

RESEARCH ARTICLE

Open Access

Defining the contribution of SNPs identified in asthma GWAS to clinical variables in asthmatic children

Asif S Tulah^{1,3}, John W Holloway² and Ian Sayers^{1*}

Abstract

Background: Asthma genome-wide association studies (GWAS) have identified several asthma susceptibility genes with confidence; however the relative contribution of these genetic variants or single nucleotide polymorphisms (SNPs) to clinical endpoints (as opposed to disease diagnosis) remains largely unknown. Thus the aim of this study was to firstly bridge this gap in knowledge and secondly investigate whether these SNPs or those that are in linkage disequilibrium are likely to be functional candidates with respect to regulation of gene expression, using reported data from the ENCODE project.

Methods: Eleven of the key SNPs identified in eight loci from recent asthma GWAS were evaluated for association with asthma and clinical outcomes, including percent predicted FEV₁, bronchial hyperresponsiveness (BHR) to methacholine, severity defined by British Thoracic Society steps and positive response to skin prick test, using the family based association test additive model in a well characterised UK cohort consisting of 370 families with at least two asthmatic children.

Results: *GSDMB* SNP rs2305480 (Ser311Pro) was associated with asthma diagnosis ($p = 8.9 \times 10^{-4}$), BHR ($p = 8.2 \times 10^{-4}$) and severity ($p = 1.5 \times 10^{-4}$) with supporting evidence from a second *GSDMB* SNP rs11078927 (intronic). SNPs evaluated in *IL33*, *IL18R1*, *IL1RL1*, *SMAD3*, *IL2RB*, *PDE4D*, *CRB1* and *RAD50* did not show association with any phenotype tested when corrected for multiple testing. Analysis using ENCODE data provides further insight into the functional relevance of these SNPs.

Conclusions: Our results provide further support for the role of *GSDMB* SNPs in determining multiple asthma related phenotypes in childhood asthma including associations with lung function and disease severity.

Keywords: Asthma, ENCODE, eQTL, GWAS, Clinical endpoints

Background

Asthma is a complex respiratory disease with susceptibility involving the interplay between genetic and environmental factors [1]. The chronic inflammation of the airways is characterised by reversible airflow obstruction and increased hyperresponsiveness of the airways. Genome-wide association studies (GWAS) are currently the preferred method for the identification of genetic determinants in complex multifactorial diseases such as asthma. These studies have the advantage of being able to identify

polymorphisms with small effect sizes and localise smaller susceptibility regions due to linkage disequilibrium (LD). The first GWAS for asthma was completed involving 994 childhood-onset asthma patients and 1,243 non-asthma controls and identified single nucleotide polymorphisms (SNPs) within chromosome 17q21 spanning the *ORMDL3/GSDMB* genes [2]. These SNPs showed highly significant association with childhood asthma and were also replicated in two independent cohorts [2]. A study in an Icelandic cohort showed significant association with asthma and a SNP in *IL1RL1* [3] and SNPs in *PDE4D* have also shown association with genome-wide significance with asthma [4]. *IL-13* and *RAD50* on 5q31.1 and the *HLA-DR/DQ* locus at 6p21.3 have been implicated by Li et al. [5].

* Correspondence: ian.sayers@nottingham.ac.uk

¹Division of Respiratory Medicine, Queen's Medical Centre, University of Nottingham, Nottingham NG7 2UH, United Kingdom
Full list of author information is available at the end of the article

Other polymorphisms spanning *IL33*, *IL18R1*, *SMAD3*, *IL2RB* and *CRB1* have been proposed by other GWAS using large scale discovery and replication approaches [6,7]. However, due to the phenotype used for case definition, which is often a physician diagnosis or self-report of asthma, there remains a need to determine the contribution of these polymorphisms to clinically relevant endpoints in asthma. This will provide a greater insight into the mechanisms altered in individuals carrying the relevant risk alleles.

One problem that arises from GWAS is that LD patterns can result in multiple polymorphisms at a locus being associated with disease phenotype, even if only one of them is causal. A problem therefore exists in the identification of potential causal variants. This can be simply analysed for polymorphisms within coding regions, as they could potentially affect the polypeptide structure and function and those SNPs in regulatory or intronic regions could affect splice site motifs, transcriptional regulation or RNA stability [8,9]. However, many GWAS have implicated polymorphisms in non transcribed regions far away from annotated genes, where their role is likely to be regulatory [10]. The ENCODE project was established to identify all the functional elements such as protein binding sites, transcripts (coding and non coding) and chromatin marks in the human genome. This represents a valuable resource to provide additional meaning to association data provided from GWAS [11,12]. Taken in consideration with LD data from the 1000 Genomes project [13], computational analysis will enable assignment of more meaningful potential function to SNPs identified from GWAS or those SNPs which are in LD with a GWAS associated SNP [10], to help better prioritise SNPs for functional analysis.

The aim of the present study was to complete a SNP for SNP analysis of the key variants identified in the Caucasian population meeting genome-wide significance and/or replication identified 2007–2011 on clinical variables underlying asthma (Table 1). As these SNPs have previously shown genome-wide association, we also sought to investigate the potential functional significance of these eleven SNPs and those in LD (with these SNPs) by mining the ENCODE dataset [14] and also investigating whether they show association with expression quantitative trait loci (eQTL).

Methods

Asthma cohort characteristics

The asthma family cohort was made up of United Kingdom families from the Southampton and Nottingham areas ($n = 370$ families were used in this study) (Table 2). From the Southampton area, 341 Caucasian families ($n = 1508$) with at least two biological siblings with physician diagnosed asthma were recruited. This cohort has been described in detail previously [15]. Baseline FEV₁ (forced expiratory volume in one second) was measured as best of three values within 5% performed using Vitalograph dry-wedge bellows spirometer (Vitalograph Ltd, Buckingham, UK). This was determined 14 days after respiratory tract infection or use of bronchodilator or anti-allergic medication. 46 Caucasian families ($n = 184$) with at least two biological siblings with physician diagnosed asthma were recruited from the Nottingham area [15]. FEV₁ was defined as the best of three values. All subjects were classified according to British Thoracic Society Step Guidelines (BTS steps, ranging from step 1 to step 5) based on physician prescribed medication [16]. Ethical approval for the

Table 1 Key SNPs identified in asthma GWAS included in the current analyses

Gene SNP	Key SNP	Non-coded Allele	Coded Allele	Location	MAF* (Control/case)	P-value	Phenotype/Reference	MAF in UK families
<i>GSDMB</i>	rs2305480	C	T	Ser311Pro	0.45	6E-23	Asthma (childhood) [2,6]	0.43
<i>GSDMB</i>	rs11078927	C	T	intron	0.45/0.40	7.4E-07	Asthma [2,6]	0.43
<i>IL33</i>	rs1342326	T	G	5' region	0.16	9E-10	Asthma [6]	0.17
<i>IL33</i>	rs3939286	G	A	5' region	0.25/0.28	5.3E-06	Asthma [3]	0.27
<i>IL18R1</i>	rs3771166	C	T	intron	0.38	3E-09	Asthma [6]	0.38
<i>IL1RL1</i>	rs1420101	G	A	intron	0.37/0.40	5.5E-12	Asthma [3]	0.38
<i>SMAD3</i>	rs744910	A	G	intron	0.49	4E-09	Asthma [6]	0.48
<i>IL2RB</i>	rs2284033	A	G	intron	0.44	1.1E-08	Asthma [6]	0.43
<i>PDE4D</i>	rs1588265	A	G	intron	0.34/0.23	4.3E-07	Asthma [4]	0.29
<i>CRB1</i>	rs2786098	C	A	intron	0.15	9.3E-11	Asthma [7]	0.22
<i>RAD50</i>	rs2244012	T	C	intron	0.21	3.05E-07	Asthma [5]	0.18

Previous GWAS associations include; a chromosome 17 locus (orm1-like protein3 (*ORMDL3*)/gasdermin B (*GSDMB*)), interleukin 33 (*IL33*), a chromosome 2 locus (interleukin 18 receptor (*IL18R*)/interleukin 1 receptor like 1 (*IL1RL1*)), mothers against decapentaplegic drosophila homolog 3 (*SMAD3*), and IL2 receptor beta (*IL2RB*), phosphodiesterase 4D, cAMP specific (*PDE4D*), a chromosome 1 locus (Crumbs homolog 1 precursor (*CRB1*)) and RAD50, *S. Cerevisiae*, homolog of (*RAD50*). MAF minor allele frequency, *a single MAF is shown where the original paper only reported the MAF in all subjects. The p-value stated is for asthma association in the original study (referenced).

Table 2 Clinical characteristics of UK families

	Pedigrees	Sibling 1	Sibling 2
Age (mean years), \pm SD	26.2 \pm 1.6	13.3 \pm 4.4	10.3 \pm 4.6
Gender (% female)	50.8	54.6	53.7
Asthma (% doctor diagnosed)	61.6	100	100
FEV ₁ (% predicted, \pm SD)	ND	95.4 \pm 15.6	96.2 \pm 14.9
Positive skin prick test (%)	62.5	73.9	65.1
Log total serum IgE (mean \pm SD)	2.04 \pm 0.7	2.33 \pm 0.7	2.33 \pm 0.7
Eczema (% questionnaire)	41.1	53.5	55.1
Hay fever (% questionnaire)	50.2	64.9	47.4
BTS (%) step 1	19.1	24.9	23.8
• Step 2	32.5	52.7	55.7
• Step 3	6.1	11.6	10.2
• Step 4	1.8	4.6	1.4
• Step 5	2.9	4.6	5.3
PC ₂₀ (4 mg/ml,%)		46.9	59.9
n	1578	370	361

The recruitment and clinical characterisation of these subjects has been extensively described elsewhere [15]. PC₂₀ (4 mg/ml, %) only available in Southampton families.

Southampton subjects was obtained from the Southampton and South West Hampshire and the Portsmouth and South East Hampshire Local Research Ethics Committees and for the Nottingham subjects, the Nottingham University Medical School Ethics Committee. Written informed consent for study participation was obtained from participants (from the parent/guardian for the children).

SNP selection and genotyping

Eleven SNPs were chosen that have been identified from asthma GWAS identified 2007–2011 (Table 1). These were in eight loci; *ORMDL3/GSDMB*, *IL33*, *IL18R1/IL1RL1*, *SMAD3*, *IL2RB*, *PDE4D*, *CRB1* and *RAD50*. SNPs were genotyped using KASPar technology by KBiosciences (Hertfordshire, UK). For quality control of genotyping data, Chi square was used to test for any deviation between the observed genotype frequencies and the expected, under the Hardy Weinberg Equilibrium.

Association analyses

Association analysis in the family cohorts between the GWAS SNPs and asthma-related traits was conducted using the family based association test (FBAT) software (version 1.5.1) [17] in the additive model. Clinically relevant endpoints including; asthma diagnosis, Forced Expiratory Volume in one second (FEV₁% Predicted), bronchial hyperresponsiveness to methacholine (BHR), British Thoracic Society (BTS) defined severity [18] and atopy (positive response to skin prick test to any one of a panel of allergens as outlined [15]). We completed power calculations based on the method of Risch & Merikangas

[19]. There was between 0.93 and 0.98 power to detect an association with asthma with a significance level of $p = 0.05$ and a relative risk of 1.5. There was between 0.51 and 0.69 power to detect an association with $p = 0.05$ and a relative risk of 1.25 (calculations based on minor allele frequency (min 0.17 to max 0.48 observed). Power curves for these MAF are provided in Additional file 1. To address multiple testing, we used Bonferonni correction, which considered $p < 9.1 \times 10^{-4}$ significant. This was based on our eleven genotyped SNPs analysed for five phenotypes (0.05/55 total tests = 9.1×10^{-4} as a threshold p-value).

Bioinformatic analysis of genotyped SNPs

HaploReg [14] (a web based tool) was used to provide information about the sequence surrounding a SNP. It has been developed to mine chromatin state information from nine human cell lines; GM12878, H1, HMEC, HSMM, HepG2, Huvec, K562, NHEK and NHLF. The tool provides information on SNPs that are in LD with the queried SNP (r^2 was set to 1 for these analyses and the population chosen for analyses was of European descent) and also gives information on which of these SNPs lie in enhancer histone marks, DNase sites and protein binding regions. It can also inform on the conservation of the relevant sequence and any changes to regulatory motifs based on the SNP allele changes. Each of the eleven genotyped SNPs were entered into HaploReg and analysis was completed with $r^2 = 1$. A more detailed analysis was completed for any SNP which showed significant association in the population study.

We used the program mRNA by SNP browser (version 1.0.1) to investigate whether any of the GWAS SNPs genotyped as part of this study were associated with regions of known asthma expression quantitative trait loci (eQTL). This program incorporates a database of eQTL from asthma studies and provides a graphical output to browse data from GWAS. It provides association data between 54,675 transcripts and 406,912 SNPs. Statistics are provided for those SNPs with association $p < 0.001$. We used the program to search for each SNP in turn by inputting either the gene name or the SNP rs number, looking to see whether the association reached genome-wide significance ($p < 10^{-8}$) [20]. We also used the eQTL browser from the University of Chicago (<http://eqtl.uchicago.edu/cgi-bin/gbrowse/eqtl/>) which considers eQTLs from different tissue sources. A second stage of this analysis was to investigate whether these same SNPs showed genome-wide association with eQTL from a genome-wide search in 1,111 human lung samples [21].

Results

Clinical characteristics and genotyping

Clinical characteristics for the asthma families are shown in Table 2. Mean age was 13.3 years and 10.3 years old,

mean FEV₁ (% Predicted) was 95.4 ± 15.6 and 96.2 ± 14.9 and % atopy (SPT) was 73.9% and 65.1% for sibling 1 and 2 respectively. These asthmatic children had predominantly mild-moderate disease as outlined by BTS guidelines. Genotyping was completed using KASPar (KBioscience) and quality control completed as outlined in previous studies [15]. All SNPs were in Hardy Weinberg equilibrium and allele frequencies in the UK families were similar to those previously reported (Table 1).

Association with asthma, FEV₁ and BHR

A *GSDMB* SNP at the 17q21 locus, rs2305480 (Ser311Pro) was associated with asthma diagnosis, with the T allele conferring risk to develop asthma as previously reported [6] ($p = 8.9 \times 10^{-4}$, $z = -3.23$) (Table 3). Another *GSDMB* SNP rs11078927 (intronic) showed similar findings with asthma but did not reach statistical significance ($p = 9.5 \times 10^{-4}$, $z = -3.23$) although this SNP was significantly associated with FEV₁ (% Predicted) ($p = 8.7 \times 10^{-4}$, $z = -3.33$). Significant association was also observed with rs2305480 (*GSDMB*, Ser311Pro) and BHR (PC₂₀ methacholine, 4 mg/ml) ($p = 8.2 \times 10^{-4}$, $z = -3.35$). No other SNP association survived correction for multiple testing (Table 3).

Association with asthma severity and atopy

Both *GSDMB* SNPs showed significant association with severity of disease defined by BTS steps (1–5), both with the same direction of association; rs2305480 ($p = 1.5 \times 10^{-4}$, $z = 3.79$) and rs11078927 ($p = 2.2 \times 10^{-4}$, $z = -3.70$). No other SNP association survived correction for multiple testing (Table 4).

Bioinformatic analysis of SNPs using data from ENCODE

We sought to investigate whether data collected from the ENCODE Consortium [12] could enrich the potential functional significance of our genotyped GWAS associated SNPs and also those SNPs which were in LD (Table 5). The two *GSDMB* SNPs were found to be in the same LD block ($r^2 = 1$), with 24 other SNPs. rs2244012 (*RAD50*/intronic) was found to be in an LD block with 50 other SNPs. The other genotyped SNPs were in smaller LD blocks. Most of our genotyped SNPs analysed were found to have at least one SNP (in their LD block) existing in an enhancer histone mark site, DNase site or a region where protein binding consensus sequences exist, highlighting a potential functional effect of these particular SNPs. There was evidence for LD SNPs altering regulatory motifs, with 14 SNPs in the *GSDMB* LD block and 25 SNPs in the *RAD50* LD block the highest numbers (Table 5).

Analyses of asthma associated *GSDMB* SNPs using ENCODE data

As the two *GSDMB* SNPs (rs2305480/Ser311Pro and rs11078927/intronic) showed significant association in our study, we sought to further investigate potential functional implications using the ENCODE dataset. These two *GSDMB* SNPs themselves were not predicted to be important from a regulatory role, however they were in a LD block with 24 other SNPs. Almost all of the SNPs in this LD block were in non-coding/intronic regions, with the exception of rs10852935, which is a synonymous polymorphism in *ZBP2* and the genotyped rs2305480/Ser311Pro (Table 6). Twenty four of these SNPs were not conserved (according to the GERP

Table 3 Association of key GWAS SNPs with asthma diagnosis, percent predicted FEV₁ and bronchial hyperresponsiveness to methacholine

SNP	Gene	Asthma	Z	P	FEV ₁	Z	P	BHR	Z	P
		Fam			Score			value		
rs2305480 C/T	<i>GSDMB</i>	265	-3.323	8.9×10⁻⁴	274	-3.247	1.1×10 ⁻³	240	-3.345	8.2×10⁻⁴
rs11078927 C/T	<i>GSDMB</i>	272	-3.305	9.5×10 ⁻⁴	278	-3.330	8.7×10⁻⁴	244	-3.238	1.2×10 ⁻³
rs1342326 T/G	<i>IL33</i>	176	+1.114	0.265	180	+1.266	0.205	156	+1.412	0.158
rs3939286 G/A	<i>IL33</i>	198	+0.412	0.680	204	+0.544	0.586	176	+0.693	0.489
rs3771166 C/T	<i>IL18R1</i>	263	-1.475	0.140	270	-1.637	0.102	228	-1.416	0.157
rs1420101 G/A	<i>IL1RL1</i>	261	+0.899	0.369	264	+0.658	0.511	223	+0.737	0.461
rs744910 A/G	<i>SMAD3</i>	264	+0.116	0.907	271	+0.823	0.411	234	-0.166	0.868
rs2284033 A/G	<i>IL2RB</i>	262	+0.825	0.409	267	+0.416	0.677	233	+0.778	0.436
rs1588265 A/G	<i>PDE4D</i>	233	-0.380	0.703	239	-0.056	0.956	205	-0.506	0.613
rs2786098 C/A	<i>CRB1</i>	202	-1.702	0.089	204	-1.324	0.185	178	-2.377	0.017
rs2244012 A/G	<i>RAD50</i>	187	+0.953	0.340	189	+0.659	0.510	164	+1.075	0.282

FBAT was used to test for association using the additive model [17]. SNP single nucleotide polymorphism, MAF minor allele frequency, Fam number of families in analysis (≥10 families). Asthma = Doctor diagnosed asthma, FEV₁ = FEV₁ percent predicted, BHR = PC₂₀ (methacholine 4 mg/ml) only available in Southampton Families, Bonferroni correction, $P < 9.1 \times 10^{-4}$ considered significant (bold).

Table 4 Association of key GWAS SNPs with asthma severity defined by BTS steps and atopy

SNP	Gene	Severity		P value	SPT		P value
		Fam	Score		Fam	Score	
rs2305480 C/T	<i>GSDMB</i>	272	-3.791	1.5×10⁻⁴	220	-3.164	1.6×10 ⁻³
rs11078927 C/T	<i>GSDMB</i>	277	-3.701	2.2 ×10⁻⁴	227	-3.133	1.7×10 ⁻³
rs1342326 T/G	<i>IL33</i>	180	+1.378	0.168	144	+2.585	9.7×10 ⁻³
rs3939286 G/A	<i>IL33</i>	204	+0.347	0.729	164	+2.146	0.032
rs3771166 C/T	<i>IL18R1</i>	269	-2.021	0.043	219	-1.683	0.092
rs1420101 G/A	<i>IL1RL1</i>	263	+1.577	0.115	218	+0.449	0.653
rs744910 A/G	<i>SMAD3</i>	268	+0.386	0.699	215	+0.157	0.875
rs2284033 A/G	<i>IL2RB</i>	268	+1.233	0.218	216	+1.265	0.206
rs1588265 A/G	<i>PDE4D</i>	235	-0.712	0.477	195	-0.360	0.719
rs2786098 C/A	<i>CRB1</i>	202	-1.021	0.307	158	-1.886	0.059
rs2244012 A/G	<i>RAD50</i>	188	+0.941	0.347	159	+1.242	0.214

FBAT was used to test for association using the additive model [17]. SNP single nucleotide polymorphism, MAF minor allele frequency, Fam number of families in analysis (≥10 families). Severity = BTS classification (1–5) [18], SPT = positive skin prick test to one or more allergen. Bonferroni correction, $P < 9.1 \times 10^{-4}$ considered significant (bold).

conservation score). Three SNPs were located in enhancer histone marks (rs12709365 and rs13380815 both *ZBP2*/intronic and identified in cell line GM12878), rs4795399 (*GSDMB*/intronic) was identified in a HepG2 cell line and also located in a DNase site identified in a human hepatocyte cell line (Huh7.5) and a human leukaemic cell line (CMK). While no SNPs were predicted to occur in a protein binding regions, 14 SNPs were predicted to alter regulatory binding motifs (Table 6).

Analysis of asthma eQTL data

We searched for the eleven SNPs genotyped as part of this study in turn using the mRNA by SNP browser resource, looking to see whether the SNP was associated with the effect meeting genome-wide significance ($p < 10^{-8}$). Information for three of the SNPs (rs230548/*GSDMB*, rs2886098/*CRB1* and rs2244012/*RAD50*) was not covered by the database and there was no information for SNPs in LD with these. Another two SNPs were also not covered by the database, however there was information for a SNP in high LD. For rs11078927/*GSDMB*, rs10445308 was in LD ($r^2 = 1.0$) and this SNP showed genome-wide significance with multiple *ORMDL3* probes (all C-allele); 223259_at ($p = 4.4 \times 10^{-13}$, LOD = 11.39, $H^2 = 14.75$) and probe 235136_at ($p = 8.0 \times 10^{-12}$, LOD = 10.16, $H^2 = 13.3$) with average *ORMDL3* ($p = 3.9 \times 10^{-22}$, LOD = 20.32, $H^2 = 26.36$). For rs3771166/*IL18R1*, rs54988956 was in LD ($r^2 = 0.96$), but this was not associated with genome-wide significance with any probe. For the other six SNPs, there was no genome-wide significance. Using the University of Chicago eQTL resource, data was available for the two *GSDMB* SNPs (rs2305480 and rs11078927) which provided some evidence that these were eQTLs and exon-QTLs in lymphoblastoid cell lines, monocytes, liver, brain

and T-cells. More specifically, both rs2305480 and rs11078927 were eQTL for *IKZF3*, *ORMDL3* and an exon-QTL for *KRT222P*, *NR1D1* and *ORMDL3*. Only three of our eleven genotyped SNPs (rs2305480/*GSDMB*, rs3939286/*IL33* and rs744910/*SMAD3*) were included in the genotyping platform for the genome-wide search for eQTL in 1,111 human lung samples [21]. None of these SNPs were found in the reported data meeting 10% False Discovery Rate (FDR). As a further analysis we also searched this resource for any of the SNPs in high LD ($r^2 = 1$) with our genotyped SNPs. rs1337167/*CRB1* met the 10% FDR. This intron SNP was predicted to alter an AP-1 transcription factor binding site according to HaploReg.

Discussion

GWAS have been successful in identifying susceptibility genes for asthma through the use of large scale population cohorts and replication approaches [2-7]. However, there remains a need to determine the contribution of these polymorphisms to clinically relevant endpoints in asthma as opposed to disease diagnosis. We adopted a two step process where we selected SNPs from key genes in the Caucasian population which have met genome-wide significance and/or with replication, namely eleven SNPs in the following genes; *ORMDL3/GSDMB*, *IL33*, *IL18R/IL1RL1*, *SMAD3*, *IL2RB*, *PDE4D*, *CRB1* and *RAD50*. These SNPs were genotyped in a UK asthma family cohort with at least two affected siblings, to test for association with clinically relevant endpoints in asthma. Our data provide support for the original *GSDMB* association; however SNPs in the other genes did not show association after correction for multiple testing. Secondly, as these SNPs have previously shown association at a genome-wide level, we also investigated the potential for these SNPs and

Table 5 Analysis of key GWAS SNPs with data from ENCODE

SNP	Gene locus/ location	Number of SNPs in LD ($r^2 = 1$) (no.)	GERP Conservation (no. of SNPs conserved)	Enhancer histone marks (no. of SNPs in regions)	DNase (no. of SNPs in regions)	Proteins bound (no. of SNPs in regions)	Motifs changed (no. of SNPs with motif changed regions)
rs2305480 C/T	<i>GSDMB</i> /Ser311Pro	26	2	3	1	0	14
rs11078927 C/T	<i>GSDMB</i> /intronic	26	2	3	1	0	14
rs1342326 T/G	<i>IL33</i> /5'-region	7	0	1	2	2	4
rs3939286 G/A	<i>IL33</i> /5'-region	1	0	0	0	0	1
rs3771166 C/T	<i>IL18R1</i> /intronic	16	0	2	1	0	9
rs1420101 G/A	<i>IL1RL1</i> /intronic	6	1	2	1	1	1
rs744910 A/G	<i>SMAD3</i> /intronic	3	0	1	2	2	1
rs2284033 A/G	<i>IL2RB</i> /intronic	3	0	0	3	0	1
rs1588265 A/G	<i>PDE4D</i> /intronic	14	2	0	1	0	5
rs2786098 C/A	<i>CRB1</i> /intronic	11	0	0	3	0	7
rs2244012 A/G	<i>RAD50</i> /intronic	51	2	6	9	5	25

The 11 SNPs genotyped as part of this study were entered into the HaploReg tool to investigate whether SNPs in the same LD block were likely to affect chromatin structure and regulatory element binding motifs.

Table 6 Analysis of the *GSDMB* LD block ($r^2 = 1$) using data from ENCODE

SNP	Gene locus/location	Chromosomal location (hg19)	GERP conservation (Y/N)	Enhancer histone marks	DNase	Proteins bound	Motifs changed
rs12709365 A/G	<i>ZPBP2</i> /intronic	17:38027400	N	GM12878	N	N	N
rs13380815 A/G	<i>ZPBP2</i> /intronic	17:38027583	N	GM12878	N	N	N
rs11870965 T/A	<i>ZPBP2</i> /intronic	17:38030205	N	N	N	N	Myb
rs10852935 C/T	<i>ZPBP2</i> /synonymous	17:38031674	Y	N	N	N	N
rs10852936 C/T	<i>ZPBP2</i> /intronic	17:38031714	N	N	N	N	Foxp1, Foxq1
rs36095411 T/G	<i>ZPBP2</i> /intronic	17:38031865	N	N	N	N	9 motifs altered
rs34189114 C/T	<i>ZPBP2</i> /intronic	17:38032460	N	N	N	N	N
rs35736272 T/C	<i>ZPBP2</i> /intronic	17:38032680	Y	N	N	N	N
rs9907088 G/A	<i>ZPBP2</i> /downstream	17:38035116	N	N	N	N	Sox
rs36038753 G/T	<i>ZPBP2</i> /downstream	17:38035370	N	N	N	N	Dobox4
rs35569035 C/T	<i>ZPBP2</i> /downstream	17:38036524	N	N	N	N	Pax4
rs9904624 A/G	<i>ZPBP2</i> /downstream	17:38036586	N	N	N	N	N
rs12232497 T/C	Intergenic	17:38040119	N	N	N	N	N
rs12232498 T/C	Intergenic	17:38040363	N	N	N	N	Lhx3
rs12941333 C/T	Intergenic	17:38040534	N	N	N	N	Pou3f3, Pou2f1
rs2872507 G/A	Intergenic	17:38040763	N	N	N	N	N
rs12936409 C/T	Intergenic	17:38043649	N	N	N	N	N
rs8069176 G/A	<i>GSDMB</i> /downstream	17:38057197	N	N	N	N	Gfi1, Mafk, Pax4
rs4795399 T/C	<i>GSDMB</i> /intronic	17:38061439	N	HepG2	Huh7.5,CMK	N	Pxx6, Pax
rs2305480 G/A	<i>GSDMB/Ser311Pro</i>	17:38062196	N	N	N	N	N
rs62067034 C/T	<i>GSDMB</i> /intronic	17:38063738	N	N	N	N	SREBP, LXR
rs11078927 C/T	<i>GSDMB</i>/intronic	17:38064405	N	N	N	N	N
rs11078928 T/C	<i>GSDMB</i> /intronic	17:38064469	N	N	N	N	Foxp1
rs4795400 C/T	<i>GSDMB</i> /intronic	17:38067020	N	N	N	N	N
rs77749396 G/C	<i>GSDMB</i> /intronic	17:38073840	N	N	N	N	ZNF219
rs9303279 G/C	<i>GSDMB</i> /intronic	17:38073968	N	N	N	N	Obox3

The two *GSDMB* SNPs genotyped as part of this study (rs2305480 and rs11078927) were part of a wider LD block containing 26 SNPs in total spanning 46,568 bp. These SNPs were analysed using HaploReg to see if they altered any binding sites for regulatory elements or other proteins and also whether they affected chromatin states and their conservation according to GERP. The CEU frequency from 1000 Genomes Project =0.48 for all SNPs. CMK: human leukaemic cell line; *GERP* genomic evolutionary rate profiling score; GM12878: lymphoblastoid cell line; *GSDMB* gasdermin-domain containing protein family B; HepG2: human hepatocellular liver carcinoma cell line; Huh7.5: human hepatocellular cell line; *ZPBP2* zona pellucid binding protein 2 gene. SNPs genotyped in the current study are shown in bold.

those in LD to affect chromatin states and alter regulatory motifs/binding sites to try to assign potential function to these associations. This analysis used recently reported data from the ENCODE project. These data support the hypothesis that SNPs associated in GWAS may be tagging SNPs for the actual causative variant, but also provide tentative evidence that these genotyped SNPs may themselves play a functional role e.g. *GSDMB* rs2305480 (Ser311Pro).

GWAS have the advantage over linkage studies in that they can identify polymorphisms with small effect sizes and localise smaller susceptibility loci as LD only generally spans <500 kb [22]. Since the identification of childhood asthma associated variants on chromosome 17q21 [2], many other GWAS have been completed in different populations implicating variants spanning different genes (see

above). One problem with these studies lies in defining the contribution of these associated variants to clinical endpoints in asthma. To this end we have completed a SNP for SNP analysis of key variants identified in the Caucasian population meeting genome-wide significance or where there was replication, focussing on GWAS from 2007–2011 (Table 1). We utilised a well characterised UK family based asthma cohort [15] to test for association with asthma diagnosis and the following clinical endpoints: FEV₁ (% predicted), BHR (to methacholine), BTS defined severity and positive SPT to one or more allergens, see [15]. These data show significant association for the two *GSDMB* SNPs; rs2305480 (Ser311Pro) and rs11078927 (intron) across more than one phenotype, with other phenotypes showing association which does not reach significance

after correction for multiple testing. These data both replicate and extend previous findings [2,6] and identify this chromosome 17 locus as containing genetic determinants underlying multiple clinical features of asthma, including FEV₁ (% Predicted) and disease severity defined by BTS. These two polymorphisms are in high LD suggesting these associations represent a single locus. The BHR association (rs2305480) is supported by different SNPs in this gene also showing association with this phenotype, albeit in different ethnic populations [23,24]. Interestingly, a recent study has found association between rs2305480 and wheezing phenotypes in asthmatic children, but did not find association with intermediate phenotypes such as atopy or lung function [25].

When interpreting the results of GWAS, the associated variant could be tagging another variant in a larger region of LD [26]. This is particularly important as it is common that only a few variants in a haplotype block are present on genotyping platforms. Methods such as imputation and using genotyping data from the 1000 Genomes project can help to further characterise these patterns [27]. However, once these variants are identified it is still necessary to assign potential function and with many associated variants from GWAS residing in non-transcribed regions, they are likely to play a regulatory role in gene expression. We used HaploReg [14] to screen chromatin data, conservation data and regulatory motif databases which was obtained from the ENCODE project to see if any biological plausible role could be assigned to the SNPs genotyped in this study and those SNPs which were in LD with $r^2 = 1$. Our results show that these SNPs in LD can potentially lead to changes in regulatory element binding motifs affecting the transcription of genes positively or negatively and there is also the suggestion that alterations in enhancer histone marks or DNase sites could be affected. The non-synonymous SNP (Ser 311 Pro) is predicted to be deleterious to protein structure (Polyphen) in Gasdermin B suggesting a potential functional mechanism altering protein function. This may be occurring at the mRNA level (altering the efficiency of splicing) as this polymorphism is predicted to be located in an exonic splicing enhancer site (FuncPred). Both *GSDMB* SNPs are in high LD with several SNPs spanning *ZBPBZ* and *GSDMB* which result in alterations in enhancer and transcription factor binding sites. However the rs2305480 and rs11078927 SNPs themselves are not predicted to have any regulatory role (ENCODE data). Gasdermin-family proteins have been implicated in TGF β 1 signalling and epithelial cell apoptosis [28], with increased airway cell apoptosis reported in severe asthma [29]. The extended chromosome 17 locus has also been implicated previously as containing SNPs associated with asthma spanning three main genes; *ZBPBZ/GSDMB/ORMDL3*

and it is still unclear which genes/polymorphisms underlie these associations. The finding that the *GSDMB* SNPs including a non-synonymous SNP are particularly relevant to clinical endpoints including lung function, BHR and severity in our childhood asthma cohort is interesting as few recent studies have genotyped this GWAS SNP or investigated clinical outcomes with the same criteria such as BTS used in this study [24,30]. However, a recent study has investigated the effect of rs7216389 (which is in linkage disequilibrium with both *GSDMB* SNPs in this study) with respect to severe asthma subjects classified with poorly controlled disease whilst taking high doses of inhaled corticosteroids, long-acting bronchodilators and short acting β_2 agonists and showed association [31]. Similarly, while not meeting genome-wide association significance, we have previously provided evidence that rs2305480 is associated with severe asthma ($p = 5.5 \times 10^{-5}$, [32]), although we do acknowledge that the criteria for BTS severity is defined in part by lung function and so this association with severe asthma may be related to a synergistic effect. These data support previous studies that identified the chromosome 17 locus as containing markers of childhood onset disease [2,6,24,30], BHR [24] and elevated total serum IgE [24] in diverse ethnic populations.

Interestingly, no other SNP survived correction for multiple testing for any phenotype analysed, including asthma diagnosis and there was no indication of any other SNP showing a trend towards significance. *CRB1* rs2786098 ($p = 0.089$ in the current study) has shown replicated association in multiple childhood cohorts, but did not reach statistical significance in our study. Perhaps this lack of association may be due to power (MAF 0.22, see earlier). One potential explanation for the lack of association for asthma and related phenotypes for the majority of SNPs is that many of the SNPs/genes were identified in adult asthma populations and the relative contribution to childhood asthma remains unclear at this time. However, in the results not meeting conventional correction for multiple testing, *IL33* SNP rs1342326G (25.7 kb from gene transcription start site (TSS)) showed a level of association with atopic asthma defined by positive skin prick test and clinical diagnosis ($p = 0.0097$, Table 4). These data are in line with the accumulating evidence suggesting a role of IL33 in allergic mechanisms in asthma (Reviewed in [33]) although we cannot exclude that this is a false positive. Similarly, while not surviving correction, suggestive evidence for SNP associations for *CRB1*/BHR, *IL18R*/severity and *IL33*/SPT were apparent ($p < 0.05$), which may represent true associations.

Studies have shown the potential for GWAS SNPs to be tagging SNPs associated with eQTL [10]. Our analyses did provide some support for this particularly for the two *GSDMB* SNPs (rs2305480 and rs11078927) when eQTL

analyses was completed across multiple cell types and tissues. These eQTL analyses identified regulation of multiple genes in the region, demonstrating the need for further work to define the biology underlying these genetic signals.

It is important to note the limitations of our study, while a SNP for SNP approach focussed to the pivotal SNPs from the GWAS has strengths to directly replicate and extend previous findings we acknowledge the assumption that the linkage disequilibrium pattern between the discovery Caucasian population and our UK population will be the same, which may at least in part explain the inability to detect associations with all variants. Similarly, we did not genotype all GWA significant SNPs identified in each of the regions. We acknowledge the reduced power of the current study using 370 families is a limitation potentially explaining while several SNPs showed nominal significance and did not meet stringent statistical criteria. A cohort of 602 families would be required to provide a power of 0.99 for the lowest MAF observed in the current study (relative risk of 1.5 and significance level of 0.05). We also acknowledge that the design of this study compared to discovery cohorts using case/control design also has reduced power, however we consider the use of an alternative association design i.e. family based to those previously published a strength. Furthermore the lack of assessing gene-environment interactions in GWAS and the current study may also account for not replicating previous observed effects due to differential environmental exposures between study populations. Finally, the potential for winners curse bias is also possible for direct replication of e.g. association with asthma diagnosis.

Conclusions

In summary, this study has provided a greater insight into the relative contribution of asthma GWAS SNPs to clinically relevant endpoints in asthmatic children and further defines the diverse role of *GSDMB* SNPs particularly in lung function and disease severity.

Additional file

Additional file 1: Power of the study based on different relative risks for minor allele frequencies of 0.17 and 0.48. Cohort size n = 370 families and p-value of 0.05. Calculation based on the method of Risch & Merikangas.

Abbreviations

BHR: Bronchial hyperresponsiveness; BTS: British Thoracic society; CRB1: Crumbs homolog 1 precursor; DENND1B: Denn/madd domain-containing 1B; ENCODE: Encyclopedia Of DNA elements; eQTL: Expression quantitative trait loci; FBAT: Family based association test; FEV₁ (% Predicted): forced expiratory volume in one second (percent predicted); *GSDMB*: Gasdermin B; GWA: Genome-wide association; GWAS: Genome-wide association study; IgE: Immunoglobulin E; IL18R: Interleukin 18 receptor; IL1RL1: Interleukin 1 receptor like 1; *IL2RB* IL2: Receptor beta; IL33: Interleukin 33; LD: Linkage

disequilibrium; MAF: Minor allele frequency; ORMDL3: Orm1-like protein 3; PDE4D: Phosphodiesterase 4D, cAMP specific; Pro: Proline (amino acid); *RAD50* S. Cerevisiae: homolog of (*RAD50*); SPT: Skin prick test; SMAD3: Mothers against decapentaplegic drosophila homolog 3; SNP: Single nucleotide polymorphism; Ser: Serine (amino acid); TGFβ1: Transforming growth factor beta 1; TSS: Transcription start site; ZBP2: Zona pellucid binding protein 2 gene.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

IS, JWH and AST designed the study and drafted the manuscript. AST completed the analyses. All authors contributed to the final version of the manuscript. All authors read and approved the final manuscript.

Acknowledgements

This work was funded by the University of Nottingham.

Author details

¹Division of Respiratory Medicine, Queen's Medical Centre, University of Nottingham, Nottingham NG7 2UH, United Kingdom. ²Human Genetics and Medical Genomics, Human Development and Health, Faculty of Medicine, University of Southampton, Southampton, UK. ³Institute of Cellular Medicine, Faculty of Medical Sciences, Newcastle University, Newcastle upon Tyne, UK.

Received: 14 June 2013 Accepted: 24 September 2013

Published: 25 September 2013

References

1. Yang IA, Savarimuthu S, Kim ST, Holloway JW, Bell SC, Fong KM: **Gene-environmental interaction in asthma.** *Curr Opin Allergy Clin Immunol* 2007, **7**(1):75–82.
2. Moffatt MF, Kabesch M, Liang L, Dixon AL, Strachan D, Heath S, Depner M, von Berg A, Bufe A, Rietschel E, et al: **Genetic variants regulating ORMDL3 expression contribute to the risk of childhood asthma.** *Nature* 2007, **448**(7152):470–473.
3. Gudbjartsson DF, Bjornsdottir US, Halapi E, Helgadóttir A, Sulem P, Jonsdóttir GM, Thorleifsson G, Helgadóttir H, Steinthorsdóttir V, Stefansson H, et al: **Sequence variants affecting eosinophil numbers associate with asthma and myocardial infarction.** *Nat Genet* 2009, **41**(3):342–347.
4. Himes BE, Hunninghake GM, Baurley JW, Rafaels NM, Sleiman P, Strachan DP, Wilk JB, Willis-Owen SA, Klanderman B, Lasky-Su J, et al: **Genome-wide association analysis identifies PDE4D as an asthma-susceptibility gene.** *Am J Hum Genet* 2009, **84**(5):581–593.
5. Li X, Howard TD, Zheng SL, Haselkorn T, Peters SP, Meyers DA, Bleeker ER: **Genome-wide association study of asthma identifies RAD50-IL13 and HLA-DR/DQ regions.** *J Allergy Clin Immunol* 2010, **125**(2):328–335 e311.
6. Moffatt MF, Gut IG, Demenais F, Strachan DP, Bouzigon E, Heath S, von Mutius E, Farrall M, Lathrop M, Cookson WO: **A large-scale, consortium-based genome-wide association study of asthma.** *N Engl J Med* 2010, **363**(13):1211–1221.
7. Sleiman PM, Flory J, Imielinski M, Bradfield JP, Annaiah K, Willis-Owen SA, Wang K, Rafaels NM, Michel S, Bonnelykke K, et al: **Variants of DENND1B associated with asthma in children.** *N Engl J Med* 2010, **362**(1):36–44.
8. Birney E, Lieb JD, Furey TS, Crawford GE, Iyer VR: **Allele-specific and heritable chromatin signatures in humans.** *Hum Mol Genet* 2010, **19**(R2):R204–209.
9. Cheung VG, Spielman RS: **Genetics of human gene expression: mapping DNA variants that influence gene expression.** *Nat Rev Genet* 2009, **10**(9):595–604.
10. Schaub MA, Boyle AP, Kundaje A, Batzoglou S, Snyder M: **Linking disease associations with regulatory information in the human genome.** *Genome Res* 2012, **22**(9):1748–1759.
11. EncodeProjectConsortium: **The ENCODE (Encyclopedia Of DNA Elements) Project.** *Science* 2004, **306**(5696):636–640.
12. Dunham I, Kundaje A, Aldred SF, Collins PJ, Davis CA, Doyle F, Epstein CB, Frietze S, Harrow J, Kaul R, et al: **An integrated encyclopedia of DNA elements in the human genome.** *Nature* 2012, **489**(7414):57–74.
13. 1000GenomesProjectConsortium: **A map of human genome variation from population-scale sequencing.** *Nature* 2010, **467**(7319):1061–1073.

14. Ward LD, Kellis M: **HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants.** *Nucleic Acids Res* 2011, **40**(Database issue):D930–934.
15. Barton SJ, Koppelman GH, Vonk JM, Browning CA, Nolte IM, Stewart CE, Bainbridge S, Mutch S, Rose-Zerilli MJ, Postma DS, *et al*: **PLAUR polymorphisms are associated with asthma, PLAUR levels, and lung function decline.** *J Allergy Clin Immunol* 2009, **123**(6):1391–1400 e1317.
16. BTS: **British guidelines on the management of asthma.** *Thorax* 2008, **63**(Suppl 4):iv1–121.
17. Horvath S, Xu X, Laird NM: **The family based association test method: strategies for studying general genotype–phenotype associations.** *Eur J Hum Genet* 2001, **9**(4):301–306.
18. BTS: *British Guideline on the Management of Asthma.* London: British Thoracic Society; 2009.
19. Risch N, Merikangas K: **The future of genetic studies of complex human diseases.** *Science* 1996, **273**(5281):1516–1517.
20. Dixon AL, Liang L, Moffatt MF, Chen W, Heath S, Wong KC, Taylor J, Burnett E, Gut I, Farrall M, *et al*: **A genome-wide association study of global gene expression.** *Nat Genet* 2007, **39**(10):1202–1207.
21. Hao K, Bosse Y, Nickle DC, Pare PD, Postma DS, Laviolette M, Sandford A, Hackett TL, Daley D, Hogg JC, *et al*: **Lung eQTLs to help reveal the molecular underpinnings of asthma.** *PLoS Genet* 2012, **8**(11):e1003029.
22. Hirschhorn JN, Daly MJ: **Genome-wide association studies for common diseases and complex traits.** *Nat Rev Genet* 2005, **6**(2):95–108.
23. Kang MJ, Yu HS, Seo JH, Kim HY, Jung YH, Kim YJ, Kim HJ, Lee SY, Hong SJ: **GSDMB/ORMDL3 variants contribute to asthma susceptibility and eosinophil-mediated bronchial hyperresponsiveness.** *Hum Immunol* 2012, **73**(9):954–959.
24. Yu J, Kang MJ, Kim BJ, Kwon JW, Song YH, Choi WA, Shin YJ, Hong SJ: **Polymorphisms in GSDMA and GSDMB are associated with asthma susceptibility, atopy and BHR.** *Pediatr Pulmonol* 2011, **46**(7):701–708.
25. Granell R, Henderson AJ, Timpson N, St Pourcain B, Kemp JP, Ring SM, Ho K, Montgomery SB, Dermitzakis ET, Evans DM, *et al*: **Examination of the relationship between variation at 17q21 and childhood wheeze phenotypes.** *J Allergy Clin Immunol* 2013, **131**(3):685–694.
26. HapMap: **A haplotype map of the human genome.** *Nature* 2005, **437**(7063):1299–1320.
27. Browning SR: **Missing data imputation and haplotype phase inference for genome-wide association studies.** *Hum Genet* 2008, **124**(5):439–450.
28. Saeki N, Kim DH, Usui T, Aoyagi K, Tatsuta T, Aoki K, Yanagihara K, Tamura M, Mizushima H, Sakamoto H, *et al*: **GASDERMIN, suppressed frequently in gastric cancer, is a target of LMO1 in TGF-beta-dependent apoptotic signalling.** *Oncogene* 2007, **26**(45):6488–6498.
29. Cohen L, X E, Tarsi J, Ramkumar T, Horiuchi TK, Cochran R, DeMartino S, Schechtman KB, Hussain I, Holtzman MJ, *et al*: **Epithelial cell proliferation contributes to airway remodeling in severe asthma.** *Am J Respir Crit Care Med* 2007, **176**(2):138–145.
30. Wu H, Romieu I, Sienra-Monge JJ, Li H, del Rio-Navarro BE, London SJ: **Genetic variation in ORM1-like 3 (ORMDL3) and gasdermin-like (GSDML) and childhood asthma.** *Allergy* 2009, **64**(4):629–635.
31. Binia A, Khorasani N, Bhavsar PK, Adcock I, Brightling CE, Chung KF, Cookson WO, Moffatt MF: **Chromosome 17q21 SNP and severe asthma.** *J Hum Genet* 2011, **56**(1):97–98.
32. Wan YI, Shrine NR, Soler Artigas M, Wain LV, Blakey JD, Moffatt MF, Bush A, Chung KF, Cookson WO, Strachan DP, *et al*: **Genome-wide association study to identify genetic determinants of severe asthma.** *Thorax* 2012, **67**(9):762–768.
33. Borish L, Steinke JW: **Interleukin-33 in asthma: how big of a role does it play?** *Curr Allergy Asthma Rep* 2011, **11**(1):7–11.

doi:10.1186/1471-2350-14-100

Cite this article as: Tulah *et al*: Defining the contribution of SNPs identified in asthma GWAS to clinical variables in asthmatic children. *BMC Medical Genetics* 2013 **14**:100.

Submit your next manuscript to BioMed Central and take full advantage of:

- **Convenient online submission**
- **Thorough peer review**
- **No space constraints or color figure charges**
- **Immediate publication on acceptance**
- **Inclusion in PubMed, CAS, Scopus and Google Scholar**
- **Research which is freely available for redistribution**

Submit your manuscript at
www.biomedcentral.com/submit

