

Article Type: Original Article

Childhood Asthma is Associated with Polymorphic Markers of *PROC* on 2q14 in Addition to 17q21 Locus

Wa Cheong Chan¹, Hing Yee Sy¹, Alice Pik-shan Kong², Chun Kwok Wong³, Lai Yin Tse¹, Kam Lun

Hon¹, Juliana Chung-ngor Chan², Gary Wing-kin Wong¹ & Ting Fan Leung¹

Departments of ¹Paediatrics, ²Medicine and Therapeutics and ³Chemical Pathology, The Chinese University of Hong Kong, Prince of Wales Hospital, Shatin, Hong Kong

Running title: *PROC* and childhood asthma

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/pai.12336

This article is protected by copyright. All rights reserved.

Correspondence:

Ting Fan Leung, Department of Paediatrics, 6/F, Lui Che Woo Clinical Sciences Building, Prince of Wales Hospital, Shatin, N.T., Hong Kong.

Tel: +852 2632 2981

Fax: +852 2636 0020

E-mail: tfleung@cuhk.edu.hk

Abstract

Background: Childhood asthma is caused by both genetic and environmental factors. The first genome-wide association study (GWAS) for asthma revealed putative candidates on nine chromosomal regions in Caucasians, with 17q21 locus being the most widely replicated one. However, there was no replication study for the other loci. This study investigated genetic associations between childhood asthma and autosomal single-nucleotide polymorphisms (SNPs) on eight loci reported in the first GWAS among Hong Kong Chinese.

Methods: 510 asthmatic children and 510 non-allergic controls were recruited. 110 tagging SNPs selected based on $r^2 \geq 0.80$ and minor allele frequency ≥ 0.05 for Han Chinese among all SNPs located 50-kb upstream and downstream of significant autosomal SNPs were genotyped by TaqMan allelic

This article is protected by copyright. All rights reserved.

discrimination assays. Transcription factor binding of SNPs was determined by electrophoretic mobility shift assay (EMSA).

Results: Asthma was significantly associated with SNPs on 17q21 and 2q14 loci. Twelve SNPs on 17q21 were associated with asthma, with rs6503527 being the most significant SNP. Five SNPs of protein C gene (*PROC*) on 2q14 were associated with asthma, with rs6755028 being the most significant SNP. Plasma protein C concentrations were higher in asthmatic patients than controls, and five *PROC* SNPs were associated with plasma protein C concentrations. EMSA showed specific differential binding of rs878461 to nuclear extracts from bronchial epithelial and hepatocarcinoma cell lines.

Conclusions: Our findings identify *PROC* on 2q14 as a novel candidate for childhood asthma, and replicate the genetic association for 17q21 locus. Rs878461 of *PROC* may increase asthma susceptibility by altering transcription factor binding.

Keywords: Childhood asthma; chromosome 2q14; genetics; protein C; transcription factor binding

Childhood asthma is a common chronic respiratory disorder caused by a complex interaction between predisposition genes and environmental exposures (1). Genome-wide association studies (GWASs) were performed to delineate the relationship between asthma and SNPs. New candidate genes and chromosome regions were identified, including ORM1-like 3/gasdermin B (*ORMDL-3/GSDMB*) (2), DENN domain-containing protein 1B (*DENND1B*) (3), Chitinase-3-like protein 1 (*CHI3L1*) (4), transducing-like enhancer of spilt 4 (*TLE4*) (5), RAD50 Homolog (*S. cerevisiae*) (*RAD50*) (6), Class I MHC-restricted T cell-associated molecule (*CRTAM*) (7), TNFAIP3 Interacting Protein 1 (*TNIP1*) (8), interleukin receptor 1/interleukin-18 receptor 1 (*IL1R1/IL18R1*), *HLADQ*, interleukin-33 (*IL-33*), *SMAD3*, interleukin 2 receptor beta (*IL2RB*) (9), *HLA-DP* (10), cAMP-specific 3',5'-cyclic phosphodiesterase 4D (*PDE4D*) (11), Kv channel interacting protein 4 (*KCNIP4*) (12), Ryanodine receptor-2 (*RYR-2*) (13), cadherin-related family member 3 (*CDHR3*) (14) and Ikaros family zinc finger protein 3 (*IKZF3*) (15). In a GWAS among Asians, asthma was associated with SNPs of *ORMDL3/GSDMB/ GSDMA*, *IL5/RAD50/IL13*, *HLA-DR/DQ* and *SMAD3*, which were also reported in European, Mexican and African populations (10). However, this study could not replicate the top SNPs of *PDE4D*, *TLE4*, *DENND1B*, *IL8R1* and *IL2RB*. Such discrepant results might be explained by differences in linkage disequilibrium (LD) patterns and minor allele frequencies (MAFs) of studied SNPs between Asian and other ethnic groups.

The first GWAS of asthma was published in 2007 in which 994 patients with history of childhood onset asthma and 1,243 non-asthmatics from Germany and UK white population are

studied. More than 317,000 single nucleotide polymorphisms (SNPs) were analyzed. Chromosome 17q21 locus was identified as the candidate region for childhood asthma (2), with multiple markers showing associations reaching genome-wide significance. The strongest association signal was found at rs7216389, which also positively associated with expression of *ORMDL3* gene. This genetic finding was subsequently replicated in various populations including our Chinese children (7-22). In addition to SNPs on chromosome 17q21, this GWAS also detected associations between asthma and seven autosomal SNPs and one X-chromosome SNP under the cut-off at 5% false discovery rate. Nonetheless, there has not been any subsequent attempt to delineate the significance of these preliminary genetic signals involving loci other than 17q21. In the present study, we sought to perform SNP-based fine mapping of these eight chromosomal regions reported in the first asthma GWAS among Chinese children in Hong Kong.

Patients and Methods

Study population

Chinese children with physician-diagnosed asthma were recruited from pediatric clinics of a university-affiliated teaching hospital. Asthma was diagnosed according to American Thoracic Society guidelines (23). Non-allergic controls without any asthma, allergic rhinitis and eczema as reported by International Study of Asthma and Allergies in Childhood questionnaire (24,25) were recruited from local community studies of childhood obesity and diabetes (26). Informed written

This article is protected by copyright. All rights reserved.

consent was obtained from subjects and their parents, and Clinical Research Ethics Committee of our university approved this study.

Selection and genotyping of SNPs

All SNPs located 50-kilobase both upstream and downstream of all autosomal SNPs with significant asthma associations as reported by Moffatt et al. (2) were retrieved from HapMap data (27). Tagging SNPs were selected from those with MAF ≥ 0.05 for Han Chinese and according to $r^2 \geq 0.8$ using HaploView version 4.2 (Daly Lab, Cambridge, MA, USA) (28). SNPs reported in original GWAS were also forced to be included during SNP selection.

Genomic DNA was extracted from EDTA-anticoagulated peripheral blood leukocytes. Ninety-six of 110 genotyped SNPs were genotyped by TaqMan Genotyping assays (Applied Biosystems, Foster City, CA, USA) on Biomark™ Dynamic Array (Fluidigm, San Francisco, CA, USA) using TaqMan Universal PCR Master Mix (Applied Biosystems). Endpoint fluorescence was captured on the BioMark™ Real-Time PCR System (Fluidigm) and analyzed using the default calling algorithm of Fluidigm SNP Genotyping Analysis software (Fluidigm) to obtain genotype calls. The remaining 14 SNPs were genotyped by TaqMan Genotyping assays on ABI 7900HT PCR System (Applied Biosystems).

Plasma protein C concentrations

Protein C (PC) concentrations in first-thawed plasma samples from a randomly selected one-third of subjects were measured by commercial enzyme-linked immunosorbent assay kit (Assaypro LLC, St. Charles, MO, USA), with the sensitivity being 0.09 µg/ml. Results were square root-transformed to achieve normality of data distribution before analysis.

Electrophoretic mobility shift assay

Two SNPs on 2q14 were predicted by web-based FuncPred program (<http://snpinfo.niehs.nih.gov/>) (29) to alter transcription factor binding, which was subsequently tested by electrophoretic mobility shift assay (EMSA). Briefly, human bronchial epithelial cell line BEAS-2B was obtained from American Type Culture Collection (Manassas, VA, USA). BEAS-2B cells were maintained in Dulbecco's Modified Eagle Medium supplied with 10% (v/v) fetal bovine serum, 100 U/ml penicillin G, and 100 µg/ml streptomycin sulfate (Invitrogen, Carlsbad, CA, USA). Cells were cultured in 37°C with 5% CO₂ until 60-70% confluence before harvesting. Human hepatocarcinoma cell line HepG2 was also cultured, following the same protocol as BEAS-2B cell line, in Minimum Essential Medium supplemented by 10% (v/v) fetal bovine serum, 100 U/ml penicillin G and 100 µg/ml streptomycin sulfate. Nuclear extracts from these cell lines were prepared using NucBuster™ Protein Extraction Kit (Merck KGaA, Darmstadt, Germany) and quantified by DC™ Protein Assay Kit (Bio-rad, Hercules,

CA, USA). EMSA was then performed using LightShift Chemiluminescent EMSA kit (Thermo Fisher, Waltham, MA, USA). Sense and antisense single-stranded 5'-biotinylated oligonucleotides (Invitrogen) were annealed to form double strands (rs878461: ACACATTTGGTTCTAGCTCCTTTGCTTGGG [C/T] CTTTCATCCATGCAGCGAAT; rs1799810: ACTCCAGGCTGTCATGGCGGCAGGACGGCG [A/T] ACTTGCAGTA TCTCCACGA). Binding reactions contained 1× binding buffer, 3 µg nuclear extracts, 0.5 µg poly (dI-dC) and 10 mM EDTA. Samples were incubated at room temperature for 20 minutes. Reaction mixtures were loaded to native 6% polyacrylamide gel electrophoresis and then transferred to a nylon membrane. The position of the oligonucleotide was detected by chemiluminescence with streptavidin-horseradish peroxidase. For competition assay, reaction included unlabeled probe with concentration ×100 and ×200 that of labelled probe.

Statistical analysis

Results were expressed as proportions or mean and standard deviation (SD) as appropriate. Patient characteristics were compared using the Student *t* test for numerical data, and Pearson χ^2 or Fisher exact test for categorical variables. Associations between SNPs and dichotomous variables were analyzed by logistic regression, adjusting for age and gender as covariates. The relationship between plasma PC concentrations and protein C gene (*PROC*) SNPs were analyzed by linear regression,

adjusting for age, sex, body mass index and asthma status as covariates. All comparisons were made two-tailed using SPSS version 17 (SPSS Inc., Chicago, IL, USA).

Significant genetic association was defined as: (a) P -value for individual SNP satisfying Bonferroni threshold; or (b) $P < 0.05$ consistently for a cluster of SNPs on the same locus. Given the *a priori* evidence for the eight tested loci (2), this study adjusted for multiple statistical comparisons by dividing conventional P -value threshold by the number of SNPs being genotyped for each locus (i.e. regional Bonferroni threshold at $0.05/n$ where n was the number of tested SNPs). Pairwise LD coefficient was calculated for each SNP pair by HaploView version 4.2 (24). Haplotype blocks were defined using confidence interval algorithm. Haplotype analysis was performed using *haplo.stats* (30) of R program version 3.03 (R Core Team; <http://www.r-project.org/>). Haplotype frequencies were predicted using *haplo.em* function. Association between asthma diagnosis and haplotypes was tested by *haplo.glm* function, with asthma diagnosis being input as binomial trait and age and sex as covariates. Web-based program SNP Function Prediction (FuncPred) from SNPinfo server (<http://snpinfom.nih.gov/>) (29) was run in the default settings to predict functions of significant SNPs on chromosome 2q14.

Interactions among SNPs on 2q14 and 17q21 loci were analyzed by generalized multifactor dimensionality reduction (GMDR) version 0.9 (<http://www.healthsystem.virginia.edu/internet/addiction-genomics/software/>), which is applicable to both continuous and dichotomous phenotypes (see details in online supplement) (31). GMDR

generates score-based statistics using maximum-likelihood estimates and classified subjects into low-risk and high-risk groups, with age and sex being adjusted as covariates in our analysis.

Results

Association between asthma and autosomal SNPs

Table 1 summarizes the clinical and laboratory characteristics of our 510 asthmatics and 510 controls.

Asthmatic children had higher plasma total IgE levels and peripheral blood eosinophil percentage than controls. One hundred and ten SNPs on eight chromosomal regions were genotyped. All SNPs were genotyped with >90% efficiency (Table S1). Table S2 summarizes the association between asthma diagnosis and the 10 SNPs on 2q14 locus. Five of these 10 genotyped SNPs showed significant associations, the strongest one being rs6755028 (OR 0.67, 95% CI 0.51-0.88, $P = 0.004$). Asthma was also associated with 12 of 16 genotyped SNPs on 17q21 locus (Table S3). Two SNPs passed the regional Bonferroni threshold, namely rs907092 (odds ratio [OR] 0.62, 95% confidence interval [CI] 0.46-0.85, $P = 0.002$) and rs6503527 (OR 0.55, 95% CI 0.40-0.75, $P = 0.0002$). Fig. 1 provides regional association plots between asthma and SNPs on 2q14 and 17q21 loci. We could not detect any asthma association in the remaining six loci (Table S4).

Fig. 2 illustrates two haplotype blocks for 2q14 and four haplotype blocks for 17q21. Table 2 summarizes the haplotypic associations with asthma diagnosis. Two protective haplotypes on 2q14 were found. The strongest haplotype ACTCC for asthma was formed from rs7599210, rs878461,

rs1799810, rs1568277 and rs671535 (OR 0.61, 95% CI 0.44-0.85; $P = 0.003$). Four protective haplotypes on chromosome 17q21 was identified. The strongest haplotype GG formed from rs2952156 and rs1058808 had OR 0.72 and 95% CI 0.57-0.91 for asthma ($P = 0.006$). GMDR analyses on 2q14 and 17q21 loci did not reveal any interaction among SNPs for asthma diagnosis (Table S5).

PROC SNPs and plasma protein C concentrations

The median (interquartile range) plasma PC concentrations in $\mu\text{g/ml}$ were 17.8 (15.5-20.8) in 177 cases and 17.1 (14.5-20.7) in 178 controls ($P = 0.028$). Five of the 10 genotyped *PROC* SNPs were significantly associated with plasma PC concentrations (Table 3).

Bioinformatics analysis and EMSA of SNPs on 2q14 locus

All genotyped SNPs positioned in chromosome 2q14 were fed into SNPinfo server. Rs1799810 and rs878461 were predicted to be part of a transcription factor binding site. Twenty transcription factor binding matrices for rs878461 and 21 transcription factor binding matrices for rs1799810 were predicted (Table S6). Fig. 3 illustrates EMSA results for rs878461. The probes representing this SNP showed binding ability towards nuclear extracts from BEAS-2B cells by producing shifted bands in EMSA. Differential binding affinity was demonstrated by drastic difference between shifted bands of

This article is protected by copyright. All rights reserved.

C allele and T allele. The T allele showed greater nuclear extract binding ability than C allele (lanes 3 and 4). The intensity of shifted probes was greatly diminished following pre-incubation of 100- and 200-fold molar excess cold probe, suggesting such binding was specific (lanes 5 to 8). EMSA performed for rs878461 on HepG2 liver cell line yielded the same results as BEAS-2B cell line, but similar experiments for rs1799810 did not reveal any specific differential binding of the probe (Fig. S1).

Discussion

This case-control genetic study of 110 tagging SNPs on all chromosomal loci associated with asthma in Caucasian children revealed five significant SNPs related to *PROC* on 2q14 locus to be associated with childhood asthma in our population. Plasma PC levels were higher among the patients, which were also associated with five *PROC* SNPs. Rs1799810 and rs878461 were predicted to be part of a transcription factor binding site, and EMSA showed differential binding affinity between C and T alleles of rs878461 for nuclear extracts from BEAS-2B and HepG2 cell lines. Two haplotypes on 2q14 were protective for asthma, supporting *PROC* to be a novel asthma candidate gene. Our findings also confirmed 17q21 to be the locus most strongly associated with asthma in our Chinese children.

Human protein C gene (*PROC*) [OMIM number 612283] is located on chromosome 2q13-14 locus. The genomic sequence, 10802-bp long, contains nine exons with the first one being untranslated (32). Studies have shown that polymorphisms on *PROC* promoter were associated with

Accepted Article

circulating PC level and thrombotic risk (33,34). Protein C is a vitamin K-dependent glycoprotein precursor. Upon activation, PC inhibits coagulation by inactivating factors Va and VIIIa (35). This system was also implicated in asthma pathophysiology (36). Activated PC to PC ratio was decreased in induced sputum of asthma patients (37). Thrombin from airway secretions of these patients was able to stimulate the proliferation of bronchial smooth muscle cells, thereby suggesting that thrombin might be involved in airway remodeling. *In vitro* studies showed that activated PC and PC inhibited eosinophil (38) and lymphocyte (39) migration towards chemoattractants. Increased thrombin level and decreased activated PC level were demonstrated in bronchoalveolar lavage from mice with ovalbumin-induced bronchial asthma (40). Inhalation of activated PC inhibited airway hyperresponsiveness and inflammation. Activated PC may protect against asthmatic inflammation by inactivating factor Va, which is the active product of thrombin. Our result of higher plasma PC levels among asthmatic children was also consistent with a Polish study in which adult patients with concomitant asthma and atopic dermatitis had higher plasma PC levels than healthy controls or patients with atopic dermatitis only (41). These findings support fibrinolytic pathway to be important in airway inflammation and asthma development. On the other hand, it is not clear how *PROC* is related to airway remodeling since no biopsies were taken from our asthmatic children.

Protein C is mostly expressed in human liver whereas human lung has low-to-moderate PC expression (42). Thus, liver cell nucleus should be the optimal source containing its relevant transcription factors for performing EMSA experiments of differential binding by different alleles of

PROC SNPs. On the other hand, such results may not be the same as what happens in bronchial epithelial cells among patients with asthma. Low PC mRNA levels were demonstrated in the lung of mice, with PC antigen being specifically detected in bronchial epithelial cells (43). PC expression was also demonstrated in human bronchial epithelial cell line (BEAS-2B cells) *in vitro* (37). Therefore, our EMSA for the functionality of two *PROC* SNPs was performed on both BEAS-2B lung cell line and HepG2 liver cell line to mimic the *in vivo* sites with airway inflammation and to provide the sufficient set of transcription factors for expression of PC. We found the same results for both liver cell line, which was shown to have the highest expression of PC, as well lung cell line which is the target tissue in patients with asthma.

In a GWAS involving African and European Americans, chromosome 2q13-14 locus was reported to be associated with plasma PC levels (44). In particular, the genetic associations for rs1158867 and rs1799810 reached genome-wide significance. These two SNPs were in high LD in this GWAS ($r^2 = 0.99$) as well as from HapMap data ($r^2 = 1.0$ in CHB population). Our study detected significant associations between plasma PC levels and five SNPs, including rs1799810, located in *PROC* on 2q14. Whereas rs1158867 is located in the DNase hypersensitive and transcription factor binding site (by ENCODE dataset: <https://genome.ucsc.edu/ENCODE/index.html>) of *PROC*, the genetic information of this functional SNP was captured by our tagging SNP rs1799810 due to their LD. Our EMSA results also supported rs878461 to be functional in having differential transcription factor binding by its C and T alleles. Besides, a number of SNPs on this locus as well as their

haplotypes had inverse associations with asthma diagnosis in our Chinese children (Tables 2 and S2). These findings suggest that genetic variants of *PROC* on 2q14 locus modulate the susceptibility for childhood asthma through differential transcription factor binding. The 3-SNP haplotype (rs7599210, rs878461, and rs1799810) associated with reduced asthma risk was located in the 5'-regulatory region of *PROC*, future studies may strengthen our finding by testing for functional differences between constructs with our identified haplotypes.

Previous replication results had consistently found significant genetic associations between SNPs on 17q21 and childhood asthma (7-22). Our results supported the importance of 17q21 locus. Six of our 16 genotyped SNPs were also reported in previous studies, and our results also showed the same direction of association and comparable ORs for asthma diagnosis (Table S7).

Plasma PC levels in our subjects were much higher than that published among White populations (4.8 µg/ml, or normalized as 100% of pooled normal plasma) (45). Zhu et al. reported normal range of plasma PC activity but not PC antigen levels among 3,493 healthy Chinese adults (46). Future studies need to confirm our finding of higher plasma PC levels in other Chinese populations and explore the reason(s) for interethnic difference in PC levels.

The most significant weakness of this study was the lack of independent replication populations to confirm our findings for the 2q14 locus. Thus, we are unclear whether such findings are generalizable to all populations or ethnic-specific, and whether *PROC* is important for childhood asthma or all patients irrespective of their age of asthma onset. Secondly, SNPs on 2q14 that were the

most strongly associated with asthma (Table S2) are not the same as those associated with plasma PC levels (Table 3). Nonetheless, three SNPs closely related to *PROC*, namely rs6755028, rs13408910 and rs10928765, were common to both phenotypes. The first two of these SNPs also had the strongest associations with asthma and PC levels respectively. These results support that childhood asthma is associated with either *PROC* or a neighboring gene with SNPs in LD with the above SNPs.

Whereas this replication study included most significant SNPs as reported by the original GWAS (2), our study was also limited by the unavailability of TaqMan Genotyping Assays for rs8179521 on 2q14 and rs3791244 on 2q22. Therefore, these two original SNPs were not genotyped (Table S4). Rs8179521 on 2q14 locus was also not selected as tagging SNP because of its low MAF in Chinese. Another limitation was that rs2069906 on 2q14 was predicted by FuncPred to be in perfect LD with our asthma-related SNP rs878461. Analysis by RegulomeDB on the ENCODE datasets revealed that rs2069906 binds foxa2, a transcription factor known to regulate liver-specific genes, in HepG2 cells (47). Nonetheless, this study did not genotype rs2069906 as it was not identified as a tagging SNP in our population. We found high atopy rate among the controls, which seemed to be a consistent finding among our several studies involving independent groups of non-allergic Chinese from hospitals and community (19,26,48-50). Our controls were older than the asthmatics, and such age difference was expected given that these controls were secondary schoolchildren who participated in our obesity and diabetic studies (26). Because some children acquire allergies only during adolescence, we deliberately chose these older controls so as to be confident that they were free from

This article is protected by copyright. All rights reserved.

any allergic disease. Lastly, as lung function was only measured in asthmatics but not controls, this study was unable to analyze any association between SNPs on 2q14 and this important intermediate phenotypes of asthma.

In conclusions, this case-control association study not only replicated the genetic association between childhood asthma and polymorphic markers on 17q21 locus but also identified novel associations between asthma and several SNPs of *PROC* on 2q14 locus in Chinese children. Our results demonstrate rs878461 to possess differential and specific binding for nuclear extracts of hepatic and bronchial epithelial cells.

Acknowledgments

This study was funded by General Research Fund (reference no. 470909) of Hong Kong Research Grants Council and Research Committee Group Research Scheme (reference no. 3110034 and 3110060) of The Chinese University of Hong Kong. We thanked Chung Yi Li, Edmund Yung, Yun Sze Wong, Man Fung Tang and Susan Shuxin Wang for helping with subject assessment, sample processing and genotyping experiments.

Conflicts of interest

All authors declare no conflict of interest.

This article is protected by copyright. All rights reserved.

References

1. Von Mutius E. Gene-environment interactions in asthma. *J Allergy Clin Immunol* 2009; **123**: 3-11.
2. Moffatt MF, Kabisch M, Liang L, et al. Genetic variants regulating *ORMDL3* expression contribute to the risk of childhood asthma. *Nature* 2007; **448**: 470-3.
3. Sleiman PM, Flory J, Imielinski M, et al. Variants of *DENND1B* associated with asthma in children. *N Engl J Med* 2010; **362**: 36-44.
4. Ober C, Tan Z, Sun Y, et al. Effect of variation in *CHI3L1* on serum YKL-40 level, risk of asthma, and lung function. *N Engl J Med* 2008; **358**: 1682-91.
5. 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.
6. Li X, Howard TD, Zheng SL, et al. Genome-wide association study of asthma identifies RAD50-IL13 and HLA-DR/DQ regions. *J Allergy Clin Immunol* 2010; **125**: 328-35.
7. Du R, Litonjua AA, Tantisira KG, et al. Genome-wide association study reveals class I MHC-restricted T cell-associated molecule gene (*CRTAM*) variants interact with vitamin D levels to affect asthma exacerbations. *J Allergy Clin Immunol* 2012; **129**: 368-73.

8. Li X, Ampleford EJ, Howard TD, et al. Genome-wide association studies of asthma indicate opposite immunopathogenesis direction from autoimmune diseases. *J Allergy Clin Immunol* 2012; **130**: 861-8.
9. Moffatt MF, Gut IG, Demenais F, et al. A large-scale, consortium-based genomewide association study of asthma. *N Engl J Med* 2010; **363**: 1211-21.
10. Noguchi E, Sakamoto H, Hirota T, et al. Genome-wide association study identifies HLA-DP as a susceptibility gene for pediatric asthma in Asian populations. *PLoS Genet* 2011; **7**: e1002170.
11. 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**: 581-93.
12. Himes BE, Sheppard K, Berndt A, et al. Integration of mouse and human genome-wide association data identifies *KCNIP4* as an asthma gene. *PLoS One* 2013; **8**: e56179.
13. Ding L, Abebe T, Beyene J, et al. Rank-based genome-wide analysis reveals the association of ryanodine receptor-2 gene variants with childhood asthma among human populations. *Hum Genomics* 2013; **7**: 16.
14. Bønnelykke K, Sleiman P, Nielsen K, et al. A genome-wide association study identifies *CDHR3* as a susceptibility locus for early childhood asthma with severe exacerbations. *Nat Genet* 2014; **46**: 51-5.

- Accepted Article
15. Galanter JM, Gignoux CR, Torgerson DG, et al. Genome-wide association study and admixture mapping identify different asthma-associated loci in Latinos: The Genes-environments & Admixture in Latino Americans study. *J Allergy Clin Immunol* 2014; **134**: 295-305.
 16. Halapi E, Gudbjartsson DF, Jonsdottir GM, et al. A sequence variant on 17q21 is associated with age at onset and severity of asthma. *Eur J Hum Genet* 2010; **18**: 902-8.
 17. Smit LM, Bouzigon E, Pin I, et al. 17Q21 variants modify the association between early respiratory infections and asthma. *Eur Respir J* 2010; **36**: 57-64.
 18. Berce V, Kozmus CE, Potočnik U. Association among *ORMDL3* gene expression, 17q21 polymorphism and response to treatment with inhaled corticosteroids in children with asthma. *Pharmacogenomics J* 2012; **13**: 1-7.
 19. Leung TF, Sy HY, Ng MC, et al. Asthma and atopy are associated with chromosome 17q21 markers in Chinese children. *Allergy* 2009; **64**: 621-8.
 20. Bisgaard H, Bønnelykke K, Sleiman PM, et al. Chromosome 17q21 gene variants are associated with asthma and exacerbations but not atopy in early childhood. *Am J Respir Crit Care Med* 2009; **179**: 179-85.
 21. Binia A, Khorasani N, Bhavsar PK, et al. Chromosome 17q21 SNP and severe asthma. *J Hum Genet* 2011; **56**: 97-8.

- Accepted Article
22. Li F-X, Tan J-Y, Yang X-X, Wu Y-S, Wu D, Li M. Genetic variants on 17q21 are associated with asthma in a Han Chinese population. *Genet Mol Res* 2012; **11**: 340-7.
 23. American Thoracic Society. Standards for the diagnosis and care of patients with Chronic Obstructive Pulmonary Disease (COPD) and asthma. *Am Rev Respir Dis* 1987; **136**: 225-44.
 24. Wong GW, Leung TF, Ko FW, et al. Declining asthma prevalence in Hong Kong Chinese schoolchildren. *Clin Exp Allergy* 2004; **34**: 1550-5.
 25. Leung TF, Yung E, Wong YS, Lam CW, Wong GW. Parent-reported adverse food reactions in Hong Kong Chinese preschoolers: epidemiology, clinical spectrum and risk factors. *Pediatr Allergy Immunol* 2009; **20**: 339-46.
 26. Leung TF, Kong AP, Chan IH, et al. Association between obesity and atopy in Chinese schoolchildren. *Int Arch Allergy Immunol* 2009; **149**: 133-40.
 27. Frazer K, Ballinger DG, Cox DR, et al. A second generation human haplotype map of over 3.1 million SNPs. *Nature* 2007; **449**: 851-61.
 28. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 2005; **21**: 263-5.

29. Xu Z, Taylor JA. SNPinfo: integrating GWAS and candidate gene information into functional SNP selection for genetic association studies. *Nucleic Acids Res* 2009; **37**: W600-5.
30. Schaid DJ, Rowland CM, Tines DE, Jacobson RM, Poland G. Score tests for association between traits and haplotypes when linkage phase is ambiguous. *Am J Hum Genet* 2002; **70**: 425-34.
31. Lou X-Y, Chen G-B, Yan L, et al. A generalized combinatorial approach for detecting gene-by-gene and gene-by-environment interactions with application to nicotine dependence. *Am J Hum Genet* 2007; **80**: 1125-37.
32. Plutzky J, Hoskins JA, Long GL, Crabtree GR. Evolution and organization of the human protein C gene. *Proc Natl Acad Sci U S A* 1986; **83**: 546-50.
33. Aiach M, Nicaud V, Alhenc-Gelas M, et al. Complex association of protein C gene promoter polymorphism with circulating protein C levels and thrombotic risk. *Arterioscler Thromb Vasc Biol* 1999; **19**: 1573-6.
34. Reiner AP, Carty CL, Jenny NS, et al. *PROC*, *PROCR* and *PROS1* polymorphisms, plasma anticoagulant phenotypes, and risk of cardiovascular disease and mortality in older adults: the Cardiovascular Health Study. *J Thromb Haemost* 2008; **6**: 1625-32.
35. Esmon CT. The roles of protein C and thrombomodulin in the regulation of blood coagulation. *J Biol Chem* 1989; **264**: 4743-6.

- Accepted Article
36. Danese S, Vetrano S, Zhang L, Poplis VA, Castellino FJ. The protein C pathway in tissue inflammation and injury: pathogenic role and therapeutic implications. *Blood* 2010; **115**: 1121-30.
37. Hataji O, Taguchi O, Gabazza EC, et al. Activation of protein C pathway in the airways. *Lung* 2002; **180**: 47-59.
38. Feistritzer C, Sturn DH, Kaneider NC, Djanani A, Wiedermann CJ. Endothelial protein C receptor-dependent inhibition of human eosinophil chemotaxis by protein C. *J Allergy Clin Immunol* 2003; **112**: 375-81.
39. Feistritzer C, Mosheimer B, Sturn DH, Riewald M, Patsch JR, Wiedermann CJ. Endothelial protein C receptor-dependent inhibition of migration of human lymphocytes by protein C involves epidermal growth factor receptor. *J Immunol* 2006; **176**: 1019-25.
40. Yuda H, Adachi Y, Taguchi O, et al. Activated protein C inhibits bronchial hyperresponsiveness and Th2 cytokine expression in mice. *Blood* 2004; **103**: 2196-204.
41. Undas A, Potaczek DP, Nishiyama C, Okumura K. Non-severe allergic asthma is associated with elevated plasma protein C and protein S. *Thromb Haemost* 2012; **107**: 1000-2.
42. Shmueli O, Horn-Saban S, Chalifa-Caspi V, et al. GeneNote: whole genome expression profiles in normal human tissues. *C R Biol* 2003; **326**: 1067-72.

43. Yamamoto K, Loskutoff DJ. Extrahepatic expression and regulation of protein C in the mouse. *Am J Pathol* 1998; **153**: 547-55.
44. Tang W, Basu S, Kong X, et al. Genome-wide association study identifies novel loci for plasma levels of protein C: the ARIC study. *Blood* 2010; **116**: 5032-6.
45. Clouse L, Comp P. The regulation of hemostasis: the protein C system. *N Engl J Med* 1986; **314**: 1298-304.
46. Zhu T, Ding Q, Bai X, et al. Normal ranges and genetic variants of antithrombin, protein C and protein S in the general Chinese population. Results of the Chinese Hemostasis Investigation on Natural Anticoagulants Study I Group. *Haematologica* 2011; **96**: 1033-40.
47. Boyle AP, Hong EL, Hariharan M, et al. Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res* 2012; **22**: 1790-7.
48. Leung TF, Liu EK, Tang NL, et al. Nitric oxide synthase polymorphisms and asthma phenotypes in Chinese children. *Clin Exp Allergy* 2005; **35**: 1288-94.
49. Ko FW, Leung TF, Wong GW, Chu JH, Sy HY, Hui DS. Determinants of, and reference equation for, exhaled nitric oxide in the Chinese population. *Eur Respir J* 2013; **42**: 767-75.
50. Wang SS, Hon KL, Kong AP, et al. Eczema phenotypes are associated with multiple vitamin D pathway genes in Chinese children. *Allergy* 2014; **69**: 118-24.

Table 1 Subjects' clinical and laboratory characteristics

Characteristic	Asthmatics (n=510)	Controls (n=510)
Age, years	10.7 ± 4.1 [†]	16.0 ± 4.6
Age at asthma diagnosis, years	3.1 ± 2.5	NA
Male, n (%)	311 (61.0) [†]	241 (47.3)
Log ₁₀ -transformed plasma total IgE levels	2.52 ± 0.65 [†]	1.81 ± 0.68
Peripheral blood eosinophils*, %	6.8 ± 4.2 [†]	3.4 ± 3.0
Atopy [§] , n (%)	404 (86.5) [†]	228 (46.2)
Spirometric parameters		
FEV ₁ , % predicted	90.3 ± 14.0	NA
FVC, % predicted	99.7 ± 16.0	NA
FEV ₁ /FVC	0.82 ± 0.11	NA
Log ₁₀ -transformed FeNO	1.77 ± 0.36	NA

FeNO, fractional exhaled nitric oxide concentration; FEV₁, forced expiratory volume in 1-second;

FVC, forced vital capacity; IgE, immunoglobulin E; NA, not applicable.

This article is protected by copyright. All rights reserved.

Results expressed in number (percentage) or mean $\pm$ standard deviation.

*As percentage of total peripheral blood leukocytes. [†] $P < 0.001$.

[§]Atopy was defined as having one or more positive results for skin prick test or allergen-specific IgE to locally prevalent aeroallergens.

Table 2 Association between asthma diagnosis and haplotypes on 2q14 and 17q21 loci

Locus	SNP combination	Haplotype	Frequencies		OR (95% CI)	<i>P</i>
			Asthma	Control		
Chr. 2q14	rs7599210,	GTACG	0.564	0.541	ref	
	rs878461,		0.111	0.160	0.61 (0.44-0.85)	0.003
	rs1799810,	ACTCC	0.249	0.230	1.05 (0.82-1.36)	0.693
	rs1568277,	GTATC				
	rs671535		0.037	0.039	1.25 (0.72-2.17)	0.429
		GTTCC				
		AG	0.570	0.557	ref	
	rs6755028,	AT	0.273	0.233	1.22 (0.96-1.56)	0.105
	rs6749002					
		GT	0.150	0.210	0.67 (0.5-0.88)	0.005
Chr.	rs2952156,	AC	0.716	0.652	ref	

17q21	rs1058808	GG	0.273	0.341	0.72 (0.57-0.91)	0.006
	rs907092,	GCGGGGA	0.681	0.630	ref	
	rs9303277,	ATGGATG	0.126	0.169	0.66(0.48-0.9)	0.008
	rs2060941,	ATGGGGA	0.011	0.014	0.69(0.26-1.8)	0.449
	rs11078924,	ATTGGTG	0.050	0.058	0.86(0.54-1.38)	0.539
	rs12150079,	GCGAGGA	0.080	0.102	0.8(0.56-1.16)	0.242
	rs11557467,					
	rs8067378					
	rs2305480,	GTC	0.777	0.745	ref	
	rs2290400,	ACT	0.185	0.223	0.78(0.59-1.02)	0.074
	rs4795405	GTT	0.015	0.014	0.83(0.35-1.98)	0.678
		GA	0.520	0.467	ref	
	rs8077456,	CA	0.376	0.404	0.81(0.64-1.03)	0.089
	rs921651	GG	0.102	0.130	0.62(0.43-0.88)	0.007

CI, confidence interval; OR, odds ratio; ref, reference.

All logistic regression analyses were adjusted for age and sex as covariates.

Table 3 Genetic associations between *PROC* SNPs and plasma protein C concentrations in 355

randomly selected subjects.

SNP	Plasma protein C levels (μg/ml)			<i>P</i> *
	Homozygous major allele	Heterozygous	Homozygous minor allele	
rs10928764	CC	GC	GG	0.180
rs7599210	GG	GA	AA	0.095
rs878461	TT	TC	CC	0.055
rs1799810	AA	TA	TT	0.029
rs1568277	CC	TC	TT	0.050
rs6710535	GG	GC	CC	0.297
rs10928765	GG	GA	AA	0.022
rs6755028	AA	GA	GG	0.007
rs6749002	GG	TG	TT	0.127
rs13408910	TT	TC	CC	0.004

PROC, protein C gene; SNP, single-nucleotide polymorphism.

Results expressed in median (interquartile range).

This article is protected by copyright. All rights reserved.

*Analyzed by linear regression adjusting for age, sex, body mass index and asthma status (yes vs no) as covariates.

Figures

Figure 1 Distribution of $-\log(p)$ values of genetic associations with asthma diagnosis for SNPs on (a) 2q14 and (b) 17q21 loci.

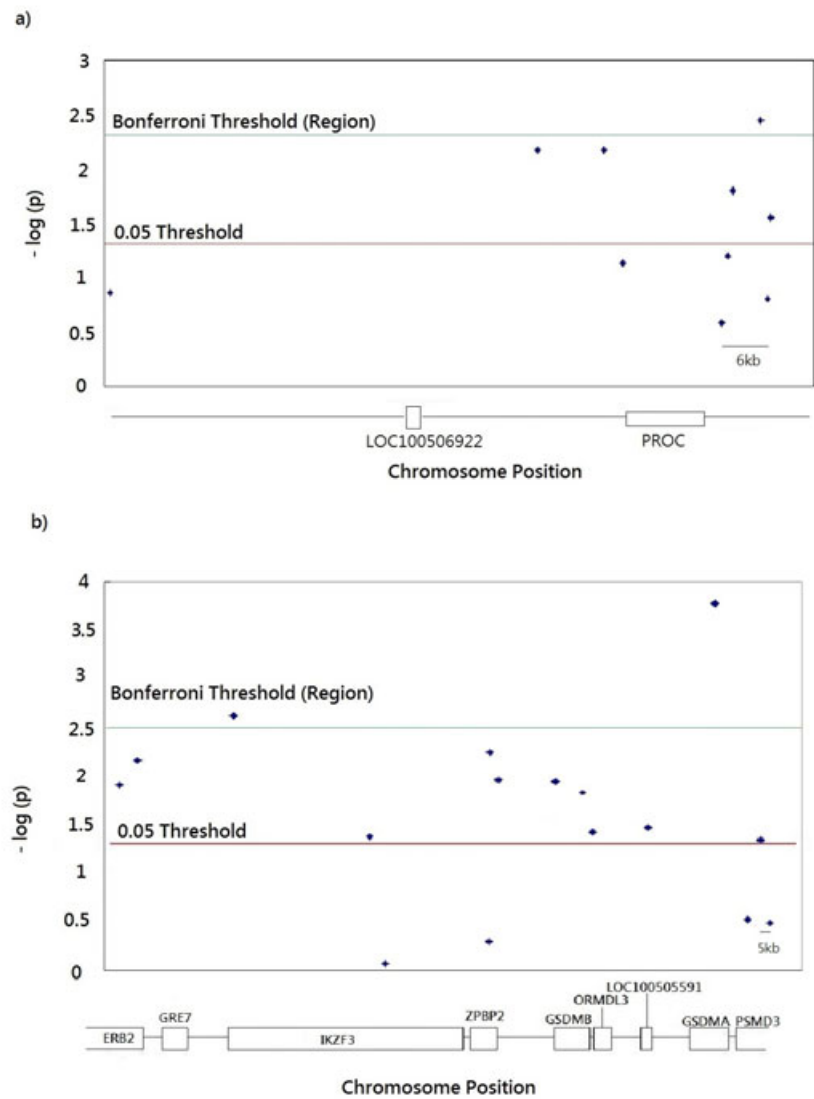
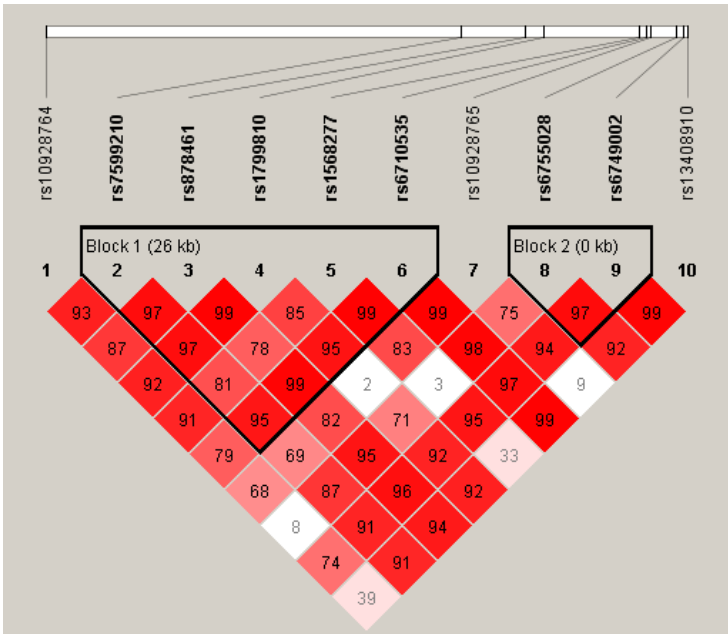


Figure 2 Linkage disequilibrium plots of SNPs on chromosome (a) 2q14 and (b) 17q21 loci. A five-SNP block flanking 26-kb and a two-SNP block flanking less than 1-kb were formed on 2q14, whereas a two-SNP block flanking 7-kb, a seven-SNP block flanking 129-kb, a three-SNP block flanking 26-kb and a two-SNP block flanking 5-kb were assigned for 17q21.

(a) Chromosome 2q14



(b) Chromosome 17q21

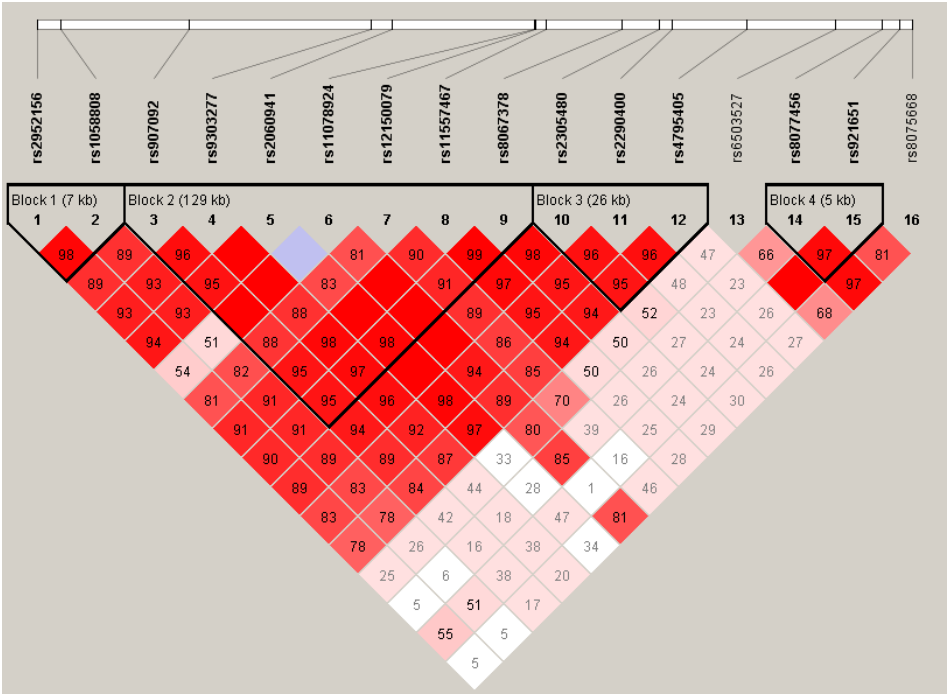
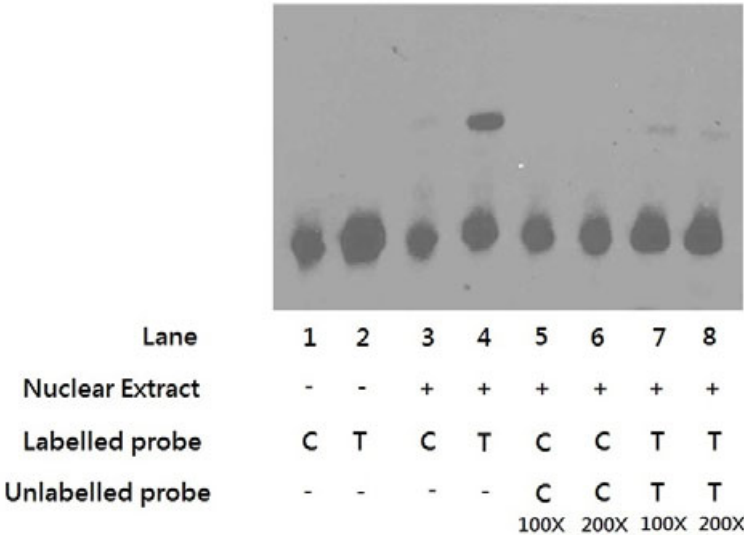


Figure 3 EMSA showing allele-specific protein-DNA binding at *PROC* involving C and T allele of rs878461 using nuclear extracts from BEAS-2B cell line. Allele C exhibits less band shift than allele T (lanes 3 and 4). Incubating excess unlabeled competitor in reaction mixture results in diminished band shift (lanes 5 to 8).



ONLINE SUPPLEMENT

Childhood Asthma is Associated with Polymorphic Markers of *PROC* on 2q14 in Addition to 17q21 Locus

Wa Cheong Chan¹, Hing Yee Sy¹, Alice Pik-shan Kong², Chun Kwok Wong³, Lai Yin Tse¹,
Kam Lun Hon¹, Juliana Chung-ngor Chan², Gary Wing-kin Wong¹ & Ting Fan Leung¹

Departments of ¹Paediatrics, ²Medicine and Therapeutics and ³Chemical Pathology, The
Chinese University of Hong Kong, Prince of Wales Hospital, Shatin, Hong Kong

GMDR analyses for SNP-SNP interactions

Generalized multifactor dimensionality reduction (GMDR, version 0.9; Department of Psychiatry and Neurobehavioral Sciences, University of Virginia, Charlottesville, VA, USA; <http://www.healthsystem.virginia.edu/internet/addiction-genomics/software/>) was used to determine high and low risk (S1). This software permits adjustment for discrete and quantitative covariates and is applicable to both continuous and dichotomous phenotypes. In brief, GMDR computes the maximum-likelihood estimates and the score values of all individuals under the null hypothesis. The cumulative score value was calculated within each multifactor cell, which would be classified either as high-risk if the average score meets or exceeds a pre-assigned threshold of 0 or as low-risk if the score is less than 0. An exhaustive search was performed to test for all possible interactions using 10-fold cross validation. Age and gender were included as covariates. In evaluation of the models, a number of outcome parameters including testing accuracy (TA), cross-validation consistency (CVC) and statistical significance of the model were considered. TA measuring how accurate individuals are classified with respect to their case/control status was used to pick the best model. A TA score of 0.5 indicates that the model predicts no better than chance whereas 1.0 indicates perfect prediction. In general, a TA score ≥ 0.55 for a model was considered satisfactory. CVC evaluates the consistency with which the selected interaction is classified as the best model among all possible combinations.

Permutation tests are considered the gold standard in multiple testing adjustment because they simultaneously provide unbiased type I error control and high power (S2). This method was applied in GMDR to identify the best interactive models. The statistical significance of a model was determined by comparing average prediction error from the observed data with the distribution of average prediction errors under the null hypothesis of no association derived empirically from 5,000 permutations. The null hypothesis was rejected when the permuted P -value was 0.05 or lower.

References

- S1. Lou XY, Chen GB, Yan L, et al. A generalized combinatorial approach for detecting gene-by-gene and gene-by-environment interactions with application to nicotine dependence. *Am J Hum Genet* 2007; **80**: 1125-37.
- S2. Gao X. Multiple testing corrections for imputed SNPs. *Genet Epidemiol* 2011; **35**: 154-8.
- S3. Moffatt MF, Kabesch M, Liang L, et al. Genetic variants regulating *ORMDL3* expression contribute to the risk of childhood asthma. *Nature* 2007; **448**: 470-3.
- S4. Halapi E, Gudbjartsson DF, Jonsdottir GM, et al. A sequence variant on 17q21 is associated with age at onset and severity of asthma. *Eur J Hum Genet* 2010; **18**: 902-8.

Table S1 Summary of genotyping efficiency for SNPs on all chromosomal regions in this study

Chromosomal locus	Number of SNPs genotyped	Successful genotyping (% of samples)				
		90-92	93-94	95-96	97-98	99-100
1q42	13	2	8	1	2	0
2p14	24	7	8	4	5	0
2q14	10	0	0	1	4	5
2q21	11	1	1	1	8	0
8p12	10	0	0	2	8	0
10p12	3	0	0	0	3	0
17q12-21	16	2	11	1	2	0
21q21	23	1	1	10	11	0
<i>TOTAL</i>	<i>110</i>	<i>13</i>	<i>29</i>	<i>20</i>	<i>43</i>	<i>5</i>

Table S2 Genetic associations between asthma and SNPs on chromosome 2q14 locus

rs number	Position on Chrom 2	Gene	Alleles	OR (95% CI)	<i>P</i> *	Model
rs10928764	128103801	<i>MAP3K2</i> Promoter	C/G	1.20 (0.94-1.52)	0.136	Additive
rs7599210	128164177	Intergenic	A/G	0.66 (0.49-0.89)	0.007	Additive
rs878461	128173417	<i>PROC</i> Promoter	C/T	0.65 (0.48-0.89)	0.007	Additive
rs1799810	128176040	<i>PROC</i> 5' UTR	A/T	0.75 (0.55-1.03)	0.073	Dominant
rs1568277	128190055	<i>PROC</i> Downstream	C/T	1.19 (0.88-1.61)	0.260	Dominant
rs6710535	128190960	<i>PROC</i> Downstream	C/G	0.70 (0.48-1.02)	0.063	Recessive
rs10928765	128191613	<i>PROC</i> Downstream	A/G	1.44 (1.07-1.94)	0.016	Dominant
rs6755028	128195442	<i>PROC</i> Downstream	A/G	0.67 (0.51-0.88)	0.004	Additive
rs6749002	128196424	<i>PROC</i> Downstream	G/T	0.77 (0.53-1.11)	0.155	Recessive
rs13408910	128196930	Intergenic	C/T	0.75 (0.58- 0.97)	0.028	Additive

CI, confidence interval; *MAP3K2*, mitogen-activated protein kinase kinase kinase 2 gene; OR, odds ratio; *PROC*, protein C gene; UTR, untranslated region.

*Analyzed by logistic regression and adjusted for age and sex as covariates.

Table S3 Association between asthma diagnosis and SNPs on chromosome 17q12-21 locus

rs number	Position on chr. 17	Gene	Alleles	OR (95% CI)	<i>P</i> *	Model
rs2952156	35130361	<i>ERBB2</i> intron	A/G	0.74 (0.59-0.94)	0.013	Additive
rs1058808	35137563	<i>ERBB2</i> P1170A	C/G	0.72 (0.57-0.92)	0.007	Additive
rs907092	35175785	<i>IKZF3</i> silent mutation	A/G	0.62 (0.46-0.85)	0.002	Dominant
rs9303277	35229995	<i>IKZF3</i> intron	C/T	0.72 (0.53-0.99)	0.043	Dominant
rs2060941	35236409	<i>IKZF3</i> intron	G/T	0.96 (0.59-1.55)	0.857	Dominant
rs11078924	35278543	<i>IKZF3</i> promoter	A/G	0.88 (0.59-1.31)	0.522	Dominant
rs12150079	35278943	<i>IKZF3</i> promoter	A/G	0.62 (0.45-0.87)	0.006	Dominant
rs11557467	35282160	<i>IKZF3</i> promoter / ZPBP2 S173I	G/T	0.66 (0.48-0.91)	0.011	Dominant
rs8067378	35304874	<i>GSDMB</i> Downstream	A/G	0.66 (0.48-0.91)	0.012	Dominant
rs2305480	35315722	<i>GSDMB</i> P302S	G/A	0.67 (0.49-0.93)	0.015	Dominant

rs2290400	35319766	<i>GSDMB</i> intron	T/C	0.72 (0.52-0.98)	0.039	Dominant
rs4795405	35341943	<i>ORMDL3</i> promoter	C/T	0.71 (0.52-0.98)	0.035	Dominant
rs6503527	35368245	<i>GSDMA</i> promoter	A/G	0.55 (0.40-0.75)	<0.001	Dominant
rs8077456	35382291	<i>GSDMA</i> intron / <i>PSMD3</i> promoter	C/G	0.84 (0.61-1.16)	0.300	Dominant
rs921651	35387448	<i>GSDMA</i> 3' UTR / <i>PSMD3</i> promoter	A/G	0.72 (0.51-0.99)	0.046	Additive
rs8075668	35391149	<i>GSDMA</i> downstream/ <i>PSMD3</i> intron	C/T	0.85 (0.61-1.18)	0.323	Dominant

CI, confidence interval; *ERBB2*, V-Erb B2 avian erythroblastic leukemia viral oncogene homolog 2 gene; *GSDMA*, gasdermin A gene; *GSDMB*, gasdermin B gene, *IKZF3*, IKAROS family zinc finger 3 gene ; *ORMDL3*, ORM1-Like 3 gene; OR, odds ratio; *PSMD3*, Proteasome (Prosome, Macropain) 26S Subunit, Non-ATPase gene; UTR, untranslated region.

*Analyzed by logistic regression and adjusted for age and sex as covariates.

Table S4 Minor allele frequencies of our genotyped SNPs among different Asian populations and *P*-values of their genetic associations with asthma diagnosis in our Chinese children

rs number	MAF*			<i>P</i>
	Our controls	JPT	CHB	
Chromosome 1q42				
rs2477599	0.116(A)	0.116(A)	0.114(A)	0.066
rs2695073	0.074(T)	0.143(T)	0.06(T)	0.464
rs2492375	0.075(C)	0.173(C)	0.091(C)	0.553
rs6699293	0.438(G)	0.333(A)	0.44(A)	0.584
rs6662005	0.367(A)	0.438(A)	0.431(A)	0.402
rs1726621	0.086(G)	0.196(G)	0.121(G)	0.757
rs10924993 [#]	0.018(G)	0.053(G)	0.055(G)	0.989
rs12033469	0.059(T)	0.167(T)	0.078(T)	0.294
rs7537831	0.418(T)	0.451(C)	0.459(T)	0.427
rs6428951	0.428(G)	0.364(G)	0.367(G)	0.103
rs12092396	0.121(T)	0.062(T)	0.095(T)	0.270
rs12134040	0.018(G)	0.056(G)	0.057(G)	0.817
rs609516	0.156(G)	0.128(G)	0.12(G)	0.410
Chromosome 2p13				
rs11695768	0.227(C)	0.261(C)	0.292(C)	0.563
rs7608529	0.278(G)	0.144(G)	0.144(G)	0.372
rs7422789	0.384(C)	0.221(C)	0.209(C)	0.208
rs12474671	0.361(T)	0.411(T)	0.444(T)	0.209
rs7586350	0.103(G)	0.089(G)	0.078(G)	0.103
rs4610084	0.281(G)	0.186(G)	0.201(G)	0.268
rs9808077	0.371(T)	0.233(T)	0.2(T)	0.464
rs4462818	0.381(T)	0.451(T)	0.427(T)	0.648
rs7557185	0.482(A)	0.438(G)	0.474(G)	0.546
rs13396876	0.370(C)	0.289(C)	0.205(C)	0.860
rs11126210	0.162(G)	0.176(G)	0.171(G)	0.159
rs12468876	0.329(T)	0.412(T)	0.361(T)	0.195
rs6716266 [#]	0.239(G)	0.031(G)	0.018(G)	0.002
rs4597566	0.168(C)	0.186(C)	0.164(C)	0.115
rs4386359	0.371(G)	0.375(G)	0.364(G)	0.307
rs6546471	0.423(C)	0.423(C)	0.412(C)	0.215
rs4608564	0.280(C)	0.344(C)	0.3(C)	0.098
rs4314026	0.245(T)	0.248(T)	0.279(T)	0.553

rs11126213	0.236(G)	0.199(G)	0.197(G)	0.742
rs11126214	0.246(T)	0.214(T)	0.192(T)	0.032
rs4255990	0.484(C)	0.469(C)	0.496(C)	0.123
rs6546477	0.437(C)	0.46(A)	0.474(C)	0.878
rs12464309	0.486(T)	0.411(T)	0.4(T)	0.321
rs12479132	0.317(T)	0.233(T)	0.256(T)	0.434

Chromosome 2q14

rs10928764	0.270(G)	0.261(G)	0.344(G)	0.136
rs7599210	0.153(A)	0.108(A)	0.13(A)	0.007
rs878461	0.143(C)	0.075(C)	0.117(C)	0.007
rs1799810	0.193(T)	0.159(T)	0.2(T)	0.073
rs1568277	0.245(T)	0.221(T)	0.318(T)	0.260
rs6710535	0.438(C)	0.378(C)	0.485(C)	0.063
rs10928765	0.262(A)	0.252(A)	0.314(A)	0.016
rs6755028	0.182(G)	0.115(G)	0.186(G)	0.004
rs6749002	0.432(T)	0.367(T)	0.482(T)	0.155
rs13408910	0.213(C)	0.168(C)	0.223(C)	0.028

Chromosome 2q22

rs4954548	0.311(A)	0.322(A)	0.189(A)	0.589
rs3113210	0.410(C)	0.279(C)	0.434(C)	0.537
rs16839974	0.180(G)	0.056(G)	0.222(G)	0.077
rs13432464	0.461(C)	0.403(T)	0.449(T)	0.295
rs6755148	0.248(G)	0.216(G)	0.2(G)	0.910
rs6430764	0.389(C)	0.35(C)	0.365(C)	0.081
rs3791242	0.087(C)	0.035(C)	0.124(C)	0.034
rs1020678	0.371(C)	0.305(C)	0.358(C)	0.112
rs17583889	0.126(A)	0.071(A)	0.12(A)	0.512
rs1020677	0.371(A)	0.261(A)	0.233(A)	0.091
rs3828168	0.254(T)	0.23(T)	0.248(T)	0.080

Chromosome 8p12

rs2919381	0.488(G)	0.451(A)	0.467(A)	0.289
rs4445183	0.136(T)	0.142(T)	0.102(T)	0.212
rs2919384	0.428(G)	0.336(G)	0.372(G)	0.213
rs16879809	0.347(T)	0.425(T)	0.42(T)	0.047
rs17669617	0.075(G)	0.111(G)	0.089(G)	0.022

rs3757932	0.326(T)	0.288(T)	0.287(T)	0.061
rs4512342 [#]	0.405(G)	0.207(T)	0.364(G)	0.413
rs16879922	0.379(T)	0.447(T)	0.449(T)	0.080
rs6992642	0.184(C)	0.119(C)	0.132(C)	0.043
rs4445184	0.360(T)	0.385(T)	0.369(T)	0.137
Chromosome10p12				
rs11012819	0.152(T)	0.071(T)	0.124(T)	0.294
rs16921965	0.202(G)	0.357(G)	0.342(G)	0.738
rs2666781 [#]	0.052(C)	0(C)	0.024(C)	0.065
Chromosome 17q12-21				
rs2952156	0.313(G)	0.386(A)	0.489(A)	0.013
rs1058808	0.311(G)	0.182(G)	0.156(G)	0.007
rs907092 [#]	0.231(A)	0.345(A)	0.328(A)	0.002
rs9303277	0.226(T)	0.35(T)	0.347(T)	0.043
rs2060941	0.060(T)	0.062(T)	0.085(T)	0.857
rs11078924	0.099(A)	0.024(A)	0.097(A)	0.522
rs12150079	0.167(A)	0.226(A)	0.245(A)	0.006
rs11557467 [#]	0.213(T)	0.277(T)	0.324(T)	0.011
rs8067378 [#]	0.216(G)	0.283(G)	0.328(G)	0.012
rs2305480 [#]	0.214(A)	0.281(A)	0.331(A)	0.015
rs2290400 [#]	0.218(C)	0.289(C)	0.345(C)	0.039
rs4795405 [#]	0.227(T)	0.29(T)	0.358(T)	0.035
rs6503527	0.275(G)	0.359(G)	0.429(G)	<0.001
rs8077456	0.392(C)	0.345(C)	0.438(C)	0.300
rs921651	0.117(G)	0.137(G)	0.106(G)	0.046
rs8075668	0.415(T)	0.339(T)	0.482(T)	0.323
Chromosome 21q21				
rs2831440	0.378(T)	0.429(T)	0.394(T)	0.263
rs974607	0.424(A)	0.442(A)	0.442(A)	0.264
rs2831447	0.295(A)	0.369(A)	0.338(A)	0.714
rs2831451	0.273(C)	0.204(C)	0.219(C)	0.051
rs2831453	0.198(A)	0.119(A)	0.219(A)	0.114
rs2205557	0.129(T)	0.178(T)	0.122(T)	0.807
rs8134384	0.139(C)	0.238(C)	0.143(C)	0.378
rs2831460	0.235(A)	0.248(A)	0.245(A)	0.252

rs4501026	0.341(C)	0.352(C)	0.256(C)	0.420
rs16981663	0.168(T)	0.178(T)	0.216(T)	0.329
rs2831461	0.208(C)	0.198(C)	0.256(C)	0.233
rs2831474	0.110(T)	0.124(T)	0.139(T)	0.022
rs2831475	0.332(T)	0.369(T)	0.324(T)	0.066
rs1997620	0.223(C)	0.1(C)	0.189(C)	0.155
rs2831482	0.074(A)	0.086(A)	0.066(A)	0.088
rs2037986 [#]	0.443(G)	0.398(G)	0.42(G)	0.183
rs2026226	0.086(A)	0.134(A)	0.087(A)	0.222
rs2831498	0.472(G)	0.302(G)	0.412(G)	0.386
rs2831505	0.210(A)	0.261(A)	0.182(A)	0.248
rs10854308	0.210(C)	0.2(C)	0.321(C)	0.252
rs1997828	0.314(A)	0.455(A)	0.267(A)	0.661
rs1883819	0.067(A)	0.068(A)	0.067(A)	0.577
rs2206881	0.122(G)	0.216(G)	0.133(G)	0.417

CHB, Han Chinese in Beijing; JPT, Japanese; MAF, minor allele frequency.

*Letters inside brackets represent minor alleles in the respective populations.

[#] Original SNPs genotyped and reported in the asthma GWAS (reference S3).

P-values that exceeded 5% significance level were highlighted in bold.

Table S5 GMDR analysis for SNP-SNP interactions on 2q14 and 17q21 loci

SNP combination	TA (%)	CVC
Chr. 17q_rs6503527	57.45	10
Chr. 2q_rs7599210, Chr. 17q_rs6503527	56.72	6
Chr. 2q_rs878461, Chr. 17q_rs1058808, Chr. 17q_rs6503527]	54.74	5
Chr. 17q_rs1058808, Chr. 17q_rs907092, Chr. 17q_rs6503527, Chr. 17q_rs8077456	53.12	3
Chr. 2q_rs10928764, Chr. 2q_rs6749002, Chr. 17q_rs1058808, Chr. 17q_rs6503527, Chr. 17q_rs8077456]	55.12	5

TA, Testing Accuracy; CVC, Cross-validation consistency.

Table S6 Transcription factors predicted binding to rs878461 and rs1799810

SNP	Allele	Transcription factor	Strand	Binding sequence	Score
rs878461	C	CDP	-	GCCTTCATCCATGCA	0.873
		CEBP	-	TTTGCTTGGGCC	0.924
		CREB	-	GGCCTTCATCCA	0.793
		EGR1	+	TTGCTTGGGCCT	0.752
		NF-1	+	TAGCTCCTTTGCTTGGGCCTTCATCCATG	0.603
		PAX2	+	GGCCTTCATCCATGCAGCG	0.728
		PAX5	-	TGCTTGGGCCTTCATCCATGCAGCGAAT	0.762
		PAX6	+	TTTGCTTGGGCCTTCATCCAT	0.543
		PAX8	+	CTCCTTTGCTTGGGC	0.760
		V\$PLZF_02	-	CTTTGCTTGGGCCTTCATCCATGCAGCGA	0.529
		PPAR-alpha	-	GGGCCTTCATCCATGCAGCG	0.541
		PPAR-gamma	-	AGCTCCTTTGCTTGGGCCTTCAT	0.452
		V\$SF1_Q6_01	+	GGGCCTTCA	0.853
		SREBP-1	-	TTGCTTGGGCCT	0.742
		V\$SZF11_01	-	GGCCTTCATCCATGC	0.647
		deltaCREB	-	GGGCCTTCATCCATG	0.723
		ZF5	-	TTGCTTGGGCCTT	0.691
	T	CDP	-	GTCTTCATCCATGCA	0.875
		CEBPgamma	-	CTTGGGTCTTCAT	0.658
		CEBP	-	TTTGCTTGGGTC	0.925
		CREB	-	GGTCTTCATCCA	0.798
		EGR1	+	TTGCTTGGGTCT	0.711

		GATA4	-	TTGCTTGGGTCT	0.796
		V\$LRP_Q2	-	TGGGTCTTC	0.902
		NF-1	+	TAGCTCCTTTGCTTGGGTCTTCATCCATG	0.608
		POU5F1C	+	GGGTCTTCATC	0.819
		PAX2	+	GGTCTTCATCCATGCAGCG	0.746
		PAX5	-	TGCTTGGGTCTTCATCCATGCAGCGAAT	0.774
		PAX6	+	TTTGCTTGGGTCTTCATCCAT	0.479
		PAX8	+	CTCCTTTGCTTGGGT	0.766
		V\$PLZF_02	-	CTTTGCTTGGGTCTTCATCCATGCAGCGA	0.531
		V\$PPARA_01	-	AGCTCCTTTGCTTGGGTCTT	0.539
		PPARgamma	+	GCTTGGGTCTTCATCCATGCAGC	0.411
		SREBP-1	-	TTGCTTGGGTCT	0.741
		V\$SZF11_01	-	GGTCTTCATCCATGC	0.647
		deltaCREB	-	GGGTCTTCATCCATG	0.721
rs1799810	A	androgen receptor	-	CGAACTTGCAGTATC	0.697
		V\$BRCA_01	-	GAACTTGC	0.839
		CUTL1	+	GAACTTGCAGTATCT	0.528
		C/EBPalpha	+	AACTTGCAGTATCT	0.892
		RXR-alpha	-	AGGACGGCGAACTTGCAGTAT	0.617
		GABP-alpha	-	CGAACTTGCAGT	0.719
		GR-alpha	-	GCGAACTTGCAGTATC	0.681
		V\$HMGY_Q6	-	AACTTGC	0.812
		HNF-4	+	GCGAACTTGCAGT	0.745
		c-Myb	-	GAACTTGCAGT	0.921

	NF-1	+	CAGGCTGTCATGGCGGCAGGACGGCGA AAC	0.553
	PAX-5	+	CAGGCTGTCATGGCGGCAGGACGGCG AA	0.767
	PAX-6	+	CGGCAGGACGGCG A ACTTGCA	0.530
	Pax-4c	-	CG A ACTTGCA G	0.690
	PPAR-gamma1	+	CATGGCGGCAGGACGGCG A ACTT	0.501
	Spz1	+	GCAGGACGGCG A ACT	0.749
	deltaCREB	+	GCGGCAGGACGGCG A	0.691
	Tel-2c	+	G A ACTTGCA G	0.739
	VDR	+	GCGGCAGGACGGCG A	0.792
	ZF5	+	GGACGGCG A ACTT	0.790
T	C/EBP	+	T ACTTGCA G TATCT	0.890
	GR-alpha	+	GCAGGACGGCG T ACTT	0.640
	NF-1	+	CAGGCTGTCATGGCGGCAGGACGGCG TAC	0.541
	PAX-5	+	CAGGCTGTCATGGCGGCAGGACGGCG TA	0.767
	PAX-6	+	CGGCAGGACGGCG T ACTTGCA	0.531
	Pax-4c	-	CG T ACTTGCA G	0.701
	PPAR-gamma1	+	CATGGCGGCAGGACGGCG T ACTT	0.526
	Spz1	+	GCAGGACGGCG T ACT	0.753
	Tel-2c	+	G T ACTTGCA G	0.802
	VDR	+	GCGGCAGGACGGCG T	0.756
	ZF5	+	GGACGGCG T ACTT	0.813

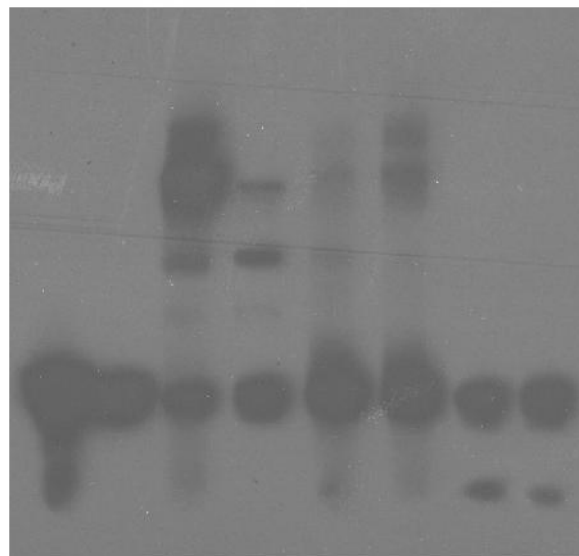
Table S7 Comparison of SNPs on chromosome 17q21 locus with published data

rs number	Reference	OR (95% CI)	Reported OR (95% CI)
	genotype		[reference]
rs907092	AA	1.64 (1.21-2.22)	1.34 [S4]
rs9303277	TT	1.44 (1.05-1.96)	1.41 (1.14-1.76) [S3]
			1.37 [S4]
rs11557467	TT	1.53 (1.12-2.10)	1.45 (1.16-1.82) [S3]
			1.35 [S4]
rs8067378	GG	1.53 (1.11-2.10)	1.35 [S4]
rs2290400	GG	1.42 (1.04-1.95)	1.47 (1.18-1.82) [S3]
			1.38 [S4]
rs4795405	TT	1.42 (1.04-1.94)	1.52 (1.20-1.89) [S3]
			1.37 [S4]

CI, confidence interval; OR, odds ratio.

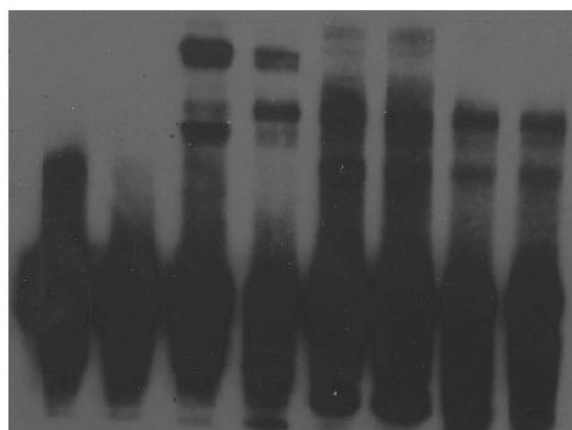
Figure S1 EMSA reactions for (a) rs878461 and (b) rs1799810 on 2q14. (a) Allele-specific protein-DNA binding at *PROC* involving C and T allele of rs878461 was shown using nuclear extracts from HepG2 cell line. Allele C exhibits less band shift than allele T (lanes 3 and 4). Incubating excess unlabeled competitor in reaction mixture resulted in diminished band shift (lanes 5 to 8). (b) Absence of protein-DNA binding at *PROC* involving A and T allele of rs1799810 was shown because incubating excess unlabeled competitor in reaction mixture did not cause reduction in band shift (lanes 5 to 8).

(a)



Lane	1	2	3	4	5	6	7	8
Nuclear Extract	-	-	+	+	+	+	+	+
Labelled probe	C	T	C	T	C	C	T	T
Unlabelled probe	-	-	-	-	C 100X	C 200X	T 100X	T 200X

(b)



Lane	1	2	3	4	5	6	7	8
Nuclear Extract	-	-	+	+	+	+	+	+
Labelled probe	A	T	A	T	A	A	T	T
Unlabelled probe	-	-	-	-	A	A	T	T
					100X	200X	100X	200X