	6.1×10^{-15}	5.2×10^{-7}
	<i>GSDMA</i>	rs2952156	37,876,835	1.3×10^{-3}	rs3859192	38,128,648	0.13	1.5×10^{-30}	1.4×10^{-24}
		rs8069176	38,057,197	1.6×10^{-8}	rs3859192	38,128,648	0.28	1.5×10^{-30}	4.8×10^{-20}

^aP-value for association of the asthma SNP with gene transcript

^bMost significant SNP (peak SNP) associated with gene transcript

^cCorrelation (r^2) between the asthma SNP and the peak SNP

^dP-value for association of the peak SNP with gene transcript; ^eP-value for the peak SNP after conditioning on the asthma SNP

Supplementary Table 13. Asthma loci reported in the GWAS catalog ($P < 5 \times 10^{-8}$) that did not reach genome-wide significance in the TAGC meta-analyses

				TAGC meta-analyses																											
				Multi-ancestry			European-ancestry			African-ancestry		Japanese		Latino	Pediatric																
Effect allele frequency ^b				Prand ^c	Pfix ^d	Phet ^e	Prand ^c	Pfix ^d	Phet ^e	Pfix ^d	Phet ^e	Pfix ^d	Phet ^e	P ^d	Prand ^c	Pfix ^d	Phet ^e	Published study ancestry, no. cases				Number and Proportion of published study cases in TAGC ^f									
REGION	Position build 37	TAGC SNP	EA ^a															EU	AF	JAP	LAT	Disc-covery	Repli-cation	Total # ca in TAGC	Multi-ancestry (%)	% ca EU	% ca AF	% ca JAP	% ca LAT		
7q22.3	105,658,451	rs6967330	G	0.16	0.22	0.10	0.12	9x10 ⁻⁴	2x10 ⁻⁴	0.03	3x10 ⁻³	6x10 ⁻⁴	0.006	0.37	0.99	0.21	0.78	NA	2x10 ⁻³	3x10 ⁻⁵	0.008	EU	8,414	1,173	7,241	6,876	0.29	0.34	NA	NA	NA
10q24.2	100,739,770	rs17111026*	G	0.02	0.02	0.17	0.02	0.16	0.16	0.65	0.02	0.02	0.77	0.45	0.29	0.87	0.32	0.91	0.72	0.74	0.41	EU AF, LAT	527	429	98	0	0	0	0	NA	0
4q12	59,213,845	rs17218161	T	0.01	0.00	0.10	0.03	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA										
9p23	12,521,826	rs16929097	G	0.01	0.10	0.00	0.02	0.81	0.81	0.56	0.62	0.67	0.37	0.31	0.91	NA	NA	0.85	0.13	0.13	0.81	EU	15,797	12,475	3,322	12,687	0.53	0.64	NA	NA	NA
1q21.3	154,426,264	rs4129267	C	0.35	0.15	0.36	0.54	8x10 ⁻⁷	8x10 ⁻⁸	0.19	5x10 ⁻⁶	6x10 ⁻⁷	0.19	0.92	0.88	2x10 ⁻³	0.13	0.68	0.08	0.02	0.05	EU, AF, LAT	7,346	359	6,987	3,011	0.13	0.11	0.21	NA	0.58
5q12.1	59,369,794	rs1588265	A	0.34	0.25	0.72	0.26	0.55	0.55	0.50	0.65	0.65	0.66	0.80	0.32	0.95	0.01	0.42	0.59	0.64	0.25	JAP	7,171	1,532	5,639	301	0.01	NA	NA	0.24	NA
12q13.2	56,412,487	rs1701704	T	0.31	0.16	0.31	0.17	2x10 ⁻³	5x10 ⁻⁶	0.01	5x10 ⁻³	6x10 ⁻⁵	0.03	0.67	0.80	5x10 ⁻⁴	0.05	0.22	0.21	0.08	0.01	JAP, KOR	2,591	938	1,653	938	0.04	NA	NA	0.60	NA
4q31.21	144,003,159	rs7686660	T	0.20	0.44	0.75	0.52	0.88	0.84	0.23	0.99	0.95	0.17	0.86	0.44	0.69	0.18	0.53	0.19	0.13	0.17										
22q12.3	37,534,034	rs2284033	G	0.42	0.39	0.59	0.37	8x10 ⁻⁵	1x10 ⁻⁴	0.15	5x10 ⁻⁵	8x10 ⁻⁵	0.11	0.58	0.33	0.49	0.62	0.52	2x10 ⁻³	2x10 ⁻³	0.82	EU	10,365	10,365	NA	8,082	0.34	0.41	NA	NA	NA
8q24.11	118,025,645	rs3019885	T	0.48	0.18	0.36	0.58	0.99	0.89	0.05	0.81	0.73	0.07	0.11	0.37	0.47	0.61	0.09	0.86	0.74	0.06	EU, AF	3,377	793	2,584	1,088	0.05	0.03	0.21	NA	NA
1q31.3	197,325,908	rs2786098	T	0.75	0.92	0.79	0.87	0.87	0.87	0.81	0.49	0.49	0.92	0.96	0.32	0.15	0.08	0.23	0.83	0.83	0.62	AF	3,759	1,612	2,147	1,257	0.05	NA	0.58	NA	NA
1q23.1	158,932,555	rs1101999	C	1.00	0.75	1.00	0.98	NA	NA	NA	NA	NA	NA	1x10 ⁻⁴	0.26	NA	NA	NA	NA	NA	NA	AF	3,759	1,612	2,147	1,257	0.05	NA	0.58	NA	NA

The five published loci with P-values $\leq 10^{-4}$ in at least one TAGC meta-analysis are shown in bold

Note: Individual studies do not have always the same sample size in the original publication and TAGC.

^aEA = effect allele as used in TAGC analyses

^bThe effect allele frequency, shown here, is from the populations of the 1000 Genomes project.

^cPrand for test of association between SNP and asthma under a random-effects model (NA when the SNP did not pass QC)

^dPfix for test of association between SNP and asthma under a fixed-effects model; P is the P-value for test of association in Latino (NA when the SNP did not pass QC).

^ePhet for test of heterogeneity across studies with the use of Cochran's Q test (NA when the SNP did not pass QC).

^fTotal no. of cases from the published study included in TAGC and proportion of cases from the published study among TAGC cases according to ancestry (NA when an ethnic group was not part of the published study).

*rs17111026 is a proxy ($r^2=1$) of the SNP reported in the GWAS catalog (rs12570188); no proxy that passed QC was found for the 4q12 SNP (rs17218161).

EU=European-ancestry, AF= African-ancestry, JAP=Japanese, LAT= Latino, KOR=Koreans.

			GWAS Catalog				
REGION	Position build 37	TAGC SNP	PMID	Studied Trait	catalog.rs	catalog.pval	reported gene
7q22.3	105,658,451	rs6967330	24241537	Asthma (child)	rs6967330	2.7x10 ⁻¹⁴	CDHR3
10q24.2	100,739,770	rs17111026*	23829686	Asthma (child)	rs12570188	5.0 x10 ⁻⁸	HPSE2
4q12	59,213,845	rs17218161	23829686		rs17218161	2.0 x10 ⁻⁸	SRIP1 - MIR548AG1
9p23	12,521,826	rs16929097	23829686		rs16929097	8.0 x10 ⁻⁹	JKAMPP1 - TYRP1
1q21.3	154,426,264	rs4129267	21907864	Asthma	rs4129267	2.0 x10 ⁻⁸	IL6R
5q12.1	59,369,794	rs1588265	19426955	Asthma	rs1588265	3.0 x10 ⁻⁸	PDE4D
12q13.2	56,412,487	rs1701704	21804548	Asthma	rs1701704	2.0 x10 ⁻¹³	IKZF4
4q31.21	144,003,159	rs7686660	21804548		rs7686660	2.0 x10 ⁻¹²	USP38, GAB1
22q12.3	37,534,034	rs2284033	20860503	Asthma	rs2284033	1.0 x10 ⁻⁸	IL2RB
8q24.11	118,025,645	rs3019885	21814517	Asthma (child)	rs3019885	5.0 x10 ⁻¹³	SLC30A8
1q31.3	197,325,908	rs2786098	20032318	Asthma (child)	rs2786098	2.0 x10 ⁻¹³	DENND1B
1q23.1	158,932,555	rs1101999	21804549	Asthma	rs1101999	4.0 x10 ⁻⁹	PYHIN1

Supplementary Table 14. Association analyses of the 1q21.3 and 22q12.3 SNPs in two sets of TAGC datasets

Region	Gene	SNP ^a	Alleles R/E ^b	EAF ^c	TAGC set	European-ancestry meta-analysis ^d <i>fixed-effects model</i>			Multi-ancestry meta-analysis ^d <i>fixed-effects model</i>		
						% of TAGC cases per set ^e	beta (se) ^f	P_{fixed} ^g	% of TAGC cases per set ^e	beta (se) ^f	P_{fixed} ^g
1q21.3	IL6R	rs4129267	C/T	0.39	Set P	64%	0.078 (0.016)	9.8×10^{-7}	53%	0.078 (0.016)	9.8×10^{-7}
					Set R	36%	0.038 (0.021)	0.07	47%	0.047 (0.018)	0.01
					P_{het} ^h			0.13			0.20
22q12.3	IL2RB	rs2284033	G/A	0.45	Set P	41%	-0.079 (0.020)	7.2×10^{-5}	34%	-0.079 (0.020)	7.2×10^{-5}
					Set R	59%	-0.031 (0.016)	0.06	66%	-0.028 (0.015)	0.06
					P_{het} ^h			0.06			0.04

The two sets of TAGC datasets are: the set P which includes TAGC datasets that were part of the original publication reporting association with asthma and the set R which includes the remaining TAGC datasets.

Note that Individual studies do not have always the same sample size in the original publication and TAGC.

^a Lead SNP reported associated with asthma in European-ancestry populations in the GWAS catalog (PMID: 21907864 for *IL6R* rs4129267 and PMID: 20860503 for *IL2RB* rs2284033).

^b R=reference allele /E=effect allele. ^cEAF=Effect allele frequency in European-ancestry populations.

^dThe results presented in this table were those obtained in a fixed-effects European-ancestry or multi-ancestry meta-analysis to test for heterogeneity of SNP effect size between the two sets.

^eThe proportions (%) of TAGC cases per set are the proportion of cases from the published study that are included in the TAGC European-ancestry or multi-ancestry meta-analysis for set P and the proportion of the remaining TAGC cases for set R.

^fBeta (se) is the regression coefficient (standard error) for a one-unit increase of the effect allele (logit scale) estimated in either one of the meta-analyses under a fixed-effects model

^g P_{fixed} for test of association between SNP and asthma

^h P_{het} for test of difference in SNP effect size between set P and set R using the Cochran's Q test.

Supplementary Table 15. Approximate conditional analysis of asthma-associated loci

Approximate conditional analysis, using the GCTA software (PMID 21167468; see URLs), is based on summary meta-analysis statistics under a fixed-effects model. This analysis was only conducted in European-ancestry populations that are assumed to share a similar LD pattern. Note that this analysis could not be done at the 9p24.1 locus which shows heterogeneity in SNP effect size across studies. At the 17q12-21, this analysis was restricted to the pediatric subgroup in which this locus does not show heterogeneity (see Online Methods for details).

							Meta-analysis results ^e				Conditional analysis results ^f		
Cytogenetic Region	Conditional SNP ^a	SNP	Position	Nearby genes ^b	Effect allele ^c	EAF ^d	Beta	SE	P _{fixed}	Beta	SE	P _{fixed}	
New asthma susceptibility loci													
5q31.3	TOP	rs7705042	141,492,419	<i>NDFIP1, GNPDA1, SPRY4</i>	A	0.63	0.08	0.01	8.5 × 10 ⁻¹⁰				
	rs7705042	rs13359221	141,035,880	<i>ARAP3, FCHSD1, PCDH1</i>	A	0.05	0.09	0.03	1.6 × 10 ⁻³	0.09	0.03	1.6 × 10 ⁻³	
6p22.1	TOP	rs1233578	28,712,247	<i>GPX5, TRIM27</i>	G	0.13	0.10	0.02	5.3 × 10 ⁻⁹				
	rs1233578	rs7509	27,418,847	<i>ZNF184, ZNF319, HIST1H2BL</i>	G	0.54	-0.04	0.01	6.1 × 10 ⁻⁴	-0.03	0.01	5.9 × 10 ⁻³	
6q15	TOP	rs2325291	90,986,686	<i>BACH2, GJA10, MAP3K7</i>	A	0.33	-0.10	0.01	8.6 × 10 ⁻¹³				
	rs2325291	rs1144157	91,184,183	<i>BACH2, MAP3K7</i>	G	0.67	-0.05	0.01	7.7 × 10 ⁻⁵	-0.05	0.01	1.2 × 10 ⁻⁴	
12q13.3	TOP	rs167769	57,503,775	<i>STAT6, NAB2, LRP1</i>	T	0.40	0.08	0.01	5.5 × 10 ⁻⁹				
	rs167769	rs10876932	57,217,358	<i>HSD17B6, SDR9C7</i>	G	0.58	-0.04	0.01	5.7 × 10 ⁻⁴	-0.05	0.01	1.2 × 10 ⁻⁴	
17q21.33	TOP	rs17637472	47,461,433	<i>ZNF652, PHB</i>	A	0.39	0.08	0.01	3.3 × 10 ⁻⁹				
	rs17637472	rs7214992	47,196,053	<i>IGF2BP1, B4GALNT2</i>	T	0.24	-0.04	0.02	3.2 × 10 ⁻³	-0.05	0.02	1.6 × 10 ⁻³	
New asthma signals at loci previously associated with asthma in ancestry-specific populations													
6p21.33	TOP*	rs2596464	31,412,961	<i>MICA, MICB</i>	C	0.48	0.09	0.01	1.8 × 10 ⁻¹³				
	rs2596464	rs2855812	31,472,720	<i>MICB, MICA, MCCD1</i>	T	0.23	0.09	0.01	2.8 × 10 ⁻¹⁰	0.07	0.01	4.3 × 10 ⁻⁷	
	rs2596464 + rs2855812	rs2844510	31,410,408	<i>MICA, MICB</i>	T	0.11	0.08	0.02	3.5 × 10 ⁻⁵	0.09	0.02	1.5 × 10 ⁻⁶	
	rs2596464 + rs2855812 + rs2844510	rs9263879	31,171,581	<i>POU5F1, HLA-C</i>	A	0.12	-0.07	0.02	4.4 × 10 ⁻⁴	-0.07	0.02	1.1 × 10 ⁻³	
10p14	TOP*	rs1663687	9,054,787	<i>GATA3, CELF2</i>	A	0.39	-0.08	0.01	1.5 × 10 ⁻¹⁰				
	rs1663687	rs2646437	8,566,311	<i>GATA3, CELF2</i>	A	0.47	-0.06	0.01	9.4 × 10 ⁻⁶	-0.05	0.01	4.8 × 10 ⁻⁵	
Asthma signals previously reported for asthma plus hay fever													
8q21.13	TOP*	rs10957979	81,289,787	<i>TPD52, ZBTB10</i>	G	0.65	-0.07	0.01	2.3 × 10 ⁻⁸				
	rs10957979	rs12675564	80,864,649	<i>MRPS28, HEY1, TPD52</i>	G	0.72	0.03	0.01	0.01	0.04	0.01	7.8 × 10 ⁻³	
16p13.13	TOP	rs17806299	11,199,980	<i>CLEC16A, DEXI, SOCS1</i>	A	0.20	-0.10	0.02	2.1 × 10 ⁻¹⁰				
	rs17806299	rs193772	11,299,077	<i>CLEC16A, SOCS1</i>	C	0.87	-0.09	0.02	2.6 × 10 ⁻⁶	-0.07	0.02	1.5 × 10 ⁻⁴	
Known asthma loci													
2q12	TOP*	rs3771180	102,953,617	<i>IL1RL1, IL1RL, IL18R1</i>	T	0.14	-0.17	0.02	1.5 × 10 ⁻²⁰				
	rs3771180	rs1420101	102,957,716	<i>IL1RL1, IL1RL2, IL18R1</i>	T	0.37	0.12	0.01	9.1 × 10 ⁻²⁰	0.08	0.01	2.2 × 10 ⁻¹¹	
5q22.1	TOP	rs10455025	110,404,999	<i>SLC25A46, TSLP</i>	C	0.34	0.14	0.01	2.0 × 10 ⁻²⁵				
	rs10455025	rs6893213	110,198,114	<i>SLC25A46, TSLP</i>	T	0.08	0.15	0.02	7.9 × 10 ⁻¹¹	0.13	0.02	2.5 × 10 ⁻⁹	
	rs10455025 + rs6893213	rs10043185	109,976,588	<i>TMEM232, MAN2A1, SLC25A46</i>	A	0.05	0.07	0.03	0.02	0.12	0.03	1.1 × 10 ⁻⁴	
	TOP	rs20541	131,995,964	<i>IL13, RAD50, IL4</i>	G	0.79	-0.12	0.02	1.4 × 10 ⁻¹⁴				
5q31	rs20541	rs6894249	131,797,547	<i>SLC22A5, IRF1</i>	G	0.39	0.09	0.01	2.2 × 10 ⁻¹¹	0.08	0.01	1.3 × 10 ⁻⁹	
	6p21.32	TOP	rs9272346	32,604,372	<i>HLADRB1, HLA-DQA1</i>	A	0.56	0.15	0.01	2.4 × 10 ⁻²⁸			
rs9272346		rs3763309	32,375,973	<i>BTNL2, HLA-DRA</i>	A	0.21	0.13	0.02	1.3 × 10 ⁻¹⁸	0.09	0.01	1.2 × 10 ⁻⁹	
rs9272346 + rs3763309		rs241429	32,803,840	<i>TAP2, HLA-DOB, PSMB8</i>	G	0.58	-0.07	0.01	2.9 × 10 ⁻⁸	-0.05	0.01	4.8 × 10 ⁻⁵	
11q13.5	TOP*	rs2155219	76,299,194	<i>C11orf30, LRRC32</i>	T	0.46	0.11	0.01	2.9 × 10 ⁻¹⁵				
	rs2155219	rs4944120	76,436,948	<i>LRRC32, TSKU</i>	G	0.24	-0.02	0.01	0.12	-0.03	0.01	2.5 × 10 ⁻²	
15q22.2	TOP	rs11071558	61,069,421	<i>RORA, NARG2, VPS13C</i>	G	0.14	-0.12	0.02	8.3 × 10 ⁻¹¹				
	rs11071558	rs339992	60,913,505	<i>RORA, NARG2, VPS13C</i>	G	0.12	-0.06	0.02	4.2 × 10 ⁻³	-0.06	0.02	3.8 × 10 ⁻³	
15q22.33	TOP*	rs17293632	67,442,596	<i>SMAD3, SMAD6, AAGAB</i>	T	0.26	0.12	0.01	8.8 × 10 ⁻¹⁶				
	rs17293632	rs2053295	67,391,739	<i>SMAD3, SMAD6, AAGAB</i>	A	0.88	0.01	0.02	0.5	0.06	0.02	7.3 × 10 ⁻⁴	
17q12-21	TOP*	rs2305479	38,062,217	<i>GSDMB, ZBP2, ORMDL3</i>	T	0.51	-0.23	0.02	2.8 × 10 ⁻²⁴				
	rs2305479	rs3859192	38,119,254	<i>GSDMA, LRRC3C, PSMD3</i>	T	0.47	0.19	0.02	1.9 × 10 ⁻¹⁶	0.09	0.02	3.4 × 10 ⁻⁵	

^aIn the conditional SNP field, "TOP" indicates we present the unadjusted results for the top SNP at each locus; otherwise the rsIDs of the top (and subsequent SNPs) used to adjust other SNPs are listed

^bGenes coding for proteins; the gene where eventually the lead SNP lies is first indicated followed by the previous and next genes

^cThe effect allele is the allele for which the regression coefficient (beta) is computed; ^dEAF= Effect allele frequency in European-ancestry populations.

^eUnder the heading meta-analysis results, beta, SE and P are the regression coefficient, standard error and P-values in fixed-effects meta-analysis of European ancestry populations

^fUnder the heading conditional analysis results, beta, SE and P are the regression coefficient, standard error and P-values following adjustment by the SNP(s) listed in the "Conditional SNP" field. The Bonferroni-corrected threshold for the total number of SNPs overall all regions investigated using GCTA is 4 × 10⁻⁶

Note that the TOP SNP in European-ancestry meta-analysis under a fixed-effects model may differ from the TOP SNP under a random-effects model (shown in Table 1); these TOP SNPs are indicated by a star. At 17q12-21 locus, the TOP SNP is the TOP SNP in the European-ancestry meta-analysis of the pediatric sample (fixed-effects model).

Supplementary Table 16. Associations between SNPs at the nine novel asthma loci and cis-gene transcripts in blood and lung

Region	Asthma SNP	Position build 37	LD (r ² /D')	with lead SNP	Alleles (R/E) ^a	Source database ^b	Tissue ^c	Gene transcript	P _{transcript} ^d	Gene transcript	P _{transcript} ^d	Gene transcript	P _{transcript} ^d	Reference PMID
New asthma susceptibility loci														
5q31.3	rs7705042	141,492,419			C/A	Blood_eQTL	Blood	NDFIP1	2.7 x 10 ⁻⁹					24013639
						MRCA	LCLs	NDFIP1	6.9 x 10 ⁻⁹					23345460
						MRCE	LCLs	NDFIP1	1.6 x 10 ⁻⁷					23345460
						GTE _x	Blood	NDFIP1	3.8 x 10 ⁻⁶					25954001
6p22.1	rs7766356	28,400,538	0.6/0.8 with rs1233578	T/C	MRCA	LCLs	ZSCAN12	3.0 x 10 ⁻⁸					23345460	
					Hao et al, 2012	Lung	ZSCAN31	6.5 x 10 ⁻¹¹					23209423	
6q15	rs10455168	90,883,525	0.9/1 with rs2325291	T/C	Blood_eQTL	Blood	BACH2	3.0 x 10 ⁻¹⁰					24013639	
12q13.3	rs167769	57,503,775		C/T	Blood_eQTL	Blood	STAT6	9.8 x 10 ⁻¹⁹⁸	NEMP1	1.6 x 10 ⁻¹¹			24013639	
					MuTHER	LCLs	STAT6	1.6 x 10 ⁻²²					22941192	
					MRCE	LCLs	STAT6	2.3 x 10 ⁻¹⁹					23345460	
					Hao et al, 2012	Lung	STAT6	3.7 x 10 ⁻³⁷					23209423	
	rs12368672	55,798,737	0.9/0.9 with rs167769		GHS_Express	Monocytes	STAT6	3.4 x 10 ⁻⁷⁷					20502693	
	rs324011	55,788,449	0.9/1 with rs167769		Blood_eQTL	Blood	STAT6	9.8 x 10 ⁻¹⁹⁸	TMEM194A	7.1 x 10 ⁻¹²			24013639	
					MuTHER	LCLs	STAT6	6.5 x 10 ⁻²⁶					22941192	
					MRCE	LCLs	STAT6	2.0 x 10 ⁻¹⁹					23345460	
					Hao et al, 2012	Lung	STAT6	1.1 x 10 ⁻³⁸					23209423	
	17q21.33	rs17637472	47,461,433		G/A	Blood_eQTL	Blood	GNGT2	2.1 x 10 ⁻⁵²	PHOSPHO1	3.2 x 10 ⁻¹²			24013639
						MuTHER	LCLs	GNGT2	2.4 x 10 ⁻¹¹					22941192
						MRCE	LCLs	GNGT2	5.3 x10 ⁻¹³					23345460
						GHS_Express	Monocytes	GNGT2	3.0 x 10 ⁻⁸					20502693
New asthma signals at loci previously reported for asthma in ancestry-specific populations														
6p21.33	rs2855812	31,472,720		G/T	Blood_eQTL	Blood	TNF	4.8 x 10 ⁻¹⁴	LTA	1.0 x 10 ⁻¹⁰			24013639	
					MuTHER	LCLs	LST1	1.0 x 10 ⁻¹³	HCG22	1.3 x 10 ⁻⁶			22941192	
					MRCE	LCLs	HLA-C	3.2 x 10 ⁻¹³					23345460	
					GTE _x	Lung	MICB	6.3 x 10 ⁻¹⁰					25954001	
	rs2596464 lead SNP in conditional analysis (fixed-effects model)	31,412,961		T/C	Hao et al, 2012	Lung	MICB	4.6 x 10 ⁻¹³	NCR3	6.2 x 10 ⁻⁷	HLA-C	5.9 x 10 ⁻⁶	23209423	
					Blood_eQTL	Blood	MICB	6.9 x 10 ⁻²¹					24013639	
					GTE _x	Blood	MICA	7.5 x 10 ⁻¹²					25954001	
					MuTHER	LCLs	HCG22	9.5 x 10 ⁻¹⁰	MICB	2.0 x 10 ⁻⁹	LST1	1.4 x 10 ⁻⁶	22941192	
					MRCE	LCLs	MICA	4.4 x 10 ⁻¹¹					22941192	
					GHS_Express	Monocytes	HLA-C	4.1 x 10 ⁻¹⁸	HCG27	5.0 x 10 ⁻¹¹	MICB	1.2 x 10 ⁻⁶	20502693	
					GTE _x	Lung	MICA	2.5 x 10 ⁻⁹					25954001	
					Hao et al, 2012	Lung	CDSN	4.8 x 10 ⁻¹⁹	MICB	1.6 x 10 ⁻⁶			23209423	
Asthma signals previously reported for asthma plus hay fever														
16p13.13	rs17806299	11,199,980		G/A	Blood_eQTL	Blood	DEXI	2.2 x 10 ⁻⁴³					24013639	
					MRCE	LCLs	SOC51	1.0 x 10 ⁻⁷					23345460	
					GHS_Express	Monocytes	DEXI	3.6 x 10 ⁻²⁰					20502693	

The eQTL databases for blood and lung were interrogated for the lead SNPs (shown in Table 1) and all SNPs having r^2 between 0.50 and 1 with the lead SNPs

The associations between asthma SNPs and transcripts, shown here, have $P_{\text{transcript}} < 10^{-5}$

^aR=reference allele /E=effect allele.

^beQTL databases are described in Online Methods.

^cBlood is whole blood; LCL = lymphoblastoid cell lines; Lung is whole lung tissue.

^d $P_{\text{transcript}}$ for test of association of SNP with gene transcript

Supplementary Table 17. Overlap between TAGC asthma association signals ($P_{\text{random}} < 10^{-4}$) and GWAS signals with specific diseases/traits in the GWAS catalog.

This Table has three parts: (i) autoimmune diseases are in the first part of the table; (ii) three allergic phenotypes in the second part; and (iii) the rest of the top 10 traits (in the number of hits in the catalog) are last. Asthma association signals are those having $P_{\text{random}} < 10^{-4}$ in the multi-ancestry meta-analysis. For each disease and trait, we selected one GWAS SNP per chromosomal band as reported in the variable “Region” in the GWAS catalog; SNPs associated with multiple diseases in the GWAS catalog could appear in counts in multiple rows.

Rank ^a	Disease/Trait	Number of GWAS catalog entries	Number of SNPs associated with asthma at $P_{\text{random}} < 10^{-4}$ in TAGC multi-ancestry meta-analysis	Enrichment P -value ^b
Auto-immune diseases				
2	Inflammatory bowel disease	97	6	9.8×10^{-16}
4	Crohn's disease	89	9	6.3×10^{-25}
8	Rheumatoid arthritis	59	4	4.5×10^{-11}
11	Ulcerative colitis	54	4	3.1×10^{-11}
14	Multiple sclerosis	49	3	1.8×10^{-8}
23	Systemic lupus erythematosus	35	2	5.9×10^{-6}
25	Type 1 diabetes	34	5	2.7×10^{-15}
30	Celiac disease	29	3	3.6×10^{-9}
38	Vitiligo	23	4	8.8×10^{-18}
Allergy-related phenotypes				
65	Self-reported allergy	16	12	Highly significant
67	Atopic dermatitis	14	5	Highly significant
99	Allergic sensitization	9	7	Highly significant
Top 10 traits in the number of hits in the GWAS catalog				
1	Height	162	3	6.8×10^{-7}
3	Blood metabolite levels	96	2	4.5×10^{-5}
5	Cholesterol, total	66	2	2.1×10^{-5}
6	HDL cholesterol	64	2	2.0×10^{-5}
7	Type 2 diabetes	61	0	1
9	Breast cancer	56	0	1
10	Prostate cancer	56	0	1

^aRank of the disease or trait in the number of unique signals in the GWAS catalog (PMID: 24316577; see URLs)

^bThe enrichment P -value is the binomial tail probability for observing the shown number of TAGC SNPs with $P_{\text{random}} < 10^{-4}$ among the SNPs reported in the GWAS catalog for a given disease or trait.

We investigated overlap with nine auto-immune diseases, three allergic phenotypes and seven other traits with large number of associations, and a conservative Bonferroni correction implies an adjusted significance threshold for enrichment of $0.05/19 = 0.003$.

Supplementary Tables 18 and 19 can be found in the Supplementary excel files

Supplementary Table 18. Overlap between TAGC asthma signals ($P_{\text{random}} < 10^{-3}$ in the multi-ancestry meta-analysis) and GWAS signals with diseases/traits in the GWAS catalog

Supplementary Table 19. Enrichment of asthma risk loci in promoter and enhancer marks by cell type. The results are for 16 out of the 18 asthma loci shown in Table 1; the 6p21.33 and 6p21.32 loci spanning the HLA complex were excluded because of high variability and LD in the region. Enhancer and promoter marks were defined using the ChromHMM 15-state model applied to functional data of 127 ROADMAP and ENCODE reference epigenomes in various cell types (PMID:25693563)

Supplementary Table 20. Enrichment of asthma risk loci in DNase I hypersensitive sites (DHSs) by cell type

The results presented in this table are for 16 out of the 18 asthma loci shown in Table 1

The 6p21.33 and 6p21.32 loci spanning the HLA complex were excluded because of high variability and LD in the region

The DHSs were available in 51 cell types using ROADMAP and ENCODE data (PMID:25693563)

EID ^a	Group ^a	Standardized Epigenome name ^a	Anatomy ^a	Type ^a	AST_loci ^b	FDR ^c
E005	ES-deriv	H1 BMP4 Derived Trophoblast Cultured Cells	ESC_DERIVED	ESCDerived	4	0.678
E006	ES-deriv	H1 Derived Mesenchymal Stem Cells	ESC_DERIVED	ESCDerived	5	0.678
E007	ES-deriv	H1 Derived Neuronal Progenitor Cultured Cells	ESC_DERIVED	ESCDerived	6	0.429
E008	ESC	H9 Cells	ESC	PrimaryCulture	10	0.024
E017	IMR90	IMR90 fetal lung fibroblasts Cell Line	LUNG	CellLine	3	0.920
E021	iPSC	iPS DF 6.9 Cells	IPSC	PrimaryCulture	8	0.251
E022	iPSC	iPS DF 19.11 Cells	IPSC	PrimaryCulture	5	0.747
E028	Epithelial	Breast variant Human Mammary Epithelial Cells (vHMEC)	BREAST	PrimaryCulture	8	0.168
E029	HSC & B-cell	Primary monocytes from peripheral blood	BLOOD	PrimaryCell	3	0.744
E032	HSC & B-cell	Primary B cells from peripheral blood	BLOOD	PrimaryCell	8	0.057
E033	Blood & T-cell	Primary T cells from cord blood	BLOOD	PrimaryCell	8	0.043
E034	Blood & T-cell	Primary T cells from peripheral blood	BLOOD	PrimaryCell	10	0.005
E046	HSC & B-cell	Primary Natural Killer cells from peripheral blood	BLOOD	PrimaryCell	8	0.048
E050	HSC & B-cell	Primary hematopoietic stem cells G-CSF-mobilized Female	BLOOD	PrimaryCell	7	0.232
E051	HSC & B-cell	Primary hematopoietic stem cells G-CSF-mobilized Male	BLOOD	PrimaryCell	0	1.000
E055	Epithelial	Foreskin Fibroblast Primary Cells skin01	SKIN	PrimaryCulture	0	1.000
E056	Epithelial	Foreskin Fibroblast Primary Cells skin02	SKIN	PrimaryCulture	0	1.000
E057	Epithelial	Foreskin Keratinocyte Primary Cells skin02	SKIN	PrimaryCulture	9	0.013
E059	Epithelial	Foreskin Melanocyte Primary Cells skin01	SKIN	PrimaryCulture	9	0.156
E080	Other	Fetal Adrenal Gland	ADRENAL	PrimaryTissue	6	0.445
E081	Brain	Fetal Brain Male	BRAIN	PrimaryTissue	7	0.393
E082	Brain	Fetal Brain Female	BRAIN	PrimaryTissue	7	0.381
E083	Heart	Fetal Heart	HEART	PrimaryTissue	9	0.156
E084	Digestive	Fetal Intestine Large	GI_INTESTINE	PrimaryTissue	4	0.710
E085	Digestive	Fetal Intestine Small	GI_INTESTINE	PrimaryTissue	4	0.568
E086	Other	Fetal Kidney	KIDNEY	PrimaryTissue	6	0.523
E088	Other	Fetal Lung	LUNG	PrimaryTissue	8	0.381
E089	Muscle	Fetal Muscle Trunk	MUSCLE	PrimaryTissue	8	0.240
E090	Muscle	Fetal Muscle Leg	MUSCLE_LEG	PrimaryTissue	9	0.161
E091	Other	Placenta	PLACENTA	PrimaryTissue	6	0.558
E092	Digestive	Fetal Stomach	GI_STOMACH	PrimaryTissue	6	0.438
E093	Thymus	Fetal Thymus	THYMUS	PrimaryTissue	8	0.130
E094	Digestive	Gastric	GI_STOMACH	PrimaryTissue	3	0.710
E097	Other	Ovary	OVARY	PrimaryTissue	5	0.541
E098	Other	Pancreas	PANCREAS	PrimaryTissue	5	0.541
E100	Muscle	Psoas Muscle	MUSCLE	PrimaryTissue	5	0.592
E109	Digestive	Small Intestine	GI_INTESTINE	PrimaryTissue	6	0.393
E114	ENCODE2012	A549 EtOH 0.02pct Lung Carcinoma Cell Line	LUNG	CellLine	3	0.839
E116	ENCODE2012	GM12878 Lymphoblastoid Cells	BLOOD	PrimaryCulture	9	0.043
E117	ENCODE2012	HeLa-S3 Cervical Carcinoma Cell Line	CERVIX	CellLine	8	0.148
E118	ENCODE2012	HepG2 Hepatocellular Carcinoma Cell Line	LIVER	CellLine	4	0.678
E119	ENCODE2012	HMEC Mammary Epithelial Primary Cells	BREAST	PrimaryCulture	9	0.076
E120	ENCODE2012	HSMM Skeletal Muscle Myoblasts Cells	MUSCLE	PrimaryCulture	6	0.541
E121	ENCODE2012	HSMM cell derived Skeletal Muscle Myotubes Cells	MUSCLE	PrimaryCulture	7	0.449
E122	ENCODE2012	HUVEC Umbilical Vein Endothelial Primary Cells	VASCULAR	PrimaryCulture	7	0.161
E123	ENCODE2012	K562 Leukemia Cells	BLOOD	PrimaryCulture	7	0.240
E124	ENCODE2012	Monocytes-CD14+ RO01746 Primary Cells	BLOOD	PrimaryCell	6	0.322
E125	ENCODE2012	NH-A Astrocytes Primary Cells	BRAIN	PrimaryCulture	8	0.168
E126	ENCODE2012	NHDF-Ad Adult Dermal Fibroblast Primary Cells	SKIN	PrimaryCulture	5	0.710
E127	ENCODE2012	NHEK-Epidermal Keratinocyte Primary Cells	SKIN	PrimaryCulture	7	0.245
E128	ENCODE2012	NHLF Lung Fibroblast Primary Cells	LUNG	PrimaryCulture	6	0.496

The method used to assess enrichment is described in detail in the Online Methods.

^aEID is the cell type number; the following variables are group, standardized epigenome name, anatomy and type for each cell type (PMID:25693563)

^bAST_loci is the number of asthma loci that co-localize with DHSs

^cFDR is the Benjamini-Hochberg false discovery rate

Supplementary Table 21. Results of GRAIL analysis for the asthma risk loci

The 18 asthma risk loci identified in this study are ordered according to P_{GRAIL} from lowest to highest (see PMID:19557189 for details on GRAIL method); no gene was found by GRAIL at 10p14 locus

Region	Start ^a	Stop ^a	P_{GRAIL}^b for the region
5q22.1	110,429,500	110,519,500	8.65E-08
9p24.1	6,134,500	6,207,000	5.39E-07
12q13.3	55,606,500	55,835,500	1.63E-06
2q12	102,282,914	102,461,414	1.75E-06
5q31	131,904,000	132,034,500	3.84E-06
6p21.32	32,447,500	32,790,500	2.64E-04
16p13.13	10,925,500	11,215,000	5.82E-04
6p21.33	30,332,000	31,536,500	1.07E-03
6p22.1	28,283,500	29,463,500	1.24E-03
17q12-q21	34,634,000	35,343,500	1.28E-02
11q13.5	75,802,500	76,016,000	1.81E-02
6q15	90,867,000	91,080,500	5.01E-02
17q21.33	44,675,500	44,834,000	5.20E-02
8q21.13	81,052,000	81,745,500	8.25E-02
15q22.2	58,803,500	58,922,000	5.80E-01
5q31.3	141,394,500	141,546,000	6.14E-01
15q22.33	65,225,500	65,255,500	6.54E-01
10p14	8,828,500	9,406,000	NA

^aStart and stop are the start and stop positions of the regions defined by GRAIL

^b P_{GRAIL} represents the GRAIL significance score which allows assessing the strength of connectivity between genes at the asthma risk loci through text-mining of PubMed abstracts; P_{GRAIL} was corrected for the number of genes assigned to a region by GRAIL.

Supplementary Note

Study descriptions, Funding and acknowledgements, and Contributing groups

Study Descriptions

All studies obtained approval from their local ethical committees. All participants provided written informed consent.

Australian Asthma Genetics Consortium (AAGC)

Subjects: Participants were drawn from two cohorts described in greater detail elsewhere¹: the Australian Asthma Genetics Consortium (AAGC) cohort (n=1,810) and the Queensland Institute of Medical Research (QIMR) GWAS cohort (n=4,157). Of the 2110 asthmatic patients, 1,269 (60%) were classified as having childhood asthma (defined by an age-of-onset at or before age 16 years) 515 (24%) patients with later onset asthma (age-of-onset after the age of 16 years) and 326 (16%) with unknown age-of-onset. 1,232 (58%) of asthmatics were atopic, as defined by a positive skin prick test response to at least one common allergen; 1,444 (68%) had at least one first-degree relative with asthma; and 760 (36%) reported lifetime smoking. The 3,857 controls included 2,030 (53%) individuals who were classified as asthma-free based on clinical examination (109 [3%]) or epidemiological questionnaires (2,592 [50%]). The remaining 1,827 (47%) individuals provided no information about their asthma status but were included in the analysis to improve power. Mean age was 34 years (SD 16.2, range 2–92) and 3,267 (55%) were women. All participants were confirmed to be unrelated and of European ancestry based on genetic data.

Genotyping and QC: The Australian GWAS included data for 4,293 (72%) individuals genotyped with Illumina 610K array and 1,674 (28%) with Illumina 370K array¹. The same single-nucleotide polymorphism (SNP) quality control filters were applied to each dataset individually, including the removal of SNPs with call rate lower than 95%, minor allele frequency (MAF) lower than 0.01, and Hardy-Weinberg equilibrium test $p < 10^{-6}$.

Data Imputation and Analysis: Autosomal SNPs passing quality control were used to impute 7.8 million variants available from the combined 1000 genomes (60 individuals from the CEU collection; March 2010 release) and HapMap 3 (955 individuals from 11 populations; February 2009 release) reference panels with IMPUTE2². The 610K and 370K datasets were imputed separately. After imputation, we excluded SNPs with imputation information < 0.3 , MAF < 0.01 , or Hardy-Weinberg equilibrium test $p < 10^{-6}$. We also excluded from analysis SNPs with significant ($p < 0.001$) allele frequency differences between the two imputation analysis groups (from case–case and control–control comparisons). After quality control, genotype data for 5.7 million SNPs were merged across the 610K and 370K datasets; only results for Hapmap2 SNPs were provided to TAGC. Individual SNPs were then tested for association with lifetime physician-diagnosed asthma with a Cochran-Mantel-Haenszel test as implemented in PLINK³, with two strata corresponding to the two imputation analysis groups.

ALLERGEN Consortium

Canadian Cohorts CAPPS and SAGE (see <http://www.genopha.ca/>)

Subjects: The Canadian Asthma Primary Prevention Study (CAPPS⁴) was initiated in 1995 to assess the effectiveness of a multifaceted intervention program on the primary prevention of asthma in high-risk infants. Infants at high risk for developing asthma and other allergic disorders were identified and their mothers were recruited during their third trimester of pregnancy. Infants at high risk were defined as those

who had a parent with asthma or two first-degree relatives with allergies or atopic disorders. The study had two recruitment centers in Canada, (Vancouver and Winnipeg). In total, there were 545 families recruited into the study (549 infants, 4 sets of twins). Loss to follow-up was minimal, with 86% of the families completing the questionnaire at the 7-year time point. There have been numerous publications on this cohort evaluating the effectiveness of the intervention strategy. For genetic analysis blood samples have been obtained from the children, and both parents.

The Study of Asthma Genes and Environment Study (SAGE⁵) is a population-based sample of 16,320 children, born in the province of Manitoba, Canada in the year of 1995. In 2002, a one-page survey was mailed to families to enquire about their health and home environment exposure. Questions included several chronic conditions in childhood, including hayfever, food allergy and asthma. Parents were asked for permission to be contacted for further study and to link the survey data to their child's health care database records. Children were stratified according to the presence of asthma (n=392), to the presence of hayfever or food allergy (n=192) and neither (n=3,002). Children with neither condition were further grouped by rural and urban location; the latter were further stratified by low and high income neighborhoods. All children in the asthma and allergy strata were invited to participate in the case-control study, as were a random sample (n=200) from the strata of children with neither condition. Following consent from parents to participate in the study, children were assessed for allergic phenotypes.

After QC a total of 766 subjects from CAPPS were retained analyses (360 cases/ 406 controls) and a total of 1022 subjects from SAGE (353 cases / 669 controls).

Genotyping and QC: Genome-wide SNP genotyping was performed using the Illumina Human610 quad array of which 821 subjects from the Canadian Cohorts (CAPPS n=414 and SAGE n=407) were included in the GABRIEL consortium GWAS analyses⁶, using the GABRIEL QC protocol. The remaining 967 subjects (CAPPS n=352 and SAGE n=615) were genotyped using the same chip (Illumina Human610 quad array) and the following QC protocols: < 97% call rate for both samples and SNPs, evaluation of Hardy-Weinberg Equilibrium (HWE), heterozygosity (3 s.d.), ethnicity checks using multidimensional scaling, checks for duplicate samples, verification of reported family relationships and Mendelian errors.

Data Imputation and Analysis: Imputation was performed using MACH with HapMap2 CEU as the reference population. Statistical analyses were performed using STATA (childhood/offspring generation) to account for sibling relationships and PLINK (case-control) for adults/parents.

Saguenay-Lac-Saint-Jean Study (SLSJ)

Subjects: The Saguenay-Lac-Saint-Jean (SLSJ) asthma study, consists of French-Canadian families ascertained through asthmatic probands⁷. Probands were included in the study if they fulfill at least two of the following criteria: 1) a minimum of three clinic visits for acute asthma within one year; 2) two or more asthma-related hospital admissions within one year; or 3) steroid dependency, as defined by either six months of oral, or one year of inhaled corticosteroid use. Individuals from 253 independent families were included in the study. Families were included in the study if at least one parent was available for phenotypic assessment, at least one parent was unaffected, and all four grandparents were of French-Canadian origin. Family members were considered asthmatics if both a self-reported history of asthma and a history of physician-diagnosed asthma were available, or by clinical evaluation following a methacholine provocation test. After QC a total of 1,199 subjects were retained in the GWAS analyses.

Genotyping and QC: Genome-wide SNP genotyping was performed using the Illumina Human610 Quad array, as part of the GABRIEL consortium (see GABRIEL paragraph below for details).

Data Imputation and Analysis: Imputation and analysis were done as part of the GABRIEL consortium (see GABRIEL paragraph below).

APCAT Consortium

The APCAT Consortium includes 21,858 European-ancestry subjects (2,055 physician-diagnosed asthmatics and 19,803 asthma-free controls) with available genome-wide data that participated in seven population based-studies, five studies from Finland (the Northern Finland Birth Cohort 1966 [NFBC66]; Helsinki Birth Cohort [HBC]; Health 2000 [H2000]; FINRISK, including FINRISK 1992, 1997, 2002 and 2007; and the Young Finns Study), the European Prospective Investigation of Cancer, Norfolk (EPIC-Norfolk) and Framingham Heart Study from the United States.

Framingham Heart Study (FHS)

Subjects: The FHS is a collection of cohorts recruited to investigate cardiovascular disease and its risk factors as described in detail elsewhere⁸. Asthma was classified based on self-report of a physician's diagnosis; 797 asthmatics and 6,463 controls were included in TAGC study.

Genotyping and QC: Genome-wide SNP genotyping was performed using Affymetrix 5.0 BeadChip. Individuals with high heterozygosity or inbreeding and missingness >0.03 were removed and only SNPs with call rate >0.95 , HWE $P > 10^{-6}$ and MAF >0.01 were used in the imputation.

Data Imputation and analysis: Imputation was carried out on 413,905 SNPs using MACH 1.0 software and HapMap2 CEU (Release 22, Build 36) as reference panel. After imputation, we used SNPs with an Rsq threshold of >0.3 and MAF threshold of $>1\%$. The association statistic for each SNP was calculated assuming an additive genetic model adjusting for sex, and age using a GEE model to take into account familial dependencies (GWAF R package).

EPIC-Norfolk Study

Subjects: EPIC-Norfolk is part of the large multi-centre Europe-wide EPIC programme looking at the connection between diet and cancer¹⁰. There were over 30,000 participants aged 45 – 70 at recruitment who lived in Norwich and the surrounding towns and rural areas. They have been contributing information about their diet, lifestyle and health through questionnaires, and through health checks carried out by EPIC nurses. In 2006, a GWAS for obesity was carried out on 3,867 individuals. A case-cohort design was used in which the subcohort (N=2,566) was a random sample of the cohort at baseline and cases were part of the remaining individuals with a value of BMI being 30 kg/m² or greater (N=1,301). Unlike controls in the commonly used case-control design, the subcohort was an unselected population sample allowing for a variety of traits to be investigated. The EPIC-Norfolk obese cases and the EPIC-Norfolk population based subcohorts were considered as two different samples in the present TAGC study. Asthmatics were defined as a positive response to the question "*Has the doctor ever told you that you have asthma?*" which was asked at baseline survey. The remaining individuals served as healthy controls unless they had bronchitis or FEV1 $< 70\%$ predicted for age, height and sex.

Genotyping and QC: Genotyping was carried out using Affymetrix 500K BeadChip. Individuals with high heterozygosity or imbreeding and missingness > 0.06 were removed and only SNPs with call rate > 0.90 , HWE $P > 10^{-6}$ and MAF > 0.01 were used in the imputation.

Data imputation and analysis: Imputation was performed on 397,438 SNPs using IMPUTE 0.3.1 and HapMap2 CEU (Release 21) as reference panel. After imputation, we used SNPs with an information threshold of > 0.4 and MAF threshold of $> 1\%$. Data analysis was conducted separately in EPIC-Norfolk Obese and EPIC-Norfolk population-based studies (number of cases and controls are shown in Supplementary Table 1). The association statistic for each SNP was calculated assuming an additive genetic model adjusting for sex, ancestry-informative principal components, and age. Association analysis between individual SNP and asthma was performed in each sub-study using SNPTEST 1.1.5.

Northern Finland Birth Cohort of 1966 (NFBC1966)

Subjects: The NFBC1966 study⁹ [<http://kelo.oulu.fi/NFBC/>] includes 12,058 live born individuals, of European descent, with expected dates of birth during 1966 in the provinces of Oulu and Lapland, in Finland. Cohort has been followed up since early pregnancy until adult age. All those living in Northern Finland or in the capital area were invited to a clinical examination and blood sampling at age 31 years. Information on asthma used in this analysis was collected at this time point; 364 asthmatics and 3,502 controls were included in TAGC study.

Genotyping and QC: Genotyping was carried out using Illumina's HumanCNV370-Duo BeadChip. Individuals with high heterozygosity or imbreeding were removed and only SNPs with call rate > 0.95 , HWE $P > 10^{-4}$ and MAF > 0.01 were used in the imputation.

Data imputation and analysis: Imputation was performed on 328,007 SNPs using IMPUTE 0.3.1 and HapMap2 CEU (Release 21, Build 35). After imputation, we used SNPs with an information threshold of > 0.4 and MAF threshold of $> 1\%$. The association statistic for each SNP was calculated assuming an additive genetic model adjusting for sex, and ancestry-informative principal components. Association analysis performed in QUICKTEST v0.94.

FINRISK Study

Subjects: This study is a population survey of risk factors for chronic diseases in Finland¹¹. The survey has been executed every five years from 1972 using independent, random and representative population samples from five geographical areas of the country. Participants have completed a health-related questionnaire and undergone a physical examination including measurement of anthropometric traits and blood draw. For the FINRISK 1992 cohort, asthmatics were individuals who have answered "YES" to the question "Have you had any of the following diseases diagnosed or treated by a doctor during the past year (last 12 months)?: Asthma of the lungs". For the FINRISK 1997, 2002 and 2007 cohorts, asthmatics were individuals who have answered "YES" to the question: "Have you ever been diagnosed with asthma?". In all FINRISK surveys the rest of the participants served as controls if their age was less than or equal to 70 years and they did not report pulmonary emphysema, or chronic bronchitis during the last 12 months.

Genotyping and QC: Participants were genotyped using Illumina 610K and Affymetrix 6.0 arrays. For QC, Individuals with high heterozygosity or imbreeding and missingness > 0.05 were removed and only SNPs with call rate > 0.95 , HWE $P > 10^{-6}$ and MAF > 0.01 were used in the imputation

Data imputation: Imputation was performed on 554,988 (Illumina) and 727,478 (Affymetrix) SNPs using MACH 1.0.16 and HapMap2 CEU (Release 22, Build 36) as reference panel. After imputation, we used SNPs with an Rsq threshold of >0.3 and MAF threshold of >1%.

Health 2000 Study (H2000)

Subjects: The study was conducted in 2000¹² and included home interview, completion of several questionnaires, laboratory and anthropometrical measurements, spirometry with bronchodilator test and clinical examination by a physician. Further information was obtained by record linkage with the National Hospital Discharge Register and the National Social Insurance Institutions register data on reimbursement of asthma medication. Information about asthma was based on the following question: *“Has a doctor ever diagnosed you with one of the following illnesses?”* One of the listed illnesses was asthma and those who responded “YES” were considered to have asthma. The rest of the participants were taken as controls if their age was less than or equal to 70 years and they never had had chronic bronchitis.

Genotyping and QC: Samples were genotyped with Illumina 610K or 370K arrays. Individuals with high heterozygosity or inbreeding and missingness > 0.05 were removed and only SNPs with call rate>0.95, HWE $P>10^{-6}$ and MAF>0.01 were used in the imputation

Data imputation: Imputation was performed on 555,388 SNPs using MACH 1.0.16 and HapMap2 CEU (Release 22, Build 36) as reference panel. After imputation, we used SNPs with an Rsq threshold of >0.3 and MAF threshold of >1%.

Helsinki Birth Cohort Study (HBCS)

The HBCS¹³ includes 8,760 subjects born in Helsinki between 1934 and 1944. Between 2001 and 2004, a representative subset of 928 males and 1,075 females participated in a clinical study focusing upon cardiovascular and metabolic outcomes and cognitive functions. Information on asthma, smoking and alcohol intake is available from questionnaires for 2,003 individuals who participated in the clinical study. Information about asthma was based on the following question: *“Have you ever had any of the following illnesses diagnosed or treated by a doctor?”* One of the listed illnesses was *“Asthma of the lungs”* and those responding “YES” to this item were considered as having asthma. Those participants who responded to the same question as having emphysema or chronic bronchitis were excluded and the others were taken as controls. All participants were less than 70 years of age.

Genotyping and QC: Genotyping was performed with the Illumina 670K array. Individuals with high heterozygosity or inbreeding and missingness > 0.05 were removed and only SNPs with call rate>0.95, HWE $P>10^{-6}$ and MAF>0.01 were used in the imputation

Data imputation: Imputation was performed on 546,814 SNPs using MACH 1.0.16 and HapMap2 CEU (Release 22, Build 36) as reference panel. After imputation, we used SNPs with an Rsq threshold of >0.3 and MAF threshold of >1%.

Young Finns Study (YFS)

Subjects: The Young Finns cohort¹⁴ is a longitudinal population study sample on the evolution of cardiovascular risk factors from childhood to adulthood. The first cross-sectional survey was conducted in 1980 in five Finnish university cities and included 3,596 participants who were in the age groups of 3, 6, 9, 12, 15, and 18 years and were randomly chosen from the national population register; equal ratios of

males and females were selected in each age group. In 2007, 2,204 subjects now aged 30 to 45 years participated in the latest follow-up study. Information about asthma was based on a following question: *“Do you have at the moment or have you had a long-term illness, handicap or injury diagnosed by a doctor?”* Those responding “YES” and specifying among the given alternatives “Asthma of the lungs” were considered as having asthma. The rest of the participants were taken as controls, except those who reported having chronic bronchitis. All YFS participants were less than 70 years of age.

Genotyping and QC: 1,963 individuals with available physician-diagnosed asthma status were genotyped with the Illumina 670K array. Individuals with high heterozygosity or inbreeding and missingness > 0.05 were removed and only SNPs with call rate > 0.95 , HWE $P > 10^{-6}$ and MAF > 0.01 were used in the imputation

Data imputation: Imputation was performed on 546,677 SNPs using MACH 1.0 and HapMap2 CEU (Release 22, Build 36) as reference panel. After imputation, we used SNPs with an Rsq threshold of > 0.3 and MAF threshold of $> 1\%$.

Statistical analysis of FINRISK, Health 2000, HBCS and YFS: The four studies were analyzed together with an adjustment term for cohort. Association analysis between individual SNP and asthma was performed using Plink v1.07. Lambda for asthma analysis was 1.02.

CARe Consortium studies contributing to TAGC

Subjects: The CARe consortium datasets which contributed to TAGC included African-American participants from four population-based studies: Atherosclerosis Risk in Communities (ARIC), Coronary Artery Risk Development in young Adults (CARDIA), Jackson Heart Study (JHS), and Multi- Ethnic Study of Atherosclerosis (MESA). The number of asthma cases and controls from each study are shown in Supplementary Table 1.

CARDIA: The CARDIA study is a prospective, multi-center investigation of the natural history and etiology of cardiovascular disease in 5,115 individuals, age 18 to 30 years at the time of initial examination (1985–1986) and drawn from four communities: Birmingham, AL; Chicago, IL; Minneapolis, MN; and Oakland, CA. Each participant’s age, race, and sex were self-reported during the recruitment phase and verified during the baseline clinic visit. Genotype data were available on 955 African-American individuals. A subset of these subjects contributed to the asthma GWAS.

JHS: The Jackson Heart Study is a prospective population-based study to evaluate common complex diseases among 5,301 African Americans, age 34 to 84 years at the time of initial examination (2000 – 2004) and drawn from the Jackson, Mississippi metropolitan area. Genotype data were available on 3,030 African-American individuals (some JHS participants are also enrolled in ARIC, and were analyzed with the ARIC dataset; 2,145 individuals are uniquely associated with JHS). Information on asthma was available for 2,162 African American individuals.

The **ARIC** and **MESA** studies contain subjects of African-ancestry and European-ancestry and are described below.

In the CARe Consortium, asthma was defined as a positive response to the question “Has a doctor ever told you that you have asthma?” or “Have you ever had asthma?” on any of the available questionnaires derived from the individual studies. The agreement between these two different questions, where available to analyze, was very good, leading us to think that subject reports of asthma stemmed from of a

medical diagnosis (kappa statistic with 95% confidence intervals for individual studies are as follows: ARIC 0.82 CI 0.80-0.85, CARDIA 0.81 CI 0.77-0.83, JHS 0.80 CI 0.78-0.83). The remaining subjects served as healthy controls if they did not report any of the following conditions: wheezing, chronic obstructive pulmonary disease, emphysema, chronic bronchitis, chronic cough, chronic sputum production, or other lung disease. Individuals were excluded from the control group if they had <70 % of the predicted FEV1 based on race- and sex-specific NHANES III prediction equations adjusting for age and height, or FEV1/FVC less than lower limits of normal for age, race and sex¹⁵.

Genotyping and QC: Samples were genotyped at the Broad Institute using the Affymetrix Genome-Wide Human SNP Array 6.0. A subset of 24 markers including a gender confirmation assay were also genotyped using the Sequenom MassArray System. Genotypes were called using Birdseed v1.33. Quality control steps were performed using the software PLINK, EIGENSTRAT and PREST-Plus (<http://fisher.utstat.toronto.edu/sun/Software/Prest/>). Multiple QC procedures were performed including: confirming genotype concordance between Sequenom iPLEX and Affy6.0; removing samples with a genome-wide genotyping success rate <95%, SNPs with genotyping success rate <90%, monomorphic SNPs, and SNPs that mapped to several genomic locations; removing poor quality DNA (identified by estimating heterozygosity rates); removing sample duplicates, contaminated samples, and cryptic relationships (identified by using genome-wide genotype data to estimate identity-by-descent between all pairwise combinations); outlier samples were removed (identified based on nearest neighbor and “clustering based on missingness” analyses in PLINK); removing SNPs with minor allele frequency (MAF) <1% or with genotyping success rate <95%; in JHS, excluding SNPs with an unusually high number of Mendel errors; and excluding SNPs that showed association with specific chemistry plates. Because several different ethnic groups were represented, with the expectation of differing genotype frequencies and admixture, no filters were applied for Hardy-Weinberg probability values.

Data Imputation and Analysis: To increase coverage and facilitate comparison with other datasets, we imputed genotype data using MACH version 1.0.16¹⁶. Imputation were done using European and African ancestry reference panels. The resulting set of SNPs was conservatively filtered to remove low imputation quality (RSQ <0.4) or minor allele frequency SNPs (MAF <0.01), for which association results would be less reliable. Analysis of association of asthma with individuals SNPs was carried out using logistic regression statistical framework in PLINK for unrelated cohorts or using R (GWAF package) scripts that model family structure for family data.

CHARGE Consortium studies contributing to TAGC

The ARIC and MESA studies include subjects of African-ancestry who are part of the CARE Consortium and European-ancestry subjects who are part of the CHARGE Consortium.

MESA: Multi-Ethnic Study of Atherosclerosis

Subjects: The Multi-Ethnic Study of Atherosclerosis (MESA) is a population-based longitudinal study of subclinical cardiovascular disease¹⁷. Between 2000 and 2002, MESA recruited 6,814 men and women 45–84 years of age from six US sites that were free of clinical cardiovascular disease. The MESA Family Study recruited 1,595 African American and Hispanic family members ages 45–84 years of age specifically for genetic analysis and the MESA Air Pollution Study recruited an additional 257 participants¹⁸. Endpoints for

asthma included self-reported history of asthma ("Have you ever had asthma?") and childhood asthma (Participants were asked at what age they "first developed symptoms." Those who reported symptoms at < 16 years of age were classified as having childhood asthma). For the current effort, genetic analyses in MESA were restricted to self-reported European-ancestry participants (MESA African-ancestry were analyzed separately through CARE). Analyses of asthma included 267 cases and 2,381 controls, while analyses of pediatric asthma included 84 cases and 1361 controls.

Genotyping and QC: All participants were genotyped using the Affymetrix Human SNP array 6.0 (Affymetrix Inc., Santa Clara, CA), with 897,981 single-nucleotide polymorphisms (SNPs) passing study-specific quality control.

Data Imputation and Analysis: IMPUTE2 was used to perform imputation of an additional ~2 million SNPs in each race/ethnic group (using HapMap Phase I and II as reference panel (release #24, NCBI Build 36) for inclusion in GWAS analyses^{19,20}. Multivariate additive genetic models were adjusted for age, sex, study site and principal components of ancestry. We examined genomic control values of all GWAS for evidence of residual population stratification, undetected family structure, or other sources of inflation in type I error.

ARIC: Atherosclerosis Risk in Communities study

The Atherosclerosis Risk in Communities (ARIC) Study, sponsored by the National Heart, Lung and Blood Institute (NHLBI), is a prospective epidemiologic study conducted in four U.S. communities.²¹ The four communities are Forsyth County, NC; Jackson, MS; the northwest suburbs of Minneapolis, MN; and Washington County, MD. Three cohorts represent the ethnic mix of their communities; the Jackson cohort, its African-American population. ARIC was designed to investigate the etiology and natural history of atherosclerosis, the etiology of clinical atherosclerotic diseases, and variation in cardiovascular risk factors, medical care and disease by race, gender, location, and date. ARIC has coordinating, ultrasound, pulmonary, and electrocardiographic centers and three central laboratories.

The Cohort began in 1987: each ARIC field center randomly selected and recruited a cohort sample of approximately 4,000 individuals aged 45-64 from a defined population in their community. A total of 15,792 participants received an extensive examination, including medical, social, and demographic data. These participants were examined with the baseline visit occurring in 1987-89, the second visit in 1990-92, the third visit in 1993-95, the fourth visit in 1996-98, and the fifth visit in 2011-13. Follow-up occurs yearly by telephone to maintain contact with participants and to assess health status of the cohort. Asthma was defined based on self-report of whether the participant had been diagnosed with asthma

Genotyping and QC: Genotyping was done using the AffymetrixGeneChip SNP Array 6.0. Quality control steps for genotyping data included exclusions for call rate <95% (participant level and SNP level), minor allele frequency <1%, HWE $P < 10^{-6}$, no chromosomal location, suspected first-degree relative of an included individual based on genotype data, or more than 8 standard deviations for any of the first ten principal components. There were 9,173 Caucasian subjects with genotyping data, asthma definition, and complete covariate information.

Data Imputation and Analysis: MACH (version 1.0.16) was used to impute all autosomal SNPs with reference to HapMap CEU (release 22, build 35) from the 669,450 SNPs. Imputed SNPs failing additional quality control criteria (monomorphism, HWE $P < 10^{-6}$, or genotype frequencies between two genotyping phases differing by $P < 10^{-6}$) were excluded, leaving 2,515,866 genotyped or imputed SNPs for analysis.

Association analysis between each individual SNP and asthma was done using logistic regression and the ProbABEL v.0.1-3 software.

CHS: Cardiovascular Health Study

Subjects: The Cardiovascular Health Study (CHS) is a population-based cohort study of risk factors for CHD and stroke in adults > 65 years conducted across four field centers. The original predominantly European ancestry cohort of 5,201 persons was recruited in 1989-1990 from random samples of the Medicare eligibility lists and an additional 687 African-Americans were enrolled subsequently for a total sample of 5,888. DNA was extracted from blood samples drawn on all participants at their baseline examination. In 2007-2008, genotyping was performed at the General Clinical Research Center's Phenotyping/Genotyping Laboratory at Cedars-Sinai on 3,980 Cardiovascular Health Study participants who were free of CVD at baseline, consented to genetic testing, and had DNA available for genotyping. Participants were excluded from analysis for sex mismatch, discordance with prior genotyping, or call rate < 95%. Non-European ancestry participants were excluded from this study. A total of 1,908 persons were excluded from the GWAS study sample due to the presence at study baseline of coronary heart disease, congestive heart failure, peripheral vascular disease, valvular heart disease, stroke or transient ischemic attack or lack of available DNA. Genotyping has been attempted to date in 3,397 self-identified white participants, and was successful in 3,291 persons. Of these, 3,237 subjects had information on asthma status: 179 self-reported a physician diagnosis of asthma (ever asthma), and the remaining 3,058 were used as controls.

Genotyping and QC: Genotyping was performed at the General Clinical Research Center's Phenotyping/Genotyping Laboratory at Cedars-Sinai using the Illumina 370CNV BeadChip system. Genotypes were called using the Illumina BeadStudio software as above. A set of 306,655 autosomal genotyped SNPs remained after exclusions for: call rate < 97%, HWE $P < 10^{-5}$, more than two duplicate errors or Mendelian inconsistency (for reference HapMap CEU) or no mapping in dbSNP.

Data Imputation and Analysis: Imputation was performed using BIMBAM (v0.99)²² with reference to HapMap2 CEU (release 22, build 36). The analysis dataset included 2,397,023 genotyped or imputed SNPs. Statistical analysis was performed using logistic regression with robust SE estimates in R²³. Adjustments were made for age at baseline, gender, and clinic site.

deCODE Genetics

Subjects: Icelandic asthma patients over 18 years of age (n=1,675) who attended an asthma clinic or emergency room at the National University Hospital of Iceland or the Icelandic Medical Center (Laeknasetrid) during the years 1977 to 2008 were recruited^{24,25}. Asthma diagnosis was based on a combination of physician's diagnosis, a positive reply to the question: "Has a doctor confirmed your asthma diagnosis?", questionnaires pertaining to asthma symptoms and ICD diagnosis when receiving emergency care^{24,25}. Atopy status (defined by at least one positive response to allergens) determined by skin prick testing was available for part of the cohort. Icelandic controls (n=33,408) were participants from various deCODE genetics programs without known asthma. The study was funded by deCODE Genetics.

Genotyping and QC: Genome-wide SNP genotyping of 38,446 Icelanders was done using Illumina HumanHap300 or HumanHapCNV370 SNP chips. After excluding SNPs with MAF < 1%, with genotype yield < 96%, or with P value for deviation from HW equilibrium < 10^{-6} , 290,447 SNPs were used for imputation.

After excluding samples with < 98% genotype yield, and individuals with asthma related phenotypes from the controls, 1,675 asthma cases and 33,408 controls were included in the association analysis.

Data Imputation and Analysis: Imputation of the HapMap Phase 2 Release 22 SNPs was performed using IMPUTE²⁶ with the CEU phased haplotypes as a reference. Tests of allelic association were performed for both genotyped and imputed SNPs using logistic regression implemented in the SNPTEST software, using the –method score option to deal imputation uncertainty. To correct for relatedness of study individuals, and for possible population stratification, the resulting chi-square test statistic was divided by the genomic inflation factor $\lambda_g = 1.26$.

EAGLE Consortium

COPSAC Study

Subjects: The COPSAC 2000 birth cohort study is a prospective clinical study of a birth cohort of 411 infants born to mothers with a history of asthma. The newborns were running at the age of 1 month, the recruitment of which was previously described in detail ²⁷⁻²⁹. Asthma was diagnosed longitudinally by the research physicians based upon a predefined algorithm as previously described ^{28,29}. Diagnostic criteria included (1) recurrent episodes of troublesome lung symptoms recorded in the daily diary cards as five episodes within 6 months, each episode lasting at least three consecutive days; (2) symptoms typical of asthma based on doctors interviews of the parents at the clinical research unit; (3) need for intermittent rescue use of inhaled β_2 -agonist; and (4) response to a 3-month course of inhaled corticosteroids and relapse when stopping treatment. Cases in the current study were defined from having an asthma diagnosis before age 7 years and controls as children with clinical follow-up to age 7 years and no asthma diagnosis.

Genotyping and QC: High throughput genome-wide SNP genotyping were performed using the Illumina Infinium™ II HumanHap550 v1, v3 or quad BeadChip platform (Illumina, San Diego), at the Children's Hospital of Philadelphia's Center for Applied Genomics. All family structures were validated by means of identity-by-descent comparison of all subjects in the data set (proportion identity by descent). Sex was checked by calculating heterozygosity rates of the X chromosome and comparing with reported sex. Low-quality samples were removed based on call rate (<97%), and low-quality markers were removed based on call rate (<98% removed), low minor allele frequency ($\leq 0.1\%$ excluded), or being out of Hardy-Weinberg equilibrium ($P < 10^{-6}$, excluded). Also, principal component analysis-based removal of non-European samples was performed (using HapMap as reference).

Data Imputation and Analysis: Imputation to the CEU panel of HapMap release 22 was performed using MACH 1.0.12³⁰ using phased haplotypes as reference. The genomic build was hg18/NCBI36. Statistical association tests were performed on 2,434,240 variants using SNPTEST v2.2.0, using a genetic additive model and taking genotype uncertainty into account.

Danish National Birth Cohort (DNBC)

Subjects: The Danish National Birth Cohort (DNBC) is a collection of data on 92,274 pregnant women recruited between 1996 and 2002 ³¹. The participating women were interviewed twice during pregnancy and the children are followed through childhood and adolescence. Children were considered to be cases for the “ever asthma” phenotype if a) they had self-reported asthma (by mother) in the DNBC

questionnaire at age 7, or b) if they had any ICD-10 diagnosis code in the chapters J45-J46 recorded in the Danish Hospital Discharge Register before age 7.

Genotyping and QC: Genotype data were available from a genome-wide association study on preterm birth nested within the DNBC³² performed with the Illumina 660 Quad chip (Illumina, Inc., San Diego, CA). The children used in the current study were all born in gestational week 37 or later. The children as well as their parents and grandparents were born in Northwestern Europe. Subjects were filtered based on < 95% call rates, and markers were filtered based on < 98% call rates, Hardy-Weinberg equilibrium p-values < 10⁻³, and minor allele frequencies < 0.01. Total number of asthma cases and controls passing QC was 113 and 850, respectively, and the total number of markers passing QC was 508,059.

Data Imputation and Analysis: Imputation of the HapMap Phase 2 Release 22 SNPs was performed using MACH version 1.0.16³⁰ with the CEU phased haplotypes as a reference. Tests of allelic association were performed by logistic regression using PLINK version 1.07 for genotyped SNPs and MACH2DAT version 1.0.10 for imputed SNPs.

GENERATION R Study

The Generation R Study³³ is a population-based prospective cohort study from fetal life onwards in Rotterdam, the Netherlands (<http://www.generationr.nl>). Assessments in pregnant women and children consisted of physical examinations, fetal ultrasounds, biological samples, and questionnaires. All children were born between April 2002 and January 2006. For the current study, physician-diagnosed ever asthma (no; yes) was assessed using questions adapted from the International Study on Asthma and Allergy in Childhood (ISAAC) at age 6 years.

Genotyping and QC: A GWA scan was performed on DNA isolated from cord blood samples using the Illumina 610 Quad array. Quality control comprised exclusion of duplicates, sex mismatches and low sample call rates. Further, SNPs with call rates <98%, low minor allele frequencies (<0.1%), and significant deviations from Hardy-Weinberg equilibrium ($P < 1 \times 10^{-6}$) were excluded.

Data Imputation and analysis: Imputation was performed using MACH with HapMap2 CEU as the reference population. Statistical analyses were performed using GRIMP³⁴.

GINAplus/LISAplus Studies

Subjects: The influence of Life-style factors on the development of the Immune System and Allergies in East and West Germany PLUS the influence of traffic emissions and genetics (LISAplus) Study is a population based birth cohort study. A total of 3,097 healthy, full term neonates were recruited between 1997 and 1999 in Munich, Leipzig, Wesel and Bad Honnef. The participants were not pre-selected based on family history of allergic diseases^{35,36}.

A total of 5,991 mothers and their newborns were recruited into the German Infant study on the influence of Nutrition Intervention PLUS environmental and genetic influences on allergy development (GINIplus) between September 1995 and June 1998 in Munich and Wesel. Infants with at least one allergic parent and/or sibling were allocated to the interventional study arm investigating the effect of different hydrolysed formulas for allergy prevention in the first year of life³⁷. All children without a family history of allergic diseases and children whose parents did not give consent for the intervention were allocated to the non-interventional arm. Detailed descriptions of the LISAplus and GINIplus studies have been published elsewhere³⁵ and³⁷, respectively. Information on ever having physician-diagnosed asthma and

related symptoms was collected using self-administered questionnaires completed by the parents. The questionnaires were completed at 6, 12, 18, 24 months, 4 and 6 years of age in the LISApplus study and 1, 2, 3, 4 and 6 years in the GINIplus study asking for physician diagnosed asthma at each year of age since the previous follow-up and wheeze in the past 12 months. DNA was collected at the age 6 and 10 years.

Genotyping and QC: Genotyping was performed using the Affymetrix Human SNP Array 5.0 and Affymetrix Human SNP Array 6.0. Genotypes were called using BRLMM-P algorithm (5.0), respectively BIRDSEED V2 algorithm (6.0). In each of the two data sets, criteria for exclusion of individuals were: a call rate below 95%, a heterozygosity outside mean $\pm 4sd$, a failure of the sex check or a failure of the similarity quality control using MDS analysis based on IBS. Criteria for exclusion of variants were: a call rate below 95%, a $MAF < 0.01$ and a HWE $p\text{-value} < 0.01$ ($n=396$) / < 0.00001 ($n=755$).

Data Imputation and Analysis: Imputation was carried out in IMPUTE2 using the HapMap release 22, NBCI build 36 of the CEU reference population. Genome-wide association analysis of asthma phenotypes was conducted using SNPTTEST V2 regressing expected allelic dosage on case-control status.

Manchester Asthma and Allergy Study (MAAS)

Subjects: The Manchester Asthma and Allergy Study is an unselected (i.e. population-based), birth cohort study³⁸⁻⁴². The setting is the maternity catchment area of Wythenshawe and Stepping Hill Hospitals, comprising of 50 square miles of South Manchester and Cheshire, UK, a stable mixed urban-rural population. Cases were children with asthma ever by age 6. All pregnant women were screened for eligibility at antenatal visits (8th-10th week of pregnancy). The study was explained to the parents, and informed consent for initial questionnaires and skin prick testing was obtained. Both parents completed a questionnaire about their and their partner's history of asthma and allergic diseases and smoking habits. If the pregnant woman's partner was not present at the antenatal clinic visit, an invitation was sent for him to attend an open-access evening clinic for skin prick testing and questionnaire. Once both parents had completed questionnaires and skin prick testing, a full explanation of the proposed future follow-up for the child was given. Of the 1,499 couples who met the inclusion criteria (<10 weeks of pregnancy, maternal age >18 years, questionnaire and skin test data available for both parents), 288 declined to take part in the study. A total of 1,185 participants had at least some evaluable data. The children have been followed prospectively, and attended review clinics at ages 1, 3, 5, 8 and 11 years.

Genotyping and QC: DNA samples were genotyping on an Illumina 610 quad chip. The Illumina genotypes were called using the Illumina GenCall application following the manufacturer's instructions. Quality control criteria for samples included: 97% call rate, exclusion of samples with an outlier autosomal heterozygosity (scree-plot visualization) gender validation and sequenome genotype concordance. Quality control criteria for SNPs included a 95% call rate, $HWE > 5.9 \times 10^{-7}$, minor allele frequency > 0.005 .

Data Imputation and Analysis: Genotypes were imputed with IMPUTE² version 2.1.2 with 1000 genomes and hapmap phase 3 reference genotypes. Association analysis was carried out using SNPTTEST v2.4.0²⁶ using frequentist with the score method.

Western Australian Pregnancy Cohort Study (RAINE)

Subjects: The Western Australian Pregnancy Cohort (Raine) Study is a prospective pregnancy cohort where 2,900 were recruited from King Edward Memorial Hospital between 1989 and 1991⁴³. Data were collected throughout pregnancy and the children have been followed-up at ages 1, 2, 3, 5, 8, 10, 14, 17, 18, 20 and

22. Individuals who were ever diagnosed by a doctor with asthma were cases, all other individuals were classified as controls.

Genotyping and QC: Genome-wide SNP genotyping of 1,593 individuals was performed on Illumina's Human660W-Quad BeadChip. Replicates were investigated and the sample with the higher proportion of missing data was excluded. Possible cryptic relatedness between individuals in the cohort was examined, and any pair of individuals who were related with a $\pi > 0.1875$ was investigated and the individual with the higher proportion of missing data was excluded. Subjects were further filtered based on $< 97\%$ call rates. SNP QC was carried out in accordance to WTCCC recommendations⁴⁴; in short, markers were filtered based on $< 95\%$ call rates, Hardy-Weinberg equilibrium p-values $< 5.7 \times 10^{-7}$, and minor allele frequencies < 0.01 . The total number of markers passing QC was 535,632.

Data Imputation and Analysis: Imputation was performed using MACH version 1.0.18c and Minimac with the HapMap2 release 22 as a reference. Tests of allelic association were performed for both genotyped and imputed SNPs using logistic regression with SNPTTEST v2.2.0.

B58C: The British 1958 Birth Cohort Study

Subjects: The British 1958 Birth Cohort (also known as the National Child Development Study) is based on all persons born in Britain during one week in 1958. They were followed up at ages 7, 11 and 16 years by parental interviews and examinations by school medical officers, and at ages 23, 33 and 42 years by means of interviews. The presence of diagnosed asthma or a history of wheezy bronchitis were ascertained by parental interviews in childhood⁴⁵. In adulthood, questionnaires were used to define asthma ever, wheezing ever, and wheezing episodes in the previous 12 months^{46,47}. During 2002-2004 all cohort members still in contact with the study team were visited for a biomedical examination⁴⁸ and invited to contribute blood samples with consent for DNA extraction. The lifetime prevalence of asthma or wheeze in the full cohort was 49.9% and in the subgroup with DNA samples was 48.8%.

For the purposes of this meta-analysis, cases (N=986) were defined as those with a history of "asthma ever" at any age up to 42 years, and controls (N=5,505) were defined as the remainder of the cohort, including those with a history of wheeze or wheezy bronchitis, but no asthma. This followed the convention set by the GABRIEL consortium⁶.

Genotyping and QC: Non-overlapping subsets of the British 1958 cohort DNA collection were genotyped as control samples by the Wellcome Trust Case-Control Consortium⁴⁹ and the Type 1 Diabetes Genetics Consortium⁵⁰, using the Illumina 550K array. From the remainder of the samples, all asthmatic cases and a sample of non-asthmatics were genotyped by the GABRIEL consortium using the Illumina 610K array⁶. These three genome-wide datasets were combined before imputation, selecting 532,203 SNPs which were common to the three deposits.

Each of the three subsets was restricted to subjects of white European ethnicity, but ancestry outliers on principal components analysis were excluded from all analyses. Samples with low average call rate ($< 98\%$) and sex mismatches were also removed. SNPs used for imputation excluded those with a call rate $< 95\%$, minor allele frequency $< 1\%$, HWE p-value < 0.0001 or discordant allele frequencies across the three deposits (p < 0.0001 for pairwise comparisons). For chromosome X, the test of HWE was restricted to

females, and additionally, SNPs with discordant allele frequencies between males and females ($p < 0.0001$) were excluded. As a result, 494,161 (93%) of the 532,203 genotyped SNPs were used for imputation.

Data Imputation and Analysis: Imputation against the HapMap 2 (release 21) CEU reference panel was carried out using MACH 1.0.16. Statistical analysis used the *p*-logist procedure in ProbABEL 0.1-3⁵¹.

EVE Consortium

STAMPEED: SNP Typing for Association with Multiple Phenotypes from Existing Epidemiologic Data

Subjects: The STAMPEED asthma project includes subjects from the Chicago Asthma Genetics Study (C.A.G.) as well as from SARP and CSGA. CAG included both European American and African American subjects collected at The University of Chicago from a) families ascertained through affected sib pairs, b) affected children and their parents, c) adults and children with severe persistent asthma, and d) non-asthmatic control subjects (over the age of 18 years). Samples a-c were recruited in the adult and/or pediatric asthma clinics at University of Chicago Hospital; controls were recruited from the medical center at large. SARP and CSGA included both European American and African American subjects. Control subjects and subjects with mild to severe asthma were recruited at the NHLBI funded Severe Asthma Research Program (SARP) centers and the NHLBI Collaborative Studies on the Genetics of Asthma (CSGA) and characterized according to asthma severity (see^{52,53}).

Genotyping and QC: Genotyping was performed on the Illumina 1Mv1 platform, with individual genotypes called using clustering algorithms as implemented in the BeadStudio software by Illumina. The resulting number of markers was 1,033,467 prior to additional QC measures. Markers were filtered based on 95% call rates, Hardy-Weinberg equilibrium p -values $> 10^{-5}$, consistency in allele frequency from the HapMap ASW (chi-square p -value $> 10^{-5}$), and < 5 heterozygous genotype calls in males for X-linked markers. The total number of markers following QC was 1,025,129. Subjects were filtered based on 95% call rates, matching genetic and reported sex (F_{stat} on X chromosome between -0.2 and 0.3 for females, and between 0.8 and 1 for males), consistency in self-reported ethnicity based on a principal component analysis in Eigenstrat (no obvious clustering with the HapMap CEU for African Americans, and the HapMap YRI for European Americans), and high or low heterozygosity ($F_{\text{stat}} < 0.5$ and > -0.2). Samples were flagged for unexpected pairwise relatedness ($\text{IBD} > 30\%$) or genetic identity ($\text{IBS} > 90\%$), with subsequent filtering performed by selecting a single sample from the pair (or group) having the highest call rate.

Data Imputation and Analysis: Imputation of the phase 2, release 21 HapMap SNPs was performed using MACH, with the consensus CEU haplotypes used as a reference for the European American cases and controls, and the combined consensus CEU and YRI haplotypes for the African American cases and controls. Tests of allelic association were performed on dosages for both genotyped and imputed SNPs using logistic regression in R (<http://CRAN.R-project.org/>). For the set of African American cases and controls, local ancestry was used as a covariate at the SNP level as estimated using the set of genotyped SNPs in the program LAMP - Local Ancestry in admixed Populations⁵⁴. Admixture in African Americans was modeled under 7 generations of admixture with a 2-population model of 81% ancestry from Africa and 19% ancestry from Europe ($\alpha = [0.81, 0.19]$) as estimated based on the first principal component in a PCA analysis using EIGENSTRAT⁵⁵). Windows were offset by a factor of 0.2, the cutoff for linkage was set to 0.1,

and a constant recombination rate was set to 10^{-8} . Local ancestry for imputed SNPs was obtained from the next closest genotyped SNP.

The Children's Health Study (CHS)

Subjects: The Children's Health Study (CHS) is an ongoing cohort study in southern California investigating both genetic and environmental factors related to childhood asthma and lung function growth. The Children's Health Study GWAS was based on a nested case-control sample of 1,249 asthmatics and 1,751 controls selected from within the cohort. All subjects in this GWAS sample were either Hispanic white (HW) (n=1,398) or non-Hispanic white (NHW) (n=1,602). Based on questionnaire responses, children were characterized as having doctor-diagnosed asthma at study entry or during active follow-up (cases), or as never having a diagnosis of asthma (controls).

Genotyping and QC: Study samples were genotyped at the USC Epigenome Center using the Illumina HumanHap550, HumanHap550-Duo or Human610-Quad BeadChip microarrays. Individuals were excluded from analysis with call rates < 90% (n=155). The HumanHap550, HumanHap550-Duo and Human610-Quad contained 366, 366 and 418 SNPs respectively that overlapped with a candidate gene study containing a large number of the subjects in this study (n=2,905). The average concordance rate between matching subjects for these overlapping SNPs was > 99.69% for > 99% of the samples having a call rate > 90%. Subjects with poor concordance with genotypes from the candidate gene study were excluded (n=19). SNPs were excluded prior to imputation if they were not concordant between Illumina genotyped HapMap samples and HapMap2_r21 (< 95% in any population) or not concordant between Illumina genotyped HapMap samples on the HumanHap550 and Human610 (< 95% in any population) (n=8,616). Additionally, SNPs were filtered if they had a call rate < 95% (n=84,381) or departed from Hardy-Weinberg equilibrium ($p < 10^{-5}$) in controls in HW and NHW samples separately (n=762 and 766 respectively).

Data Imputation and Analysis: Imputation was performed with MACH v1.0.16 using the HapMap phase 2 release 21 consensus CEU haplotypes as a reference for the NHW subjects and CEU+ASN haplotypes for the HW subjects. Tests of allelic association were performed on dosages for both genotyped and imputed SNPs using logistic regression using R (www.R-project.org), adjusting for age, community of residence, sex, and ancestry covariates derived from the program STRUCTURE⁵⁶ applied to 557 ancestrally informative markers.

GRAAD: Genomic Research on Asthma in the African Diaspora

Subjects: The Genomic Research on Asthma in the African Diaspora study consists of 498 asthma cases and 500 non-asthmatic controls from the Baltimore-Washington, D.C. metropolitan area who self-reported as African American. Because asthma is often characterized by onset during childhood, there was a deliberate decision to favor adults in the control group to minimize including controls with the potential for developing asthma. A standardized questionnaire based on either the American Thoracic Society⁵⁷ or International Study of Asthma and Allergy in Childhood (ISAAC)⁵⁸ was administered by a clinical coordinator. Asthma was defined as both a reported history of asthma, and a documented history of physician-diagnosed asthma (past or current). All controls (except 50, see below) were likewise administered a standardized questionnaire and were determined to be negative for a history of asthma. Asthma status on 50 controls participating in a study of the genetics of human pigmentation⁵⁹ was not explicitly determined, although "known clinical disease" was among the exclusion criteria. The study

protocol was approved by the institutional review board at either the Johns Hopkins University or Howard University.

Genotyping and QC: Details of genotyping and quality control (QC) have been previously described⁶⁰. Briefly, genotypes were generated at the Center for Inherited Disease Research (CIDR) for 665,352 SNPs on the Illumina HumanHap650Y Versions 1 and 3 BeadChips and the Illumina Infinium II assay protocol. Genotypes were released for 994 GRAAD samples for SNPs with < 5% missing data, no replicate error among other standard QC protocols. Relationships between individuals within each study were verified using PLINK and RELPAL⁶¹, and marker-level QC parameters (including minor allele frequency, differential missing rates between cases and controls, and Hardy Weinberg Equilibrium) were evaluated in PLINK. The genetic structure of African American cases and controls was evaluated using unrelated individuals from the three HapMap “continental” ancestral populations (CEU, YRI, and ASN; www.hapmap.org) using 416 SNPs identified as ancestry informative markers (AIMs) selected for maximal difference between African and European populations in STRUCTURE.

Data Imputation and Analysis: Genotypes were imputed at all HapMap phase 2, release 21 SNPs using MACH, and using a combined panel of the HapMap CEU, YRI and ASN phased haplotypes. For the case-control sample (GRAAD), a two sample t-test was used to compare the allele dosage between cases and controls.

SAPPHERE: The Study of Asthma Phenotypes and Pharmacogenomic Interactions by Race-Ethnicity

Subjects: Study participants received their care through a large health system serving southeast Michigan. Cases were part of the Study of Asthma Phenotypes and Pharmacogenomic Interactions by Race-Ethnicity (SAPPHERE) (n=149), and met the following criteria: age 12-56 years, a prior clinical diagnosis of asthma, no recorded diagnosis of chronic obstructive pulmonary disease (COPD) or congestive heart failure (CHF), a baseline FEV₁ between 40-90% predicted, >12% baseline bronchodilator reversibility, no smoking in the preceding year or <10 pack-year smoking history total, no oral or inhaled corticosteroid use in the 4 weeks preceding screening, and not pregnant at the time of enrollment and not intending to get pregnant during the study period. Controls were obtained from a separate cohort, the Wayne County Health Environment Allergy and Asthma Longitudinal Study. All were women ≥ 21 years of age who also resided in the Detroit metropolitan area and received their care from the same health system. The analytic group for the meta-analysis was restricted to individuals who reported being African American. The Institutional Review Board of Henry Ford Health System approved all of the components of this study.

Genotyping and QC: Genotyping for the cases was performed on the Affymetrix Genome-Wide Human SNP Array 6.0, whereas control individuals were genotyped on either the Mapping 500K Array and the Genome-Wide Human SNP Array 5.0 (Affymetrix, Inc., Santa Clara, California). Genotyping calls and the chip quality control (QC) call rate were assessed using Affymetrix Genotyping Console. Subjects with the following were excluded from the analysis: missing information for one chip (i.e., for those genotyped on the 2 chip Mapping 500K Array), genetic sex inconsistent with reported sex, chips which did not meet the manufacture’s recommended QC call rate, and <90% overall call rate. In order to appropriately match cases and controls, genetic ancestry was estimated in all individuals using markers informative for African and European ancestry and the program PSMIX⁶². We restricted the control set to individuals with ≥30% African ancestry (n=132), which was the lower limit of African ancestry among the cases. Single nucleotide

polymorphisms (SNPs) common to all genotyping platforms were selected and subjected to the following additional criteria: call rate $\geq 95\%$ and an exact Hardy-Weinberg equilibrium test with p -value $>10^{-5}$ among the controls.

Data Imputation and Analysis: A total of 404,088 SNPs were used to impute the HapMap phase 2, release 21 SNPs using MACH with the phased HapMap CEU and YRI haplotypes as a reference. Case/control association tests for asthma status were performed using logistic regression in R (<http://CRAN.R-project.org/>) on genotype dosages, and adjusting for the first principal component from EIGENSTRAT.

Japanese Adult Asthma Research Consortium (JAARC)

Subjects: We selected subjects with childhood-onset asthma from the case subjects who participated in a genome-wide association study (GWAS) for adult asthma in the Japanese population⁶³. A total of 301 subjects with a history of childhood-onset asthma (asthma onset by 15 years of age; age [mean $\pm$ sd], 39.4 $\pm$ 14.8) years; female [%], 47.8) were recruited by physicians' interviews. All subjects with asthma were diagnosed by physicians according to the American Thoracic Society criteria. A total of 3,304 control subjects for the GWAS (age [mean $\pm$ sd], 52.5 $\pm$ 15.1) years; female [%], 44.4) were recruited as described⁶³.

Genotyping and QC: For the GWAS, we genotyped 301 cases using the Illumina HumanHap610-Quad BeadChip and 3,304 controls using the Illumina HumanHap550 BeadChip. We conducted allele sharing analysis and performed principal component analysis for the genotype data of the samples along with European (CEU), African (YRI) and East Asian (Japanese [JPT], and Han Chinese [CHB]) individuals obtained from the Phase II HapMap database by using SMARTPCA. We excluded related samples and outliers who were away from the east Asian cluster. We also excluded SNPs with minor allele frequencies of less than 0.01 in both cases and controls. SNPs having call rates $\geq 99\%$ in both cases and controls were used for the association study. We conducted exact Hardy-Weinberg equilibrium (HWE) analysis, and SNPs with P values less than the cut-off values of the HWE test ($P < 10^{-6}$ in controls) were also excluded from the analysis.

Data Imputation and Analysis: Imputation of HapMap2, Release21 SNPs was performed using MACH with the JPT+CHB haplotypes as a reference. SNPs with P values of HWE in the control group $\geq 1 \times 10^{-6}$ were imputed. In the GWAS, the statistical significance of the association with each SNP was assessed using a logistic regression analysis.

Japan Pediatric Asthma Consortium (JPAC)

Subjects: The part of the study population, genotyping and quality control have been previously described⁶⁴. All subjects with asthma were child or child-onset (<15 years old) asthmatics in Japan. Patients were recruited from 3 pediatric hospitals and 1 pediatric clinic, and the diagnosis of the asthma in all patients was confirmed by specialists in pediatric allergology on the basis of the criteria of the National Institutes of Health, USA, with minor modifications. The control for the GWAS were healthy Japanese adult subjects from Tokyo who had no current history of asthma. All subjects were self-reported Japanese origin.

Genotyping and QC: Genome-wide SNP genotyping of 978 pediatric asthmatics and 694 controls was performed on Illumina's HumanHap550v3/610 Genotyping BeadChip (Illumina, Inc., San Diego, CA). We examined the potential genetic relatedness on the basis of pairwise identity by state for all of the

successfully genotyped samples using the EIGENSTRAT software, and we excluded samples belonging to Han Chinese or Ryukyu, an island located in southern Japan. Subjects were filtered based on < 95% call rates, and markers were filtered based on < 95% call rates, Hardy-Weinberg equilibrium p-values < 10^{-4} , and minor allele frequencies < 0.01. Total number of pediatric asthmatics and controls passing QC was 938 and 672, respectively, and the total number of markers passing QC was 482,803 for HumanHap 610 and 435,199 for HumanHap550v3.

Data Imputation and Analysis: Imputation of the HapMap Phase 2 Release 21 SNPs was performed using MACH version 1.0.17 with the JPT+CHB phased haplotypes as a reference. Tests of allelic association were performed for both genotyped and imputed SNPs using logistic regression with MACH2DAT software version 1.1.9.

GABRIEL Consortium

The GABRIEL Consortium includes 19 studies that contributed to TAGC. Seven of the 19 studies were split into two datasets according to age of onset of asthma (before or after 16 years of age); the INDUSTRIAL subgroup included four studies and the German MAS and MAGICs studies were combined.

EGEA: Briefly, the EGEA study (N=2,120 subjects) combines a case-control and a family-based study of asthma cases with three surveys over 20 years (EGEA1: 1991-1995, EGEA2: 2003-2007 and EGEA3: 2011-2013). The whole study population included 388 asthmatic probands recruited in chest clinics and their 1,317 family members (probands' parents and/or siblings) plus 415 population-based controls⁶⁵. The probands (asthmatics and controls) were between 7 and 70 years old at time of study. All subjects were of European ancestry and were born in France. Data collected through face-to-face interviews and examination included extensive phenotypic characterization (detailed clinical data based on standardized questionnaire, skin prick tests, lung function tests, bronchial responsiveness, blood samples, white blood cell counts, total IgE), data on risk factors (environmental exposures, diet, physical activity, hormone-related events) and drug consumption. For the present study, cases with asthma were defined by a positive answer to the question "Have you ever had asthma attacks?" at EGEA1 or EGEA2. Detailed description of the study can be found in Kauffmann *et al*^{65,66} and at <http://egeanet.vjf.inserm.fr>. After QC, a total of 1,928 EGEA subjects were retained in the GWAS analyses.

ALSPAC: is a population-based, prospective, longitudinal, birth-cohort study that was recruited during pregnancy. Pregnant women resident in Avon, United Kingdom and with estimated dates of delivery 1/4/1991-31/12/1992 were recruited through antenatal clinics.⁶⁷(www.bristol.ac.uk/alspac) Of 14,451 pregnancies recruited, there were 14,072 live births and 13,988 children remained alive at age 1 year. Children have been followed from birth using a combination of self-completion questionnaires sent at regular intervals to their mothers and hands-on assessments at annual dedicated research clinics from age 7 years. Asthma in children was defined as a positive response to the question, "Has a doctor ever told you that your child has asthma?" in a questionnaire sent to their mothers at 91 months after birth. A total of 5,231 children had data on asthma at 91 months with DNA available for genotyping. Of these, 650 (12.4%) had asthma. An equal number of controls were selected at random from the remaining population with questionnaire responses at 91 months and DNA available. After QC a total of 1,216 subjects were retained in the GWAS analyses.

ECRHS: Sixteen centres (eight countries) in the ECRHS contributed samples to the GWAS.^{68,69} (www.ecrhs.org) In each centre, a representative community-based sample of at least 3000 adults aged 20-44 years were invited to complete a brief postal questionnaire asking about respiratory symptoms (ECRHS I - Stage 1) between 1991-1993. A random sample of these (600 per centre) underwent intensive further investigation (ECRHS I - Stage 2 – random sample). Participants who had symptoms highly suggestive of asthma but who had not been selected at random to take part in Stage 2 were also invited to undergo intensive investigations (ECRHS I - Stage 2- enriched sample). About ten years later all adults who had taken part in Stage 2 were re-contacted (ECRHS II) and again asked about respiratory symptoms. Samples suitable for DNA extraction were collected. For the GWAS initiative all cases of asthma were identified (participants from the random or enriched sample who said yes to the question ‘Have you ever had asthma?’ at either ECRHS Stage I or Stage II). Controls were a random sample (of the random sample) who answered ‘no’ to the same question in both surveys. After quality control (QC) criteria a total of 2,101 ECRHS subjects were retained in the GWAS analyses.

BAMSE: Between 1994 and 1996, 4,089 newborn infants were recruited in the BAMSE study, and questionnaire data on baseline study characteristics were obtained.^{70,71} The catchment area included central and north-western parts of Stockholm. At approximately one, two, four, and eight years of age, parents completed questionnaires on their children’s symptoms related to asthma and other allergic diseases. The response rates were 96%, 94%, 92% and 84%, respectively. At eight years of age, all children of the BAMSE study were invited to clinical testing, and blood samples were obtained from 2,480 children. DNA was extracted from 2,033 samples after exclusion of samples with too little blood, lack of questionnaire data, or if parental consent to genetic analysis of the sample was not obtained. From these samples, all children with a doctor’s diagnosis of asthma (ever) were selected as cases and children with no history of asthma or other allergic diseases were selected as controls. After QC a total of 485 subjects were retained in the GWAS analyses.

BUSSELTON: Residents of the town of Busselton in the southwest of Western Australia have been involved in a series of health surveys since 1966. (www.busseltonhealthstudy.com) The population is predominantly of European origin. In 1994/95 there was a follow-up study involving a subset of those who had attended any of the previous surveys. Cases of asthma were defined as those who reported doctor-diagnosed asthma at any survey that they attend from 1966 to 1994 (answer ‘Yes’ to ‘Has your doctor ever told you that you had asthma?’).⁷² Controls are those who have consistently answered ‘No’ to ‘Has your doctor ever told you that you had asthma?’ at all previous surveys that they have attended from 1996 to 1994. For the GWA study, a case control sample of unrelated individuals was selected. After QC a total of 1,207 subjects were retained in the GWAS analyses.

GABRIEL ADVANCED SURVEYS: Are cross-sectional population-based surveys conducted in rural areas of Austria, Germany, and Switzerland. In total, 135,359 children aged 6-12 years were addressed through schools. In a first stage in fall/winter 2006, asthma, allergic disease, and contact to farming environments were assessed using a short parental questionnaire (n=79,888). In a second stage in spring/summer 2007, 9,668 children were selected among families consenting in writing to blood sampling, genetic testing and collection of environmental samples by stratified random sampling to ensure representation of children with high exposure to farming environments.⁷³ Genomic DNA was purified from blood samples using the

Puregene chemistry (QIAGEN, Hilden, Germany) on an Autopure LS instrument (QIAGEN, Hilden, Germany). Genomic DNA and questionnaire data were available for 7,303 children of whom 862 cases and 865 controls were selected for genotyping. A case was defined as a parental report of asthma diagnosed by a doctor at least once or asthmatic bronchitis diagnosed at least twice during lifetime. To account for the stratified random sampling, probability weights were introduced in the statistical analyses. After QC a total of 1,692 subjects were retained in the GWAS analyses.

KSMU: is a population-based case-control study of adult cases of asthma and controls matched for age and sex.^{74,75} A total of 429 unrelated subjects were recruited in this study, (215 patients with asthma and 214 controls). The study subjects were of Russian origin from Central Russia. All patients were recruited at the Department of Pulmonology, Kursk Regional Clinical Hospital between 2003 and 2004. Additional adult patients with asthma and healthy subjects (>200 samples) from the same population were recruited between 2007 and 2008 specially in order to increase final sample size for the GWAS initiative. All patients were diagnosed with asthma by the presence of characteristic symptoms, reversibility of airway obstruction or airway hyperresponsiveness to methacholine. All control subjects were enrolled in accordance with the following criteria: (1) no symptoms and history of allergic diseases, (2) normal total serum IgE levels, (3) and normal pulmonary function test results. Personal data, including smoking status and age of the disease onset, was collected through in-person interviews. After QC a total of 568 subjects were retained in the GWAS analyses.

MRCA-UK: is a collection of nuclear families (207 families total number of subjects 783) ascertained through a proband with $\geq$ Step 3 asthma according to British Thoracic Society guidelines⁷⁶ and four hundred and thirty seven non-asthmatic UK controls (UK-C)⁷⁷. Probands and family members all answered a nurse-administered comprehensive respiratory questionnaire with asthma being defined as positive response to the question “Has your doctor ever told you that you have asthma?”. The children from this collection together with those from MAGICs (below) formed the case control dataset genotyped in the first GWAS for asthma⁷⁷.

MAGICs consists of 655 asthmatic children (asthma diagnosed by a pulmonologist) were derived from the Multicenter Asthma Genetics In Childhood Study with a further 73 asthmatics and 694 unaffected individuals (randomly selected) from Phase One of the International Study of Asthma and Allergy in Childhood (ISAAC)^{78,79}. For the 73 asthmatics from ISAAC, asthma was physician diagnosed and/or reports by parent of recurrent spastic or asthmatic bronchitis. MAGICs together with the children from the MRCA collection formed the case control dataset genotyped in the first GWAS for asthma⁷⁷.

MAS: Consists of 1,314 healthy mature children born in 5 German cities in the year 1990. All children were followed at the age of 1, 3, 6, 12, 18, and 24 months and at yearly intervals thereafter until age 15 years.^{80,81} Clinical and environmental assessment included standardized interviews, questionnaires, physical examinations, and environmental studies. Total and specific IgE antibodies to hen’s egg, cow’s milk, wheat, and soy were determined at the age of 1, 2, 3, 5, 7, and 10 years. DNAs from all children with a doctor diagnosis of asthma were provided for the GWA study. This dataset was combined with the MAGICs dataset.

PIAMA: Is a birth cohort study consisting of two parts: a placebo controlled intervention study in which the effect of mite impermeable mattress covers was studied and a natural history study in which no

intervention took place. In this study, only data from the natural history part of the study are presented. Details of the study design have been published previously.⁸²⁻⁸⁴ Recruitment took place in 1996-1997. A screening questionnaire was distributed to 10,232 pregnant women visiting one of 52 prenatal clinics. 10,232 pregnant women completed a validated screening questionnaire at their prenatal health care clinic. Based on this screening, 7,862 women were invited to participate, of whom 4,146 women agreed and gave informed consent. Mothers reporting a history of asthma, current hay fever or allergy to pets or house dust mite were defined as allergic. Follow-up of the children took place at 3 months of age and yearly from 1 to 8 years of age. The response rates to the annual questionnaires ranged from 3030 (92%) at age 1 to 2732 (83%) at age 8 years. For the GWA study, DNA from childhood asthma cases and a set of matched non-asthmatic controls were provided. After QC a total of 359 subjects were retained in the GWAS analyses.

SAPALDIA: Contributed all self-reported asthma cases as well as a random sample of controls. These subjects were obtained from among 6,055 SAPALDIA cohort subjects that participated in both, the baseline (1991) and follow-up (2002) examinations and agreed to providing blood for genetic analysis. SAPALDIA is a population-based cohort that originally recruited subjects aged 18 to 60 from population registries in eight Swiss communities representing the three largest language groups (German, French, Italian) as well as different levels of air pollution, altitude and degrees of urbanization.⁸⁵⁻⁸⁷ At both baseline and follow-up examination subjects underwent spirometry as well as a detailed interview on respiratory health, smoking history, lifestyle factors and anthropometry. At follow-up, 8,047 of 9,651 baseline subjects re-participated in at least one part of the study and a formal biobank was established. SAPALDIA questions about smoking and asthma status were equivalent to those used by the ECRHS. Asthma status was defined by an affirmative answer to the question “Have you ever had asthma” at baseline and/or follow-up interview. After QC a total of 1,521 subjects were retained in the GWAS analyses.

TOMSK: Is a population-based family study conducted by the Research Institute of Medical Genetics and Siberian State Medical University (Tomsk, Russia) from 1998 onwards.^{88,89} Both nuclear families and extended pedigrees were recruited through atopic bronchial asthmatic probands. All participants were Russians or of a mixed ethnic origin due to marriages between Russians and major East Slavonic populations (Ukrainians, Byelorussians). Altogether, 196 families were studied, out of which 150 families were recruited in Tomsk Region Children Hospital and Tomsk Region Hospital (Tomsk, Russia), and 46 families were recruited in the city of Irkutsk hospitals by the staff of the Irkutsk State Institute of Doctor’s Advanced Training (Irkutsk, Russia). Both probands and their relatives were clinically examined to establish diagnosis of asthma and atopy by the GINA criteria (Global Initiative for Asthma: Global Strategy for Asthma Management and Prevention. <http://www.ginasthma.org>). Besides the clinical examination, laboratory and functional testing were conducted to assess common IgE levels (solid-phase immune-enzyme assay), specific sensitization (skin-prick tests), lung volumes (spirometry), and airway responsiveness (bronchoprovocative tests with methacholine). After QC a total of 681 subjects were retained in the GWAS analyses.

UFA: Is a population-based case-control study of asthma cases and controls matched on age and sex. Cases are unrelated patients with physician-diagnosed asthma and controls disease-free. Subjects are of different ethnic origins (Russians, Tatars and Bashkirs) from Volga-Ural region of Russian Federation. The

Volga-Ural region is located at the border between Europe and Asia and has been the arena of permanent genetic exchanges among Siberian, European, Central Asian and other populations. Anthropologically, Russians, Tatars and Bashkirs are Caucasians and have a varying Asian component. Recruitment of cases and controls was carried out by the Hospital of the Bashkir Medical State University and Ufa Municipal Hospital №21 between 1999 and the year 2007. Asthma patients were diagnosed by pulmonologists on the basis of clinical examination, family and medication history, objective tests of lung function. The controls were healthy subjects who met all the following criteria: (1) no symptoms or history of asthma or other pulmonary diseases; (2) no symptoms or history of atopy; and (3) absence of first-degree relatives with a history of asthma or atopy. After QC a total of 678 subjects were retained in the GWAS analyses.

INDUSTRIAL (the Industrial Cohorts Research Group): The study base consists of several pooled surveys on occupational asthma from Denmark and The Netherlands. The populations come from industries with exposure to known major allergens and irritants; high molecular weight sensitizers (wheat flour, fungal alpha amylase, animal allergens), low molecular weight sensitizers (isocyanate monomers and oligomers) and irritants and inflammatory agents (isocyanates, organic dust, endotoxins, wood). Prevalent cases and industry, gender and age matched study-specific controls were included for the GWAS initiative. Briefly, cases and controls originate from the following population based or prospective cohort studies in agricultural dust and/or animal exposed environments: the Danish SUS study^{90,91}, the Dutch Omega^{92,93} and Veterinarian's Health Study⁹⁴; in wheat and amylase exposed environments: the bakers from the Dutch Bakers health surveillance project⁹⁵; in wood dust exposed environments: workers in the Danish furniture industry of the Danish Wood Dust Cohort;^{96,97} and for isocyanate exposed workers: a population of Dutch industrial spray painters.^{98,99} Asthma, respiratory symptoms and potential confounders like smoking were primarily assessed from questionnaires. More objective health measures were also collated including lung function measurements, a specific bronchial hyperresponsiveness and specific IgE measurements. Exposure was assessed by taking job histories from questionnaires. For all studies extensive measurement series have been collected which have been used to create generic or study specific Job Exposure Matrices by which the exposure can be estimated quantitatively on the basis of job title and sometimes tasks performed¹⁰⁰. After QC a total of 1,227 subjects were retained in the GWAS analyses.

SEVERE: Severe/refractory cases of asthma were recruited through three specialist severe asthma clinics; adult and childhood clinics based at the Royal Brompton Hospital, London and an adult clinic at the Glenfield Hospital, Leicester. Patients attending the Glenfield Hospital clinics had full characterisation and were deemed to have severe/refractory asthma according to a specialised protocol involving parameters of airway inflammation, airway physiology, as well as quality of life and control of symptoms.^{101,102} Those attending Royal Brompton Hospital adult clinics were also fully characterised, with severe asthma defined according to the ATS and ERS definition of severe asthma.^{103,104}

Severe asthma in the pediatric clinic was defined as one or more of the following criteria: (1) Persistent (most days, for at least 3 months) chronic symptoms (the necessity because of symptoms for short-acting β -2 agonists at least three times/week) of airways obstruction despite high dose inhaled corticosteroids (Beclomethasone equivalent 800 mcg/day) and trials of every add-on medication available in the country of residence (these would include, if available, long acting β -2 agonist, leukotriene receptor antagonist,

oral theophylline in the low, anti-inflammatory dose). This group includes Type 1 brittle asthma. (2) Recurrent severe asthma exacerbations despite attempts with medication including trials of allergen avoidance, low dose daily inhaled corticosteroids or intermittent high dose inhaled corticosteroids: *either* at least one admission to an intensive care unit, or at least two hospital admissions requiring intravenous medication/s, *or* ≥ 2 courses of oral steroids during the last year, despite the above therapy. This group includes Type 2 brittle asthma. (3) Persistent airflow obstruction: post oral steroid, post-bronchodilator Z score < -1.96 for FEV₁, with appropriate normative data despite the above therapy. (4) The necessity of prescription of alternate day or daily oral steroids to achieve control of asthma. Children were evaluated in detail to exclude as far as possible non-adherence to therapy, significant co-morbidity (for example, rhinosinusitis and gastroesophageal reflux), psychosocial issues and adverse environmental circumstances as contributing factors to the severity of asthma.¹⁰⁴⁻¹⁰⁷ Because of the relatively small numbers of severe asthmatics, this group were not subdivided into childhood onset and adult onset groups. After QC a total of 290 subjects were retained in the GWAS analyses.

For the SEVERE cohorts, control data for analysis was obtained through the WTCCC (<http://www.wtccc.org.uk>). Data used was from 974 samples genotyped using an Illumina 1.2M custom chip.

Genotyping and QC for all GABRIEL datasets: All Gabriel consortium datasets, except for the MRCA and MAGICs datasets, were genotyped at Centre National de Génotypage (CNG, Evry, France) using the Illumina Human610-Quad array⁶. The MRCA and MAGICs datasets were genotyped using Illumina Sentrix Human-1 and Sentrix HumanHap300 BeadChips, as part of the first asthma GWAS⁷⁷. QC of individuals and SNPs genotyped at CNG was done in each dataset following the same protocol⁶. Briefly, individuals were removed from analysis if they were not of European descent (principal component analysis of each GABRIEL datasets with all HapMap populations), had a low genotyping call rate ($<95\%$) or were discrepant or ambiguous for genetic sex. Single Nucleotide Polymorphisms (SNPs) with call rate lower than 95% or minor allele frequency (MAF) lower than 0.01, or with Hardy-Weinberg equilibrium P-value $<10^{-4}$ were removed. QC for MRCA and MAGICs is detailed Moffatt *et al.*⁷⁷

Imputation and analysis of all GABRIEL datasets: In each dataset, genome-wide imputations were performed using MACH 1.0 software and HapMap Phase 2 (Release 21) as reference panel. SNPs with imputation quality score (rsq) ≥ 0.5 and minor allele frequency $\geq 1\%$ were kept for analysis. Association analysis between asthma and individual SNPs was performed using a logistic regression model that included allele dosage for each SNP and principal components to account for population structure. In family data, a robust sandwich estimation of the variance and a family cluster were used to take into account familial dependencies. Moreover, due to the complex sampling design of the GABRIELA study, survey regression techniques were used for this study to estimate robust standard errors ('svy' command in Stata). All analyses were performed using Stata® version 10 (distributed by Stata Corporation, College Station, Texas, USA).

NTR: The Netherlands Twin Register Study

Subjects: NTR ^{108,109} participants are ascertained because of the presence of twins or triplets in the family and consist of multiples, their parents, siblings and spouses. Twins are born in all strata of society and NTR

represents a general sample from the Dutch population. In this study, 2,867 individuals were included (451 cases and 2,416 controls) with asthma definition: “ever doctor diagnosed asthma ignoring age of onset”.

Genotyping and QC: [La table sup 2 indique “Illumina 660K, Illumina Omni Express 1 mil, Affymetrix Perlegen and Affymetrix 6.0”] Genotype data were checked for European ancestry, Mendelian errors, gender inconsistencies and high genome-wide homozygosity. Checks for relatedness were carried out and only unrelated individuals were used in the analysis. Genotype data were further checked based on Hardy-Weinberg Equilibrium (1×10^{-6} HWE), minor allele frequencies (MAF 0.01), SNP call rate (95% threshold of subjects with missing genotypes per SNP) and sample call rate (75% threshold of missing SNPs per subject).

Data Imputation and Analysis: The reference set used for imputation is HapMap phase II CEU data NCBI build 36 (UCSC hg18). Imputation was carried out using IMPUTE. Genome-wide association analysis was conducted using logistic regression (under an additive model) and including sex, age at examination and 4 principal components as covariates. Association analysis was conducted in PLINK, taking the uncertainty of the imputed genotypes into account.

Rotterdam Study (RS-I, RS-II, RS-III)

Subjects: The Rotterdam Study is a prospective population based cohort study ongoing since 1990 in the city of Rotterdam in The Netherlands. Study design and objectives have been described elsewhere¹¹⁰. The first cohort (RS I) consists of 7,983 participants, aged 55 years and over. The second cohort (RS II) was recruited in 2000 with the same inclusion criteria and consists of 3,011 participants. The third cohort (RS III) consists of 3,932 participants, aged 45 years and over and was recruited in 2006. As of 2008, 14,926 subjects aged 45 years or over comprise the Rotterdam Study cohort. Asthma cases were collected by chart review. Cases were ascertained by 1) Doctor’s diagnosis of Asthma **and** 2) no conflicting respiratory diagnosis. Additionally, prescriptions for respiratory medicines were used for case finding. The number of cases and controls by study is shown in the Supplementary Table 1.

Genotyping and QC: Genotyping was done using the following arrays: Illumina Infinium II HumanHap550 chip v3.0 array (RS-I), the HumanHap550 Duo Arrays and the Illumina Human610-Quad Arrays (RS-II), and the Human 610 Quad Arrays Illumina (RS-III). Exclusions were based on a call rate < 98%, Hardy-Weinberg p-value < 10^{-6} and MAF < 0.01%. SNP exclusions were based on a call rate < 98%, Hardy-Weinberg p-value < 10^{-6} and MAF < 0.01%.

Data Imputation and Analysis: We used the Markov Chain Haplotyping (MACH) package version 1.0.15 software (Rotterdam, The Netherlands; imputed to plus strand of NCBI build 36, HapMap release #22) for the analyses of RS-I and RS-II. RS-III was imputed using MACH 1.0.16. For the analyses we used a MACH2DAT implemented in GRIMP, a high-speed pipeline for Genome Wide Association analyses³⁴.

Dutch Asthma Genetics Consortium (DAGC)

Subjects : The Dutch Asthma GWAS (DAG) cohort consists of in total 915 asthma cases and 986 controls, all from the northern of the Netherlands. The DAG cohort was genotyped in two phases and meta-analysed afterwards. For the first phase, 463 cases were selected from a trio and family study. The 469 controls were non-asthmatic spouses or pseudo-controls of untransmitted alleles in our trio design (GWAS I). For the second phase (GWAS II), 452 asthmatics were selected from previous clinical and genetic studies

performed by our research institute. The 517 controls were selected from the COPACETIC study, a geographically matched population-based study on lung cancer screening in male smokers. All asthmatics had a physician's diagnosis of asthma, asthma symptoms, and BHR to either histamine or methacholine. BHR was measured with a metacholine or histamine challenge test, and defined as PC20 histamine. Controls had no asthma or COPD, nor any evidence of significant airway obstruction.¹¹¹

Genotyping: Genotyping was performed on two platforms, 1. the Illumina HAPMAP 317K platform and 2. HAPMAP Illumina 370 Duo Chip. Quality control was applied; subjects were removed for analysis if they were not of Caucasian descent, had a low genotyping call rate (< 95%) or were discrepant or ambiguous for genetic sex. SNPs were deleted if the call rates were low (95%), if they were not in Hardy-Weinberg Equilibrium ($p < 10E-04$), or if the minor allele frequency was < 0.05.

Data Imputation and Analysis: Imputation was performed using BEAGLE and the HapMap2 CEU reference panel. All statistical analyses have been performed using PLINK v1.07.

Funding and Acknowledgements

Australian Asthma Genetics Consortium (AAGC)

The Australian Asthma Genetics Consortium (AAGC) was funded by the National Health and Medical Research Council of Australia (grant 613627).

ALLERGEN Consortium

AllerGen NCE (DD, AB, AK, CL), is supported by the Canadian Institutes for Health Research (CIHR) (DD, AB, MCY, AK, CL), Canadian Research Chairs (DD), Michael Smith Foundation for Health Research (DD), Inspiraplex NCE (AB), Children's Hospital Foundation of Manitoba (AB), Manitoba Government (AB).

Saguenay-Lac-Saint-Jean Study (SLSJ): The Canada Research Chair held by C Laprise and the funding supports from Canadian Institutes of Health Research (CIHR) enabled the maintenance and continuation of the SLSJ asthma study. C. Laprise is the director of the Asthma Strategic Group of the Respiratory Health Network of the Fonds de la recherche en santé du Québec (FRSQ) and member of Allergen network.

APCAT Consortium

The Framingham Heart Study is supported by the National Heart, Lung and Blood Institute's Framingham Heart Study (Contract No. N01-HC-25195) and its contract with Affymetrix, Inc for genotyping services (Contract No. N02-HL-6-4278). A portion of this research utilized the Linux Cluster for Genetic Analysis (LinGA-II) funded by the Robert Dawson Evans Endowment of the Department of Medicine at Boston University School of Medicine and Boston Medical Center.

The EPIC-Norfolk Study is supported by programme grants from the Medical Research Council, and Cancer Research UK and with additional support from the European Union, Stroke Association, British Heart Foundation, Research into Ageing, Department of Health, The Wellcome Trust and the Food Standards Agency. We acknowledge the contribution of the staff and participants of the EPIC-Norfolk Study.

Northern Finland Birth Cohort of 1966 (NFBC 1966) : NFBC1966 received financial support from the Academy of Finland (project grants 104781, 120315, 129269, 1114194, 24300796, Center of Excellence in Complex Disease Genetics and SALVE), University Hospital Oulu, Biocenter, University of Oulu, Finland (75617), NHLBI grant 5R01HL087679-02 through the STAMPEED program (1RL1MH083268-01), NIH/NIMH (5R01MH63706:02), ENGAGE project and grant agreement HEALTH-F4-2007-201413, and the Medical Research Council, UK (G0500539, G0600705, G1002319, PrevMetSyn/SALVE,) and the European Union's Horizon 2020 research and innovation programme (DynaHEALTH, grant agreement No 633595). The DNA extractions, sample quality controls, biobank up-keeping and aliquotting was performed in the National Public Health Institute, Biomedicum Helsinki, Finland and supported financially by the Academy of Finland and Biocentrum Helsinki. A. Couto Alves acknowledges the European Commission, Framework 7, grant number 223367. We thank the late Professor Paula Rantakallio (launch of NFBC1966), and Ms Outi Tornwall and Ms Minttu Jussila (DNA biobanking). The authors would like to acknowledge the contribution of the late Academician of Science Leena Peltonen.

The FINRISK Study is mainly funded through budgetary funds of THL. VS was supported by the Academy of Finland, grant number 139635 and by the Finnish Foundation for Cardiovascular Research

The Health 2000 Study is funded by the National Institute for Health and Welfare (THL), the Finnish Centre for Pensions (ETK), The Social Insurance Institution of Finland (KELA), The Local Government Pensions Institution (KEVA) and other organizations listed on the website of the survey (<http://www.terveys2000.fi>). GWAS genotyping was supported by the Wellcome Trust Sanger Institute.

The Helsinki Birth Cohort Study (HBCS) has been supported by grants from the Academy of Finland, the Finnish Diabetes Research Society, Folkhälsan Research Foundation, Novo Nordisk Foundation, Finska Läkaresällskapet, Signe and Ane Gyllenberg Foundation, University of Helsinki, Ministry of Education, Ahokas Foundation, Emil Aaltonen Foundation.

The Young Finns Study (YFS) has been financially supported by the Academy of Finland: grants 286284, 134309 (Eye), 126925, 121584, 124282, 129378 (Salve), 117787 (Gendi), and 41071 (Skidi); the Social Insurance Institution of Finland; Competitive State Research Financing of the Expert Responsibility area of Tampere, Turku and Kuopio University Hospitals (grant X51001); Juho Vainio Foundation; Paavo Nurmi Foundation; Finnish Foundation for Cardiovascular Research; Finnish Cultural Foundation; Tampere Tuberculosis Foundation; Emil Aaltonen Foundation; Yrjö Jahnsson Foundation; Signe and Ane Gyllenberg Foundation; and Diabetes Research Foundation of Finnish Diabetes Association.

CARe and CHARGE Consortium studies

CARDIA: The CARDIA study was supported by the American Asthma Foundation (to JNH) and NIH T32 HD040128 (to DBK).

MESA: MESA and the MESA SHARe project are conducted and supported by the National Heart, Lung, and Blood Institute (NHLBI) in collaboration with MESA investigators. Support for MESA is provided by contracts HHSN268201500003I, N01-HC-95159, N01-HC-95160, N01-HC-95161, N01-HC-95162, N01-HC-95163, N01-HC-95164, N01-HC-95165, N01-HC-95166, N01-HC-95167, N01-HC-95168, N01-HC-95169, UL1-TR-001079, UL1-TR-000040, and DK063491. Funding for SHARe genotyping was provided by NHLBI Contract N02-HL-64278. Genotyping was performed at Affymetrix (Santa Clara, California, USA) and the Broad Institute of Harvard and MIT (Boston, Massachusetts, USA) using the Affymetrix Genome-Wide Human SNP Array 6.0. MESA Lung acknowledges NIH/NHLBI (R01 HL077612). MESA Lung SHARe acknowledges NIH/NHLBI (RC1 HL100543).

ARIC: The Atherosclerosis Risk in Communities Study is carried out as a collaborative study supported by National Heart, Lung, and Blood Institute contracts (HHSN268201100005C, HHSN268201100006C, HHSN268201100007C, HHSN268201100008C, HHSN268201100009C, HHSN268201100010C, HHSN268201100011C, and HHSN268201100012C), R01HL087641, R01HL59367 and R01HL086694; National Human Genome Research Institute contract U01HG004402; and National Institutes of Health contract HHSN268200625226C. Infrastructure was partly supported by Grant Number UL1RR025005, a component of the National Institutes of Health and NIH Roadmap for Medical Research. Work specifically for this paper was supported in part by the Intramural Research Program of the NIH, National Institute of Environmental Health Sciences (Z01 ES43012).

The authors thank the staff and participants of the ARIC study for their important contributions. The authors also thank Jianping Jin PhD of Westat, Inc (RTP, NC) and Frank Day, PhD of NIEHS for expert computational assistance.

Cardiovascular Health Study (CHS): This Cardio Vascular Health Study research was supported by NHLBI contracts HHSN268201200036C, HHSN268200800007C, N01HC55222, N01HC85079, N01HC85080, N01HC85081, N01HC85082, N01HC85083, N01HC85086; and NHLBI grants U01HL080295, R01HL087652, R01HL105756, R01HL103612, R01HL120393, and R01HL130114 with additional contribution from the National Institute of Neurological Disorders and Stroke (NINDS). Additional support was provided through R01AG023629 from the National Institute on Aging (NIA). A full list of principal Cardio Vascular Health Study investigators and institutions can be found at CHS-NHLBI.org. The provision of genotyping data was supported in part by the National Center for Advancing Translational Sciences, CTSI grant UL1TR000124, and the National Institute of Diabetes and Digestive and Kidney Disease Diabetes Research Center (DRC) grant DK063491 to the Southern California Diabetes Endocrinology Research Center.

deCODE Genetics

We thank all participants who provided data for the deCODE study and our valued colleagues who contributed to data collection and phenotypic characterization of clinical samples, genotyping and analysis of genome sequence data.

EAGLE Consortium

COPSAC is funded by private and public research funds all listed on www.copsac.com. The Lundbeck Foundation (Grant no R16-A1694); The Danish Ministry of Health (Grant no 903516); Danish Council for Strategic Research (Grant no 0603-00280B); *The Danish Council for Independent Research (Grant no 10-082884 and 271-08-0815)* and The Capital Region Research Foundation (No grant number) have provided core support for COPSAC.

We express our deepest gratitude to the children and families of the COPSAC₂₀₀₀ cohort study for all their support and commitment. We acknowledge and appreciate the unique efforts of the COPSAC research team.

The Danish National Birth Cohort (DNBC) study is a result of major grants from the Danish National Research Foundation, the Danish Pharmacists' Fund, the Egmont Foundation, the March of Dimes Birth Defects Foundation, the Augustinus Foundation, and the Health Fund of the Danish Health Insurance Societies. The DNBC biobank is a part of the Danish National Biobank resource, which is supported by the Novo Nordisk Foundation. The generation of GWAS genotype data for the DNBC samples was carried out within the Gene Environment Association Studies (GENEVA) consortium with funding provided through the National Institutes of Health's Genes, Environment, and Health Initiative (U01HG004423; U01HG004446; U01HG004438). Dr. Bjarke Feenstra was supported by an Oak Foundation fellowship.

GENERATION R: The Generation R study is made possible by financial support from the Erasmus Medical Centre, Rotterdam, the Erasmus University Rotterdam and the Netherlands Organization for Health Research and Development. Dr Vincent Jaddoe received grants from the Netherlands Organization for Health Research and Development (ZonMw-VIDI 016.136.361) and the European Research Council (ERC-2014-CoG-64916). Dr Liesbeth Duijts received funding from the European Union's Horizon 2020 co-funded programme ERA-Net on Biomarkers for Nutrition and Health (ERA HDHL) (ALPHABET project (no 696295;

2017), ZonMW The Netherlands (no 529051014; 2017). The project received funding from the European Union's Horizon 2020 research and innovation programme (LIFECYCLE project, grant agreement no 733206; 2016). The researchers are independent from the funders. The Generation R Study is conducted by the Erasmus Medical Center in close collaboration with the School of Law and Faculty of Social Sciences of the Erasmus University Rotterdam, the Municipal Health Service Rotterdam area, Rotterdam, the Rotterdam Homecare Foundation, Rotterdam and the Stichting Trombosedienst & Artsenlaboratorium Rijnmond (STAR-MDC), Rotterdam. We gratefully acknowledge the contribution of children and parents, general practitioners, hospitals, midwives and pharmacies in Rotterdam. The generation and management of GWAS genotype data for the Generation R Study were done at the Genetic Laboratory of the Department of Internal Medicine, Erasmus MC, The Netherlands. We would like to thank Karol Estrada, Dr. Tobias A. Knoch, Anis Abuseiris, Luc V. de Zeeuw, and Rob de Graaf, for their help in creating GRIMP, BigGRID, MediGRID, and Services@MediGRID/D-Grid, (funded by the German Bundesministerium fuer Forschung und Technology; grants 01 AK 803 A-H, 01 IG 07015 G) for access to their grid computing resources. We thank Mila Jhamai, Manoushka Ganesh, Pascal Arp, Marijn Verkerk, Lizbeth Herrera and Marjolein Peters for their help in creating, managing and QC of the GWAS database. Also, we thank Karol Estrada and Carolina Medina-Gomez for their support in creation and analysis of imputed data.

GINIplus/LISAplus: The authors thank all families for participation in the studies and the LISAplus and GINIplus study teams for their excellent work. The GINIplus study team wishes to acknowledge the following: Institute of Epidemiology I, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg (Heinrich J, Brüske I, Schulz H, Flexeder C, Zeller C, Standl M, Schnappinger M, Sußmann M, Thiering E, Tiesler C); Department of Pediatrics, Marien-Hospital, Wesel (Berdel D, von Berg A); Ludwig-Maximilians-University of Munich, Dr von Hauner Children's Hospital (Koletzko S); Child and Adolescent Medicine, University Hospital rechts der Isar of the Technical University Munich (Bauer CP, Hoffmann U); IUF- Environmental Health Research Institute, Düsseldorf (Schikowski T, Link E, Klümper C). The LISAplus study team wishes to acknowledge the following: Helmholtz Zentrum München, German Research Center for Environmental Health, Institute of Epidemiology I, Munich (Heinrich J, Schnappinger M, Brüske I, Sußmann M, Lohr W, Schulz H, Zeller C, Standl M); Department of Pediatrics, Municipal Hospital "St. Georg", Leipzig (Borte M, Gnoldtke E); Marien Hospital Wesel, Department of Pediatrics, Wesel (von Berg A, Berdel D, Stiers G, Maas B); Pediatric Practice, Bad Honnef (Schaaf B); Helmholtz Centre of Environmental Research – UFZ, Department of Environmental Immunology/Core Facility Studies, Leipzig (Lehmann I, Bauer M, Röder S, Schilde M, Nowak M, Herberth G, Müller J, Hain A); Technical University Munich, Department of Pediatrics, Munich (Hoffmann U, Paschke M, Marra S); Clinical Research Group Molecular Dermatology, Department of Dermatology and Allergy, Technische Universität München (TUM), Munich (Ollert M).

Manchester Asthma and Allergy Study (MAAS): We would like to thank the children and their parents for their continued support and enthusiasm. We greatly appreciate the commitment they have given to the project. We would also like to acknowledge the hard work and dedication of the study team (post-doctoral scientists, research fellows, nurses, physiologists, technicians and clerical staff). This report includes independent research supported by National Institute for Health Research Respiratory and Allergy Clinical Research Facility at University Hospital of South Manchester NHS Foundation Trust. The

views expressed in this publication are those of the author(s) and not necessarily those of the NHS, the National Institute for Health Research or the Department of Health. This research was supported by the Asthma UK Grants No 301 (1995-1998), No 362 (1998-2001), No 01/012 (2001-2004), No 04/014 (2004-2007), The BMA James Trust, The Moulton Charitable Foundation (2004-current), The Medical Research Council (MRC) Grants G0601361, MR/K002449/1 and MR/L012693/1, and Angela Simpson is supported by the NIHR Manchester Biomedical Research Centre. The authors would like to acknowledge the North West Lung Centre Charity for supporting this project

The Western Australian Pregnancy Cohort Study (RAINE): The Raine study was supported by the National Health and Medical Research Council of Australia [grant numbers 403981 and 003209] and the Canadian Institutes of Health Research [grant number MOP-82893]. The authors are grateful to the Raine Study participants and their families, and to the Raine Study research staff for cohort coordination and data collection. The authors gratefully acknowledge the NH&MRC for their long term contribution to funding the study over the last 20 years and also the following Institutions for providing funding for Core Management of the Raine Study: The University of Western Australia (UWA), Curtin University, Raine Medical Research Foundation, UWA Faculty of Medicine, Dentistry and Health Sciences, the Telethon Kids Institute, the Women and Infants Research Foundation (King Edward Memorial Hospital and Edith Cowan University). This work was supported by resources provided by the Pawsey Supercomputing Centre with funding from the Australian Government and the Government of Western Australia. The authors gratefully acknowledge the assistance of the Western Australian DNA Bank (National Health and Medical Research Council of Australia National Enabling Facility).

The British 1958 Birth Cohort Study (B58C)

We acknowledge use of phenotype and genotype data from the British 1958 Birth Cohort DNA collection, funded by the Medical Research Council grant G0000934 and the Wellcome Trust grant 068545/Z/02. Genotyping for the B58C-WTCCC subset was funded by the Wellcome Trust grant 076113/B/04/Z. The B58C-T1DGC genotyping utilized resources provided by the Type 1 Diabetes Genetics Consortium, a collaborative clinical study sponsored by the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), National Institute of Allergy and Infectious Diseases (NIAID), National Human Genome Research Institute (NHGRI), National Institute of Child Health and Human Development (NICHD), and Juvenile Diabetes Research Foundation International (JDRF) and supported by U01 DK062418. B58C-T1DGC GWAS data were deposited by the Diabetes and Inflammation Laboratory, Cambridge Institute for Medical Research (CIMR), University of Cambridge, which is funded by Juvenile Diabetes Research Foundation International, the Wellcome Trust and the National Institute for Health Research Cambridge Biomedical Research Centre; the CIMR is in receipt of a Wellcome Trust Strategic Award (079895). The B58C-GABRIEL genotyping was supported by a contract from the European Commission Framework Programme 6 (018996) and grants from the French Ministry of Research.

EVE Consortium

This work was supported by grants from the Office of the Director, NIH to C.O. and D.L.N. and the National Heart, Lung, and Blood Institute (RC2 HL101651 to C.O. and D.L.N.; HL087665 to D.L.N.;

HL070831, HL072414 and HL049596 to C.O.; HL064307 and HL064313 to F.D.M.; HL075419, HL65899, HL083069, HL066289, HL087680, HL101543 and HL101651 to S.T.W.; HL079055 and HL118267 to L.K.W. and the National Institute of Diabetes and Digestive and Kidney Diseases to L.K.W. (DK064695 and DK113003).; HL087699, HL49612, HL075417, HL04266 and HL072433 to K.C.B.; HL061768 and HL076647 to F.D.G.; HL087680 to W.J.G.; HL078885 and HL088133 to E.G.B.; HL87665 to D.A.M.; and HL69116, HL69130, HL69149, HL69155, HL69167, HL69170, HL69174 and HL69349 to D.A.M., E.R.B., W.W.B., W.J.C., M.C., K.F.C., S.C.E., E.I. and S.E.W.), the National Institutes of Allergy and Infectious Disease (AI070503 to C.O.; AI079139 and AI061774 to L.K.W.; AI50024, AI44840 and AI41040 to K.C.B.; and AI077439 to E.G.B.), the National Institute of Diabetes and Digestive and Kidney Diseases to L.K.W. (DK064695); the National Institutes of Environmental Health Sciences (ES09606, ES018176 and ES015903 to K.C.B.; ES007048, ES009581, R826708, RD831861 and ES011627 to F.D.G.; ES015794 to E.G.B.; and the Division of Intramural Research, Z01 ES049019, to S.J.L.); the National Center for Research Resources (RR03048 to K.C.B).

Japanese Adult Asthma Research Consortium (JAARC)

We thank all the patients and volunteers for participating in the study and also thank the collaborating physicians, Dr. Mayumi Tamari, and Dr. Tomomitsu Hirota for helping with sample and data collection. We are grateful to the members of BioBank Japan and the Rotary Club of Osaka-Midosuji District 2660 Rotary International in Japan for supporting our study. This work was conducted as a part of the BioBank Japan Project supported by of Ministry of Education, Culture, Sports, Science and Technology, Japan. This work was also partly supported by grants from the Ministry of Health, Labour and Welfare, Japan.

Japanese Pediatric Asthma Consortium (JPAC)

This project was funded by a grant from the National Institute of Biomedical Innovation (ID05-41) to K Matsumoto and by Grant-in-Aids for Scientific Research from the Ministry of Education, Culture, Sports, Science, and Technology (18018003) to E Noguchi. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

GABRIEL Consortium

Genotyping of the Gabriel data was supported by grants from the European Commission (No. LSHB-CT-2006-018996-GABRIEL) and the Wellcome Trust (WT084703MA). Funding for individual GABRIEL studies can be found in Moffatt *et al.*⁶

INDUSTRIAL: We would like to acknowledge the other team member of the various Industrial surveys including: C. Hjort (Danish SUS cohort), I. Schaumburg, G. Jacobsen (Danish Wood Dust cohort), E. Meijer, E. Suarhana (Dutch Bakers Health Surveillance Project), L.A.M. Smit, G. Doekes (Dutch OMEGA study) and A. Pronk, G. Doekes, J-W. Lammers, L. Preller (Dutch Isocyanate study). We would also like to thank the following for their support: the Dutch Asthma Foundation; Ceffic, Long-range research initiative, Belgium; the Dutch Ministry of Social Affairs and Employment; The Dutch Product Board for Grains, Seeds and Pulses; TNO Quality of Life; The Danish Agency for Science Technology and Innovation, The Danish Medical Research Council, The Danish Agricultural Research Council, Helsefonden, The P.C. Petersen Foundation,

The Danish Working Environment Research Fund, The Danish Research Council Aarhus University and The Danish Lung Association. The Danish Work Environment Foundation; Viborg County; The Danish Medical Research Council; The Health Insurance Fund; The Danish Lung Association.

SAPALDIA is funded by the Swiss National Science Foundation grants no 33CS30-148470/1&2, 33CSCO-134276/1, 33CSCO-108796

THE NETHERLANDS TWIN REGISTER (NTR) Study

Funding was obtained from the Netherlands Organization for Scientific Research (NWO) and The Netherlands Organisation for Health Research and Development (ZonMW) grants 904-61-090, 985-10-002, 904-61-193, 480-04-004, 400-05-717, Addiction-31160008, Middelgroot-911-09-032, Spinozapremie 56-464-14192, Biobanking and Biomolecular Resources Research Infrastructure (BBMRI –NL, 184.021.007). VU Institute for Health and Care Research (EMGO+); the European Community's Seventh Framework Program (FP7/2007-2013), ENGAGE (HEALTH-F4-2007-201413); the European Research Council (ERC Advanced, 230374, ERC Starting grant 284167), Rutgers University Cell and DNA Repository (NIMH U24 MH068457-06), the Avera Institute, Sioux Falls, South Dakota (USA) and the National Institutes of Health (NIH, R01D0042157-01A, MH081802; R01 DK092127-04, Grand Opportunity grants 1RC2 MH089951). Part of the genotyping and analyses were funded by the Genetic Association Information Network (GAIN) of the Foundation for the National Institutes of Health. Computing was supported by BiG Grid, the Dutch e-Science Grid, which is financially supported by NWO.

Rotterdam Study (RS-I, RS-II, RS-III)

The generation and management of GWAS genotype data for the Rotterdam Study is supported by the Netherlands Organization of Scientific Research NWO Investments (nr. 175.010.2005.011, 911-03-012). This study is funded by the Research Institute for Diseases in the Elderly (014-93-015; RIDE2), the Netherlands Genomics Initiative (NGI)/Netherlands Organisation for Scientific Research (NWO) project nr. 050-060-810. We thank Pascal Arp, Mila Jhamai, Marijn Verkerk, Lizbeth Herrera and Marjolein Peters for their help in creating the GWAS database, and Karol Estrada and Maksim V. Struchalin for their support in creation and analysis of imputed data. The Rotterdam Study is funded by Erasmus Medical Center and Erasmus University, Rotterdam, Netherlands Organization for the Health Research and Development (ZonMw), the Research Institute for Diseases in the Elderly (RIDE), the Ministry of Education, Culture and Science, the Ministry for Health, Welfare and Sports, the European Commission (DG XII), and the Municipality of Rotterdam. The authors are grateful to the study participants, the staff from the Rotterdam Study and the participating general practitioners and pharmacists. We would like to thank Karol Estrada, Dr. Fernando Rivadeneira, Dr. Tobias A. Knoch, Anis Abuseiris, Luc V. de Zeeuw, and Rob de Graaf (Erasmus MC Rotterdam, The Netherlands), for their help in creating GRIMP, and BigGRID, MediGRID, and Services@MediGRID/D-Grid, (funded by the German Bundesministerium fuer Forschung und Technology; grants 01 AK 803 A-H, 01 IG 07015 G) for access to their grid computing resources.

Dutch Asthma Genetics Consortium (DAGC)

The Dutch Asthma GWAS was supported by Netherlands Asthma Foundation grants AF 95.09, AF 98.48, AF 3.2.02.51 and AF 3.2.07.015 and a grant from the University Medical Center Groningen.

Contributing groups for TAGC and by study

TAGC Consortium coordinating group and writing group: Florence Demenais, Kathleen C. Barnes, William OC Cookson, Miriam F. Moffatt, Carole Ober, Dan L Nicolae

TAGC Analysis Team: Florence Demenais, Patricia Margaritte-Jeannin, Dan Nicolae, Myriam Brossard, Sébastien Letort, Hamida Mohamdi, Amaury Vaysse

Australian Asthma Genetics Consortium (AAGC)

Study-PIs: Manuel A Ferreira, Melanie C Matheson, Colin F Robertson, Philip J Thompson;

Data collection or analysis: Manuel A Ferreira, Melanie C Matheson, Colin F Robertson, Graham Jones and Philip J Thompson.

Collaborators of the Australian Asthma Genetics Consortium (AAGC)

David L. Duffy^a, Dale R. Nyholt^a, John Beilby^{b-d}, Svetlana Baltic^e, Loren Price^e, Faang Cheah^e, Desiree Mészáros^f, Scott D. Gordon^a, Melissa C. Southey^g, Margaret J. Wright^a, James Markos^h, Li P. Chung^e, Anjali K. Henders^a, Graham Gilesⁱ, Suzanna Temple^e, John Whitfield^a, Brad Shelton^e, Chalermchai Mitrpant^e, Minh Bui, PhD,^j Mark Jenkins^j, Haydn Walters^f, Michael J. Abramson^k, Michael Hunter^{l,d}, Bill Musk^{l,d,m,n}, Peter Le Souëf,^o Guy Marks^p, Shyamali C. Dharmage^j, Grant W. Montgomery,^a Alan James,^{c,m,d}, Nicholas G. Martin^a

^a QIMR Berghofer Medical Research, Brisbane, Australia.

^b PathWest Laboratory Medicine of Western Australia (WA), Nedlands, Australia.

^c School of Pathology and Laboratory Medicine, The University of WA, Nedlands, Australia.

^d Busselton Population Medical Research Foundation, Sir Charles Gairdner Hospital, Perth, Australia.

^e Institute for Respiratory Health, Harry Perkins Institute of Medical Research, Nedlands, Australia

^f Menzies Research Institute, Hobart, Australia.

^g Department of Pathology, The University of Melbourne, Melbourne, Australia.

^h Launceston General Hospital, Launceston, Australia.

ⁱ Cancer Epidemiology Centre, The Cancer Council Victoria, Melbourne, Australia.

^j Melbourne School of Population and Global Health, The University of Melbourne, Melbourne, Australia.

^k Department of Epidemiology & Preventive Medicine, Monash University, Melbourne, Australia

^l School of Population Health, The University of WA, Nedlands, Australia

^m School of Medicine and Pharmacology, University of Western Australia, Nedlands, Australia

ⁿ Department of Respiratory Medicine, Sir Charles Gairdner Hospital, Perth, Australia

^o School of Paediatrics and Child Health, Princess Margaret Hospital for Children, Perth, Australia

^p Woolcock Institute of Medical Research, University of Sydney, Sydney, Australia

ALLERGEN Consortium

CAPPS and SAGE studies:

Study-PIs: Allan B Becker obtained funding and was responsible for study subjects assessment, biologic samples collection; Moira Chan-Yeung obtained funding and was responsible for study subjects assessment; Denise Daley obtained funding, performed statistical design and statistical analysis of CAPPS and SAGE, wrote descriptions of cohort and statistical analyses and is AllerGen representative to TAGC consortium; Anita L Kozyrskyj was responsible for study design and obtained funding of SAGE study. Statistical analysis: Denise Daley and Julie E Park conducted statistical analysis of CAPPS and SAGE.

SLSJ Study : Study-PIs: Catherine Laprise & Thomas J Hudson; Study design and management: C.L.

APCAT Consortium

APCAT-PIs: Joel N. Hirschhorn, Marjo-Riitta Jarvelin, Veikko Salomaa

FHS: Study-PI: George T O'Connor; Data collection or analysis: Sailaja Vedantam, Zofia Gajdos

EPIC-Norfolk: Study-PI: Nick J Wareham; Data collection or analysis: Jing Hua Zhao, Robert Scott

NFBC1966: Study-PI: Marjo-Riitta Jarvelin; Data collection or analysis: Alexessander Couto Alves, Adaikalavan Ramasamy

FINRISK: Study-PI: Veikko Salomaa; Data collection or analysis: Mikko Kuokkanen, Tarja Laitinen

H2000: Study-PIs: Markku Heliövaara, Pekka Jousilahti; Data collection or analysis: Mikko Kuokkanen, Tari Haahtela

HBCS: Study-PI: Johan G Eriksson; Data collection or analysis: Elisabeth Widén, Aarno Palotie

YFS: Study-PI: Olli T Raitakari; Data collection or analysis: Terho Lehtimäki, Mika Kähönen

CARe Consortium

CARe-PIs: Joel N. Hirschhorn, Stephen S. Rich

Data collection or analysis : Wendy B White, Lewis J. Smith, Rajesh Kumar, Kristin M. Burkart, Cameron D Palmer, David B Kantor

MESA

MESA-PIs: R. Graham Barr (data collection, design and evaluation of pulmonary data in CARe)

and Stephen S. Rich (evaluation of MESA genetic QC, pulmonary data and analysis results)

Data integration: Kathleen M Donohue; Statistical analysis of MESA European-American data: Ani Manichaikul.

ARIC

Study-PI: Stephanie J London

Data collection or analysis: Stephanie J London, Laura R Loehr

Cardiovascular Health Study (CHS)

Study-PIs: Sina A. Gharib (study design, data analysis), Susan R. Heckbert (study design, data analysis)

Data analysis: Guo Li

deCODE Genetics

Study PIs: Kari Stefansson (study design and management); Ingileif Jonsdottir (study design, data collection, genotyping analyses); Gudmar Thorleifsson (study design, statistical analysis); Unnur Thorsteinsdottir (study design, genotyping management); Daniel F. Gudbjartsson (study design, statistical analysis)

Data collection and analysis: Ingileif Jonsdottir, Gudmar Thorleifsson, Daniel F. Gudbjartsson

Phenotype data ascertainment, coordination of subject recruitment: Unnur Steina Bjornsdottir

EAGLE CONSORTIUM

EAGLE-PI: Hans Bisgaard

COPSAC: Study-PIs, design of the study, planning of phenotyping: Hans Bisgaard, Klaus Bønnelykke; Data analysis: Eskil Kreiner, Johannes Waage

DNBC - Danish National Birth Cohort: Study-PI: Mads Melbye; Phenotyping, Data analysis: Bjarke Feenstra, Frank Geller

GENERATION R: Study-PI: Johan C. de Jongste; Analysis, data collection: Ralf JP van der Valk, Liesbeth Duijts, Vincent W.V. Jaddoe

GINIplus/LISAplus: Study-PI, Phenotyping : Joachim Heinrich; Statistical Analysis: Marie Standl;
Genotyping : Carla MT Tiesler
MAAS: Study-PIs: Angela Simpson, Adnan Custovic; Statistical analysis: John A Curtin
RAINE: Study-PI: Patrick Holt; Statistical analysis : Wei Ang, Craig E. Pennell

B58C: The British 1958 Birth Cohort Study

PI and data analyst: David P Strachan

EVE Consortium

Study-PIs: Carole Ober, Dan L Nicolae, Kathleen C Barnes, Eugene Bleecker, Esteban Burchard, Jim Gauderman, Frank Gilliland, Stephanie J London, Fernando J Martinez, Deborah Meyers, Isabelle Romieu, Scott T. Weiss, L. Keoki Williams, Benjamin A Raby
Data collection or analysis: Dan L Nicolae, Jim Gauderman, Stephanie J London, Deborah Meyers, Dara G Torgerson, Benjamin A. Raby, Blanca E. Himes, Penelope E Graves, Muhammad T. Salam, Celeste Eng, Blanca E Del-Rio-Navarro, James J Yang, Albert M. Levin, Rachel A Myers, Rasika A. Mathias, Terri H. Beaty

JAARC Consortium

Study-PI: Tatsuhiko Tsunoda,
Data collection or analysis: Tatsuhiko Tsunoda, Atsushi Takahashi, Michiaki Kubo,

JPAC Consortium

Management of the study and statistical analysis: Emiko Noguchi
Data collection for the pediatric asthma cases: Kenji Matsumoto
Data collection for the controls: Hiroshi Hirose

GABRIEL Consortium

GABRIEL-PIs: William OC Cookson, Erika Von Mutius
GABRIEL genotyping: Mark Lathrop, Ivo Gut
GABRIEL analysis team: Emmanuelle Bouzigon, Florence Demenais, Martin Farrall, David P Strachan
GABRIEL writing group: Miriam Moffatt, Ivo Gut, Florence Demenais, David Strachan Emmanuelle Bouzigon, Simon Heath, Erika Von Mutius, Martin Farrall, Mark Lathrop, William OC Cookson
EGEA: Valérie Siroux (Study-PI), Florence Demenais (PI-genetic component), Mark Lathrop (genomics), Emmanuelle Bouzigon (statistical analysis)
ALSPAC: John Henderson (Study-PI), Raquel Granell (statistical analysis), Wendy L. McArdle (genotyping)
ECRHS: Deborah Jarvis (Study-PI), Christer Janson, Joachim Heinrich
BAMSE: Erik Melén, Magnus Wickman & Goran Pershagen (Study-PIs)
BUSSELTON: Arthur W Musk (Study-PI, phenotype data collection and collection of DNA), Alan James (Study-PI, phenotype data collection and collection of DNA from Busselton), Jennie Hui (overseeing Busselton Population Medical Research Institute genetic collaborations; genotyping), John Beilby (Study-PI, curator of genotypic collection for all Busselton studies)
GABRIEL Advanced Studies: Erika Von Mutius (Study-PI), Michael Kabesch (phenotyping), Jon Genuneit (statistical analysis)
KSMU: Alexey Polonikov (Study-PI), Maria Solodilova (assessment questionnaire of patients/controls, blood/DNA samples preparation), Vladimir Ivanov (organizing of clinical data collection)
MRCA-UKC: William OC Cookson and Miriam Moffatt (Study-PIs), Liming Liang (analysis of eQTL data)
MAGICS : Michael Kabesch (Study-PI), Andrea von Berg, Sven Michel
MAS: Young-Ae Lee (Study-PI), Susanne Lau, Ingo Marenholz

PIAMA: Gerard H Koppelman & Dirkje S Postma (Study-PIs, clinical data collection), Ulrike Gehring (clinical data collection)

SAPALDIA: Nicole Probst-Hensch (Study-PI, data collection), Medea Imboden (data collection, genotyping), Ashish Kumar (genotyping, imputation)

TOMSK: Ludmila M Ogorodova & Valery P Puzyrev (Study-PIs), Maxim B Freidin (analysis), Polina A Selivanova (data collection)

UFA: Elza Khusnutdinova (Study-PI), Alexandra S Karunas (clinical data and DNA collection), Yuliya Fedorova (blood collection and DNA extraction)

INDUSTRIAL: Dick Heederik and Torben Sigsgaard (Study-PIs), Inge M Wouters, Vivi Schlünssen

SEVERE: Andrew Bush, Kian Fan Chung, Christopher E. Brightling (Study-PIs)

The Netherlands Twin Register Study (NTR)

Study-PI: Dorret I Boomsma; Phenotyping: Gonneke Willemsen; Genotyping and Imputation: Jouke Jan Hottenga; Data analysis: Hamdi Mbarek

Rotterdam study

Study-PIs: Project conception, design and management: Albert Hofman, Bruno H Stricker, Guy G Brusselle
Genotyping: André G Uitterlinden; Phenotype collection and data management: Guy G Brusselle, Bruno H Stricker, Daan W Loth, Lies Lahousse; Data Analysis: Daan W Loth, Lies Lahousse.

Dutch Asthma Genetics Consortium (DAGC)

Study-PI, clinical data collection, genotyping: Gerard H Koppelman; Study-PI, data collection: Dirkje S Postma; Genotyping: Janine Altmüller; Statistical analysis: Maartje A.E. Nieuwenhuis, Judith M Vonk

Supplementary References

1. Ferreira, M.A. *et al.* Identification of IL6R and chromosome 11q13.5 as risk loci for asthma. *Lancet* **378**, 1006-14 (2011).
2. Howie, B.N., Donnelly, P. & Marchini, J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet* **5**, e1000529 (2009).
3. Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* **81**, 559-75 (2007).
4. Chan-Yeung, M. *et al.* The Canadian Childhood Asthma Primary Prevention Study: outcomes at 7 years of age. *J Allergy Clin Immunol* **116**, 49-55 (2005).
5. Kozyrskyj, A.L. *et al.* A novel study design to investigate the early-life origins of asthma in children (SAGE study). *Allergy* **64**, 1185-93 (2009).
6. Moffatt, M.F. *et al.* A large-scale, consortium-based genomewide association study of asthma. *N Engl J Med* **363**, 1211-21 (2010).
7. Laprise, C. The Saguenay-Lac-Saint-Jean asthma familial collection: the genetics of asthma in a young founder population. *Genes Immun* **15**, 247-55 (2014).
8. Splansky, G.L. *et al.* The Third Generation Cohort of the National Heart, Lung, and Blood Institute's Framingham Heart Study: design, recruitment, and initial examination. *Am J Epidemiol* **165**, 1328-35 (2007).
9. Canoy, D. *et al.* Early growth and adult respiratory function in men and women followed from the fetal period to adulthood. *Thorax* **62**, 396-402 (2007).
10. Day, N. *et al.* EPIC-Norfolk: study design and characteristics of the cohort. European Prospective Investigation of Cancer. *Br J Cancer* **80 Suppl 1**, 95-103 (1999).
11. Vartiainen, E. *et al.* Thirty-five-year trends in cardiovascular risk factors in Finland. *Int J Epidemiol* **39**, 504-18 (2010).
12. Aromaa, A. & Koskinen, S. (eds.). *Health and functional capacity in Finland. Baseline Results of the Health 2000 Health Examination Survey*, (National Public Health Institute KTL, Helsinki, 2004).
13. Barker, D.J., Osmond, C., Forsen, T.J., Kajantie, E. & Eriksson, J.G. Trajectories of growth among children who have coronary events as adults. *N Engl J Med* **353**, 1802-9 (2005).
14. Raitakari, O.T. *et al.* Cohort profile: the cardiovascular risk in Young Finns Study. *Int J Epidemiol* **37**, 1220-6 (2008).
15. Kantor, D.B. *et al.* Replication and fine mapping of asthma-associated loci in individuals of African ancestry. *Hum Genet* **132**, 1039-47 (2013).
16. Lettre, G. *et al.* Genome-wide association study of coronary heart disease and its risk factors in 8,090 African Americans: the NHLBI CARE Project. *PLoS Genet* **7**, e1001300 (2011).
17. Bild, D.E. *et al.* Multi-ethnic study of atherosclerosis: objectives and design. *Am J Epidemiol* **156**, 871-81 (2002).
18. Kaufman, J.D. *et al.* Prospective study of particulate air pollution exposures, subclinical atherosclerosis, and clinical cardiovascular disease: The Multi-Ethnic Study of Atherosclerosis and Air Pollution (MESA Air). *Am J Epidemiol* **176**, 825-37 (2012).
19. Manichaikul, A. *et al.* Analysis of family- and population-based samples in cohort genome-wide association studies. *Hum Genet* **131**, 275-87 (2012).
20. Soler Artigas, M. *et al.* Genome-wide association and large-scale follow up identifies 16 new loci influencing lung function. *Nat Genet* **43**, 1082-90 (2011).
21. THE ARIC INVESTIGATORS. The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. The ARIC investigators. *Am J Epidemiol* **129**, 687-702 (1989).

22. Guan, Y. & Stephens, M. Practical issues in imputation-based association mapping. *PLoS Genet* **4**, e1000279 (2008).
23. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. (2014).
24. Gudbjartsson, D.F. *et al.* Sequence variants affecting eosinophil numbers associate with asthma and myocardial infarction. *Nat Genet* **41**, 342-7 (2009).
25. Halapi, E. *et al.* A sequence variant on 17q21 is associated with age at onset and severity of asthma. *Eur J Hum Genet* **18**, 902-8 (2010).
26. 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* **39**, 906-13 (2007).
27. Bisgaard, H. The Copenhagen Prospective Study on Asthma in Childhood (COPSAC): design, rationale, and baseline data from a longitudinal birth cohort study. *Ann Allergy Asthma Immunol* **93**, 381-9 (2004).
28. Bisgaard, H., Hermansen, M.N., Loland, L., Halkjaer, L.B. & Buchvald, F. Intermittent inhaled corticosteroids in infants with episodic wheezing. *N Engl J Med* **354**, 1998-2005 (2006).
29. Bisgaard, H. *et al.* Childhood asthma after bacterial colonization of the airway in neonates. *N Engl J Med* **357**, 1487-95 (2007).
30. Li, Y., Willer, C.J., Ding, J., Scheet, P. & Abecasis, G.R. MaCH: using sequence and genotype data to estimate haplotypes and unobserved genotypes. *Genet Epidemiol* **34**, 816-34 (2010).
31. Olsen, J. *et al.* The Danish National Birth Cohort--its background, structure and aim. *Scand J Public Health* **29**, 300-7 (2001).
32. Ryckman, K.K. *et al.* Replication of a genome-wide association study of birth weight in preterm neonates. *J Pediatr* **160**, 19-24 e4 (2012).
33. Jaddoe, V.W. *et al.* The Generation R Study: design and cohort update 2012. *Eur J Epidemiol* **27**, 739-56 (2012).
34. Estrada, K. *et al.* GRIMP: a web- and grid-based tool for high-speed analysis of large-scale genome-wide association using imputed data. *Bioinformatics* **25**, 2750-2 (2009).
35. Heinrich, J. *et al.* Allergens and endotoxin on mothers' mattresses and total immunoglobulin E in cord blood of neonates. *Eur Respir J* **20**, 617-23 (2002).
36. Heinrich, J. *et al.* [Two German Birth Cohorts: GINIplus and LISAplus]. *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz* **55**, 864-74 (2012).
37. Berg, A. *et al.* Impact of early feeding on childhood eczema: development after nutritional intervention compared with the natural course - the GINIplus study up to the age of 6 years. *Clin Exp Allergy* **40**, 627-36 (2010).
38. Custovic, A., Simpson, B.M., Murray, C.S., Lowe, L. & Woodcock, A. The National Asthma Campaign Manchester Asthma and Allergy Study. *Pediatr Allergy Immunol* **13 Suppl 15**, 32-7 (2002).
39. Lowe, L.A. *et al.* Wheeze phenotypes and lung function in preschool children. *Am J Respir Crit Care Med* **171**, 231-7 (2005).
40. Murray, C.S. *et al.* Lung function at one month of age as a risk factor for infant respiratory symptoms in a high risk population. *Thorax* **57**, 388-92 (2002).
41. Nicolaou, N.C. *et al.* Exhaled breath condensate pH and childhood asthma: unselected birth cohort study. *Am J Respir Crit Care Med* **174**, 254-9 (2006).
42. Nicolaou, N.C. *et al.* Day-care attendance, position in sibship, and early childhood wheezing: a population-based birth cohort study. *J Allergy Clin Immunol* **122**, 500-6 e5 (2008).
43. Newnham, J.P., Evans, S.F., Michael, C.A., Stanley, F.J. & Landau, L.I. Effects of frequent ultrasound during pregnancy: a randomised controlled trial. *Lancet* **342**, 887-91 (1993).
44. Wellcome Trust Case Control, C. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* **447**, 661-78 (2007).

45. Anderson, H.R., Pottier, A.C. & Strachan, D.P. Asthma from birth to age 23: incidence and relation to prior and concurrent atopic disease. *Thorax* **47**, 537-42 (1992).
46. Strachan, D.P., Butland, B.K. & Anderson, H.R. Incidence and prognosis of asthma and wheezing illness from early childhood to age 33 in a national British cohort. *BMJ* **312**, 1195-9 (1996).
47. Butland, B.K. & Strachan, D.P. Asthma onset and relapse in adult life: the British 1958 birth cohort study. *Ann Allergy Asthma Immunol* **98**, 337-43 (2007).
48. Strachan, D.P. *et al.* Lifecourse influences on health among British adults: effects of region of residence in childhood and adulthood. *Int J Epidemiol* **36**, 522-31 (2007).
49. Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* **447**, 661-78 (2007).
50. Barrett, J.C. *et al.* Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. *Nat Genet* **41**, 703-7 (2009).
51. Aulchenko, Y.S., Struchalin, M.V. & van Duijn, C.M. ProbABEL package for genome-wide association analysis of imputed data. *BMC Bioinformatics* **11**, 134 (2010).
52. Moore, W.C. *et al.* Characterization of the severe asthma phenotype by the National Heart, Lung, and Blood Institute's Severe Asthma Research Program. *J Allergy Clin Immunol* **119**, 405-13 (2007).
53. Moore, W.C. *et al.* Identification of asthma phenotypes using cluster analysis in the Severe Asthma Research Program. *Am J Respir Crit Care Med* **181**, 315-23 (2010).
54. Sankararaman, S., Sridhar, S., Kimmel, G. & Halperin, E. Estimating local ancestry in admixed populations. *Am J Hum Genet* **82**, 290-303 (2008).
55. Price, A.L. *et al.* Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* **38**, 904-9 (2006).
56. Pritchard, J.K., Stephens, M. & Donnelly, P. Inference of population structure using multilocus genotype data. *Genetics* **155**, 945-59 (2000).
57. American Thoracic Society. Standards for the diagnosis and care of patients with chronic obstructive pulmonary disease (COPD) and asthma. This official statement of the American Thoracic Society was adopted by the ATS Board of Directors, November 1986. *Am Rev Respir Dis* **136**, 225-44 (1987).
58. The International Study of Asthma and Allergies in Childhood (ISAAC) Steering Committee. Worldwide variations in the prevalence of asthma symptoms: the International Study of Asthma and Allergies in Childhood (ISAAC). *Eur Respir J* **12**, 315-35 (1998).
59. Bonilla, C. *et al.* The 8818G allele of the agouti signaling protein (ASIP) gene is ancestral and is associated with darker skin color in African Americans. *Hum Genet* **116**, 402-6 (2005).
60. Mathias, R.A. *et al.* A genome-wide association study on African-ancestry populations for asthma. *J Allergy Clin Immunol* **125**, 336-346 e4 (2010).
61. Epstein, M.P., Duren, W.L. & Boehnke, M. Improved inference of relationship for pairs of individuals. *Am J Hum Genet* **67**, 1219-31 (2000).
62. Wu, B., Liu, N. & Zhao, H. PSMIX: an R package for population structure inference via maximum likelihood method. *BMC Bioinformatics* **7**, 317 (2006).
63. Hirota, T. *et al.* Genome-wide association study identifies three new susceptibility loci for adult asthma in the Japanese population. *Nat Genet* **43**, 893-6 (2011).
64. Noguchi, E. *et al.* Genome-wide association study identifies HLA-DP as a susceptibility gene for pediatric asthma in Asian populations. *PLoS Genet* **7**, e1002170 (2011).
65. Kauffmann, F. *et al.* Epidemiological study of the genetics and environment of asthma, bronchial hyperresponsiveness, and atopy: phenotype issues. *Am J Respir Crit Care Med* **156**, S123-9 (1997).
66. Kauffmann, F. *et al.* EGEA (Epidemiological study on the Genetics and Environment of Asthma, bronchial hyperresponsiveness and atopy)-- descriptive characteristics. *Clin Exp Allergy* **29 Suppl 4**, 17-21 (1999).

67. Boyd, A. *et al.* Cohort Profile: the 'children of the 90s'--the index offspring of the Avon Longitudinal Study of Parents and Children. *Int J Epidemiol* **42**, 111-27 (2013).
68. ECRHS. The European Community Respiratory Health Survey II. *Eur Respir J* **20**, 1071-9 (2002).
69. Burney, P.G., Luczynska, C., Chinn, S. & Jarvis, D. The European Community Respiratory Health Survey. *Eur Respir J* **7**, 954-60 (1994).
70. Wickman, M., Kull, I., Pershagen, G. & Nordvall, S.L. The BAMSE project: presentation of a prospective longitudinal birth cohort study. *Pediatr Allergy Immunol* **13 Suppl 15**, 11-3 (2002).
71. Melen, E. *et al.* Interactions between glutathione S-transferase P1, tumor necrosis factor, and traffic-related air pollution for development of childhood allergic disease. *Environ Health Perspect* **116**, 1077-84 (2008).
72. James, A.L. *et al.* Changes in the prevalence of asthma in adults since 1966: the Busselton health study. *Eur Respir J* **35**, 273-8 (2010).
73. Genuneit, J. *et al.* The GABRIEL Advanced Surveys: study design, participation and evaluation of bias. *Paediatr Perinat Epidemiol* **25**, 436-47 (2011).
74. Polonikov, A.V. *et al.* Promoter polymorphism G-50T of a human CYP2J2 epoxygenase gene is associated with common susceptibility to asthma. *Chest* **132**, 120-6 (2007).
75. Polonikov, A.V. *et al.* The relationship between polymorphisms in the glutamate cysteine ligase gene and asthma susceptibility. *Respir Med* **101**, 2422-4 (2007).
76. British guideline on the management of asthma. *Thorax* **58 Suppl 1**, i1-94 (2003).
77. Moffatt, M.F. *et al.* Genetic variants regulating ORMDL3 expression contribute to the risk of childhood asthma. *Nature* **448**, 470-3 (2007).
78. Weiland, S.K. *et al.* Prevalence of respiratory and atopic disorders among children in the East and West of Germany five years after unification. *Eur Respir J* **14**, 862-70 (1999).
79. Weiland, S.K. *et al.* Phase II of the International Study of Asthma and Allergies in Childhood (ISAAC II): rationale and methods. *Eur Respir J* **24**, 406-12 (2004).
80. Lau, S. *et al.* The development of childhood asthma: lessons from the German Multicentre Allergy Study (MAS). *Paediatr Respir Rev* **3**, 265-72 (2002).
81. Nickel, R. *et al.* How should a birth cohort study be organised? Experience from the German MAS cohort study. *Paediatr Respir Rev* **3**, 169-76 (2002).
82. Zuidgeest, M.G. *et al.* Persistence of asthma medication use in preschool children. *Respir Med* **102**, 1446-51 (2008).
83. Wijga, A.H. *et al.* Cohort profile: the prevention and incidence of asthma and mite allergy (PIAMA) birth cohort. *Int J Epidemiol* **43**, 527-35 (2014).
84. Brunekreef, B. *et al.* The prevention and incidence of asthma and mite allergy (PIAMA) birth cohort study: design and first results. *Pediatr Allergy Immunol* **13 Suppl 15**, 55-60 (2002).
85. Downs, S.H. *et al.* Reduced exposure to PM10 and attenuated age-related decline in lung function. *N Engl J Med* **357**, 2338-47 (2007).
86. Castro-Giner, F. *et al.* TNFA -308G>A in two international population-based cohorts and risk of asthma. *Eur Respir J* **32**, 350-61 (2008).
87. Imboden, M. *et al.* Decreased PM10 exposure attenuates age-related lung function decline: genetic variants in p53, p21, and CCND1 modify this effect. *Environ Health Perspect* **117**, 1420-7 (2009).
88. Freidin, M.B., Puzyrev, V.P., Ogorodova, L.M., Kobiakova, O.S. & Kulmanakova, I.M. [Polymorphism of interleukins and interleukin receptor genes: population distribution and association with atopic bronchial asthma]. *Genetika* **38**, 1710-8 (2002).
89. Freidin, M.B., Kobyakova, O.S., Ogorodova, L.M. & Puzyrev, V.P. Association of polymorphisms in the human IL4 and IL5 genes with atopic bronchial asthma and severity of the disease. *Comp Funct Genomics* **4**, 346-50 (2003).

90. Sigsgaard, T., Hjort, C., Omland, Ø., Miller, M.R. & Pedersen, O.F. Respiratory Health and Allergy Among Young Farmers and Non-Farming Rural Males in Denmark. *Journal of Agromedicine* **4**, 63-78 (1997).
91. Omland, O., Sigsgaard, T., Hjort, C., Pedersen, O.F. & Miller, M.R. Lung status in young Danish rurals: the effect of farming exposure on asthma-like symptoms and lung function. *Eur Respir J* **13**, 31-7 (1999).
92. Smit, L.A. *et al.* Exposure-response analysis of allergy and respiratory symptoms in endotoxin-exposed adults. *Eur Respir J* **31**, 1241-8 (2008).
93. Smit, L.A. *et al.* Ex vivo cytokine release reflects sensitivity to occupational endotoxin exposure. *Eur Respir J* **34**, 795-802 (2009).
94. Samadi, S. *et al.* Allergy among veterinary medicine students in The Netherlands. *Occup Environ Med* **69**, 48-55 (2012).
95. Jacobs, J.H., Meijster, T., Meijer, E., Suarhana, E. & Heederik, D. Wheat allergen exposure and the prevalence of work-related sensitization and allergy in bakery workers. *Allergy* **63**, 1597-604 (2008).
96. Schlunssen, V., Schaumburg, I., Taudorf, E., Mikkelsen, A.B. & Sigsgaard, T. Respiratory symptoms and lung function among Danish woodworkers. *J Occup Environ Med* **44**, 82-98 (2002).
97. Jacobsen, G., Schlunssen, V., Schaumburg, I., Taudorf, E. & Sigsgaard, T. Longitudinal lung function decline and wood dust exposure in the furniture industry. *Eur Respir J* **31**, 334-42 (2008).
98. Pronk, A. *et al.* Respiratory symptoms, sensitization, and exposure response relationships in spray painters exposed to isocyanates. *Am J Respir Crit Care Med* **176**, 1090-7 (2007).
99. Pronk, A. *et al.* Different respiratory phenotypes are associated with isocyanate exposure in spray painters. *Eur Respir J* **33**, 494-501 (2009).
100. Basinas, I. *et al.* Sensitisation to common allergens and respiratory symptoms in endotoxin exposed workers: a pooled analysis. *Occup Environ Med* **69**, 99-106 (2012).
101. Haldar, P. *et al.* Cluster analysis and clinical asthma phenotypes. *Am J Respir Crit Care Med* **178**, 218-24 (2008).
102. Haldar, P. *et al.* Mepolizumab and exacerbations of refractory eosinophilic asthma. *N Engl J Med* **360**, 973-84 (2009).
103. Proceedings of the ATS workshop on refractory asthma: current understanding, recommendations, and unanswered questions. American Thoracic Society. *Am J Respir Crit Care Med* **162**, 2341-51 (2000).
104. Chung, K.F. *et al.* Difficult/therapy-resistant asthma: the need for an integrated approach to define clinical phenotypes, evaluate risk factors, understand pathophysiology and find novel therapies. ERS Task Force on Difficult/Therapy-Resistant Asthma. European Respiratory Society. *Eur Respir J* **13**, 1198-208 (1999).
105. Ayres, J.G., Miles, J.F. & Barnes, P.J. Brittle asthma. *Thorax* **53**, 315-21 (1998).
106. Okaneya, T., Mizusawa, H., Taguchi, I. & Yoneyama, T. [Clinical evaluation of tandem PSA for the diagnosis and follow-up of prostate cancer]. *Hinyokika Kiyo* **42**, 861-7 (1996).
107. Stanojevic, S. *et al.* Reference ranges for spirometry across all ages: a new approach. *Am J Respir Crit Care Med* **177**, 253-60 (2008).
108. Boomsma, D.I. *et al.* Genome-wide association of major depression: description of samples for the GAIN Major Depressive Disorder Study: NTR and NESDA biobank projects. *Eur J Hum Genet* **16**, 335-42 (2008).
109. Willemsen, G. *et al.* The Netherlands Twin Register biobank: a resource for genetic epidemiological studies. *Twin Res Hum Genet* **13**, 231-45 (2010).
110. Hofman, A. *et al.* The Rotterdam Study: 2012 objectives and design update. *Eur J Epidemiol* **26**, 657-86 (2011).

111. Nieuwenhuis, M.A. *et al.* Combining genomewide association study and lung eQTL analysis provides evidence for novel genes associated with asthma. *Allergy* **71**, 1712-1720 (2016).