npg

ARTICLE

Association between *ORMDL3*, *IL1RL1* and a deletion on chromosome 17q21 with asthma risk in Australia

Manuel AR Ferreira*,1, Allan F McRae[1], Sarah E Medland[1], Dale R Nyholt[1], Scott D Gordon[1],
Margaret J Wright[1], Anjali K Henders[1], Pamela A Madden[2], Peter M Visscher[1], Naomi R Wray[1],
Andrew C Heath[2], Grant W Montgomery[1], David L Duffy[1] and Nicholas G Martin[1]

Genome-wide association studies followed by replication provide a powerful approach to map genetic risk factors for asthma. We sought to search for new variants associated with asthma and attempt to replicate the association with four loci reported previously (*ORMDL3*, *PDE4D*, *DENND1B* and *IL1RL1*). Genome-wide association analyses of individual single nucleotide polymorphisms (SNPs), rare copy number variants (CNVs) and overall CNV burden were carried out in 986 asthma cases and 1846 asthma-free controls from Australia. The most-associated locus in the SNP analysis was *ORMDL3* (rs6503525, $P=4.8\times10^{-7}$). Five other loci were associated with $P<10^{-5}$, most notably the chemokine CXC motif ligand 14 (*CXCL14*) gene (rs31263, $P=7.8\times10^{-6}$). We found no evidence for association with the specific risk variants reported recently for *PDE4D*, *DENND1B* and *ILR1L1*. However, a variant in *IL1RL1* that is in low linkage disequilibrium with that reported previously was associated with asthma risk after accounting for all variants tested (rs10197862, gene wide $P=0.01$). This association replicated convincingly in an independent cohort ($P=2.4\times10^{-4}$). A 300-kb deletion on chromosome 17q21 was associated with asthma risk, but this did not reach experiment-wide significance. Asthma cases and controls had comparable CNV rates, length and number of genes affected by deletions or duplications. In conclusion, we confirm the association between asthma risk and variants in *ORMDL3* and identify a novel risk variant in *IL1RL1*. Follow-up of the 17q21 deletion in larger cohorts is warranted.
European Journal of Human Genetics advance online publication, 8 December 2010; doi:10.1038/ejhg.2010.191

INTRODUCTION

Asthma currently affects 10% of the Australian population and is responsible for 1 million work days lost, 36 000 hospital admissions, 402 deaths and $606 million a year in direct costs.[1] A complete understanding of the genetic and environmental risk factors for asthma is critical to design new treatments or prevention strategies that can reduce the disease burden. This is a challenging task, as asthma is a complex disease that involves the interplay between multiple physiological processes.[2]

Genome-wide association studies (GWAS) currently provide the most powerful approach to identify genetic risk factors for complex diseases. In the past few years, hundreds of variants with replicated associations have been identified for a range of complex diseases or traits using this approach.[3,4] As hoped, several of these discoveries have identified aetiologic pathways not previously implicated in these traits, such as the autophagy pathway in Crohn's disease[5] and the HLA-C locus in control of viral load in HIV infection.[6]

Seven GWAS of asthma have been published recently.[7–13] From these, three loci have emerged with reproducible effects on asthma risk in independent samples: the *GSDMB/ORMDL3* locus (henceforth referred to as *ORMDL3*) on chromosome 17q12-21;[7] the *PDE4D* gene on chromosome 5q12[9] and the *DENND1B* on chromosome 1q31.[11] Three GWAS of asthma biomarkers were also published recently.[14–16] Quantitative trait loci were identified for the respective phenotype analysed in each study, but for only one of these there

was also a convincing association with asthma, the *IL1RL1* gene on chromosome 2q12.[16] A variant at this locus was a significant predictor of eosinophil levels while also reproducibly increasing asthma risk.

Although single nucleotide polymorphisms (SNPs) are now routinely tested for association in GWAS for many common diseases, structural variants, a widespread class of genetic polymorphisms that range in size from only a few base pairs to whole chromosomal rearrangements,[17] remain much less well studied. The Welcome Trust Case Control Consortium recently performed a GWAS of common copy number variants (CNVs) for eight common diseases, but failed to identify any new risk loci that had not been previously reported through the analysis of SNPs.[18] These results suggest that common CNVs are unlikely to contribute greatly to the genetic basis of common diseases, but leave open the possibility that rarer structural variants may have a greater impact on risk, as shown for obesity,[19] autism[20] and schizophrenia.[21] The only published genome-wide association analysis between CNV data and asthma identified a region on chromosome 7p14 that appeared to be associated with asthma risk, although closer inspection of the region could not rule out the possibility that this was a false-positive association as a result of a biological artefact.[22]

Therefore, the identified variants in *ORMDL3*, *PDE4D*, *DENND1B* and *IL1RL1* arguably represent the most convincing associations between genetic variants and asthma risk reported to date. In this

[1]Queensland Institute of Medical Research, Brisbane, QLD, Australia; [2]Washington University School of Medicine, St Louis, MO, USA
*Correspondence: Dr MAR Ferreira, Genetic Epidemiology Laboratory, Queensland Institute of Medical Research, P.O. Royal Brisbane Hospital, Brisbane, QLD 4029, Australia.
Tel: +61 7 3845 3552; Fax: +61 7 3362 0101; E-mail: manuel.ferreira@qimr.edu.au

study, we analysed whole-genome genotype data from 986 asthma cases and 1846 disease-free controls of European descent from the Australian population to (1) search for new sequence or structural variants with strong effects on asthma and (2) confirm the previously reported associations with *ORMDL3*, *PDE4D*, *DENND1B* and *IL1RL1*.

MATERIALS AND METHODS
Case and control samples
The samples analysed herein are a subset of a larger cohort of 16 140 individuals ascertained in several waves from the Australian general population and genotyped recently as described in detail elsewhere.[23] Of these, we selected for analysis 986 unrelated cases and 1846 controls based on the presence (cases) or absence (controls) of self-reported lifetime asthma, as reported in at least one lifestyle or health questionnaire completed as part of six epidemiological studies conducted at the Queensland Institute of Medical Research (QIMR) (Supplementary Table 1). Using SOLAR,[24] we estimate that this phenotype is 61% heritable, based on a polygenic model fitted to 2054 individuals for whom phenotype information was also available for 3154 of their relatives (parents and/or siblings).

Whole-genome genotyping and imputation of HapMap SNPs
Genotype data were obtained with Illumina 317K, 370K or 610K arrays and stringent quality control (QC) procedures applied to ensure high data quality as described in detail previously[23] and summarised in Supplementary Table 2. After QC, genotype data were available for 16 140 individuals and 561 815 SNPs. To increase whole-genome coverage, we inferred unmeasured HapMap SNPs using MACH (http://www.sph.umich.edu/csg/abecasis/MACH/) and the phased haplotype data from the CEU HapMap samples (phase I+II, release22, build 36). Following imputation, we restricted our analysis to 2.38 million SNPs that were genotyped or imputed with high confidence ($r^2 > 0.3$) and had an MAF > 0.01, Hardy–Weinberg equilibrium test *P*-value > 10^{-6} in controls and < 5% missing data in the selected sample of 2832 individuals with asthma information. Individuals were confirmed as unrelated through pair wise whole-genome identity-by-descent analysis.

Population substructure and whole-genome SNP association analyses
All analyses were performed with PLINK.[25] To address whether subtle population substructure could have an impact on the association results, we performed a multidimensional scaling (MDS) analysis of identity-by-state (IBS) distances calculated between all pairs of individuals using a subset of SNPs in linkage equilibrium. For reference, we included the 11 HapMap III populations.[26] To test whether cases and controls were well matched with respect to ancestry, we applied a permutation test for between-group IBS differences. A significant test would imply that cases and controls are unlikely to belong to the same population, which could lead to spurious association results. The main association analysis was performed using a standard allelic test. Age-of-onset information was only available for 20% of cases, and so was not considered for analysis. Lifetime smoking status, which was available for 96% of individuals (38% ever smokers), was not a significant predictor of asthma in this sample (*P*=0.82) and so was not included as a covariate.

Whole-genome CNV analysis
CNV segments were identified for a subset of 759 individuals that participated in the longitudinal Brisbane Adolescent Twin Study (270 asthma cases and 489 controls; Supplementary Table 1) using the program QuantiSNP v1.1[27] based on the SNP and CNV probes present in the Illumina 610K arrays. The default program settings were used in addition to a maximum copy number of four and GC correction. To minimise false-positive CNV calling, we restricted our analysis to large (> 100 kb and < 1 Mb) deletions or duplications called with high confidence (log Bayes' factor > 10 and with > 10 probes). We also dropped common CNVs (MAF > 0.05), as these were likely well tagged by SNPs.

We performed two sets of association analyses between CNVs and asthma risk. First, we investigated whether the frequency of deletions or duplications at specific loci was associated with asthma status. This is equivalent to the genome-wide SNP analysis, except that a small number of rare structural variants are tested for association instead of a large number of common SNPs. Significance was estimated through the analysis of 10 000 permutations.

Second, we tested the hypothesis that the presence of multiple deletions or duplications across an individual's genome (ie CNV burden) may significantly increase the risk of developing asthma. Specifically, we addressed the following questions: (A) do asthma cases have on average more CNVs than controls? In those individuals having at least one CNV, do cases have on average (B) larger CNVs than controls or (C) more genes affected by CNVs than controls? For (A) and (B), the test applied was a logistic regression of disease status on the number of observed CNVs and CNV length (mean or total), respectively, while accounting for potential Illumina batch and plate effects. To determine whether more genes were affected by CNVs in cases than controls (C), we downloaded the genomic coordinates for 18 212 genes from the Mar 2006 UCSC browser (based on the longest isoform ± 50 kb) and determined for each individual how many genes in total were intersected by a CNV (gene count). The test was then a logistic regression of disease status on gene count, while accounting for an individual's number of CNVs, mean CNV length, batch and plate effects. Significance was estimated from 10 000 permutations.

Follow-up of top SNP and CNV associations
We attempted to replicate the top SNP and CNV associations in an independent cohort of 604 individuals, including 391 doctor-diagnosed asthmatics (Supplementary Table 3). These individuals were ascertained from the same QIMR studies that contributed data to the GWAS (Supplementary Table 2) and were recently typed with Illumina 610K arrays as part of an additional, as yet unpublished wave of genotyping conducted in the first semester of 2010. All samples are of European descent and passed the standard QC filters described above for the GWAS. Data from this wave were available at this stage only for replication of specific associations.

RESULTS
Population substructure analyses
We first performed an MDS analysis of IBS distances to identify subtle ancestry differences between samples that could potentially have an impact on the association results. As expected, most of the 2832 Australian samples analysed clustered closely with the CEU HapMap reference panel, indicating that most individuals have a homogeneous European ancestry (Supplementary Figure 1A). However, two groups of individuals clustered close to, but just outside this larger group of homogeneous European ancestry: one group clustered along the European–African ancestry axis, while the second group clustered along the European–Asian axis. The first group likely represents samples with a southern European ancestry (eg Maltese Australians), as indicated by the partial overlap with the HapMap samples from Tuscany in Italy (TSI) (Supplementary Figure 1B). The second group consists of samples with a predominant European ancestry that is partially mixed with Asian ancestry, as reflected by the small overlap with the HapMap Mexican ancestry samples from Los Angeles (MEX). Importantly, however, this subtle population substructure was found to be independent of case–control disease status (*P*=0.775; Supplementary Figure 1C), suggesting that it would not significantly bias whole-genome association results. Consistent with this observation, the genomic inflation factor for the main association analysis was 1.006 (Supplementary Figure 1D), thus confirming that population substructure or other technical artefacts had a minimal impact on the results.

Whole-genome SNP association results
Given that the 986 asthma cases and 1846 asthma-free controls were well matched with respect to ancestry, we then applied a standard allelic test of association between individual SNPs and lifetime disease

Table 1 Regions most associated with asthma ($P < 10^{-5}$) in the genome-wide SNP analysis

Chromosome, bp position	SNP, reference allele[a]	Closest gene, distance kb[b]	Allele frequency		Odds ratio	P-value
			Cases	Controls		
17, 35348700	rs6503525, C	*ORMDL3*, 11	0.50	0.43	1.33	4.8×10^{-7}
6, 103075599	rs12206223, C	*GRIK2*, 450	0.12	0.08	1.54	3.0×10^{-6}
7, 38792860	rs17700752, G	*VPS41*[#], 62	0.15	0.20	0.71	4.4×10^{-6}
22, 34352607	rs4622835, C	*LOC284912*, 378	0.53	0.46	1.29	4.6×10^{-6}
10, 130181728	rs7893667, A	*MKI67*, 367	0.35	0.29	1.31	6.1×10^{-6}
5, 134996861	rs31263, C	*CXCL14*, 53	0.34	0.28	1.31	7.8×10^{-6}

[a]The SNPs listed were imputed with high confidence using MACH. With the exception of rs7893667, nearby genotyped SNPs provided comparable evidence for association (illustrated in Figure 1a for the 17q12 region).
[b]Distance to first base pair of the first exon or last base pair of the last exon, whichever is smallest. A pound (#) indicates that the SNP is located within the gene.
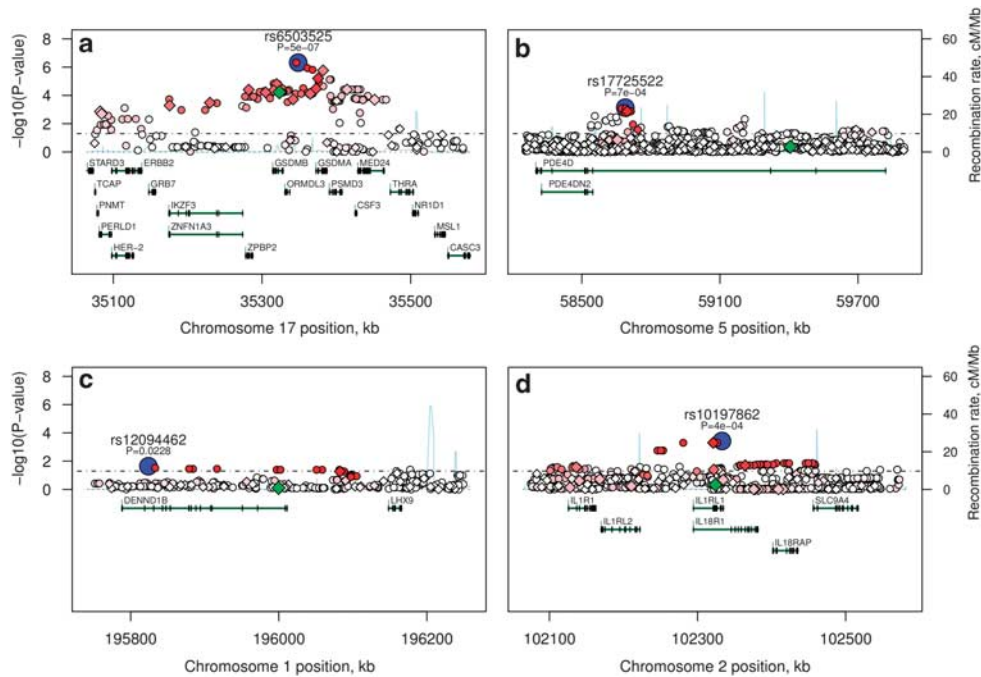
Figure 1 Association plots for the four loci previously reported to associate with asthma risk (a–d). The most-associated SNP for each region is shown in blue, and the colour of the remaining markers reflects the linkage disequilibrium (r^2) with the top SNP in each panel (increasing red hue associated with increasing r^2). The most-associated SNP reported in the original study for each gene is shown in green. Imputed SNPs are represented by circles and genotyped SNPs by diamonds. The recombination rate (second *y* axis) is plotted in light blue and is based on the CEU HapMap population. Exons for each gene are represented by vertical bars.

status to search for new susceptibility loci for asthma. We acknowledge that given the modest sample size, power was only adequate (80%) at the genome-wide level ($\alpha = 5 \times 10^{-8}$) to detect loci with strong effects (eg an OR > 1.4, for an allele frequency of 40%). The loci most associated with asthma risk ($P < 10^{-5}$) in this analysis are listed in Table 1, with the full genome-wide results displayed in Supplementary Figure 2. No variant exceeded the widely used threshold for genome-wide significance of 5×10^{-8}, but the most-associated SNP was located in the *ORMDL3* region (rs6503525, OR=1.33 for the C allele, $P = 4.8 \times 10^{-7}$), a locus previously reported to have strong and reproducible effects on doctor-diagnosed asthma.[7] Multiple variants in this region were associated with disease risk in our analysis (Figure 1a). The rs6503525 SNP is in linkage disequilibrium (LD, $r^2 = 0.61$) with the top *ORMDL3* variant (rs7216389) reported in the original study by Moffat *et al*;[7] both predisposing alleles (rs6503525:C and

rs7216389:T) are in phase. We found no evidence for additional independent risk variants in this region (not shown), consistent with Moffatt *et al*.[7] These results thus confirm that a genetic variant in the *ORMDL3* locus significantly influences asthma risk in the Australian population.

Among the five remaining regions most associated with asthma (Table 1), only for chromosome 5q31 there was an obvious candidate gene under the association peak, the chemokine CXC motif ligand 14 (*CXCL14*) gene, which is thought to be a potent chemoattractant and activator of dendritic cells.[28]

Analysis of loci reported in previous GWAS of asthma
Despite low power to detect risk loci for asthma with genome-wide significance, our whole-genome data set nonetheless provided an opportunity to attempt to replicate previously identified loci, given

Table 2 SNP association results for *ORMDL3*, *PDE4D*, *DENND1B* and *IL1RL1*

| Region | Gene | Literature SNP, allele[a] | Reported OR | Power, %[b] | Allele frequency | | OR | 95% CI OR | P-value |
					Cases	Controls			
(a) Most-associated variant reported in the original studies									
17q12-21	*ORMDL3*	rs7216389, T	1.45	100	0.53	0.47	1.25	1.12–1.40	6.0×10^{-5}
5q12	*PDE4D*	rs1588265, G	0.85	78	0.31	0.30	1.05	0.93–1.18	0.4334
1q31	*DENND1B*	rs1775456, G	0.75	93	0.21	0.20	1.02	0.89–1.17	0.7862
2q12	*IL1RL1*	rs1420101, T	1.16	77	0.39	0.38	1.04	0.93–1.17	0.4649

| Region | Gene | Best SNP, allele[a] | N SNPs tested[c] | LD with literature SNP, r[b] | Allele frequency | | OR | P-value | |
					Cases	Controls		Uncorrected	Corrected[d]
(b) Most-associated variant in this study									
17q12-21	*ORMDL3*	rs6503525,C	232	0.61	0.50	0.43	1.33	4.8×10^{-7}	0.0002
5q12	*PDE4D*	rs17725522,C	1503	0.00	0.06	0.08	0.68	0.0007	0.1948
1q31	*DENND1B*	rs12094462,G	171	0.03	0.10	0.12	0.81	0.0229	0.3238
2q12	*IL1RL1*	rs10197862,G	182	0.09	0.12	0.16	0.75	0.0004	0.0122

[a]All four SNPs in (a) were present in the Illumina chips used and passed QC filters. The top SNPs listed in (b) were imputed with high confidence with MACH, with nearby genotyped SNPs providing comparable evidence for association (cf. Figure 1).
[b]Power to detect the initial association at α=0.05 in this study.
[c]Number of SNPs located in or within 50 kb of each gene. For *ORMDL3*, we considered all SNPs located between 50 kb downstream of *GSDMB* and 50 kb upstream of *GSDMA*.
[d]Significance of the best SNP after accounting for the number of and LD between all SNPs tested in the respective gene, estimated from 10 000 permutations.

the less stringent threshold required for significance. We, therefore, investigated whether there was any evidence for association with SNPs located in the three remaining loci that have previously been reported to have reproducible effects on asthma: *PDE4D*, *DENND1B* and *IL1RL1*. For completeness, we also report results from these analyses for *ORMDL3*.

We first focused on the individual variant with strongest association reported in the original study for each locus. We found consistent evidence for association with the leading literature SNP for *ORMDL3* (rs7216389, OR=1.25, $P=6\times10^{-5}$), but not for the three remaining loci, despite reasonable power (Table 2a).

To address the possibility that allelic heterogeneity could explain the negative results for *PDE4D*, *DENND1B* and *IL1RL1*, we then extended the analysis to all variants available in our study for each of these loci ($\pm$ 50 kb). A variant in *IL1RL1* was found to be associated with asthma after accounting for the number of and LD between SNPs in the gene (Table 2b); in contrast, there was no overall evidence for association with *PDE4D* or *DENND1B* (Table 2b and Figure 1).

The peak variant for *IL1RL1* was rs10197862 (OR=0.75 for the G allele, P=0.0004, corrected P=0.01), which is in low LD (r^2=0.09) with the rs1420101 variant reported by Gudbjartsson *et al*.[16] Haplotype analysis of both variants indicates that the rs10197862:G protective allele is found exclusively on a haplotype that also carries the rs1420101:C protective allele (not shown). Although this GC haplotype could be a proxy for a single functional variant that explains both associations, this is unlikely given that with a high-risk allele frequency of 84%, the associated effect size for such variant would have to be high (OR of $\sim$1.8) to match the results reported for rs1420101 (OR=1.15) by Gudbjartsson *et al*.[16] As rs10197862 was also in low LD (r^2<0.1) with two nearby variants (rs2310173 and rs917997) reported recently to associate with other immune-related diseases,[29,30] our results suggest that rs10197862 represents or tags an independent functional variant in the 2q12 region with significant effects on asthma risk.

Given that the *IL1RL1* and *ORMDL3* loci have also been reported to associate with celiac[29] and Crohn's disease,[31] respectively, we investigated whether there was any evidence for association between asthma and other confirmed loci for these diseases. Among these, no single

locus was associated with asthma risk after accounting for the number of SNPs considered (Supplementary Tables 3 and 4).

We extended our genome-wide search for asthma risk loci by considering CNV data that were obtained for 759 adolescents, including 270 doctor-diagnosed asthma cases and 489 controls.

Given that common CNVs were likely tagged by SNPs tested in the genome-wide analysis,[18] we focused on large (>100 kb and <1 Mb), uncommon (MAF <0.05) deletions or duplications identified with high confidence. A total of 2681 such segments were detected, with a median of three CNVs per individual (mean=3.9, range 1–21); the median CNV size was 198 kb (mean 252 kb, range 100–998 kb). Of these, 1621 (61%) were deletions and 1060 were duplications (39%). These CNVs mapped to 244 (deletions) and 230 (duplications) non-overlapping regions of the genome, with a median region length of 259 kb (range 102 kb–4.0 Mb) and 281 kb (range 102 kb–3.6 Mb), respectively.

We performed two sets of CNV association analyses. First, we tested whether the frequency of deletions or duplications at individual loci were associated with asthma status. After accounting for all the regions tested, no single locus was significantly associated with asthma; the two most-associated regions are shown in Table 3, including a 102–775 kb deletion of chromosome 1q21 and a 129–432 kb deletion of chromosome 17q21. To visually validate the CNV calls made by QuantiSNPv1.1 for these two regions, we inspected the local log R ratio values for the individual samples identified by the software as carrying a CNV. Log R ratio patterns for the 1q21 region were ambiguous (Supplementary Figure 3) and so this region may represent a technical artefact. In contrast, the log R ratio patterns for all 38 individuals identified as carrying the chromosome 17q21 deletion are indeed consistent with a CNV in this region (Supplementary Figure 4), which is known to contain a common inversion in populations of European ancestry.[32] The telomeric break point for this deletion was located in exon 13 of the *NSF* gene for 36 of the 38 individuals that carried this deletion, all of which belonged to the H2 haplotype as defined by the inversion-informative rs1800547 variant.[32]

Second, we tested the association between genome-wide CNV burden and asthma risk. We hypothesised that the cumulative effect of multiple deletions or duplications across an individual's genome

could significantly increase the risk of developing asthma. After correcting for the number of hypotheses tested, we found no evidence that cases and controls had significant differences in the total number of CNVs, length of CNVs or in the number of genes affected by a CNV, either when considering deletions or duplications (Table 4).

Follow-up of top SNP and CNV associations

Finally, we attempted to replicate the most-associated loci identified in the whole-genome SNP and CNV analyses, as well as the novel *IL1RL1* association, in an independent cohort of 391 doctor-diagnosed asthmatics and 213 controls (Supplementary Table 5). Power to nominally replicate these associations was relatively modest (30–60%), but we observed a convincing association between the *IL1RL1* variant rs10197862 and asthma (OR=0.58, $P=1.2\times10^{-5}$), which confirms the presence of a novel independent risk allele in this region (Supplementary Table 6). Also consistent with our initial observation, the 17q21 deletion was more frequent in asthma cases (5/391) than controls (1/213) (Supplementary Figure 5), but these effects were not significant (Supplementary Table 6). The different frequency estimates for this deletion between the discovery (5%) and replication (1%) panels suggests that its prevalence is underestimated in the latter (as all deletions were visually validated) and highlights the technical

limitations of CNV detection using genotyping arrays, even when the same array (Illumina 610K) and CNV calling pipeline are used across projects. Further analyses in larger cohorts are thus warranted to confirm this new putative association.

DISCUSSION

Our analysis of whole-genome SNP data in 986 self-reported asthma cases and 1846 controls confirms that variants in *ORMDL3*, a locus previously reported to reproducibly associate with asthma in European and North American populations,[7] influences disease risk in the Australian population. The most-associated variant in our analysis was rs6503525, which lies 39 kb upstream of *ORMDL3*. Although gene expression results point to *ORMDL3* as the most likely causal candidate at 17q12, we note that other nearby genes should not be neglected in future experimental work conducted under this broad region of high LD. Specifically, on a recent analysis of genetic determinants of peripheral blood lymphocyte cell levels,[33] we found that a synonymous change (rs907092) in the Ikaros Family Zinc Finger 3 (*IKZF3*), which is located 57 kb downstream of *ORMDL3*, was a significant predictor of B cell levels in adolescents. The G allele that associated with increased numbers of B cells in that analysis had a predisposing effect for asthma in this study (OR=1.22, $P=5.1\times10^{-4}$). This variant is in partial LD with rs6503525 ($r^2=0.52$) and, although it may simply tag the same underlying causal variant, it may also represent an additional risk locus in this region. Further studies that explore this possibility are warranted.

Our results also demonstrate that genetic variation in *IL1RL1* significantly influences asthma risk, but the specific variant identified in our study (rs10197862) is independent of that reported previously to influence eosinophil levels and asthma by Gudbjartsson *et al*.[16] These results thus indicate that multiple independent variants in *IL1RL1* can contribute to asthma risk and further emphasise the need to apply gene-based association tests that can effectively account for such effects. Two additional variants near the neighbouring interleukin genes *IL18R1* and *IL1R2* have recently been reported to associate with celiac disease[29] and ankylosing spondylitis,[30] but these were in low LD with rs10197862. Together, these data indicate that the 2q12 region has an important and diverse function in the regulation of immune responses.

PDE4D and *DENND1B* were recently identified as asthma susceptibility genes.[9,11] Our analyses do not support these findings, either when focusing on the individual variants reported by the original studies or after extending to all the genetic variation tested in each gene in our study. Failure to replicate the reported associations with PDE4D and DENND1B could represent a false-negative finding in our study (eg limited power to detect an actual effect size that is smaller than that estimated in the original studies), but could also indicate

Table 3 Individual regions with strongest association between CNVs and asthma risk

Chromosome	1q21	17q21
CNV type	Deletion	Deletion
Average CNV span (range) (kb)	283 (102–775)	320 (129–432)
Range shared region (bp)	146397003–146438291	42008822–42137359
Length shared region (kb)	41	129
N cases with CNV/total	8/270	20/270
N controls with CNV/total	2/489	18/489
OR	7	2
P-value	0.0051	0.0196
Corrected *P*-value[a]	0.4403	0.9947
Genes in region ± 50 kb	*NBPF1, NBPF10, NBPF12, NBPF14, PPIAL4A*	*ARL17, ARL17P1, ARL17A, ARL17B, NSF*
SNP most correlated with CNV (r^2)	rs7554500 (0.05)	rs7225002 (0.30)
Visual validation of CNVs[b]	Ambiguous	Yes

[a]Significance after correcting for all deletions or duplications tested across the genome, estimated from 10 000 permutations.
[b]Local log R ratio patterns for individuals identified as carrying a deletion were inspected to validate the respective CNV calls produced by QuantiSNPv1.1. Intensity patterns for the 17q21 but not for the 1q21 regions were consistent with the CNV calls made (Supplementary Figures 3 and 4).

Table 4 Association results between genome-wide CNV burden and asthma risk in 270 asthma cases and 489 controls

	Deletions				Duplications			
	Cases	*Controls*	P-value[a]		*Cases*	*Controls*	P-value[a]	
			Uncorrected	*Corrected*			*Uncorrected*	*Corrected*
N CNVs detected	564	1057	—	—	326	734	—	—
N CNVs/individual	2.09	2.16	0.8305	0.9989	1.21	1.50	0.0658	0.2352
Total CNV span (kb)	693.5	695.2	0.9371	1.0000	476.7	572.2	0.0282	0.0994
Average CNV span (kb)	223.3	237.3	0.4115	0.8181	259.1	269.1	0.5835	0.9623
N genes spanned by CNVs	6.67	7.31	0.4318	0.8165	4.49	4.01	0.0675	0.2180

[a]Two-sided *P*-value estimated from 10 000 permutations. Corrected *P*-values were obtained after adjusting for the overall number of hypotheses tested (N=8) through the analysis of the permuted data sets.

that these genes represent false-positive discoveries. Further analyses in large cohorts are thus warranted to confirm both genes as true asthma susceptibility loci.

We also show that CNVs with strong effects on asthma could not be detected with experiment-wide significance in a subset of 270 cases and 489 controls. However, it is noteworthy that the chromosome 17q21 deletion that associated with increase asthma risk in our analysis overlaps with the known 900 kb inversion[32] that was recently reported to associate with corticosteroid response in asthma.[34] We thus provide some further support for the association between structural variants in chromosome 17q21 and asthma, but further studies are required to carefully characterise this complex region of the genome. Last, results from our genome-wide CNV burden analysis suggest that asthma cases do not have increased numbers of uncommon large chromosomal deletions or duplications when compared with unaffected individuals. The number of genes affected by CNVs was also comparable between both groups. These results contrast with recent reports for psychiatric diseases, namely schizophrenia.[21] In our analysis, we observed significant confounding effects of technical variables, such as array batch effects. Ignoring these may result in false-positive results, particularly when cases and controls have been genotyped in different batches or at different times.

There are a number of limitations that should be considered when interpreting our results. First, our definition of lifetime asthma status is somewhat loose, as it was based on self-reported information provided in surveys that often included comparable but not identical questions, and did not exclude co-morbid diseases such as COPD. However, we note that despite these differences, the resulting phenotype is heritable (61%) and allowed us to unambiguously confirm the association with *ORMDL3*, effectively a positive control for asthma. Second, the sample size analysed was small for a GWAS of a complex disease. These results will now be combined with other ongoing studies to provide a more powerful and comprehensive survey of asthma risk factors. Third, age-at-onset information was only available for a subset of cases and so it was not considered in our analysis. Nonetheless, we note that to date the only locus with confirmed age-at-onset effects is *ORMDL3*, a locus that we replicate in this study. Last, our analysis of rare structural variants relied heavily on the ability to identify such segments from SNP and CNV probe intensity patterns, an approach that can be influenced by a number of technical artefacts; the sample size used was also very modest. Thus, confirmation of our results by independent studies are warranted, particularly the putative association with the 17q21 deletion.

In conclusion, we report a convincing association between asthma risk and both novel and previously reported sequence variants in two loci, *IL1RL1* and *ORMDL1*; we also identify a 300 kb deletion on chromosome 17q21 with putative effects on disease risk. We found no evidence for association with *PDE4D* or *DENND1B*, and no indication that asthmatics have more or larger CNVs than non-asthmatics. These results may help direct future studies of asthma genetics.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

1 Australian Centre for Asthma Monitoring. *Asthma in Australia 2008*. AIHW Asthma Series No. 3. Cat. No. ACM 14. Canberra: AIHW, 2008.
2 Lemanske Jr RF, Busse WW: Asthma: clinical expression and molecular mechanisms. *J Allergy Clin Immunol* 2010; **125**: S95–102.
3 Altshuler D, Daly MJ, Lander ES: Genetic mapping in human disease. *Science* 2008; **322**: 881–888.
4 Hindorff LA, Sethupathy P, Junkins HA *et al*: Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc Natl Acad Sci USA* 2009; **106**: 9362–9367.
5 Rioux JD, Xavier RJ, Taylor KD *et al*: Genome-wide association study identifies new susceptibility loci for Crohn disease and implicates autophagy in disease pathogenesis. *Nat Genet* 2007; **39**: 596–604.
6 Fellay J, Shianna KV, Ge D *et al*: A whole-genome association study of major determinants for host control of HIV-1. *Science* 2007; **317**: 944–947.
7 Moffatt MF, Kabesch M, Liang L *et al*: Genetic variants regulating ORMDL3 expression contribute to the risk of childhood asthma. *Nature* 2007; **448**: 470–473.
8 Choudhry S, Taub M, Mei R *et al*: Genome-wide screen for asthma in Puerto Ricans: evidence for association with 5q23 region. *Hum Genet* 2008; **123**: 455–468.
9 Himes BE, Hunninghake GM, Baurley JW *et al*: Genome-wide association analysis identifies PDE4D as an asthma-susceptibility gene. *Am J Hum Genet* 2009; **84**: 581–593.
10 Hancock DB, Romieu I, Shi M *et al*: Genome-wide association study implicates chromosome 9q21.31 as a susceptibility locus for asthma in Mexican children. *PLoS Genet* 2009; **5**: e1000623.
11 Sleiman PM, Flory J, Imielinski M *et al*: Variants of DENND1B associated with asthma in children. *N Engl J Med* 2010; **362**: 36–44.
12 Mathias RA, Grant AV, Rafaels N *et al*: A genome-wide association study on African-ancestry populations for asthma. *J Allergy Clin Immunol* 2010; **125**: 336–346.e4.
13 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**: 328–335.e11.
14 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**: 1682–1691.
15 Weidinger S, Gieger C, Rodriguez E *et al*: Genome-wide scan on total serum IgE levels identifies FCER1A as novel susceptibility locus. *PLoS Genet* 2008; **4**: e1000166.
16 Gudbjartsson DF, Bjornsdottir US, Halapi E *et al*: Sequence variants affecting eosinophil numbers associate with asthma and myocardial infarction. *Nat Genet* 2009; **41**: 342–347.
17 Scherer SW, Lee C, Birney E *et al*: Challenges and standards in integrating surveys of structural variation. *Nat Genet* 2007; **39**: S7–15.
18 The Wellcome Trust Case Control Consortium: Genome-wide association study of CNVs in 16 000 cases of eight common diseases and 3000 shared controls. *Nature* 2010; **464**: 713–720.
19 Bochukova EG, Huang N, Keogh J *et al*: Large, rare chromosomal deletions associated with severe early-onset obesity. *Nature* 2010; **463**: 666–670.
20 Weiss LA, Shen Y, Korn JM *et al*: Association between microdeletion and microduplication at 16p11.2 and autism. *N Engl J Med* 2008; **358**: 667–675.
21 The International Schizophrenia Consortium: Rare chromosomal deletions and duplications increase risk of schizophrenia. *Nature* 2008; **455**: 237–241.
22 Ionita-Laza I, Perry GH, Raby BA *et al*: On the analysis of copy-number variations in genome-wide association studies: a translation of the family-based association test. *Genet Epidemiol* 2008; **32**: 273–284.
23 Medland SE, Nyholt DR, Painter JN *et al*: Common variants in the trichohyalin gene are associated with straight hair in Europeans. *Am J Hum Genet* 2009; **85**: 750–755.
24 Blangero J, Almasy L: Multipoint oligogenic linkage analysis of quantitative traits. *Genet Epidemiol* 1997; **14**: 959–964.
25 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**: 559–575.
26 Frazer KA, Ballinger DG, Cox DR *et al*: A second generation human haplotype map of over 3.1 million SNPs. *Nature* 2007; **449**: 851–861.
27 Colella S, Yau C, Taylor JM *et al*: QuantiSNP: an Objective Bayes Hidden-Markov Model to detect and accurately map copy number variation using SNP genotyping data. *Nucleic Acids Res* 2007; **35**: 2013–2025.

28 Shurin GV, Ferris RL, Tourkova IL *et al*: Loss of new chemokine CXCL14 in tumor tissue is associated with low infiltration by dendritic cells (DC), while restoration of human CXCL14 expression in tumor cells causes attraction of DC both *in vitro* and *in vivo*. *J Immunol* 2005; **174**: 5490–5498.

29 Dubois PC, Trynka G, Franke L *et al*: Multiple common variants for celiac disease influencing immune gene expression. *Nat Genet* 2010; **42**: 295–302.

30 Australo-Anglo-American Spondyloarthritis Consortium (TASC)Reveille JD, Sims AM, Danoy P *et al*: Genome-wide association study of ankylosing spondylitis identifies non-MHC susceptibility loci. *Nat Genet* 2010; **42**: 123–127.

31 Barrett JC, Hansoul S, Nicolae DL *et al*: Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. *Nat Genet* 2008; **40**: 955–962.

32 Stefansson H, Helgason A, Thorleifsson G *et al*: A common inversion under selection in Europeans. *Nat Genet* 2005; **37**: 129–137.

33 Ferreira MA, Mangino M, Brumme CJ *et al*: Quantitative trait loci for CD4: CD8 lymphocyte ratio are associated with risk of type 1 diabetes and HIV-1 immune control. *Am J Hum Genet* 2010; **86**: 88–92.

34 Tantisira KG, Lazarus R, Litonjua AA, Klanderman B, Weiss ST: Chromosome 17: association of a large inversion polymorphism with corticosteroid response in asthma. *Pharmacogenet Genomics* 2008; **18**: 733–737.

Supplementary Information accompanies the paper on European Journal of Human Genetics website (http://www.nature.com/ejhg)