Allergic asthma is a heritable chronic disease that affects >10% of children in developed countries (1), many of whom do not respond to the few established treatments (2). In spite of a large public health burden, our knowledge of the aetiological mechanisms underlying asthma, both genetic and environmental, is still very limited. One of the most promising approaches to further our understanding of the disease mechanisms involved is the identification of genetic variation that contributes to asthma risk.

As demonstrated for other complex diseases (3), genome-wide association studies (GWAS) are the most successful gene mapping approach currently available to identify systematically all pathways at which genetic variation contributes to disease risk. However, given the multiple testing burden associated with GWAS and the small effect sizes that are predicted for most risk loci, the success of this approach rests on the ability to analyse data for many thousands of samples, for both the initial discovery and replication stages.

While adequately powered datasets are being collected for such large-scale hypothesis-free analyses, smaller, more focused association studies can still provide a valuable framework to investigate specific genetic hypotheses. In this report, we screened eight genes that were selected based on their potential functional relevance from four regions previously reported to be linked to asthma traits. Amongst these were known candidate genes for asthma, namely the \textit{CD28}, cytotoxic T-lym-phocyte-associated protein 4 (\textit{CTLA4}) and inducible T-cell co-stimulator (\textit{ICOS}) genes, but also new genes not previously tested in the literature, including ADAM23, ADAMTSL1, MS4A2, CDH26 and HRH3.

### Background:
Linkage studies have implicated the 2q33, 9p21, 11q13 and 20q13 regions in the regulation of allergic disease. The aim of this study was to test genetic variants in candidate genes from these regions for association with specific asthma traits.

### Methods:
Ninety-five single nucleotide polymorphisms (SNP) located in eight genes (\textit{CD28}, \textit{CTLA4}, \textit{ICOS}, \textit{ADAM23}, \textit{ADAMTSL1}, \textit{MS4A2}, \textit{CDH26} and \textit{HRH3}) were genotyped in >5000 individuals from Australian (\textit{n} = 1162), Dutch (\textit{n} = 99) and Danish (\textit{n} = 303) families. Traits tested included doctor-diagnosed asthma, atopy, airway obstruction, total serum immunoglobulin (Ig) E levels and eosinophilia. Association was tested using both multivariate and univariate methods, with gene-wide thresholds for significance determined through simulation. Gene-by-gene and gene-by-environment analyses were also performed.

### Results:
There was no overall evidence for association with seven of the eight genes tested when considering all genetic variation assayed in each gene. The exception was \textit{MS4A2} on chromosome 11q13, which showed weak evidence for association with IgE (gene-wide \(P < 0.05\), rs502581). There were no significant gene-by-gene or gene-by-environment interaction effects after accounting for the number of tests performed.

### Conclusions:
The individual variants genotyped in the 2q33, 9p21 and 20q13 regions do not explain a large fraction of the variation in the quantitative traits tested or have a major impact on asthma or atopy risk. Our results are consistent with a weak effect of \textit{MS4A2} polymorphisms on the variation of total IgE levels.
Ferreira et al.

metallopeptidase domain 23 (ADAM23) and histamine receptor H3 (HRH3). We employed novel association methods and tested for gene-by-gene and gene-by-environment interaction effects to assess the influence of genetic variants in these genes with both self-reported and clinical measures of asthma in 5470 individuals.

**Material and methods**

Selection of genes for analysis

Our aim was to investigate the association between specific asthma traits and genes located in four chromosomal regions identified through previous linkage studies (Table 1). Genes were selected from these regions for single nucleotide polymorphism (SNP) genotyping based on their potential functional relevance and included CD28, CTLA4, ICOS and ADAM23 on chromosome 2q33; ADAMTS-like 1 (ADAMTS1L1) on 9p21; membrane-spanning 4-domains, subfamily A, member 2 (MS4A2) previously named β chain of the high-affinity IgE receptor, FCER1B) on 11q13; and cadherin-like 26 (CDH26) and HRH3 on 20q13.

**Samples**

We genotyped samples drawn from two studies conducted at the Queensland Institute of Medical Research (QIMR), Australia, a study conducted at the VU University (VU), the Netherlands, and a study conducted at the Bispebjerg Hospital (BH), Denmark. The first QIMR study (AS) included data for 735 families that are part of a larger collection of pedigrees ascertained through an asthma twin proband and described elsewhere (4). A subset of this sample was used in the analysis that reported linkage of multiple asthma traits to 20q13 (5). For the QIMR second study (TM), data were included for 303 families ascertained through an asthma twin proband and described elsewhere (4). A subset of this sample was used in the analysis that reported linkage of multiple asthma traits to 20q13 (5). For the TM study when twins were aged 12, 14 and 16 as described previously (6). For the 2q33 region, the phenotypes selected for analysis were Dpter, FEV1/FVC, IgE and eosinophil levels, as these were the traits previously reported to be linked (6, 9, 10) or associated (11, 12) to this region. For the TM study, we selected the age 12 eosinophil data as this was the time point at which significant linkage was detected to chromosome 2q33 (6). Similarly, we selected IgE for analysis with 9p21 (M. A. R. Ferreira, D. L. Duffy, N. G. Martin, unpublished data), IgE and Atopy with 11q13 (11, 13, 14), and Dpter, atopy, BHR, FEV1, asthma and FEV1/FVC for 20q13 (5).

Continuous traits were normalized and adjusted for significant covariates, which included age, gender, height (FEV1 only) and regular smoking (FEV1/FVC). As age showed a strong positive correlation with FEV1 for individuals aged 6–20 years (younger cohort) and a negative correlation for those aged >20 years, an interaction term with a cohort indicator was also used to correct for this effect.

SNP genotyping and quality control (QC)

A set of 108 SNPs reported to be highly polymorphic in the NCBI SNP Database and evenly spaced across each of the eight genes was originally selected for genotyping in the AS study using the MassARRAY Sequenom system as described previously (15). A subset of these was also genotyped in the TM (n = 39), VU (n = 41) and BH (n = 41) studies. Genotyping was performed in five batches, with QC analyses performed for each batch separately (see Table S2 for details) using PLINK (16) and PEDSTATS (17). Twenty SNPs were genotyped in two or three (n = 1) batches to assess concordance between genotyping events. After individual QC, data from all batches were merged, resulting in a dataset with 5535 individuals genotyped for 95 SNPs. The concordance rate between different genotyping events was 0.995. After data were merged, we compared genotype calls for self-reported monoyzotic (MZ) twin pairs (n = 351, concordance 0.992) and (i) deleted all individuals (n = 65) from families (n = 22) with MZ twins having >1 discordant genotype and (ii) set to missing any discordant genotype for the remaining MZ pairs. We then used MERLIN (18) to detect and delete unlikely genotypes (n = 1227) and to infer missing genotypes using identical by descent (IBD) information (n = 17 755). Excluding 1250 ungenotyped founders for whom MERLIN inferred a small fraction of genotypes (a median of 5.3%), the final dataset consisted of 5470 individuals genotyped for 95 SNPs.
Main association analyses

We tested the following four hypotheses: (i) variants in the selected 2q33 genes are associated with Dpter response, FEV1/FVC, total serum IgE or eosinophil levels; (ii) variants in ADAMTS11 (9p21) are associated with IgE levels; (iii) variants in MS4A2 (11q13) are associated with Atopy or IgE levels; and (iv) variants in CDH26 and HRH3 genes (20q13) are associated with Dpter response, atopy, BHR, FEV1, asthma or FEV1/FVC. Hypothesis (i) was tested in the AS, VU, BH and TM (eosinophilis) studies, whereas hypotheses (ii), (iii) and (iv) were tested in the AS study only.

For hypotheses (i), (iii) and (iv), the association between individual SNPs and all traits was tested using a recently described multivariate test of association implemented in PLINK (19). For these analyses, phenotype data for one twin from MZ pairs was discarded (that with milder asthma symptoms). Association between eosinophil levels and 2q33 SNPs in the TM study was tested with MERLIN while accounting for IBD allele sharing (21) because of the strong linkage signal present in this dataset (logarithm of odds, LOD > 3). The same approach was used to test hypothesis (ii).

Gene-wide significance thresholds and power for the main analyses

We adopted the concept of the gene as the unit of association (22). To assess the overall evidence for association with a particular gene given the genetic variation assayed, we used MERLIN to simulate 1000 datasets under the null hypothesis of no linkage or association, while preserving the marker informativeness, spacing, missing data patterns and allele frequency and linkage disequilibrium (LD) between markers (21). Each dataset was analysed as described above and the most significant P-value of all SNPs tested in each gene and analysis was retained. The threshold for gene-wide significance (i.e. gene-wide P = 0.05) corresponded to the 95th percentile of the 1000 sorted P-values retained for each gene and analysis. Thresholds ranged between 0.002 and 0.028.

To estimate the power of the proposed analyses, we simulated 1000 datasets consisting of an SNP marker and an associated quantitative trait (40% heritable) with MERLIN, while preserving the missing data patterns of the original data. Based on these analyses, power (α = 0.001) to detect univariate association with the AS (or AS + VU + BH) study was 0.034 (0.055) for a quantitative trait locus (QTL) explaining 0.1% of the trait variance, 0.500 (0.752) for a 0.5% QTL, 0.938 (0.981) for a 1% QTL and 1.00 (1.00) for a 2% QTL. Power for the TM study and the same QTL variances was respectively 0.01, 0.05, 0.18 and 0.55.

Secondary analyses

We performed gene-by-gene and gene-by-environment interaction analyses to address two additional hypotheses. The first considered the possibility that under specific genetic models with strong epistatic effects and weak or no main effects, risk loci could be detected through interaction but not association analyses (23, 24). Second, that an association between the eight genes tested and atopy might only be detected when considering individuals exposed to certain environmental factors.

To investigate whether strong gene-by-gene effects could be detected, we focused on the Atopy trait, as this was the general clinical phenotype reported to be linked with all four regions tested. We selected a subset of 350 unrelated cases and 184 unrelated controls from the AS study and tested pairwise SNP interactions between the eight genes using logistic regression as implemented in PLINK. A total of 3733 pair-wise SNP interactions were tested. Given the sample size and number of tests performed, power was only adequate to detect strong interaction effects. For example, assuming a disease prevalence of 30%, a genotype relative risk of 0.6 for each of the four possible interaction parameters (additive × additive, additive × dominant, dominant × additive, dominant × dominant), no main effects and a minor allele frequency of 0.25 for both loci, power to detect epistasis was 0.80 (with α = 0.05/3733 = 1.3 x 10^-5), while power to detect association to either loci was 0.18 (α = 0.05/95 = 5.3 x 10^-4).

To investigate the presence of gene-by-environment effects, we specifically tested whether SNPs were associated with atopy status in individuals: (i) exclusively breastfed under the age of six months; (ii) exposed to parental smoking under the age of 2 years; (iii) with frequent chest infections under the age of two and (iv) who lived in a carpeted house under the age of two. After appropriately defining cases (e.g. breastfed atopics) and controls (e.g. breastfed nonatopics) within each family, association was tested using the DFAM test implemented in PLINK, with an adaptive permutation procedure used to assess significance.

Results

Previous linkage and association studies implicated the 2q33 region in the regulation of Dpter response, FEV1/FVC, IgE and eosinophil levels. We tested whether variants in the 2q33 genes CD28, CTLA4, ICOS or ADAM23 were associated with any of these four phenotypes in the combined data from the AS, VU and BH studies using a multivariate approach. The 70 SNPs genotyped tagged (r^2 > 0.8) between 44% and 79% of the common [minor allele frequency (MAF) > 0.05] variation reported for these genes in the HapMap CEPH population. No single SNP exceeded the respective threshold for gene-wide significance (Table 2). The most associated SNP was rs3181096 (multivariate P = 0.028), which is located in the CD28 promoter region. The main phenotype associated with rs3181096 was FEV1/FVC (Table S3). A subset of 30 SNPs was also tested for association with eosinophil levels in the TM study. This SNP set provided a similar coverage for the CD28/CTLA4/ICOS cluster (40-62%), but considerable less for ADAM23 (7%). In spite of evidence for linkage to this region (best LOD = 3.77, P = 1.5 x 10^-5, rs3732079), no single SNP was associated with eosinophilia at a gene-wide significant level (Table S4). Of note, the most associated SNP was rs3181098 (univariate P = 0.02), which is in near complete LD (r^2 = 0.99) with rs3181096 (P = 0.034).

Next, we investigated whether any of the six SNPs genotyped in ADAMTS11 (13% coverage) on chromosome 9p21 were associated with total serum IgE levels in the AS study. This region and gene were selected based on previous linkage results for a subset of samples from this study. There was some evidence for linkage (best LOD = 0.85, P = 0.024, rs1889007) but not association (best P = 0.103) with the markers tested (Table S5).
**Table 2. Summary of multivariate association results between four genes in the 2q33 region and four asthma traits (Dpter, FEV1/FVC, total IgE and eosinophils), in the combined analysis of the AS, VU and BH studies**

<table>
<thead>
<tr>
<th>SNP set</th>
<th>CD28</th>
<th>CTLA4</th>
<th>ICOS</th>
<th>ADAM23</th>
</tr>
</thead>
<tbody>
<tr>
<td>N HapMap†</td>
<td>204, 239</td>
<td>204, 401</td>
<td>204, 470</td>
<td>206, 977</td>
</tr>
<tr>
<td>N tested (in HapMap)</td>
<td>1626</td>
<td>1626</td>
<td>1626</td>
<td>1626</td>
</tr>
<tr>
<td>Coverage‡</td>
<td>0.79</td>
<td>0.62</td>
<td>0.48</td>
<td>0.44</td>
</tr>
<tr>
<td>Gene-wide significance threshold§</td>
<td>0.004</td>
<td>0.006</td>
<td>0.003</td>
<td>0.003</td>
</tr>
<tr>
<td>N with multivariate P &lt; 0.05</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

**SNP, single nucleotide polymorphism.**
*Start position and length are based on the b36 coordinates for the first and last transcribed bases ±4 kb.
†Number of described common (minor allele frequency >0.05) SNPs located within the target region in CEU HapMap population. Also includes SNPs genotyped in this study that are not in the HapMap.
‡Proportion of CEU HapMap SNPs captured with $r^2 \geq 0.8$ by the SNPs tested.
§Estimated from 1000 simulated datasets using MERLIN. See Methods for details.

Chromosome 11q13 and the *MS4A2* gene in particular were amongst the first loci implicated by linkage studies in the regulation of IgE levels and the atopic response. We applied the same multivariate approach to test the association between *MS4A2* SNPs and variation in IgE levels or Atopy in the AS study. Of the six SNPs tested in this gene (42% coverage), five were significant at the $P < 0.05$ level (all in strong LD, $r^2 > 0.8$), with the most associated SNP (rs502581, multivariate $P = 0.010$) just reaching gene-wide significance ($P < 0.012$). The observed multivariate association signal to rs502581 was driven by variation in total serum IgE levels but not atopy status (Table S6), a result that we confirmed using standard univariate analysis ($P = 0.015$ and $P = 0.857$, respectively). This SNP explained approximately 0.4% of the phenotypic variation of IgE in the AS study. There were no significant maternal effects at this locus for either trait (not shown).

Finally, following our earlier report of linkage between chromosome 20q13 and Dpter, Atopy, BHR, FEV1/FVC, Asthma and FEV1/FVC, we tested these six phenotypes for association with 13 SNPs genotyped in the *CD26* (13% coverage) and *HRH3* (19%) genes in the AS study. There was no overall evidence for association with either gene after accounting for the number of SNPs tested (Table S7). The most associated SNP was rs6062144 (multivariate $P = 0.036$), located approximately 8 kb upstream of *HRH3*.

Given the largely negative results with most genes tested, we investigated whether strong gene-by-gene or gene-by-environment effects could be detected in the AS study. We focused on the Atopy trait, as this was the general clinical phenotype reported to be linked with the four regions tested. No significant pairwise interactions were detected between the eight genes, after accounting for the number of tests performed (Table S8). Similarly, no overall significant associations were detected with any gene after considering breastfeeding under six months and smoking exposure, the occurrence of chest infections or living in a carpeted house under the age of 2 years (not shown).

**Discussion**

Functional (25–27) and linkage studies (6, 9, 10) pointed towards a potential role of genetic variation in the *CD28/CTLA4/ICOS* gene cluster on the development of allergic disorders and many association studies have followed-up on this hypothesis (Table S9). However, most of these focused on a small set of variants that were tested in relatively underpowered samples. In fact, the only reported associations that would remain significant after a strict adjustment for multiple testing were between two independent variants ($r^2 < 0.1$) in *CTLA4* (rs5742909 and rs231775) and IgE levels (11, 12) and between a variant in *ICOS* (rs11883722) and cockroach skin prick response (28). In our analyses, there was no evidence for association between rs5742909 or rs11883722 and the four asthma traits tested (Dpter, FEV1/FVC, IgE and eosinophils). Similarly, there was no association with a variant (rs231779) that is in complete LD with rs231775.

In addition to these negative results to specific variants previously reported in the literature, there was no overall evidence for association with seven of the eight genes tested and the respective phenotypes when we considered all genetic variation assayed in each gene. The only exception was for the gene encoding the $\beta$ chain of the high-affinity IgE receptor (*MS4A2* on chromosome 11q13, which showed weak evidence for association with total serum IgE at the gene-wide level. Although the observed association would not remain significant at the experiment-wide level (given the number of genes tested), it is consistent with the hypothesis that genetic variation in this gene may have an effect on the variation of IgE levels, as suggested by others (11, 14).

Given the power provided by the sample size used and analyses performed, we conclude that the individual genetic variation assayed in the 2q33, 9p21 and 20q13 regions does not explain a large fraction (i.e. >1%) of the variation in the quantitative traits tested or have a major impact on asthma or atopy risk. However, given the caveats of the present study, namely, the small number of genes tested per region, the low SNP coverage provided for some genes, the low power to detect weaker but more realistic effect sizes and the limited number of environmental risk factors considered, the possibility remains that risk variants in these regions may only be identified using larger studies with improved power. Alternatively, it is possible that some or all of these regions represented false-positive results from previous linkage-analyses.
this respect, combining data from published candidate-gene studies and emerging GWAS will provide ample opportunities to address some of these limitations.

Acknowledgments

The authors thank all families who participated in this study. The authors also thank Dr Louise O’Gorman, Dr Peter Le Souef, Dr Paul R Burton, Dr Brett G Toelle, Dr Colin F Robertson, Dr Claire Wainwright, Barbara Haddon, Dixie Statham, Wendy Oddy, Sunaleen Devadason, Dr Nigel Dore, Marita Dalton, Brigitte Willemsie, Elena Belousova, Janet Li and Dr Jennifer Peat for their roles in phenotype collection, and Renee Mayne for DNA preparation and genotyping. Phenotype data analysed in this study were originally collected with funding from Sequana Therapeutics Inc and later by AxyS Pharmaceutical Inc. Genotype collection and data analysis were supported by Project Grant 290274 and a Sidney Sax fellowship (to MARF) from the National Health and Medical Research Council (NHMRC), Australia.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Supplementary Methods.

Table S1. Characteristics of the study participants for the four individual studies.

References


Association mapping under asthma linkage peaks

Table S2. Details of the QC filters applied to the SNP data.
Table S3. Multivariate association results between four asthma traits and CD28, CTLA4 and ICOS SNPs on chromosome 2q33, for the combined analysis of the AS, VU and BH studies.
Table S4. Univariate association results for the TM study between eosinophil levels and 128 CD28, CTLA4 and ICOS SNPs on chromosome 2q33.
Table S5. Univariate association results for the AS study between IgE levels and 131 ADAMTS11 SNPs on chromosome 9p21.
Table S6. Multivariate association results for the AS study between MS4A2 SNPs and 134 atopy and total IgE levels.
Table S7. Multivariate association results for the AS study between 20q13 SNPs and six 142 asthma traits.
Table S8. Gene-gene interaction results for atopy in the AS study between all eight genes 152 tested.
Table S9. Summary of main results from published ‘candidate-gene’ association studies of asthma traits and CD28, CTLA4 and ICOS polymorphisms.

Supplementary References.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.


