ORIGINAL ARTICLE

Genome-wide association study to identify genetic determinants of severe asthma

Y I Wan,[1,*] N R G Shrine,[2,*] M Soler Artigas,[2] L V Wain,[2] J D Blakey,[1] M F Moffatt,[3] A Bush,[3] K F Chung,[3] W O C M Cookson,[3] D P Strachan,[4] L Heaney,[5] B A H Al-Momani,[6] A H Mansur,[7] S Manney,[7] N C Thomson,[8] R Chaudhuri,[8] C E Brightling,[9] M Bafadhel,[9] A Singapuri,[9] R Niven,[10] A Simpson,[10] J W Holloway,[11,12] P H Howarth,[12,13] J Hui,[14] A W Musk,[14] A L James,[14] the Australian Asthma Genetics Consortium[15] M A Brown,[16] S Baltic,[17] M A R Ferreira,[18] P J Thompson,[17] M D Tobin,[2] I Sayers,[1] I P Hall[1]

► Additional materials are published online only. To view these files please visit the journal online (http://dx.doi.org/10.1136/thoraxjnl-2011-201262).

For numbered affiliations see end of article.

Correspondence to
Yize I Wan, Division of Therapeutics and Molecular Medicine, University Hospital of Nottingham, Nottingham NG7 2UH, UK;
yize.wan@nottingham.ac.uk

YIW and NRGS are joint first authors.
*A full list of collaborators is available in the online appendix.

Received 19 October 2011
Accepted 1 March 2012
Published Online First
5 May 2012

ABSTRACT

Background The genetic basis for developing asthma has been extensively studied. However, association studies to date have mostly focused on mild to moderate disease and genetic risk factors for severe asthma remain unclear.

Objective To identify common genetic variants affecting susceptibility to severe asthma.

Methods A genome-wide association study was undertaken in 933 European ancestry individuals with severe asthma based on Global Initiative for Asthma (GINA) criteria 3 or above and 3346 clean controls. After standard quality control measures, the association of 480 889 genotyped single nucleotide polymorphisms (SNPs) was tested. To improve the resolution of the association signals identified, non-genotyped SNPs were imputed in these regions using a dense reference panel of SNP genotypes from the 1000 Genomes Project. Then replication of SNPs of interest was undertaken in a further 231 cases and 1345 controls and a meta-analysis was performed to combine the results across studies.

Results An association was confirmed in subjects with severe asthma of loci previously identified for association with mild to moderate asthma. The strongest evidence was seen for the *ORMDL3/GSDMB* locus on chromosome 17q12-21 (rs4794820, $p=1.03\times10^{(-8)}$ following meta-analysis) meeting genome-wide significance. Strong evidence was also found for the *IL1RL1/IL18R1* locus on 2q12 (rs9807989, $p=5.59\times10^{(-8)}$ following meta-analysis) just below this threshold. No novel loci for susceptibility to severe asthma met strict criteria for genome-wide significance.

Conclusions The largest genome-wide association study of severe asthma to date was carried out and strong evidence found for the association of two previously identified asthma susceptibility loci in patients with severe disease. A number of novel regions with suggestive evidence were also identified warranting further study.

Key messages

What is the key question?

► The aim of this study was to identify genetic determinants of severe asthma and to evaluate whether susceptibility to severe asthma differs from that of mild to moderate asthma.

What is the bottom line?

► The first genome-wide association study of severe asthma was undertaken which identified the contribution of some but not all genetic loci previously associated with mild to moderate disease. Suggestive evidence for a number of novel loci associated with severe disease is also reported.

Why read on?

► Novel loci, which may be specific to severe asthma, potentially provide further insight into disease mechanisms and warrant further study.

INTRODUCTION

Asthma is a chronic inflammatory condition of the airways characterised by recurrent episodes of reversible airway obstruction and increased bronchial hyper-responsiveness.[1] Approximately 10% of patients with asthma are prone to severe exacerbations and remain symptomatic despite treatment with high-dose inhaled corticosteroids (ICS) and long-acting β2-adrenergic receptor agonists.[2] This subgroup of patients disproportionately consume healthcare resources related to asthma and contribute the largest proportion of morbidity and mortality.[3]

The genetic basis for developing asthma has been extensively investigated; numerous candidate genes have been studied for association with asthma due to the potential biological effects on airway function of the relevant gene products, although replication has been inconsistent.[4–10] In addition, recent genome-wide association studies (GWAS) in asthma have identified a widely replicated locus on chromosome 17q12-21 containing genes *ORMDL3*,

CCL11 and *GSDML*, and additional genes including *CHI3L1*, *IL1RL1* and *WDR36* on chromosomes 1q31, 2q12, and 5q22 respectively.[11–13]

Recently, the biggest collaborative effort thus far investigating the genetic determinants of asthma was published by the GABRIEL consortium. This study consisted of 10 365 case subjects and 16 110 controls and observed genome-wide significance between asthma and single nucleotide polymorphisms (SNPs) within previously reported loci and genes, including genes *IL18R1*, *HLA-DQ*, *IL33*, and chromosome 17q12-21, the latter specific to childhood-onset disease.[14] Subsequently, a large Australian collaborative effort carried out a GWAS in 2669 cases and 4528 controls. A selected number of identified loci were then followed up in a replication analysis following a meta-analysis with the results of the GABRIEL study. This paper provided independent support for loci reported by GABRIEL: *IL18R1*, *IL33*, *ORMDL3* and *IL2RB*, and identified a further two loci reaching genome-wide significance in the combined analysis of all studies (n=57 800): *IL6R* and chromosome 11q13.5.[15]

Furthermore, the EVE consortium conducted a meta-analysis of North-American GWAS (n=5416 in meta-analysis, n=12 649 in replication) inclusive of individuals of European American, African American or African Caribbean, and Latino ancestry. This study reported that previously identified loci on 17q21, near *IL1RL1*, *TSLP* and *IL33*, were robust to ethnic differences showing significant association in all three ethnic groups.[16]

In addition, a single small GWAS (473 cases, 1892 controls) was conducted in 2009 on a population of patients with severe or difficult to treat asthma from The Epidemiology and Natural History of Asthma: Outcomes and Treatment Regimes (TENOR) study and identified association with multiple SNPs in the *RAD50-IL13* and *HLA-DR/DQ* regions, although no loci reached conventional GWAS significance criteria.[17]

These studies have generally involved subjects with mild asthma. The aims of the current study were first to identify genetic determinants of severe asthma, and second to evaluate whether susceptibility to severe asthma differs from that of mild to moderate asthma.

METHODS

Participants

Discovery cohort

We genotyped 1026 individuals of European ancestry with severe asthma based on the Global Initiative for Asthma (GINA) criteria.[18] Only subjects in classes 3–5 were included, recruited across UK-based centres. Subjects were selected from individuals participating in the Difficult/Severe Asthma (BTS) study (n=290), supplemented with subjects from other centres. A total of 3353 control subjects without history of asthma or wheeze (clean controls) were collected from the UK and Western Australia, all of whom were of European ancestry.

Replication cohort

A replication cohort of 231 cases with more severe asthma based on clinical examination by a respiratory physician and treatment steps (ie, receiving ICS ≥400 μg in combination with a long- and short-acting β2-adrenergic receptor agonists and short-acting beta-agonist) and 1345 controls without asthma were recruited by the Australian Asthma Genetics Consortium (AAGC) study. All individuals from Australia were of European ancestry.

Baseline characteristics and participant recruitment of all study populations per centre are described in the online repository.

Genotyping and procedures

Participants were genotyped using the Illumina HumanHap550K, 610K, 660K, and 1.2M SNP chip platforms (Illumina, San Diego, California, USA). To minimise bias due to the use of different genotyping platforms in case and control cohorts, only the 490 303 SNPs in common across all six platforms were used. Furthermore, the use of principal components analysis (PCA) covariates provides some correction for assay effects.[19] We also excluded any SNPs that showed a significant difference in allele frequencies between the three control groups ($p<10^{(-6)}$). However, it is not possible to completely eliminate bias due to genotyping platform and centre and hence we sought replication. A table describing the number of samples genotyped on each platform has been included (online table E1). Within each study, individuals with <90% of SNPs called were excluded and SNPs were excluded if they had low call rates (proportion of genotypes called <90%), were not in Hardy-Weinberg equilibrium (HWE, $p<10^{(-4)}$), had a low minor allele frequency (MAF<1%) or with differential missingness between cases and controls ($p<10^{(-4)}$). PCA was carried out to detect outlying samples and to correct for residual population structure using EIGENSOFT V.3.0 (online figure E1).[19]

Statistical analyses

Association tests of genotyped SNPs were carried out using PLINK V.1.07 with an additive genetic model with the first 10 principal components as covariates.[20] We tested association with 480 889 SNPs present across all cohorts; 21 SNPs showing a significant difference in allele frequencies ($p<10^{(-6)}$) between the three control groups were removed. We identified regions of interest as those with a sentinel SNP showing association with asthma (at a threshold of $p<5\times10^{(-5)}$) with at least one additional SNP within 500 Kb also reaching a threshold of $p<5\times10^{(-5)}$.

1000 Genomes imputation

Imputation was used to improve the resolution of regions identified for association from genotyped data. Genotyped SNPs were used for imputation to 6.9 million SNPs using the June 2010 release of the 1000 Genomes CEU reference panel comprising 120 individuals genotyped at 6 858 242 SNPs.[21] Haplotypes were phased by comparing genotypes across our 4279 cases and controls with all alleles defined on the positive strand using MaCH.[22] Imputation of genotypes was carried out by comparing haplotype blocks in our phased samples with those in the 1000 Genomes reference panel using minimac.[23] A measure of confidence in the imputation is given by the metric r^2_{imp} which is an estimate of the correlation between imputed and true genotypes ranging from 0 to 1 (1 for genotyped SNPs). Quality control was carried out to exclude SNPs with MAF<1% or imputation quality r^2_{imp} <0.3 (recommended r^2_{imp} filter to exclude 70% of poorly imputed SNPs).[22] Association tests were performed with ProbABEL using a logistic model with the dose of the effect allele (on a continuous scale between 0 and 2 reflecting imputation uncertainties) as the independent variable and 10 ancestry principal components derived from our genotyped SNPs as covariates.[24] Post-association filters were applied to remove SNPs showing significant association in control–control comparisons ($p<10^{-6}$) leaving 6 103 628 SNPs.

Replication and meta-analysis

Subjects from the AAGC study were used to test replication of 24 SNPs identified in the discovery GWAS and subsequent imputation analyses. This SNP list consisted of six genotyped SNPs from the regions identified in genotype analyses, four

imputed SNPs in the same regions with a lower p value than the original genotyped SNPs, two SNPs with $p>10^{(-5)}$ responsible for secondary peaks in regions with known asthma genes, and 12 SNPs from new regions identified through imputation. In order to select the best candidate imputed SNPs for replication we used a lower p value of $<10^{(-5)}$ and a stricter imputation quality ($r^2_{imp}\geq0.7$) than used for fine-mapping around genotyped SNPs. Statistical significance for replication was assessed using a 5% significance threshold and results of inverse-variance weighted meta-analysis assessed using conventional criteria for genome-wide significance ($p=5\times10^{(-8)}$).[25]

Evaluation of GABRIEL loci

We also investigated the contribution of polymorphisms identified for mild to moderate asthma by the GABRIEL consortium in our severe asthma population. We examined regions within 500 Kb of SNPs reported to be associated with asthma in the GABRIEL study, including both regions reported to have reached genome-wide significance ($p\leq7.2\times10^{(-8)}$) and those providing suggestive evidence of association ($p\leq5\times10^{(-7)}$) in GABRIEL.

Comparison of severe versus mild to moderate asthma

A total of 1028 individuals of European ancestry with a history of doctor diagnosed asthma at GINA steps 1 or 2 were collected from the WTCCC2, T1DGC and Busselton populations. A GWAS was then carried out comparing the 1026 patients with severe asthma against the 1028 patients with mild to moderate asthma. All genotyping and quality control procedures were conducted as above for the severe asthma versus clean control analyses. We tested association with 488 889 SNPs present across all cohorts. Imputation was used to improve resolution of identified regions and replication was assessed in the AAGC study using a comparison of the 231 severe cases versus 1085 patients with mild to moderate asthma identified as never having received steroid medication in their lifetime.

RESULTS

Genotype data for 933 cases and 3346 controls were available for the primary discovery analysis after quality control. Replication of identified SNPs was assessed in 231 cases and 1345 controls. Characteristics of the study cohorts are summarised in the online appendix. The test statistic inflation factor λ for the discovery GWAS was modest (λ=1.04). Results shown by the quantile–quantile plot suggest the presence of multiple loci with modest effects (figure 1A).

In the initial analysis of genotyped data in the discovery cohort, no SNPs met genome-wide significance for association with severe asthma using a conservative cut-off defined by the Bonferroni correction: $p=1.04\times10^{(-7)}$ (figure 1B). We therefore went on to evaluate other potential loci with statistical significance below this threshold. A total of eight SNPs were identified with $p<5\times10^{(-5)}$, with at least one other SNP within 500 kb with $p<5\times10^{(-5)}$. Assessment for supporting evidence within the region for these SNPs suggests that six of these loci may contain susceptibility genes for severe asthma: rs3771166 within *IL18RL1* on 2q12.1 ($p=1.93\times10^{(-5)}$), rs11745587 in the 3′ untranslated region (UTR) of *C5orf56* on 5q31.1 ($p=2.09\times10^{(-6)}$), rs9382936 on 6p23 ($p=5.61\times10^{(-6)}$) tagging *CD83*, rs12699949 on 7p21.1 ($p=1.19\times10^{(-5)}$) tagging *PRPS1L1*, rs2496764 within an intergenic region on 13q31.1 ($p=7.86\times10^{(-6)}$), and rs1810132 within *ERBB2* on 17q12-21 ($p=1.73\times10^{(-5)}$) (table 1, figure 2). We did not follow up those loci characterised by a single SNP within 500 kb showing an association ($p<5\times10^{(-5)}$); these loci are listed in online table E4.

Imputation was used to refine association signals within all six loci. This analysis identified a further four SNPs with a lower p value than the original genotyped SNP in five regions: rs9807989 on 2q12.1 ($p=5.20\times10^{(-6)}$) tagging *IL18RL1*, rs12699948 on 7p21.1 ($p=4.84\times10^{(-6)}$) tagging *PRPS1L1*, rs9547037 intergenic on 13q31.1 ($p=6.60\times10^{(-6)}$), and rs9972882 within *STARD3* on 17q12-21 ($p=5.17\times10^{(-6)}$). Two imputed SNPs produced secondary peaks in regions with known asthma-associated genes: rs13035227 on 2q12.1 ($p=8.91\times10^{(-5)}$) tagging *IL1RL1*, and rs847 in the 3′ UTR of *IL13* on 5q31 ($p=4.05\times10^{(-5)}$). An additional 12 SNPs with $p<10^{(-5)}$ and $r^2_{imp}=0.7$ produced signals in new regions identified through imputation ($1.19\times10^{(-5)}\leq p\leq2.82\times10^{(-7)}$) (table 1, online figure E2).

Two loci were replicated in the AAGC study cohort. The first of these was on 17q12-21 by rs4794820 tagging the *ORMDL3* locus ($p=0.002$). The second was on 2q12.1 by a cluster of three SNPs: rs3771166 within *IL18R1* ($p=0.001$), rs9807989 tagging *IL18R1* ($p=0.003$), and rs13035227 tagging *IL1R1* ($p=0.002$). In non-replicated regions, a consistent direction of effect for the minor allele was seen across studies for 18 out of the 22 remaining SNPs. Following meta-analysis, the signal on *ORMDL3* met conventional genome-wide significance ($p=1.03\times10^{(-8)}$) and the signal on *IL18RL1* approached this threshold ($p=5.59\times10^{(-8)}$) (table 1, figure 2).

Next, we proceeded to test all SNPs reported in the GABRIEL study for both genome-wide significance and suggestive evidence for association with mild to moderate asthma to assess the degree of association with severe asthma in Asthma UK Genetics of Severe Asthma (AUGOSA) (online table E2 and E3 and figure E2). In general, as might be expected, we also found an association with these loci apart from rs2284033 on chromosome 22q12 ($p=0.105$) and rs11071559 on chromosome 15q22 ($p=0.159$).

A comparison of patients with severe asthma versus those with mild to moderate asthma was carried out (online figure E3). The test statistic inflation factor λ for this GWAS was again modest (λ=1.04). A single SNP met genome-wide significance in this analysis: rs981516 intergenic on 4p32.1 ($p=3.34\times10^{(-8)}$, OR 1.50 95% CI 1.30 to 1.73). However, this was not replicated in the AAGC study ($p=0.451$) although the same direction of effect was seen for the minor allele.

DISCUSSION

We conducted the largest severe asthma GWAS to date in a cohort of 933 cases defined by GINA steps three or above for severity and 3346 clean controls to determine if there are common genetic polymorphisms contributing to susceptibility to severe asthma.

Overall, we did not identify any novel SNPs meeting genome-wide significance. We carried out further analysis of results for polymorphisms just below this threshold to look for regions which did not meet standard genome-wide significance but had supporting evidence with at least one additional SNP with $p<5\times10^{(-5)}$ within 500 Kb. Using this criterion, we identified six loci with suggestive evidence for association. Two of these loci, chromosomes 2q12 ($p=5.20\times10^{(-6)}$) and 17q12-21 ($p=5.17\times10^{(-5)}$) implicating the *IL1RL1/IL18R1* and *ORMDL3/GSDMB* loci respectively have been previously reported by GWAS for association with mild to moderate asthma.[11 13–16] Both of these loci replicated in a second cohort of 231 severe asthma cases and 1345 controls (2q12, $p=0.001$; 17q12-21, $p=0.002$). Evidence for the 17q12-21 locus became genome-wide significant following meta-analysis and was just below this threshold for the 2q12 locus, highlighting a potentially important role for these loci in asthma irrespective of severity.

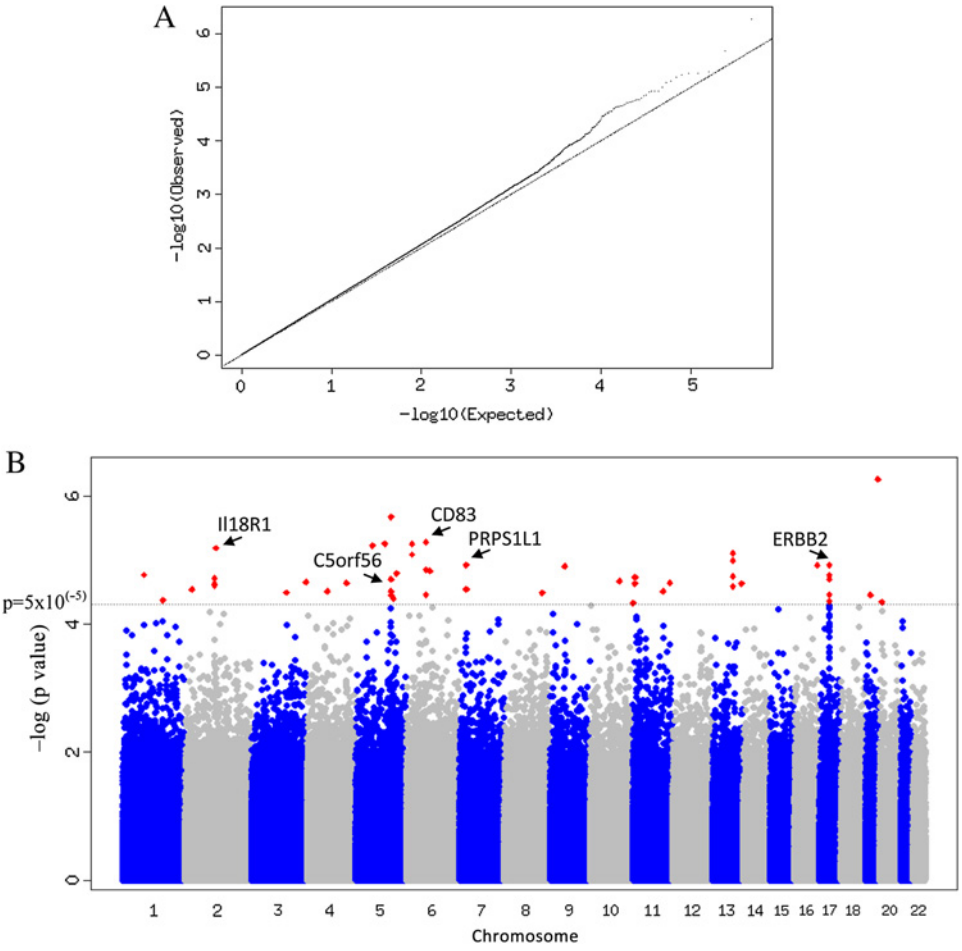


Figure 1 Quantile–quantile plot and Manhattan plot of study results. (A) Quantile–quantile plot of observed versus expected log10 p values for all tested genotyped single nucleotide polymorphisms (SNPs). (B) Manhattan plot of − log10 p values for all tested SNPs against genomic position. *Note: the signal on chromosome 13q31.1 was intergenic.

A previous GWAS in which the main phenotype was blood eosinophil counts identified evidence for an association with asthma for SNPs in *IL1RL1* and its ligand, *IL33*.[13] Subsequently, asthma association of both loci have been shown to be robust to differences in ancestry.[16] We report association with the same SNP, rs3771166 within *IL18R1*on chromosome 2q12. Rs3771166 was also reported by the GABRIEL consortium showing the same direction of effect for association with the minor allele. A stronger protective effect size is observed for this polymorphism within our severe asthma cohort (OR=0.79) compared with that shown in the GABRIEL study (OR=0.87).[14] With current data, we are unable to determine if association is driven by the *IL1RL1* or the *IL18R1* gene. However, both genes are plausible biological candidates in the inflammatory cascade in the pathway to asthma pathogenesis.[26 27]

Associations with asthma and SNPs located on chromosome 17q12-21 have been reported and replicated across multiple study populations.[11 14 28 29] Despite this, the region on 17q12-21 harbours a number of genes of poorly understood function. In combination with the complex linkage disequilibrium structure in this locus, it is difficult to be sure which genetic variants are causal. Furthermore, 17q12-21 has previously been suggested to be exclusive to childhood-onset disease.[11 14] Although, we have not assessed age of onset in the current study, we found significant association with this locus despite studying predominantly adult subjects (mean age of disease onset: 21 years), thus challenging this hypothesis.

We report an additional four loci as showing suggestive evidence of association with severe asthma. The most significant result on chromosome 6p23 implicates *CD83*, the gene encoding for the CD83 antigen expressed on dendritic cells and which may play a role in immune modulation in the airways.[30] The SNP on chromosome 5q31 lies within the 3′ UTR of *C5orf56*, however it is downstream of the T_h2 cytokine genes *IL4* and *IL13*, and *RAD50*, a region previously reported for suggestive association with severe asthma by the TENOR study.[17] The remaining two regions downstream of *PRPS1L1* on 7p21 and intergenic on 13q31 are potentially novel with unknown function in asthma and warrant further study.

We list in online table E4 those loci not followed up in the AAGC study due to not meeting our strict predefined threshold for supporting SNPs in the region (online table E4). Interestingly, these loci include *TSLP*, which was identified as a potential asthma locus previously.[16] We then undertook a MAGENTA[31] pathway analysis on the full GWAS dataset to look for enrichment of association in known biological pathways from six databases (Gene Ontology, Ingenuity Pathway, KEGG, PANTHER Pathways, PANTHER Molecular Function and PANTHER Biological Processes). No pathway was significant after correction for multiple testing, although the Fc Epsilon RI Signalling pathway reached nominal significance (p=0.0014) and includes *IL13*.

We then went on to assess the contribution of previously identified asthma susceptibility loci in patients with severe disease reported by the recent GABRIEL study.[14] As the GABRIEL study is currently the largest published association study investigating the genetic determinants of asthma, we aimed to determine if these signals also contribute to disease susceptibility in our severe asthma cohort. Individuals in AUGOSA were included in the total case subjects in the GABRIEL study. However, as they only constitute a relatively

Table 1 Single nucleotide polymorphisms (SNPs) showing highest association signals for severe asthma

				AUGOSA (933 cases, 3346 controls)			AAGC (231 cases, 1345 controls)			Meta-analysis	
Chromosome	**Locus**	**SNP**	**Position**	**r^2_{imp}**	**OR (95% CI)**	**p Value**	**r^2_{imp}**	**OR (95% CI)**	**p Value**	**OR (95% CI)**	**p Value**
Six genotyped SNPs from regions identified in genotyped SNP analyses											
2	IL18R1	rs3771166	102352654	GENO	0.79 (0.71 to 0.88)	**$1.93\times10^{(-5)}$**	GENO	0.71 (0.57 to 0.87)	**0.001**	0.77 (0.70 to 0.85)	$1.24\times10^{(-7)}$
5	C5orf56	rs11745587	131824821	GENO	1.30 (1.17 to 1.45)	**$2.09\times10^{(-6)}$**	GENO	1.13 (0.92 to 1.39)	0.25	1.26 (1.15 to 1.39)	$2.13\times10^{(-6)}$
6	CD83	rs9382936	14173097	GENO	1.31 (1.17 to 1.48)	**$5.61\times10^{(-6)}$**	GENO	1.00 (0.79 to 1.25)	0.98	1.24 (1.12 to 1.38)	$5.68\times10^{(-5)}$
7	PRPS1L1	rs12699949	18010787	GENO	0.77 (0.69 to 0.87)	**$1.19\times10^{(-5)}$**	GENO	0.89 (0.72 to 1.10)	0.27	0.80 (0.72 to 0.88)	$1.19\times10^{(-5)}$
13	*Intergenic*	rs2496764	84477159	GENO	1.34 (1.18 to 1.52)	**$7.86\times10^{(-6)}$**	GENO	0.99 (0.78 to 1.27)	0.96	1.26 (1.12 to 1.41)	$8.03\times10^{(-5)}$
17	ERBB2	rs1810132	35119531	GENO	1.28 (1.14 to 1.43)	**$1.73\times10^{(-5)}$**	GENO	1.07 (0.86 to 1.32)	0.56	1.23 (1.11 to 1.36)	$4.54\times10^{(-5)}$
Four imputed SNPs in regions above with a lower p value than the original genotyped SNP in imputed SNP analyses											
2	IL18R1	rs9807989	102337632	0.91	0.76 (0.67 to 0.85)	**$5.20\times10^{(-6)}$**	0.91	0.72 (0.58 to 0.89)	**0.003**	0.75 (0.68 to 0.83)	$5.59\times10^{(-8)}$
7	PRPS1L1	rs12699948	18010735	0.92	0.75 (0.66 to 0.85)	**$4.84\times10^{(-6)}$**	0.93	0.92 (0.74 to 1.15)	0.46	0.79 (0.71 to 0.88)	$1.44\times10^{(-5)}$
13	*Intergenic*	rs9547037	84476839	0.90	1.38 (1.20 to 1.58)	**$6.60\times10^{(-6)}$**	0.90	1.00 (0.78 to 1.29)	0.97	1.28 (1.13 to 1.44)	$8.01\times10^{(-5)}$
17	STARD3	rs9972882	35061224	0.98	1.32 (1.17 to 1.49)	**$5.17\times10^{(-6)}$**	0.99	1.07 (0.85 to 1.34)	0.55	1.26 (1.13 to 1.40)	$1.64\times10^{(-5)}$
Two SNPs with $p>10^{(-5)}$ responsible for secondary peaks in regions with known asthma genes											
2	IL1R1	rs13035227	102130269	0.96	1.36 (1.16 to 1.58)	**$8.91\times10^{(-5)}$**	GENO	1.53 (1.17 to 1.99)	**0.002**	1.40 (1.22 to 1.59)	$6.69\times10^{(-7)}$
5	IL13	rs847	132024568	0.86	1.35 (1.17 to 1.55)	**$4.05\times10^{(-5)}$**	0.94	1.12 (0.88 to 1.43)	0.37	1.29 (1.14 to 1.45)	$6.43\times10^{(-5)}$
Twelve SNPs from new regions identified by imputation with $p<10^{(-5)}$ and $r^2_{imp}=0.7$ (includes two secondary peaks in previously identified 17q12-21 region)											
2	*Intergenic*	chr2:211694960	211694960	0.70	1.76 (1.40 to 2.21)	**$1.27\times10^{(-6)}$**	0.78	1.11 (0.73 to 1.70)	0.62	1.59 (1.30 to 1.94)	$6.80\times10^{(-6)}$
5	FLJ37543	rs7715669	60972053	GENO	0.74 (0.65 to 0.84)	**$5.92\times10^{(-6)}$**	GENO	0.93 (0.74 to 1.18)	0.56	0.78 (0.69 to 0.87)	$2.29\times10^{(-5)}$
5	NDFIP1	rs6867913	141426164	0.99	0.72 (0.63 to 0.82)	**$1.74\times10^{(-6)}$**	0.97	0.90 (0.69 to 1.15)	0.39	0.75 (0.67 to 0.85)	$3.82\times10^{(-6)}$
6	GCLC	rs9395865	53415653	GENO	0.76 (0.67 to 0.85)	**$5.27\times10^{(-6)}$**	GENO	1.15 (0.93 to 1.42)	0.19	0.84 (0.76 to 0.93)	$9.27\times10^{(-4)}$
6	*Intergenic*	rs6922932	73384242	0.78	0.70 (0.60 to 0.82)	**$7.36\times10^{(-6)}$**	0.83	1.02 (0.79 to 1.31)	0.88	0.78 (0.68 to 0.89)	$1.99\times10^{(-4)}$
8	DUSP4	rs650230	29322727	0.78	1.37 (1.19 to 1.57)	**$6.40\times10^{(-6)}$**	0.76	1.04 (0.83 to 1.30)	0.73	1.27 (1.13 to 1.42)	$5.74\times10^{(-5)}$
9	ACO1	rs10970976	32423526	0.98	0.76 (0.67 to 0.86)	**$9.14\times10^{(-6)}$**	0.96	0.84 (0.67 to 1.05)	0.13	0.78 (0.70 to 0.86)	$3.69\times10^{(-6)}$
11	OR52E4	rs4453217	5883859	0.87	1.43 (1.25 to 1.64)	**$2.82\times10^{(-7)}$**	0.88	0.94 (0.72 to 1.22)	0.63	1.31 (1.16 to 1.48)	$1.50\times10^{(-5)}$
11	ETS1	rs7125574	127866404	0.93	1.34 (1.18 to 1.51)	**$5.28\times10^{(-6)}$**	0.93	0.92 (0.73 to 1.16)	0.49	1.23 (1.10 to 1.37)	$2.13\times10^{(-4)}$
17	STAC2	rs9897185	34642935	GENO	1.31 (1.16 to 1.47)	**$1.19\times10^{(-5)}$**	GENO	0.95 (0.75 to 1.19)	0.64	1.22 (1.10 to 1.35)	$2.50\times10^{(-4)}$
17	ORMDL3	rs4794820	35342870	0.94	0.76 (0.68 to 0.85)	**$1.52\times10^{(-6)}$**	0.93	0.72 (0.59 to 0.89)	**0.002**	0.75 (0.69 to 0.83)	**$1.03\times10^{(-8)}$**
19	ZNF665	rs16984547	58373854	GENO	1.53 (1.30 to 1.82)	**$5.45\times10^{(-7)}$**	GENO	1.06 (0.76 to 1.49)	0.74	1.43 (1.23 to 1.66)	$3.54\times10^{(-6)}$

Effect allele refers to the minor allele. GENO denotes genotyped SNP. Significant p values $p\leq0.05$ in replication, $p<5\times10^{(-8)}$ in meta-analysis are shown in bold. The gene tagged by the SNP due to linkage disequilibrium (LD) block structure within the locus is listed.
AAGC, Australian Asthma Genetics Consortium; AUGOSA, Asthma UK Genetics of Severe Asthma.

small proportion (8.9%), an association signal within GABRIEL for mild to moderate asthma would be unlikely to originate only from AUGOSA individuals.

For genome-wide significant results compared with mild to moderate asthma from GABRIEL, effect sizes for the risk allele are greater in our cohort of patients with severe asthma for rs3771166 on chromosome 2q21: the observed OR for the protective allele was 0.79 compared with 0.87 in GABRIEL for rs2305480 on chromosome 17q12-21; the observed OR for the protective allele was 0.80 compared with 0.85 for rs3894194 on chromosome 17q12-21; the OR for the risk allele was 1.25 compared with 1.17. Further to this, we show supporting evidence for loci reported by previous studies such as the GABRIEL study, including *IL13* on 5q31 ($p=9.30\times10^{(-5)}$) and *SMAD3* on 15q22.33 ($p=2.88\times10^{(-4)}$) containing weaker association signals which fell just below the threshold for significance chosen for our primary analyses.

The explanations behind the increased effect size of risk alleles in severe asthma are potentially twofold. The most likely explanation is that the contribution of genetic effects driven by variation in these genes is greater in patients with more severe asthma. However, it is possible that in populations with milder asthma misclassification of cases and controls may result in an underestimate of true effect sizes. Furthermore, differences in population sizes used in analyses may also reflect our varying ability to determine robust effect sizes.

A comparison of severe versus mild to moderate asthma was carried out. This identified a potentially novel locus on 4p32.1 meeting genome-wide significance which may be specific to the development of severe as opposed to milder forms of asthma. The sentinel SNP rs981516 is in an intergenic region but in linkage disequilibrium with rs17291045 ($r^2=0.423$), a SNP previously reported to be strongly associated with progression in HIV-1 infection and may have functional effects on viral control.[32] The identified SNP was not replicated in the AAGC study, however given the modest effect size seen in the discovery GWAS (OR=1.50), a much larger follow-up cohort would have been necessary to reliably assess replication.

Although the current study is the largest effort so far to determine genetic determinants of severe asthma, we are still limited by the numbers of subjects in being able to generate enough statistical power to detect all variants with modest effects. While we can probably exclude major effects being driven by a single gene as a specific risk for severe asthma, our data suggest there may be a number of loci which may be specific for severe asthma but with relatively small overall contributions to the risk of developing severe disease. The obvious solution to resolving this issue is to undertake further replication studies in much larger severe asthma populations. However, these populations by their very definition are hard to recruit: the current study included subjects recruited from eight major centres in the UK and replication in a second study consisting of subjects recruited from two major centres in Australia. Hence, obtaining suitable replication populations to take this work forward will require additional international efforts to establish appropriately large populations with severe disease.

In summary, we provide evidence to support an enhanced role for known genetic risk factors for asthma in the development of severe disease, and also have identified novel loci which may be specific to the development of severe as opposed to milder forms

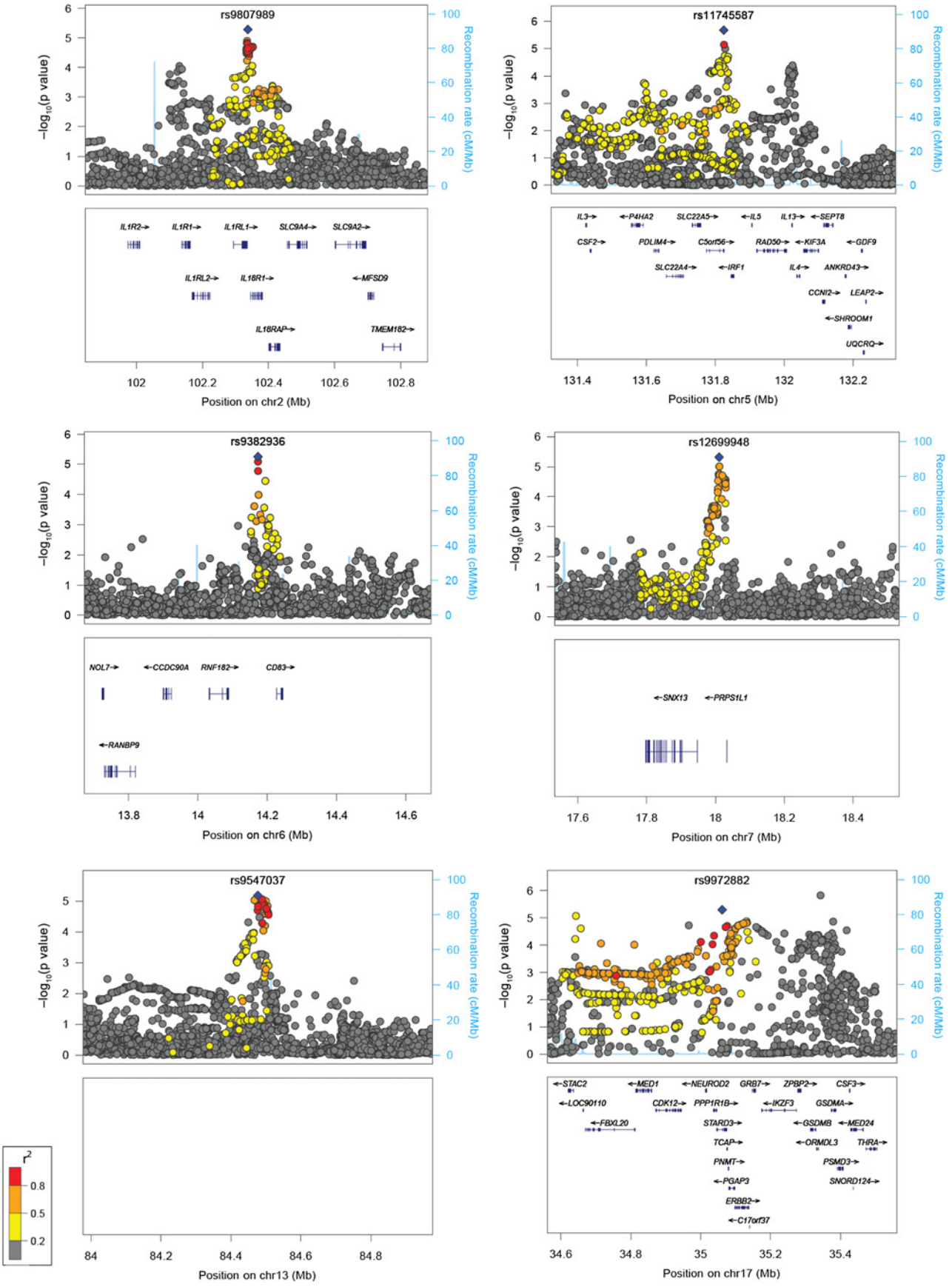


Figure 2 Region plots for suggestive loci. The region plots following imputation show the statistical significance of each single nucleotide polymorphism (SNP) on the −log10 scale as a function of chromosome position (National Center for Biotechnology Information (NCBI) build 36). The sentinel SNP is shown in blue and the correlation (r^2) of each of the surrounding SNPs to the sentinel SNP is shown by their colour (see key). Fine scale recombination rate is plotted in blue.

of asthma. These results potentially provide insight into the biological mechanisms that underlie the regulation of severe asthma and might help in the discovery of novel therapeutic targets for disease.

Author affiliations
[1]Therapeutics and Molecular Medicine, University of Nottingham, Nottingham, UK
[2]Department of Health Sciences, University of Leicester, Leicester, UK
[3]National Heart and Lung Institute, Imperial College, London, UK
[4]Division of Community Health Sciences, St George's, University of London, London, UK
[5]Centre for Infection and Immunity, Queen's University of Belfast, Belfast, UK
[6]School of Pharmacy, Queen's University of Belfast, Belfast, UK
[7]Respiratory Medicine, Birmingham Heartlands Hospital and University of Birmingham, Birmingham, UK
[8]Respiratory Medicine, Institute of Infection, Immunity and Inflammation, University of Glasgow, Glasgow UK
[9]Institute for Lung Health, University of Leicester, Glenfield Hospital, Leicester, UK
[10]The University of Manchester, Manchester Academic Health Science Centre, NIHR Translational Research Facility in Respiratory Medicine, Manchester, UK
[11]Human Genetics and Medical Genomics, Human Development and Health University of Southampton Faculty of Medicine, Southampton, UK
[12]Clinical and Experimental Sciences, University of Southampton Faculty of Medicine, Southampton, UK
[13]Southampton NIHR Respiratory Biomedical Research Unit, University of Southampton Faculty of Medicine, Southampton, UK
[14]Department of Pulmonary Physiology, West Australian Sleep Disorders Research Institute, Western Australia, Australia
[15]A full list of collaborators is available in the web appendix.
[16]The University of Queensland Diamantina Institute, Brisbane, Australia
[17]Lung Institute of Western Australia and Centre for Asthma, Allergy and Respiratory Research, University of Western Australia, Perth, Australia
[18]The Queensland Institute of Medical Research, Brisbane, Australia

Acknowledgements We would like to thank all members of study cohorts taking part in this study, particularly those who provided consent to use of their DNA for genetic epidemiologic analyses and the research staff who contributed to the successful completion of all field studies.

Contributors YIW, IS and IPH wrote the manuscript. Statistical analyses were carried out by YIW, NGS, MSA, LVW and MDT. Subjects were recruited and phenotype data collected by all authors from individual centres. We would like to acknowledge genotyping by CNG, Paris. The study was conceived by IPH and WOCMC. All authors have read and approved the final version.

Funding We acknowledge support of Asthma UK, the Wellcome Trust and the Medical Research Council. Specifically, 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 (http://www.b58cgene.sgul.ac.uk/). Genotyping for the B58C-WTCCC subset was funded by the Wellcome Trust grant 076113/B/04/Z. The B58C-T1DGC genotyping utilised 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). CEB is funded as a Wellcome Trust Senior Clinical Fellow. MDT has been supported by MRC fellowships G0501942 and G0902313. The Australian Asthma Genetics Consortium is funded by the National Health and Medical Research Council of Australia (613627).

Competing interests None.

Ethics approval Ethics approval was provided by individual study centres during study recruitment.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1. **Anon.** 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* 1987;**136**:225—44.
2. **Wenzel SE.** Severe asthma in adults. *Exp Lung Res* 2005;**31**(Suppl 1):22.
3. **Antonicelli L,** Bucca C, Neri M, *et al*. Asthma severity and medical resource utilisation. *Eur Respir J* 2004;**23**:723—9.
4. **Beghé B,** Barton S, Rorke S, *et al*. Polymorphisms in the interleukin-4 and interleukin-4 receptor alpha chain genes confer susceptibility to asthma and atopy in a Caucasian population. *Clin Exp Allergy* 2003;**33**:1111—17.
5. **Nakao F,** Ihara K, Kusuhara K, *et al*. Association of IFN-gamma and IFN regulatory factor 1 polymorphisms with childhood atopic asthma. *J Allergy Clin Immunol* 2001;**107**:499—504.
6. **Salam MT,** Gauderman WJ, McConnell R, *et al*. Transforming growth factor-1 C-509T polymorphism, oxidant stress, and early-onset childhood asthma. *Am J Respir Crit Care Med* 2007;**176**:1192—9.
7. **Van Eerdewegh P,** Little RD, Dupuis J, *et al*. Association of the ADAM33 gene with asthma and bronchial hyperresponsiveness. *Nature* 2002;**418**:426—30.
8. **Wang JY,** Liou YH, Wu YJ, *et al*. An association study of 13 SNPs from seven candidate genes with pediatric asthma and a preliminary study for genetic testing by multiple variants in Taiwanese population. *J Clin Immunol* 2009;**29**:205—9.
9. **Kormann MS,** Depner M, Hartl D, *et al*. Toll-like receptor heterodimer variants protect from childhood asthma. *J Allergy Clin Immunol* 2008;**122**:86—92, 92.e1—8.
10. **Hopes E,** McDougall C, Christie G, *et al*. Association of glutamine 27 polymorphism of beta 2 adrenoceptor with reported childhood asthma: population based study. *BMJ* 1998;**316**:664.
11. **Moffatt MF,** Kabesch M, Liang L, *et al*. Genetic variants regulating ORMDL3 expression contribute to the risk of childhood asthma. *Nature* 2007;**448**:470—3.
12. **Ober C,** Tan Z, Sun Y, *et al*. Effect of variation in CHI3L1 on serum YKL-40 level, risk of asthma, and lung function. *N Engl J Med* 2008;**358**:1682—91.
13. **Gudbjartsson DF,** Bjornsdottir US, Halapi E, *et al*. Sequence variants affecting eosinophil numbers associate with asthma and myocardial infarction. *Nat Genet* 2009;**41**:342—7.
14. **Moffatt MF,** Gut IG, Demenais F, *et al*; GABRIEL Consortium. A large-scale, consortium-based genomewide association study of asthma. *N Engl J Med* 2010;**363**:1211—21.
15. **Ferreira MA,** Matheson MC, Duffy DL, *et al*. Identification of IL6R and chromosome 11q13.5 as risk loci for asthma. *Lancet* 2011;**378**:1006—14.
16. **Torgerson DG,** Ampleford EJ, Chiu GY, *et al*. Meta-analysis of genome-wide association studies of asthma in ethnically diverse North American populations. *Nat Genet* 2011;**43**:887—92.
17. **Li X,** Howard TD, Zheng SL, *et al*. Genome-wide association study of asthma identifies RAD50-IL13 and HLA-DR/DQ regions. *J Allergy Clin Immunol* 2010;**125**:328—35.e11.
18. **Masoli M,** Fabian D, Holt S, *et al*; Global Initiative for Asthma (GINA) Program. The global burden of asthma: executive summary of the GINA Dissemination Committee report. *Allergy* 2004;**59**:469—78.
19. **Patterson N,** Price AL, Reich D. Population structure and eigenanalysis. *PLoS Genet* 2006;**2**:e190.
20. **Purcell S,** Neale B, Todd-Brown K, *et al*. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007;**81**:559—75.
21. **1000 Genomes Project Consortium.** A map of human genome variation from population-scale sequencing. *Nature* 2010;**467**:1061—73.
22. **Li Y,** Willer CJ, Ding J, *et al*. MaCH: using sequence and genotype data to estimate haplotypes and unobserved genotypes. *Genet Epidemiol* 2010;**34**:816—34.
23. *Minimac—Genome Analysis Wiki*. http://genome.sph.umich.edu/wiki/Minimac (accessed 9 Apr 2012).
24. **Aulchenko YS,** Struchalin MV, van Duijn CM. ProbABEL package for genome-wide association analysis of imputed data. *BMC Bioinformatics* 2010;**11**:134.
25. **McCarthy MI,** Abecasis GR, Cardon LR, *et al*. Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nat Rev Genet* 2008;**9**:356—69.
26. **Schmitz J,** Owyang A, Oldham E, *et al*. IL-33, an interleukin-1-like cytokine that signals via the IL-1 receptor-related protein ST2 and induces T helper type 2-associated cytokine. *Immunity* 2005;**23**:479—90.
27. **Fukao T,** Matsuda S, Koyasu S. Synergistic effects of IL-4 and IL-18 on IL-12-dependent IFN-gamma production by dendritic cells. *J Immunol* 2000;**164**:64—71.
28. **Bisgaard H,** Bønnelykke K, Sleiman PM, *et al*. Chromosome 17q21 gene variants are associated with asthma and exacerbations but not atopy in early childhood. *Am J Respir Crit Care Med* 2009;**179**:179—85.
29. **Galanter J,** Choudhry S, Eng C, *et al*. ORMDL3 gene is associated with asthma in three ethnically diverse populations. *Am J Respir Crit Care Med* 2008;**177**:1194—200.
30. **van den Berg TK,** Nath D, Ziltener HJ, *et al*. Cutting edge: CD43 functions as a T cell counterreceptor for the macrophage adhesion receptor sialoadhesin (Siglec-1). *J Immunol* 2001;**166**:3637—40.
31. **Segrè AV,** Groop L, Mootha VK, *et al*; DIAGRAM Consortium, MAGIC Investigators. Common inherited variation in mitochondrial genes is not enriched for associations with type 2 diabetes or related Glycemic traits. *PLoS Genet* 2010;**6**:e1001058.
32. **Fellay J,** Ge D, Shianna KV, *et al*. Common genetic variation and the control of HIV-1 in humans. *PLoS Genet* 2009;**5**:e1000791.

ONLINE REPOSITORY

Y I Wan, PhD[1*], N R G Shrine, PhD[2*], M Soler Artigas, BSc MSc[2], L V Wain, PhD[2], J D Blakey, BMBS PhD[1], M F Moffatt, DPhil[3], A Bush, MD FRCP FRCPCH[3], K F Chung, MD DSc[3], W O C M Cookson, MD DPhil[3], D P Strachan, MD[4], L Heaney, MD MRCP[5], B A H Al-Momani, BSc MSc[6], A H Mansur, PhD FRCP[7], S Manney, PhD[7], N C Thomson, MD FRCP[8], R Chaudhuri, MD[8], C E Brightling, PhD MRCP[9], M Bafadhel, MB ChB[9], A Singapuri, BSc[9], R Niven, MD FRCP[10], A Simpson, MD PhD[10], J W Holloway, PhD[11, 12], P H Howarth, DM FRCP[12, 13], J Hui, PhD[14], A W Musk, PhD FRACP[14], A L James, MD FRACP[14], the Australian Asthma Genetics Consortium[15], M A Brown, PhD[16], S Baltic, PhD[17], M A R Ferreira, PhD[18], P J Thompson, MD FRACP[17], M D Tobin, PhD FFPH[2], I Sayers, PhD[1], I P Hall, DM FRCP[1]

[1]Therapeutics & Molecular Medicine, Nottingham Respiratory BRU, University of Nottingham, Nottingham, UK. [2]Department of Health Sciences, University of Leicester, Leicester, UK. [3]National Heart and Lung Institute, Imperial College, London, UK. [4]Division of Community Health Sciences, St George's, University of London, UK. [5]Centre for Infection and Immunity, Queen's University of Belfast, Belfast, Northern Ireland UK. [6]School of Pharmacy, Queen's University of Belfast, Belfast, Northern Ireland UK. [7]Respiratory Medicine, Birmingham Heartlands Hospital and University of Birmingham, Birmingham, UK. [8]Respiratory Medicine, Institute of Infection, Immunity and Inflammation, University of Glasgow, Glasgow UK. [9]Institute for Lung Health, University of Leicester, Glenfield Hospital, Leicester, UK. [10]The University of Manchester, Manchester Academic Health Science Centre, NIHR Translational Research Facility in Respiratory Medicine, Manchester, UK. [11]Human Genetics & Medical Genomics, Human Development and Health University of

Southampton Faculty of Medicine, Southampton, UK. [12]*Clinical & Experimental Sciences, University of Southampton Faculty of Medicine, Southampton, UK.* [13]*Southampton NIHR Respiratory Biomedical Research Unit, University of Southampton Faculty of Medicine, Southampton.* [14]*Department of Pulmonary Physiology, West Australian Sleep Disorders Research Institute, Western Australia, Australia.* [15]*A full list of collaborators is available in the web appendix.* [16]*The University of Queensland Diamantina Institute, Brisbane, Australia.* [17]*Lung Institute of Western Australia and Centre for Asthma, Allergy and Respiratory Research, University of Western Australia, Perth, Australia.* [18]*The Queensland Institute of Medical Research, Brisbane, Australia.*

*Joint First Authors

Correspondence to:

Yize I Wan,

Division of Therapeutics and Molecular Medicine,

University Hospital of Nottingham,

Nottingham, NG7 2UH,

U.K.

Tel: +44 (0)115 8231068

Fax: +44 (0)115 8231059

Email: yize.wan@nottingham.ac.uk

POPULATION CHARACTERISTICS

Severe asthma case cohorts

The Asthma UK Genetics of Severe Asthma (AUGOSA) Study

This study consists of 750 individuals of European ancestry with severe asthma classified as steps 3 or above based on the Global Initiative for Asthma (GINA) criteria recruited across 8 UK-based centres: Nottingham QMC, Nottingham City Hospital, Belfast, Glasgow, Leicester, Manchester, Birmingham, and Southampton. Analyses included 682 individuals recruited for the AUGOSA study. Subjects from Belfast, Glasgow, Leicester and Manchester fulfilled the American Thoracic Society (ATS) definition of refractory asthma and were recruited as part of the British Thoracic Society (BTS) National Difficult Asthma Registry [1]. Clinical characteristics for subjects from each centre are described below.

Nottingham QMC, Nottingham City Hospital:

A total of 149 individuals were recruited from Nottingham QMC and Nottingham City Hospital, of which 57.3% were female. All individuals had a physician diagnosis of severe asthma and were judged to be at GINA criteria step 3 or above. The mean age was 39.8 years (SD 11.8, range 16–62) and the age of onset of asthma was ≤16 years in 53.7%. The mean FEV1 was 2.61L (SD 0.8) and the mean % predicted FEV1 was 83.2% (SD 22.8). In this population, 13.4% were lifetime smokers and 9.4% were current smokers.

Belfast:

A total of 99 individuals were recruited from Belfast City Hospital, of which 61.6% were

female. All individuals had a physician diagnosis of severe asthma and required treatment at step 4/5 of BTS/SIGN guideline [2]. The mean age was 48.6 years (SD 12.6, range 18–78) and the age of onset of asthma was ≤16 years in 47.2%. The mean FEV1 was 2.1L (SD 0.73) and the mean % predicted FEV1 was 74.4% (SD 24.4). In this population, 65.7% were never smokers, 31.3% were ex-smokers, & 3% were current smokers.

Glasgow:

A total of 249 individuals were recruited from Glasgow, of which 62.2% were female. All individuals had a physician diagnosis of asthma and were judged to be at GINA criteria step 3 or above. The mean age was 51.8 years (SD 12.8, range 18-85) and the age of onset of asthma was ≤16 years in 58.4%. The mean FEV1 was 1.98L (SD 0.8) and the mean % predicted FEV1 was 70.6% (SD 21.5). In this population, 49.4% were never smokers, 12.6% ever smokers and 38.0% were current smokers.

Leicester:

A total of 63 individuals were recruited from The Institute for Lung Health, University of Leicester, of which 79% were female. All individuals had a physician diagnosis of severe asthma and were judged to be at GINA criteria step 3 (18%), 4 (48%) or 5 (34%). The mean age was 47 years (SD 17) and the age of onset of asthma was ≤16 years in 39.7%. The mean FEV1 was 2.4L (SD 0.9) and the mean % predicted FEV1 was 78% (SD 22). In this population, 17% were current smokers and a further 5% were ever smokers (>10 pack year history).

Manchester:

A total of 65 individuals (69% females) were recruited by the Manchester Academic Health Science Centre. All individuals had a physician diagnosis of severe asthma and were judged to be at GINA criteria step 3 or above. The mean age was 45.7 years (SD 12.8) and the age of onset of asthma was ≤16 years in 47.1%. The mean FEV1 was 1.97L (SD 0.9) and the mean % predicted FEV1 was 67% (SD 29). In this population, 32% were lifetime smokers and 11% were current smokers.

Birmingham:

A total of 47 individuals (15% males) were recruited from Birmingham Heartlands Hospital. All individuals were put through a specialist clinic structured severe asthma protocol and had a physician diagnosis of severe asthma and were judged to be at GINA criteria step 4 or above (70% were on maintenance or intermittent oral corticosteroids treatment). The mean age was 44.6 years (range 16–71) and the age of onset of asthma was ≤16 years in 59.2%. The mean FEV1 was 1.95L (SD=0.88) and the mean % predicted FEV1 was 56.7% (SD=28). In this population, 20.5% were lifetime smokers and 2.5% were current smokers."

Southampton:

A total of 78 individuals were recruited from Southampton General Hospital of which 75.6% were female. All individuals had a physician diagnosis of severe asthma and were managed at step 4/5 of the BTS/SIGN guidelines [2]. Their mean age was 42.7 years (SD 12.1, range 17–74) and the age of onset of asthma was ≤16 years in 50.1%. The mean FEV1 was 2.04L (SD 0.7) and the mean % predicted FEV1 was 69.3% (SD 19.1). In this population, 29.9% were lifetime smokers and 15.6% were current smokers.

A further 344 individuals with severe asthma were recruited from three specialist 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 [2, 3]. 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 [4, 5].

Severe asthma in the paediatric 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 β2 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 β2 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* $\geq$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 FEV_1, 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

147 gastroesophageal reflux), psychosocial issues and adverse environmental circumstances as
148 contributing factors to the severity of asthma [5-8].
149
150 The primary analysis dataset of 1,059 cases and 3,345 controls had 80% power ($\alpha = 0.05$) to
151 detect an OR of approximately 1.19 for a SNP with MAF 10%; 1.14 for a SNP with MAF
152 25%; and 1.12 for a SNP with MAF 40%.
153
154 **Replication Cohort**
155 *Australian Asthma Genetics Consortium (AAGC) study*
156 This study includes 7,197 unrelated individuals of European ancestry from Australia. Of
157 these, we selected for the present analysis 231 cases recruited by the Lung Institute of
158 Western Australia (LIWA) who (a) were diagnosed with asthma by clinical examination by a
159 respiratory physician and (b) were on inhaled steroids at 400μg or higher plus a LABA. All
160 recruited subjects had a <10 pack year smoking history. In addition all patients were
161 interviewed by a chest physician, had spirometry and reversibility and/or bronchial reactivity
162 measured. They also had a clinical history compatible with asthma including a history of
163 variability and exacerbations from common asthma triggers. Clinical characteristics for these
164 subjects are summarised in Table E1. All samples were genotyped with Illumina 610K
165 arrays.
166
167 **AAGC samples included in the follow-up analysis.**

Attribute	Asthma cases	Asthma controls
N	231	1345
% Female	61.0	60.6
Mean age (SD, range)	51 (16.3, 11-87)	32 (15.0, 12-89)
% with family history of asthma	65.3	0
% SPT+	85.7	-
% Asthma onset <= 16 years	54.5	-

Mean FEV1, L (SD)	2.21 (0.88)	-
Mean FVC, L (SD)	3.23 (1.09)	-
Mean FEV1/FVC (SD)	0.68 (0.13)	-
% lifetime smoker	48.1	-
% current smoker	7.4	-
% ever admitted to hospital for asthma	53.7	-

We restricted our analysis to 1,345 asthma-free controls also genotyped with 610K arrays recruited through two studies: LIWA (n=35) and the Queensland Institute of Medical Research (QIMR) studies (n=1,310). The latter reported never having had asthma in questionnaires completed as part of five epidemiological studies previously conducted at QIMR and described in more detail elsewhere [9]. All subjects were confirmed to be unrelated and of European ancestry through the analysis of genome-wide allele sharing.

Standard SNP QC filters were applied, including the removal of SNPs with call rate <95%, minor allele frequency (MAF) < 0.01 and Hardy-Weinberg equilibrium test *P*-value < 10^{-6}. Autosomal SNPs passing QC were then used to impute up to 7.8 million variants available from the combined 1000 Genomes (CEU, Mar 2010 release) and HapMap 3 (all 11 populations, Feb 2009 release) reference panels using Impute2 [10]. After imputation, we excluded SNPs with low imputation accuracy (information < 0.3), MAF < 0.01, Hardy-Weinberg equilibrium test *P*-value < 10^{-6}. After QC, genotype data was for 5.7 million SNPs that were tested for association using a standard case-control allelic test. The genomic inflation factor for this analysis was 1.019.

The replication cohort dataset of 231 cases and 1,345 controls had 80% power to detect an OR 1.46 (MAF=10%), 1.31 (MAF=25%), and 1.28 (MAF=40%).

For the severe versus mild analyses, a total of 1,085 mild asthmatics were identified as having never received steroid medication in their lifetime.

Full list of collaborators:

M C Matheson[1], D L Duffy[2], G B Marks[3], P Danoy[4], M J Abramson[5], C F Robertson[6], J Hui[7-10], P Le Souef[11], N G Martin[2], S C Dharmage[1], G W Montgomery[2], L P Chung[12], F Chea[12], L Price[12], W Mitrpant[12], S Temple[12], C B Lim[12].

[1]Centre for Molecular, Environmental, Genetic and Analytic Epidemiology, University of Melbourne, Melbourne, Australia. [2]The Queensland Institute of Medical Research, Brisbane, Australia. [3]Woolcock Institute of Medical Research, University of Sydney, Sydney, Australia. [4]University of Queensland Diamantina Institute, Princess Alexandra Hospital, Brisbane, Australia. [5]Department of Epidemiology & Preventive Medicine, Monash University, Melbourne, Australia. [6]Department of Respiratory Medicine, Royal Children's Hospital, Parkville, Australia. [7]PathWest Laboratory Medicine of Western Australia (WA), Nedlands, Australia. [8]School of Population Health, The University of WA, Nedlands, Australia. [9]School of Pathology and Laboratory Medicine, The University of WA, Nedlands, Australia. [10]Busselton Population Medical Research Foundation, Sir Charles Gairdner Hospital, Perth, Australia. [11]School of Paediatrics and Child Health, Princess Margaret Hospital for Children, Perth, Australia. [12]Lung Institute of Western Australia and Centre for Asthma, Allergy and Respiratory Research, University of Western Australia, Perth, Australia.

References:

1. Heaney LG, et al. Thorax 2010. 65(9) : 787-794.
2. Levy ML, et al. Prim Care Respir J 2009;18 Suppl 1:S1-16.

3. Haldar P, et al. Am J Respir Crit Care Med 2008; 78(3):218-24.
4. Haldar P, el at. N Engl J Med 2009; 360(10):973-84.
5. Proceedings of the ATS Workshop on Refractory Asthma. Am J Respir Crit Care Med 2000; 162(6):2341-51.
6. Chung KF, et al. Eur Resp J 1999. 13(5): 1198-208.
7. Ayres JG, et al. Thorax 1998. 53(4): 315-21
8. Bacharier LB, et al. J Allergy Clin Immunol 2008. 122(6): 1127-1135.e8.
9. Stanojevic S, et al. Am J Resp Crit Care Med 2008. 177(3): 253-60.
10. Ferreira MA, et al. Eur J Hum Genet 2010. 19(4):458-64.
11. Howie BN, et al. PLoS Genet 2009; 5(6): e1000529

TABLE E1

	Platform	SNP markers	Samples
Cases			
AUGOSA 610	llumina 610K	582,892	113
AUGOSA 660	llumina 660W	557,124	530
GABRIEL	llumina 610K	582,892	290
`otal			**933**
Controls			
3HS	llumina 610K	82,892	65
VTCCC2	llumina 1.2M	,157,986	481
`1DGC	llumina 550K	61,303	300
`otal			**346**
SNPs in common		90,303	

Samples genotyped on each platform.

FIG E1

Principal components analysis (PCA) of study populations was carried out to correct for population structure based on the covariance of effect allele loadings on a representative sample of 57,213 SNPs in low LD using EIGENSOFT. A) No clear separation of clusters for each cohort suggests homogeneous populations. B) The PCA analysis identified 119 outliers with variances >6σ.

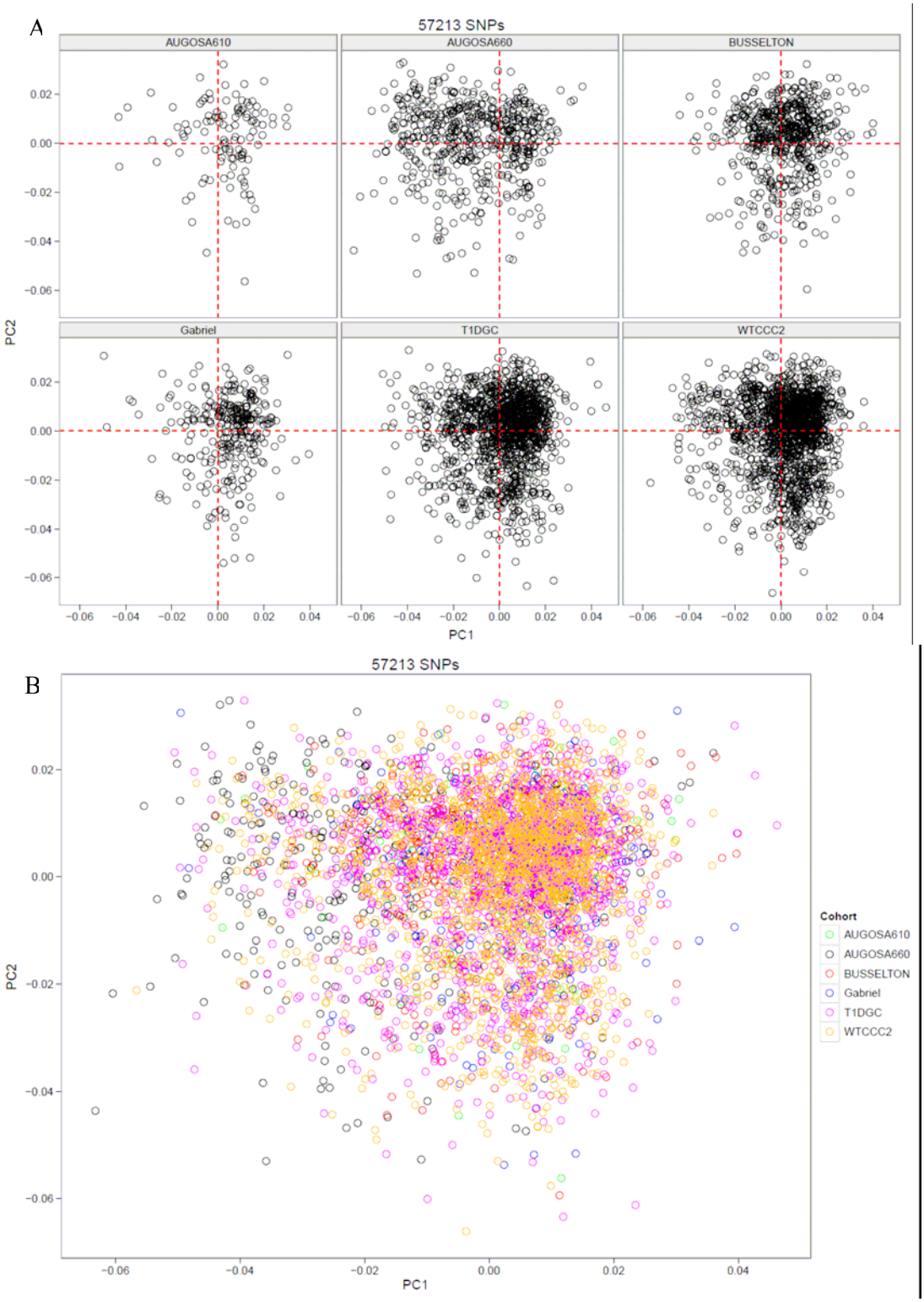

A
57213 SNPs
AUGOSA610
AUGOSA660
BUSSELTON
Gabriel
T1DGC
WTCCC2
PC2
PC1
B
57213 SNPs
PC2
PC1
Cohort
AUGOSA610
AUGOSA660
BUSSELTON
Gabriel
T1DGC
WTCCC2

TABLE E2

Location				AUGOSA (933 cases, 3,346 controls)				GABRIEL (10,365 cases 16,110 controls)			
Chr	Locus	SNP	Position	Ref	Alt	OR (95% CI)	P Value	Ref	Alt	OR (95% CI)	P Value
Genome-wide significant loci ($p \leq 7.2 \times 10^{-8}$) in the GABRIEL study											
2	IL18R1	rs3771166	102352654	G	A	0.79 (0.71-0.88)	1.93×10^{-5}	G	A	0.87 (0.83-0.91)	3.40×10^{-9}
6	HLA-DQ	rs9273349	32733847	-	-	-	-	T	C	1.18 (1.13-1.24)	7.00×10^{-14}
9	IL33	rs1342326	6180076	T	C	1.18 (1.03-1.35)	0.018	T	C	1.20 (1.13-1.28)	9.20×10^{-10}
15	SMAD3	rs744910	65233839	A	G	1.21 (1.09-1.35)	2.88×10^{-4}	G	A	0.89 (0.86-0.92)*	3.90×10^{-9}
17	GSDMB	rs2305480	35315722	G	A	0.80 (0.72-0.89)	5.56×10^{-5}	G	A	0.85 (0.81-0.90)	9.60×10^{-8}
17	GSDMA	rs3894194	35375519	G	A	1.25 (1.12-1.39)	4.39×10^{-5}	G	A	1.17 (1.11-1.23)	4.60×10^{-9}
22	IL2RB	rs2284033	35863980	G	A	0.92 (0.82-1.02)	0.105	G	A	0.89 (0.86-0.93)	1.20×10^{-8}
Suggestive loci ($p \leq 5 \times 10^{-7}$) in the GABRIEL study											
5	SLC22A5	rs2073643	131751187	C	T	1.15 (1.04-1.28)	0.009	T	C	0.90 (0.87-0.94)*	2.20×10^{-7}
5	IL13	rs1295686	132023742	C	T	1.29 (1.14-1.47)	9.30×10^{-5}	T	C	0.87 (0.83-0.92)*	1.40×10^{-7}
15	RORA	rs11071559	58857280	C	T	0.89 (0.76-1.05)	0.159	C	T	0.85 (0.80-0.90)	1.10×10^{-7}

* Opposite coding allele

Association results in AUGOSA of SNPs identified with significant or suggestive evidence of effects on the risk of asthma by the GABRIEL Consortium. Odds ratios were calculated by designating alternative alleles (Alt) as effect alleles. Ref denotes reference allele, Alt alternative allele and CI confidence interval. We were unable to test for association with the *HLA-DQ* locus due to reduced coverage on the Illumina genotyping platform, no proxy SNPs in LD were identified. The SNP in closest proximity was genotyped rs2187668 (32713862), p=0.26.

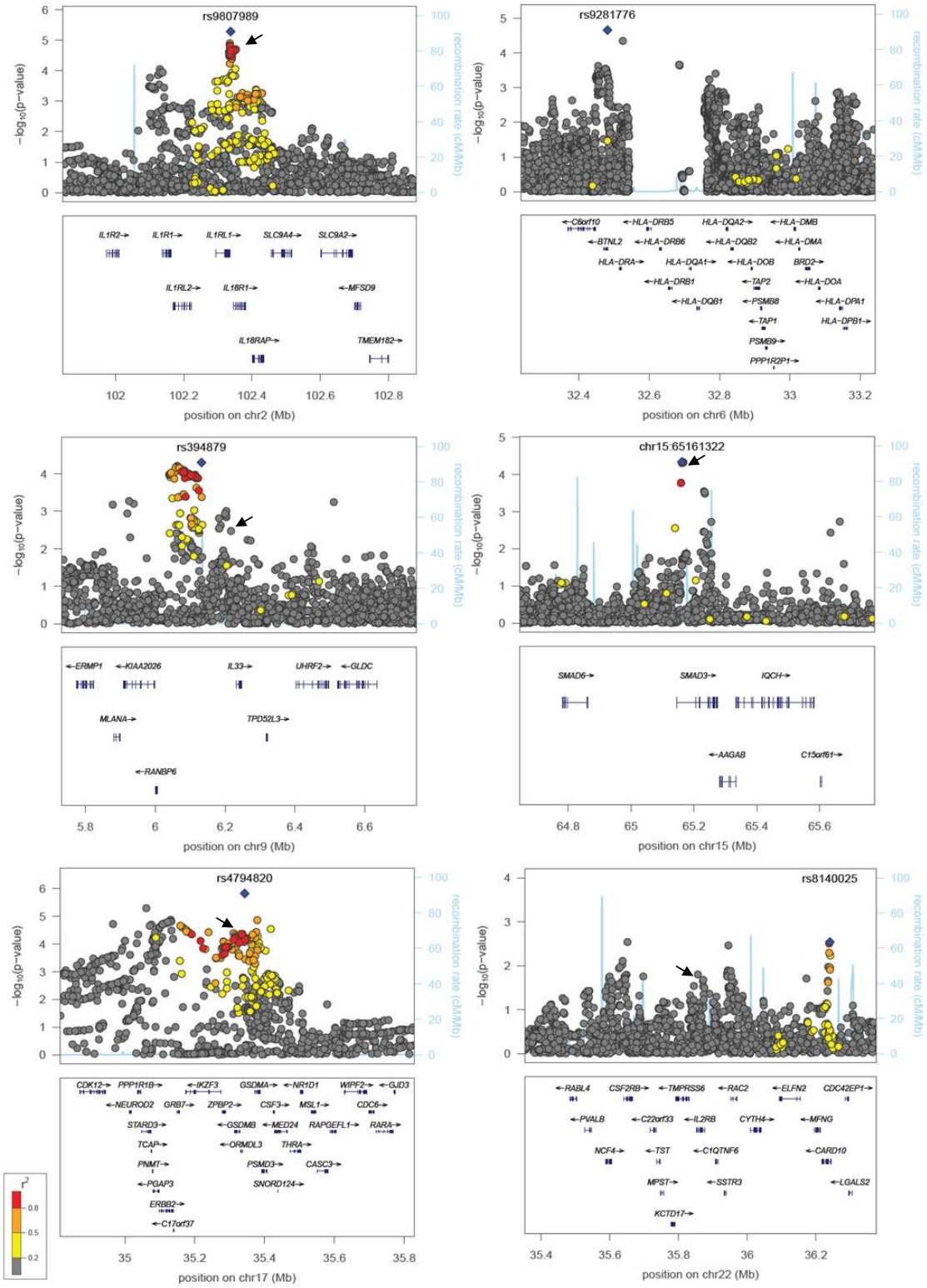
FIG E2

Regions plots were generated for the highest significance SNP in AUGOSA within GABRIEL identified loci (reported SNP ±500Kb), listed in Table E2.

The Region plots show the statistical significance of each SNP on the –log10 scale as a function of chromosome position (NCBI build 36). The pivotal SNP is shown in blue and the correlation (r^2) of each of the surrounding SNPs to the pivotal SNP is shown by their colour (see key). Fine scale recombination rate is plotted in blue. SNPs reported by GABRIEL are identified by the arrow shown.

We were unable to test for association with the *HLA-DQ* locus due to reduced coverage on the Illumina genotyping platform, no proxy SNPs in LD were identified.

Association results in AUGOSA of genome-wide significant loci ($p \leq 7.2x10^{(-8)}$) in the GABRIEL study



Association results in AUGOSA of suggestive loci ($p \leq 5x10^{(-7)}$) in the GABRIEL study

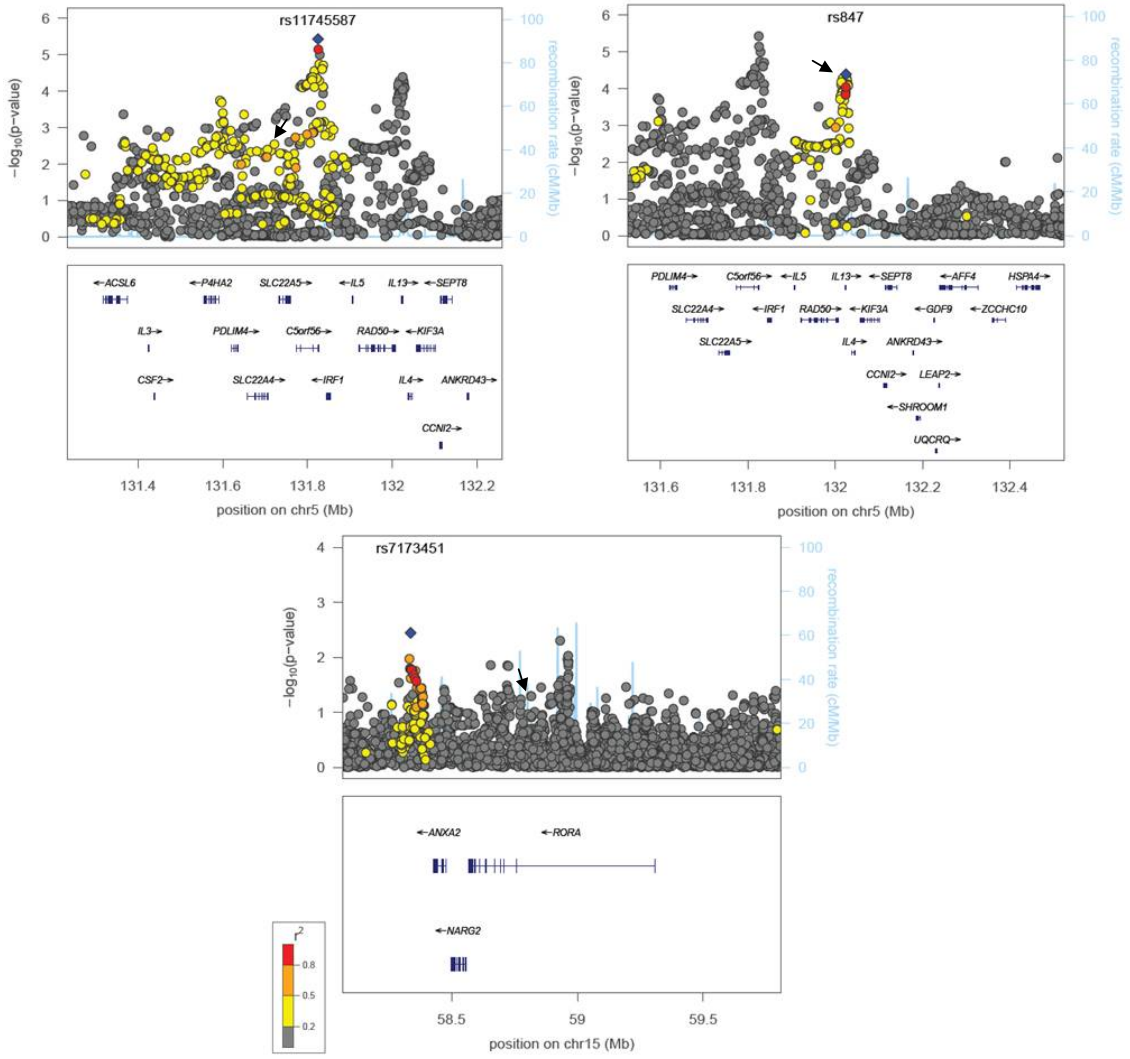


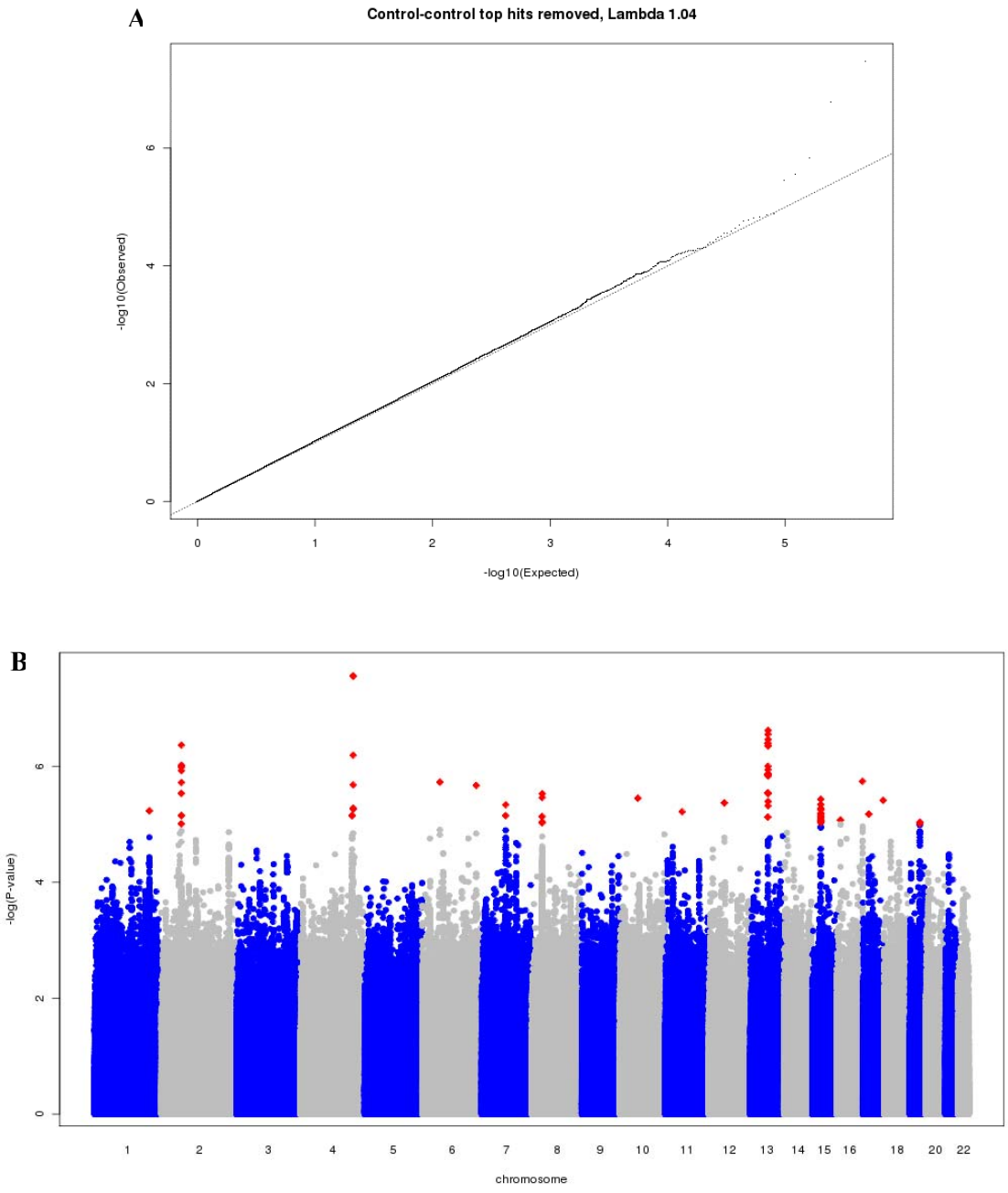
TABLE E3

Chr	Locus	GABRIEL SNP	AUGOSA top SNP in region (±500kb)	OR (95% CI)	P Value
Genome-wide significant loci ($p \leq 7.2x10^{(-8)}$) in the GABRIEL study					
2	IL18R1	rs3771166	rs9807989	0.76 (0.67-0.85)	$5.20x10^{(-6)}$
6	HLA-DQ	rs9273349	rs9281776	1.72 (1.34-2.22)	$2.18x10^{(-5)}$
9	IL33	rs1342326	rs394879	0.74 (0.64-0.86)	$5.01x10^{(-5)}$
15	SMAD3	rs744910	chr15:65161322	1.66 (1.30-2.11)	$4.61x10^{(-5)}$
17	GSDMB	rs2305480	rs4794820	0.76 (0.68-0.85)	$1.52x10^{(-6)}$
17	GSDMA	rs3894194	rs4794820	0.76 (0.68-0.85)	$1.52x10^{(-6)}$
22	IL2RB	rs2284033	**rs8140025**	**1.17 (1.06-1.30)**	**0.003**
Suggestive loci ($p \leq 5x10^{(-7)}$) in the GABRIEL study					
5	SLC22A5	rs2073643	**rs11745587**	**1.30 (1.17-1.45)**	**$2.09x10^{(-6)}$**
5	IL13	rs1295686	rs847	1.35 (1.17-1.55)	$4.05x10^{(-5)}$
15	RORA	rs11071559	rs7173451	0.85 (0.76-0.95)	0.004

Highest significance SNPs in AUGOSA within GABRIEL identified loci. Genotyped SNPs are shown in bold, imputed SNPs are shown in non-bold.

FIG E3

Quantile-Quantile plot and Manhattan plot of severe versus mild-to-moderate asthma results. A) Quantile-Quantile (Q-Q) plot showing GWA results for genotyped SNPs. The straight line shows the distribution of 488,809 SNPs analysed in 1,026 severe asthmatics and 1,028 mild asthmatics under the null hypothesis. B) Manhattan plot showing GWA results for 488,809 genotyped SNPs analysed in 1,026 severe asthmatics and 1,028 mild asthmatics under analysis.



330

331

332

333

334

335

TABLE E4

Chr	SNP	Position	A1	BETA	SE	OR (95% CI)	p value	Gene
5	rs1837253	110429771	T	-0.29	0.06	0.75(0.66-0.85)	$5.52x10^{(-6)}$	*TSLP* (upstream)
3	rs11711981	195373706	A	-0.23	0.05	0.80 (0.72-0.88)	$2.19x10^{(-5)}$	
13	rs1414320	109333093	C	0.35	0.08	1.42 (1.21-1.68)	$2.33x10^{(-5)}$	
2	rs333236	107970399	C	0.29	0.06	1.33 (1.18-1.51)	$6.47x10^{(-6)}$	*SLC5A7* (intronic)
16	rs8057431	83397455	G	-0.27	0.06	0.76 (0.67-0.86)	$1.20x10^{(-5)}$	
6	rs12200468	73408591	G	-0.31	0.07	0.73 (0.64-0.84)	$1.47x10^{(-5)}$	*KCNQ5* (intronic)
5	rs6876572	150719195	G	-0.28	0.06	0.76 (0.67-0.86)	$1.60x10^{(-5)}$	
1	rs6424762	78522496	G	-0.23	0.05	0.79 (0.71-0.88)	$1.70x10^{(-5)}$	
10	rs2068888	94829632	A	0.23	0.05	1.26 (1.13-1.40)	$2.12x10^{(-5)}$	
4	rs981516	161784820	A	0.24	0.06	1.27 (1.14-1.43)	$2.27x10^{(-5)}$	
2	rs896733	20682797	C	-0.28	0.07	0.75 (0.66-0.86)	$2.87x10^{(-5)}$	*HS1BP3* (intronic)
4	rs13150370	80106067	A	0.22	0.05	1.25 (1.13-1.39)	$3.05x10^{(-5)}$	
11	rs597872	107071795	A	-0.23	0.06	0.79 (0.71-0.88)	$3.08x10^{(-5)}$	
3	rs1343700	125054444	G	0.23	0.06	1.26 (1.13-1.40)	$3.20x10^{(-5)}$	*MYLK* (intronic)
8	rs4870880	125020708	C	-0.27	0.07	0.76 (0.67-0.87)	$3.25x10^{(-5)}$	*FER1L6* (intronic)
19	rs2241351	18294225	T	0.32	0.08	1.37 (1.18-1.60)	$3.50x10^{(-5)}$	*LSM4* (intronic)
1	rs6701588	171619782	T	0.30	0.07	1.35 (1.17-1.56)	$4.21x10^{(-5)}$	
20	rs2326614	5195393	C	0.23	0.06	1.25 (1.12-1.40)	$4.53x10^{(-5)}$	
11	rs7480563	1091649	T	0.22	0.05	1.24 (1.12-1.38)	$4.68x10^{(-5)}$	*MUC2* (missense)

Additional SNPs with $p< 5\times10^{-5}$ not already reported due to having no second supporting SNP with $p<5\times10^{-5}$.