ORIGINAL RESEARCH

Airway Epithelial Expression Quantitative Trait Loci Reveal Genes Underlying Asthma and Other Airway Diseases

Wei Luo[1,2]*, Ma'en Obeidat[3,4]*, Antonio Fabio Di Narzo[2,5], Rong Chen[2,5], Don D. Sin[3,4], Peter D. Paré[3,4], and Ke Hao[2,5,6]

[1]College of Computer Science and Technology, Huaqiao University, Xiamen, China; [2]Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, New York; [3]The University of British Columbia Center for Heart Lung Innovation, St. Paul's Hospital, Vancouver, British Columbia, Canada; [4]Respiratory Division, Department of Medicine, University of British Columbia, Vancouver, British Columbia, Canada; [5]Icahn Institute of Genomics and Multiscale Biology, Icahn School of Medicine at Mount Sinai, New York, New York; and [6]Department of Respiratory Medicine, Shanghai Tenth People's Hospital, Tongji University, Shanghai, China

Abstract

Genome-wide association studies (GWASs) have identified loci that are robustly associated with asthma and related phenotypes; however, the molecular mechanisms underlying these associations need to be explored. The most relevant tissues to study the functional consequences of asthma are the airways. We used publically available data to derive expression quantitative trait loci (eQTLs) for human epithelial cells from small and large airways and applied the eQTLs in the interpretation of GWAS results of asthma and related phenotypes. For the small airways ($n = 105$), we discovered 660 eQTLs at a 10% false discovery rate (FDR), among which 315 eQTLs were not previously reported in a large-scale eQTL study of whole lung tissue. A large fraction of the identified eQTLs is supported by data from Encyclopedia of DNA Elements (ENCODE) showing that the eQTLs reside in regulatory elements (57.5 and 67.6% of *cis*- and *trans*-eQTLs, respectively). Published pulmonary GWAS hits were enriched as airway epithelial eQTLs (9.2-fold). Further, genes regulated by asthma GWAS loci in epithelium are significantly enriched in immune response pathways, such as IL-4 signaling (FDR, 5.2×10^{-4}). The airway epithelial eQTLs described in this study are complementary to previously reported lung eQTLs and represent a powerful resource to link GWAS-associated variants to their regulatory function and thus elucidate the molecular mechanisms underlying asthma and airway-related conditions.

Keywords: airway epithelium; asthma; expression quantitative trait loci; airway diseases

Clinical Relevance

This work systematically characterized the architecture of genetic control of gene expression in airway epithelium. By comparing with a large lung tissue expressional quantitative trait loci (eQTLs) study, the authors demonstrated the common and unique features in gene expression regulation for these two important respiratory tissues. Further, the authors used the epithelium eQTLs to reveal genes and pathways underlying asthma.

Recent genome-wide association studies (GWASs) have identified loci that harbor susceptibility genes for many respiratory diseases, including asthma and related phenotypes (1–16). Many of the most significant GWAS hits are at loci with unknown function that had not been previously considered biologically plausible candidates for disease pathogenesis. Extensive linkage disequilibrium within many of these loci makes it difficult to identify the casual susceptibility variant,

(Received in original form October 1, 2014; accepted in final form June 22, 2015)

*These authors contributed equally to this work.

This work was supported by National Institutes of Health grant R01HL118542 (M.O., A.F.D.N., P.D.P., and K.H.), by National Natural Science Foundation of China grant no. 21477087 (K.H.), by postdoctoral Fellowship of the Michael Smith Foundation for Health Research and the Canadian Institute for Health Research Integrated and Mentored Pulmonary and Cardiovascular Training program (M.O.), and by a Fujian Province Overseas Studying Fellowship (W.L.).

Author Contributions: Conception and design: W.L., M.O., D.D.S., P.D.P., and K.H. Analysis and interpretation: W.L., M.O., A.F.D.N., R.C., D.D.S., P.D.P., and K.H. Drafting the manuscript for important intellectual content: W.L., M.O., D.D.S., P.D.P., and K.H.

Correspondence and requests for reprints should be addressed to Ke Hao, Sc.D., Icahn School of Medicine at Mount Sinai, Genetics and Genomic Sciences, 1425 Madison Avenue, Icahn Building L3-44, New York, NY 10029. E-mail: ke.hao@mssm.edu

This article has an online supplement, which is accessible from this issue's table of contents at www.atsjournals.org

Am J Respir Cell Mol Biol Vol 54, Iss 2, pp 177–187, Feb 2016
Copyright © 2016 by the American Thoracic Society
Originally Published in Press as DOI: 10.1165/rcmb.2014-0381OC on June 23, 2015
Internet address: www.atsjournals.org

let alone which genes or proteins they influence. Moreover, the associated polymorphisms can only explain a relatively small proportion of the variability of the phenotype in human populations (9, 17). Some of the missing heritability may be due to the limited power of GWASs, which miss disease susceptibility single-nucleotide polymorphisms (SNPs) of small-to-moderate effect sizes. Integrative genomics is a promising new approach to identify causal genes and variants, with improved statistical power. By using gene expression as a phenotype and examining how DNA polymorphisms contribute to both gene expression (expression quantitative trait loci [eQTLs]) and disease phenotypes, true causal relationships can be discovered (18–20).

A recently published large study of lung expression quantitative trait loci (eQTLs) (21) has been used to inform respiratory GWAS signals (22, 23). Because many eQTLs have been shown to be tissue- and cell-type specific (18, 24) and because the airway epithelium plays an important role in the pathogenesis of respiratory diseases (25–29), we used publically available genetic and genomic data to discover airway epithelial cell eQTLs. In this study, we report eQTLs from small and large airways, which will complement lung tissue eQTLs. Further, we demonstrate the utility of the airway epithelial eQTLs in mining publically available GWAS data (e.g., asthma GWASs) and identifying causal pathways.

Materials and Methods

Datasets

Airway epithelial gene expression and genotype datasets (30, 31) were retrieved from NCBI Gene Expression Omnibus (GEOaccession: GSE5057 and GSE40364). Gene expression data were generated on Affymetrix HG-U133 Plus 2.0 microarray using a standard protocol (31). The gene expression levels were quantified from Affymetrix CEL files using the latest version of the CDF file, where probes that harbor frequent SNPs were masked. The Affymetrix expression array includes 11 separate probes for each probeset, and the robust multiarray analysis algorithm was applied for expression level quantification, where outlier probes have little impact. The genotype data were generated from blood DNA using Affymetrix 5.0 microarrays (31). The published GWAS result catalog, which contained 11,598 unique SNPs, was downloaded from the National Human Genome Research Institute (NHGRI) website (32) on November 1, 2013. We manually curated this table and determined that 538 SNPs were associated with pulmonary diseases and related phenotypes. The GABRIEL asthma meta-analysis GWAS results were retrieved from the publication website (9). The ENCODE regulatory elements database was downloaded from regulomeDB (33). Human lung tissue eQTLs were retrieved from the published study's online supplement (21).

Data Quality Control

We normalized the gene expression data from small ($n = 112$) and large ($n = 40$) airways separately using a robust multiarray average method. Sex was inferred by identifying expression for the Y-linked gene *RPSY41*. We also performed quality control on genetic data. We excluded SNPs that had a low call rate (<0.9) or deviated significantly from Hardy–Weinberg equilibrium ($P < 1 \times 10^{-6}$). Genotype imputation was performed using 1000G cosmopolitan reference and the MaCH pipeline (34). We removed all imputed SNPs that had a low imputation score (i.e., MaCH $r^2 < 0.3$).

eQTL Discovery

We used a previously described Bayesian method (35) to determine whether the paired RNA and DNA samples were derived from the same individual. Discordance was found in seven small airway samples and in one large airway sample, which were then excluded from this analysis, reducing the sample size to 105 and 39 small and large airway samples, respectively. Common SNPs (minor alleles that were observed at least five times in dataset) were used to identify *cis*- and *trans*-eQTLs following a method similar to that described previously (19).

The gene expression data were adjusted for the top five transcriptome-derived principal components (PCs) and sex using a robust linear model to mitigate the effects of confounders and outliers. Residuals from this linear model were subjected to inverse normal transformation and then used for eQTL construction. One potential limitation of this approach was that the PCs may have consumed some of the true genetic variance signal, resulting in false negatives. To investigate this possibility, five additional GWAS analyses were performed, in which the top five PCs were treated as phenotypes (i.e., one PC for each GWAS). We did not observe any associations that surpassed the genome significance threshold of $P < 5 \times 10^{-8}$; thus, it was unlikely that PCs contained a significant amount of true genetic signal. We used both a linear model and Efficient Mixed-Model Association eXpedited methods (36) to evaluate the relationship between genotypes and transcriptomic expression. These two methods provided similar findings. For parsimony, we present data from linear models in this report. The R statistical package (version 3.02) and meta-eQTL software suite were used in data analysis (37).

The magnitude of the enrichment of GWAS SNPs as airway eSNPs (SNP associated with gene expression level) was calculated as the proportion of the eSNPs among GWAS SNPs divided by the proportion of the eSNPs among all SNPs on Affymetrix SNP 5.0 microarrays. The significance level of the enrichment was assessed by resampling, where we randomly sampled the same number of SNPs as the GWAS SNPs and examined the proportion of airway eSNPs in the random SNPs.

GWAS Datasets

GWAS summary statistics (association *P* values and coefficients) of five studies were retrieved from web resources described in the GWAS reports, including asthma (GABRIEL meta GWAS) (9), Alzheimer's disease (38), type 2 diabetes (39), obesity (40), and schizophrenia (41, 42).

Rank–Rank Plot

For a GWAS dataset (e.g., GABRIEL asthma meta GWAS), we were interested in examining whether a subset of the studied SNPs (e.g., airway eSNP) was enriched for small GWAS *P* values. First, we ranked all SNPs by GWAS *P* value ascendingly; the SNP with the smallest *P* value got the lowest rank value of 1. Using this method, each eSNP would be assigned a rank value, and we were interested in whether eSNPs were enriched for low-ranked *P* values. We standardized ranks for each SNP as standardized rank = (rank)/(number of SNPs in the GWAS study). For example, the SNP with smallest *P* value had a standardized rank value of $1/N$, where N is the number of SNP surveyed in GWAS. In the rank–rank

plot, the y axis is the observed standardized rank value of the eSNPs, and the x axis is the expected standardized rank value of the eSNPs if the eSNPs were uniformly distributed in the ranked SNP list.

Results

Genome-Wide Association Analysis for eQTLs

We identified *cis*- and *trans*-acting eQTLs and quantified the false discovery rate (FDR) through imputation using established methods (19). For parsimony, we assumed that each individual transcript could have no more than one *cis*-eQTL, which was defined as a significant association between a SNP and its transcriptomic expression within 500 kb upstream or downstream of the putative gene. *trans*-eQTLs were defined as association signals from SNPs located more than 500 kb from the probeset or on a different chromosome. Peak eSNP for a particular probeset was defined as the SNP that was most significantly associated (i.e., having the lowest P value) with the expression trait. A summary of the eQTLs identified at a 10% FDR is shown in Table E1 in the online supplement. We identified 616 and 44 *cis*- and *trans*-eQTLs, respectively, in the small airway epithelium. Despite the modest sample size ($n = 105$), we discovered a large number of eQTLs because the eQTL effect sizes were generally large, with a median r^2 value of 0.34 (Figure 1A). Consistent with a previous observations (21), the average *trans*-eQTLs' r^2 value (0.70) was larger than that for *cis*-eQTLs (0.31), likely owing to the higher threshold required to achieve statistical significance for *trans*-eQTLs. Because of the smaller sample size for the large airway analysis ($n = 39$), we had limited statistical power for eQTL discovery in this dataset. Only four *cis*- and three *trans*-eQTLs were discovered at a 10% FDR (Table E1). All of the four *cis*-eQTLs in the large airways were also *cis*-eQTLs in the small airways.

We found that 315 (48%) of the airway epithelial eQTLs were not discovered in lung tissue samples of a previous large study that evaluated lung eQTLs (Table E2) (21). However, the latter study did not carefully phenotype the lung tissue (e.g., airway, parenchyma, blood vessels, extracellular matrix, and inflammatory cells), which could have obscured airway-specific eQTLs.

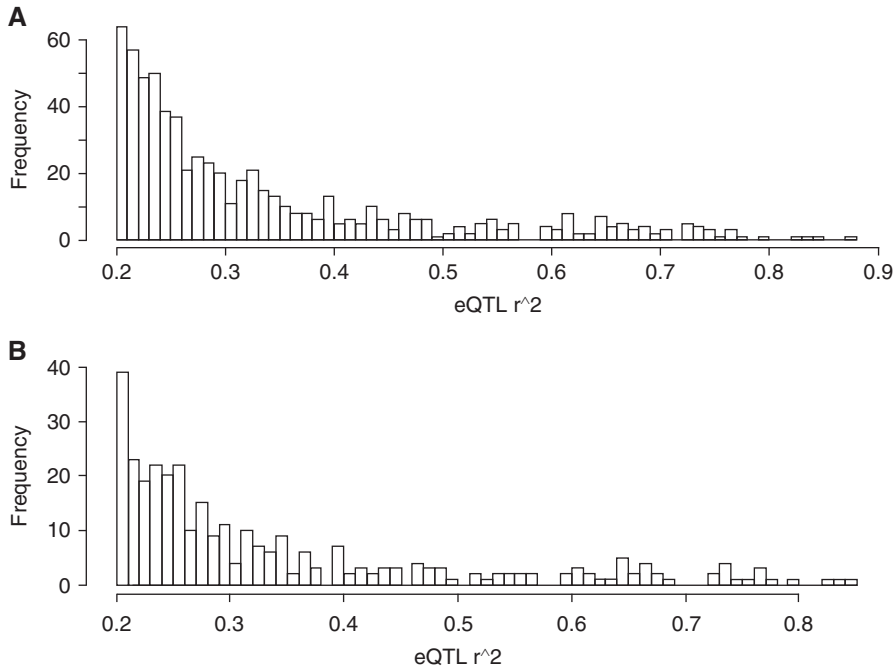

Figure 1. Histogram of r^2 values for expressional quantitative trait loci (eQTLs) (fraction of gene expression level explained by the peak single nucleotide polymorphism) for all small airway eQTLs (*A*; median r^2 = 0.335) and for small airway specific eQTLs (*B*, median r^2 = 0.339).

For example, *IL-6R*, which has been previously implicated in asthma (43), was regulated by a very strong eQTL in small airway epithelium (with a large effect size; $r^2 = 0.24$) but not in lung tissue (21).

A notable feature of airway-specific eQTLs was that they had, in general, large effect sizes. For example, glutathione *S*-transferase Theta 2 expression was strongly controlled by a *cis*-eQTL (rs9620341; $r^2 = 0.68$ and $P = 5.8 \times 10^{-30}$) in small airway epithelium. Overall, the median effect size (r^2) of the airway-specific eQTLs was 0.34 (Figure 1B), which was very similar to that of all eQTLs (regardless of airway specificity) discovered in small airway epithelium (Figure 1A). Some airway eQTLs discovered in the present study (Table E2) have been previously reported to be eQTLs in other tissues and organs (e.g., liver [44], lymphoblastoid cell lines [Electronic Database ED1] [45], and whole blood [46]). This indicates that some eQTLs in small airways may also be regulatory in other tissues.

Using Airway Epithelial eQTLs to Interpret Public GWAS Results

One important use of eQTLs is to determine the potential functional relevance of published GWAS findings. To this end, we retrieved a list of GWAS hits from the NHGRI public catalog (32, 47) using stringent thresholds based on guidelines for statistical significance for replication (32). From this GWAS hit list, we also identified a subset of SNPs related to pulmonary diseases and related phenotypes, which we called "pulmonary GWAS hits" (Table 1). This list included studies of asthma, chronic obstructive pulmonary disease (COPD), physiological measurements of lung function such as FEV_1 and the FEV_1/FVC ratio, and biomarkers of pulmonary or systemic inflammation, such as C-reactive protein. We then integrated the SNPs from the GWAS hit lists with the airway eQTLs using a 10% and 20% FDR, respectively. We found that "all GWAS hits" were significantly enriched as airway epithelial eQTLs. At a 10% FDR, 129 out of 11,598 GWAS hits were airway eSNPs (representing 9.2-fold enrichment). The significant level was assessed using a random sampling method. After 200 iterations, we determined that "all GWAS hits" were significantly enriched as airway epithelial eQTLs at $P < 0.005$. Although there was a similar magnitude of enrichment for the "pulmonary GWAS hits" as airway eQTLs (9.2- and 9.5-fold enrichment at a 10 and 20% FDR, respectively), the enrichment was not statistically significant in the random

Table 1. Intersecting National Human Genome Research Institute Genome-Wide Association Studies Results with Small Airway Expressional Quantitative Trait Loci*

SNP Sets	GWAS Hits (*n*)	10% FDR			20% FDR		
		cis eQTLs (*n*)	*trans* eQTLs (*n*)	Enrichment fold	*cis* eQTLs (*n*)	*trans* eQTLs (*n*)	Enrichment fold
All diseases	11,598	129	0	9.2†	171	0	7.5†
Airway diseases	538	6	0	9.2‡	6	4	9.5‡

Definition of abbreviations: eQTL, expressional quantitative trait locus; FDR, false discovery rate; GWAS, genome-wide association study; SNP, single-nucleotide polymorphism.
*At 10 and 20% FDR, we detected 660 and 1,066 small airway eQTLs, respectively.
†Significant at the <0.005 level based on resampling test.
‡Not statistically significant based on resampling-based test.

resampling analysis. This was probably due to the small number of "pulmonary GWAS hits" and reduced statistical power. Airway eSNP enrichment was similar between "all GWAS hits" and "pulmonary GWAS hits" most likely because some airway eQTLs are eQTLs in other tissue types (48, 49) (Table E2) and may contribute to different disease phenotypes.

Ten significant SNPs from the pulmonary GWAS studies, which were airway eQTLs, are shown in Table 2. An example was SNP rs13233571, which has been previously shown to be associated with circulating C-reactive protein levels (50). Interestingly, this SNP resides in a region not known to play a significant role in chronic inflammation, and its genetic association with C-reactive protein was originally assigned to the *BCL7B* gene (50). Here we showed that rs13233571 is an eSNP controlling the expression of the Claudin 4 (*CLDN4*) gene ($r^2 = 0.16$; $P = 3.2 \times 10^{-5}$), which is adjacent to the *BCL7B* gene.

We found that two SNPs that had been previously associated with lung function (one with baseline FEV_1 and the second with FEV_1 decline [51–53]) were strong *trans*-eQTLs. SNP rs17331332 resides in an intergenic region on Chr4q24 and is associated with FEV_1 ($P = 1 \times 10^{-16}$) (52). It controls the expression level of PDZ and LIM domain 3 (*PDLIM3*) on chr4q35.1 in small airways with a substantial effect size ($r^2 = 0.35$). The second SNP, rs16856186, resides in an intergenic region of chr1q32.1 and has been shown previously to associate with FEV_1 decline in patients with asthma ($P = 9 \times 10^{-6}$) (51). It controls the expression levels of the *ARV1* homolog (*Saccharomyces cerevisiae*) (*ARV1*) (chr1q42.2) and cystatin SN (*CST1*) (chr20p11.21) genes, both in *trans* with a r^2 value of 0.30.

Whereas most GWAS publications report only the top signals, the GABRIEL consortium (9) released the asthma genetic association results for all SNPs investigated, allowing an in-depth analysis. We examined various *P* value cutoffs for the GWASs and queried whether the SNPs passing these thresholds were eQTLs (Table 3). At all cutoffs (1×10^{-2}, 1×10^{-3}, 1×10^{-4}, and 1×10^{-5}), the SNPs associated with asthma were enriched as airway eQTLs. At a *P* value cutoff of 1×10^{-4}, the enrichment was the largest (33.8-fold and 55.9-fold for 10 and 20% FDR, respectively). Enrichment was the lowest at a *P* value cutoff of 1×10^{-2}, indicating reduced signal-to-noise ratio. Importantly, we observed higher enrichment at *P* value of 1×10^{-4} and 1×10^{-5} cutoffs when using a 20% FDR for eQTLs (Table 3), suggesting that some GWAS signals may be mediated by relatively weak eSNPs, which can only be discovered at a less stringent threshold (especially when the sample size is small to moderate). Among the 567,589 SNPs investigated in the GABRIEL GWAS meta-analysis, 1,458 were identified as eSNPs at a 10% FDR in small airways. We did not restrict the analysis to the peak eSNP for each eQTL but rather included all underlying eQTLs as eSNPs as long as their *P* values survived the 10% FDR threshold. The 1,458 eSNPs were significantly enriched for hits in the GABRIEL meta GWAS data (Figure 2). This is consistent with a previous report showing that SNPs associated with complex traits are more likely to be eQTLs (21). In Figure 2, we used a rank–rank plot to test for enrichment of small *P* values among airway eSNPs in the GABRIEL study (*see* Materials and Methods). The rank–rank plot showed an upward deviation from the diagonal line, indicating that airway epithelial eSNPs were enriched as hits in GABRIEL meta-GWAS (Figure 2).

In addition, we interrogated well-documented asthma candidate genes (21, 54) (Table E3) and found that these genes were also significantly more likely to be airway eSNPs. In Figure E1A in the online supplement, the rank–rank curve deviated above the diagonal line, indicating that airway epithelial eSNPs in asthma candidate genes were enriched for being significantly associated with asthma in the GWAS. The magnitude of deviation of the curve above the diagonal line in Figure E1A was larger than the deviation in Figure 2, suggesting that the eSNPs in asthma candidate genes are more enriched for GWAS signal than airway eSNPs as a whole. Consistent with this notion, the rank–rank plot curve in Figure E1B deviated above the diagonal line, suggesting the asthma GWAS associations for the airway epithelial eSNPs that influence asthma candidate genes harbor more GWAS signal than all airway epithelial eSNPs. Although the contribution of candidate genes from the pre-GWAS era to today's GWAS findings has been modest, these results suggest that these candidate genes have relevance to the pathogenesis of asthma.

Importantly, many of the most significant hits in the GABRIEL study were airway eSNPs (Table 4). Chromosomes 2q12 and 17q21 both contained strong GWAS hits for asthma. However, these loci harbored a number of genes that could potentially explain these associations (9–11). We identified genes in these GWAS regions that were regulated by airway epithelial eQTLs and have biological plausibility. For example, rs10192157, which reached genome-wide significance in

Table 2. Pulmonary Diseases Genome-Wide Association Study Results Informed by Small Airway Expressional Quantitative Trait Loci*

SNP rs ID	Expression Probeset ID	SNP Location (*hg19*)	Airway eQTL Statistics				GWAS Reports	
			P Value	β	r^2	eQTL-Regulated Gene†	Proposed Gene‡	GWAS Disease/Traits
Cis-eQTLs								
rs13233571	201428_at	7:72971231	3.2×10^{-5}	0.400	0.155	*CLDN4*	*BCL7B*	C-reactive protein
rs2658782	1554071_at	11:93166731	1.0×10^{-4}	−0.343	0.137	*CCDC67*	*CCDC67*	Pulmonary function decline
rs412658	210697_at	19:22359440	1.1×10^{-4}	−0.089	0.136	*ZNF257*	*ZNF676*	Telomere length
rs12188164	226125_at	5:428236	1.1×10^{-4}	−0.505	0.136	*LOC100288152*	*AHRR*	Cystic fibrosis severity
rs9272346	209480_at	6:32604372	2.6×10^{-34}	−2.529	0.767	*HLA-DQB1*	*HLA-DQA1*	Asthma, Type 1 diabetes
rs9272346	213831_at	6:32604372	1.1×10^{-27}	−2.516	0.686	*HLA-DQA1*	*HLA-DQA1*	Asthma, Type 1 diabetes
rs9272346	236203_at	6:32604372	9.9×10^{-14}	−0.910	0.417	*HLA-DQA1*	*HLA-DQA1*	Asthma, Type 1 diabetes
	212999_x_at	6:32604372	2.7×10^{-26}	−1.192	0.666	*HLA-DQB1*	*HLA-DQA1*	Asthma, Type 1 diabetes
rs9268905	221491_x_at	6:32432077	7.8×10^{-14}	1.248	0.420	*HLA-DRB1*, *HLA-DRB3*, *HLA-DRB4*, *HLA-DRB5*	*HLA-DRA*	Cystic fibrosis severity
rs9268905	209728_at	6:32432077	2.5×10^{-26}	−2.515	0.667	*HLA-DRB4*	*HLA-DRA*	Cystic fibrosis severity
Trans-eQTLs								
rs16856186	238140_at	1:205678126	1.65×10^{-9}	−0.743	0.299	*ARV1*	*SLC45A3*, *NUCKS1*	Pulmonary function decline
rs16856186	206224_at	1:205678126	1.90×10^{-9}	−1.405	0.297	*CST1*	*SLC45A3*, *NUCKS1*	Pulmonary function decline
rs7927044	233413_at	11:127761666	1.37×10^{-9}	−1.299	0.301	—	*KIRREL3-AS3*, *ETS1*	Asthma (childhood onset)
rs4452212	232233_at	2:137015991	7.69×10^{-11}	1.734	0.338	*SLC22A16*	*CXCR4*	Telomere length
rs4452212	232232_s_at	2:137015991	4.93×10^{-10}	1.926	0.314	*SLC22A16*	*CXCR4*	Telomere length
rs4452212	205048_s_at	2:137015991	2.78×10^{-9}	1.985	0.292	*PSPH*	*CXCR4*	Telomere length
rs17331332	210170_at	4:106808107	4.41×10^{-11}	−0.606	0.345	*PDLIM3*	*INTS12*, *GSTCD*, *NPNT*	Pulmonary function

Definition of abbreviations: eQTL, expressional quantitative trait locus; GWAS, genome-wide association study; SNP, single nucleotide polymorphism.
*A 20% false discovery rate was applied in eQTL discovery. In this analysis, eQTL discovery is carried out only in GWAS SNPs, which greatly reduced the number of tests and multiple testing penalty. Therefore, the *P* value threshold for 10 or 20% false discovery rate is much more relaxed than a genome-wide eQTL search.
†The eQTL gene is the gene regulated by the airway epithelium eQTL.
‡The proposed gene is the gene underlying GWAS hit, as proposed in the original publication.

the asthma GWAS ($P = 8.12 \times 10^{-12}$) (9), was associated with IL-1 receptor, type II (*IL-1R2*) expression level in the airway epithelium ($P = 7.48 \times 10^{-4}$). SNP rs10192157 was also associated with the expression level of cyclin-dependent kinase 12 (*CDK12*) ($P = 1.95 \times 10^{-3}$) and StAR-related lipid transfer domain containing 3 (*STARD3*) ($P = 4.22 \times 10^{-4}$) in airway epithelia.

Beyond the GABRIEL asthma meta GWAS, we also surveyed four large GWAS data on nonairway diseases: Alzheimer's disease, schizophrenia, obesity, and type 2 diabetes (*see* Materials and Methods). We explored whether the airway eSNPs were enriched for association signal in these four diseases using rank–rank plot methods (Figure E2). For all diseases, a portion of the points was above the line of identity, suggesting that there is enrichment for eSNPs among susceptibility alleles; this was especially so for schizophrenia and Alzheimer's disease. This suggests there is pleiotropy for eSNPs because a percentage of them are likely to be eSNPs in multiple tissues. However, the extent of deviation was less than that shown in Figure 2 for asthma, indicating that the airway eSNPs are more likely to be susceptibility alleles for asthma than for nonairway diseases. We also noted that an eSNP might exist in more than one tissue types (Tables E1 and E2).

Airway eQTLs Are Enriched in Regulatory Elements

A large fraction of GWAS hits do not change protein sequence but do affect regulatory elements. The ENCODE Project has mapped out open chromatin and protein binding regions for a large number of factors

Table 3. Intersecting GABRIEL Asthma Meta Genome-Wide Association Study Results with Small Airway Expressional Quantitative Trait Loci

		10% FDR*				20% FDR			
P value cutoff	GABRIEL Hits	*cis*-eQTLs (n)	*trans*-eQTLs (n)	Enrichment Fold	Enrichment *P* Value	*cis*-eQTLs (n)	*trans*-eQTLs (n)	Enrichment Fold	Enrichment *P* Value
1×10^{-2}	7,204	52	3	6.3	1.9×10^{-27}	77	3	5.7	1.4×10^{-35}
1×10^{-3}	1,186	31	4	24.4	7.2×10^{-38}	37	3	17.3	7.6×10^{-37}
1×10^{-4}	348	13	3	38.0	4.3×10^{-22}	36	2	55.9	1.9×10^{-51}
1×10^{-5}	175	4	0	18.9	2.8×10^{-6}	16	0	46.8	9.9×10^{-24}

Definition of abbreviations: eQTL, expressional quantitative trait locus; FDR, false discovery rate.
*At 10 and 20% FDR we detected small airway 660 and 1,066 eQTLs, respectively.

across many different cell types (33, 54). These data serve as powerful orthogonal evidence to complement GWAS and eQTL findings by demonstrating the potential functional consequences of variants, which are outside of the coding regions. We queried the regulomeDB databases (33) and intersected the asthma GWAS and airway eQTL data together with (*1*) ChIP-seq information for a variety of important regulatory factors across a diverse set of cell types, and (*2*) chromatin state information across over 100 cell types. A total of 372 (56.4%) of probesets influenced by the airway eQTLs were supported by regulatory element data, and 55.9% of the airway eSNPs, which were on the "all GWAS hit" list, were supported by regulatory element data (enrichment $P = 3.3 \times 10^{-35}$). Further, we investigated the enrichment of ENCODE entries for eSNPs among GABRIEL top hits (Table E4). Notably, rs10192157 (GABRIEL meta GWAS $P = 8.1 \times 10^{-12}$) was an eSNP-controlling *IL1R2* and was supported by ENCODE data, showing that this SNP spans a DNase hypersensitivity site in a number of cell types and also spans a transcription factor binding site for the transcription factor CCAAT/Enhancer Binding Protein Beta (CEBPB) and CCCTC-Binding Factor (CTCF) as determined by ChIP-seq assays. In Table E5, we list the leading SNP for each 10% FDR eQTL that was annotated in RegulomeDB (33). In brief, we queried the RegulomeDB database by eQTL chromosome position and identified the regulatory elements where at least one eQTL resided. The regulatory elements are small DNA regions that modulate expression through various mechanisms. Their detailed position information can be retrieved from [Electronic Database ED2] (55) by searching for SNP position or using dbSNP ID. For example, rs10954213 influences expression of *IRF5*, which has been linked to asthma (56). Footprinting evidence showed that rs10954213 is located in the *HOXD3* motif in several cell types (Helas3, Hsmmt, and Huh7). The detailed location of the regulatory site can be found at [Electronic Database ED3] (57).

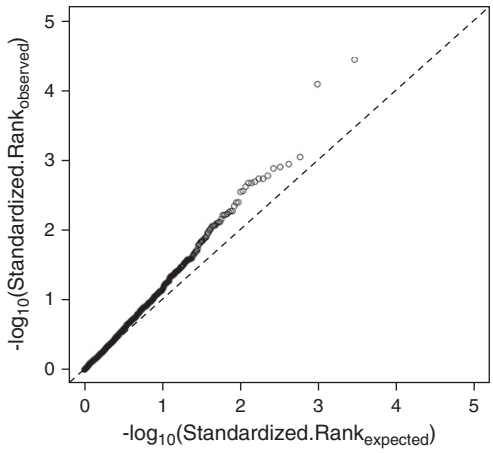

Figure 2. Rank–rank plot (*see* Materials and Methods) were used to examine whether airway SNP associated with gene expression levels (eSNPs) were enriched for small *P* values in the GABRIEL meta genome-wide association study (GWAS). Among all SNPs surveyed by the GABRIEL meta GWAS, 1,458 were eSNPs in the airway epithelia at 10% false discovery rate. *y*-Axis: Observed standardized rank of airway eSNPs among all SNPs in the GABRIEL meta GWAS. *x*-Axis: Expected standardized rank value if airway eSNPs were randomly distributed in ranked GABRIEL SNPs. SNP, single-nucleotide polymorphism.

BioProcess Enrichment of Asthma GWAS Genes

Airway epithelial eQTLs serve as an empirical "bridge" between GWAS-associated SNPs and the genes underlying these associations. First, we summarized "asthma SNPs" with globally significant associations (e.g., NHGRI catalog) and suggestive association in GABRIEL. Second, we obtained a list of 88 genes that were controlled by "asthma SNPs" through eQTL in airway epithelium. Last, we tested the 88 genes for enrichment in biological processes using the BioProcess gene sets. A number of BioProcess gene sets were significantly enriched (Figure 3), including responses to IFN-γ, immune response signaling, proliferation of various types of immune cells, and IL-4 signaling (noting IL-4 or its receptor is not directly identified by GWAS SNPs).

Discussion

In this paper, publically available data on genetic variation and gene expression in a collection of human small and large airway epithelial cell samples were used to explore the genetic influence on gene expression in this tissue. With a relatively small sample size, we discovered 660 and 1,066 eQTLs at a 10 and 20% FDR, respectively, in the small

Table 4. GABRIEL Meta Genome-Wide Association Study Results Informed with Expressional Quantitative Trait Loci in Small Airway but Not Lung Tissue*

SNP rs ID	Expression Probeset ID	SNP Location (*hg19*)	GWAS *P* Value	Small Airway eQTL: eQTL *P* Value	β	r^2	eQTL-Regulated Gene	eQTL Presence in Other Tissues†
rs10192157	205403_at	2:102968356	8.12×10^{-12}	7.48×10^{-4}	0.509	0.105	*IL-1R2*	NYY
rs2596464	205905_s_at	6:31412961	5.05×10^{-6}	7.12×10^{-5}	−0.167	0.143	*MICA*, *MICB*	NNN, YYY
rs2596464	205904_at	6:31412961	5.05×10^{-6}	9.99×10^{-4}	−0.178	0.100	*MICA*	NNN
rs3939286	206426_at	9:6210099	4.53×10^{-14}	7.19×10^{-4}	−0.060	0.106	*MLANA*	NNN
rs6485713	203409_at	11:46962474	9.24×10^{-6}	2.31×10^{-3}	−0.124	0.087	*DDB2*	NYN
rs7117590	203363_s_at	11:47005633	8.84×10^{-6}	2.27×10^{-3}	0.126	0.087	*ATG13*	NNN
rs12453682	219226_at	17:37770005	3.32×10^{-7}	4.04×10^{-4}	0.166	0.115	*CDK12*	NNN
rs4252627	225165_at	17:37868715	4.77×10^{-8}	1.82×10^{-3}	0.227	0.090	*PPP1R1B*	NNY
rs12150079	205766_at	17:38025417	2.04×10^{-9}	1.52×10^{-3}	0.194	0.093	*TCAP*	NNN
rs11557467	202991_at	17:38028634	2.49×10^{-17}	4.22×10^{-4}	0.157	0.114	*STARD3*	NNN
rs11557467	225694_at	17:38028634	2.49×10^{-17}	1.95×10^{-3}	−0.202	0.089	*CDK12*	NNN

For definition of abbreviations, *see* Table 2.
*A 20% false discovery rate was applied in eQTL discovery.
†Whether the small airway eQTL-regulated gene presents in other tissue types: liver, lymphoblastoid cell lines, and whole blood, respectively. NNN denotes a gene is regulated by eQTL in none of lymphoblastoid cell lines, liver, and whole blood. YYY denotes a gene is regulated by eQTL in all of lymphoblastoid cell lines, liver, and whole blood. NYY denotes a gene is regulated by eQTL in lymphoblastoid cell lines and whole blood, but not liver. NNY denotes a gene is regulated by eQTL in whole blood, but not liver or lymphoblastoid cell lines, and NYN denotes a gene is regulated by eQTL in lymphoblastoid cell lines, but not liver or whole blood.

airways. These eQTLs are genetic variants that directly or indirectly govern gene expression in the airway epithelium and represent a unique resource for respiratory diseases research. We provide a few examples of how this dataset can be used to elucidate new molecular drivers of asthma. Because of the relatively small sample size, we were only able to capture eQTLs that were strong. Weaker eQTLs may have been missed by this study. Larger sample sizes will be needed to fully characterize the genetic architecture of gene expression in airway epithelium.

Chromosome 17q21 is one of the most consistent loci associated with asthma (9, 10, 58–67). Based on an eQTL study performed in lymphoblastoid cell lines and lung tissue, *ORMD3L*, *GSDMB*, and *GSDMA* have been suggested to be prime candidates driving the GWAS signal (9, 10, 21, 61). In Table 4, we show that the GWAS hits at this locus controlled expression levels of additional genes: *CDK12*, *PPP1R1B*, *TCAP*, and *STARD3* in the airway epithelium. Interestingly, these genes were missed in a large-scale lung tissue eQTL study most likely owing to the heterogeneity of cell types in the lung tissue (21). The cyclin-dependent kinase (*CDK*) family, of which *CDK12* is a member, is a group of serine/threonine kinases that regulate cell cycle events through phosphorylation of transcription factors and tumor suppressor proteins required for DNA replication and cell division (66). CDKs are possible therapeutic targets to dampen inflammation (67). The CDK inhibitor R-roscovitine has been shown to induce apoptosis of human eosinophils (66). However, little is known about a possible role for titin-cap (*TCAP*), and protein phosphatase 1 regulatory subunit 1B (*PPP1R1B*, also known as dopamine- and cAMP-regulated neuronal phosphoprotein (*DARPP-32*) in asthma.

Chromosome 2q12 is another locus consistently associated with asthma (9). The previous study of eQTLs in whole lung tissues showed that the expression of *IL-1RL1* was controlled by a GWAS hit (rs13431828) (21). Herein, we showed that the GWAS hits at this locus also control expression levels of *IL-1R2* in airway epithelial cells, indicating that this is an airway eQTL. It is noteworthy that *IL-1R2* was detectable in microarrays of whole lung tissue (*IL-1R2* expression was "present" in 98.8% of subjects). *IL-1R2* has been shown to be associated with aspirin-induced asthma, both at the level of gene expression in human nasal polyps and as SNP associations (68). In another study, *IL-1R2* SNPs were associated with atopy and showed interaction with early childhood virus infection (69). *IL-1R2* encodes type II receptor, which acts as a soluble decoy receptor for IL-1 and thus is a negative regulator of the IL-1 pathway. IL-1 cytokine prolongs the *in vitro* survival of polymorphonuclear cells through IL-1R2. IL-4 antagonizes the action of IL-1 by inducing the expression and release of IL-1R2 (70, 71). Furthermore, dexamethasone promotes the expression and release of IL-1R2 in polymorphonuclear cells (70, 72). In addition to corticosteroids and IL-4, IL-13 induces the expression and release of IL-1R2 (73, 74), whereas bacterial LPS causes rapid shedding, followed by inhibition of the transcripts' expression (75–77). Interestingly, data from ENCODE suggest that the *IL-1R2* eSNP is located at a binding site for CEBPB, a member of the CEBP family of basic leucine zipper transcription factors that regulates inflammatory protein expression in several cell types, including lung epithelial cells (78). These data on IL-1R2 show the utility of airway epithelial eQTLs in explaining GWAS signals, especially when this information is integrated with complementary resources such as ENCODE.

SNP rs16856186 resides in an intergenic region of chr1q32.1 and has been previously associated with FEV_1 decline in patients with asthma ($P = 9 \times 10^{-6}$) (51). In epithelial cells, it controls the expression levels of *ARV1* Homolog (*S. cerevisiae*) (*ARV1*) on chr1q42.2 and Cystatin SN (*CST1*) on chr20p11.21 both in *trans* with r^2 values of 0.30. Recently, differential expression of *CST1* was reported in airway cells when comparing with asthma and exercise-induced bronchoconstriction with

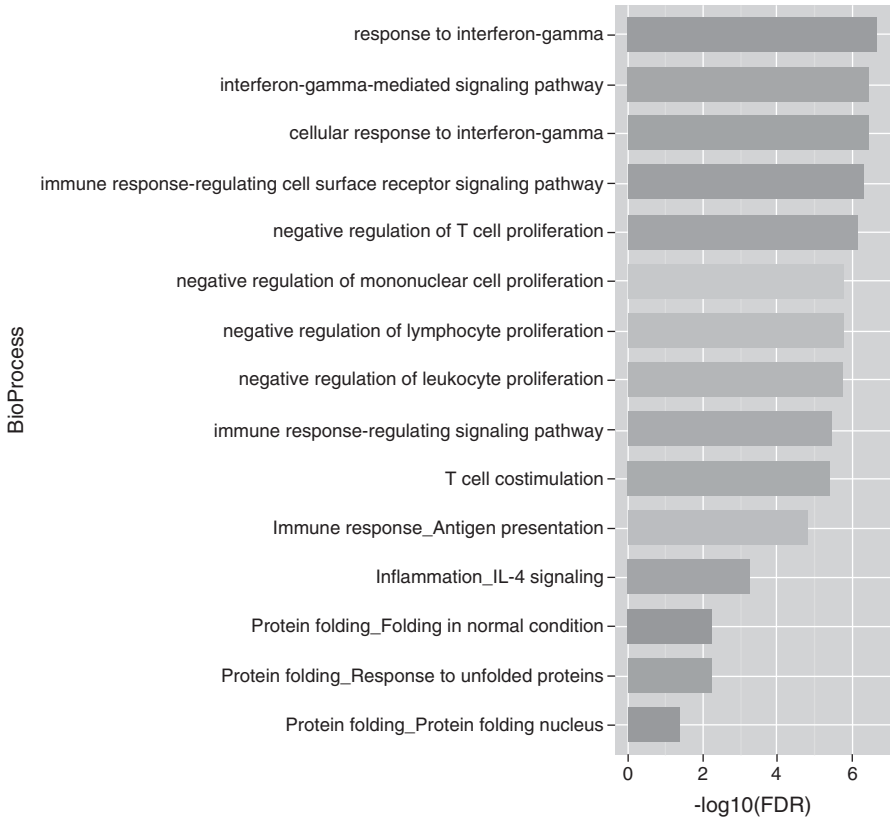

Figure 3. BioProcess enrichment of genes transcriptionally regulated by asthma genome-wide association study single-nucleotide polymorphisms in airway epithelium. FDR, false discovery rate.

patients with asthma without exercise-induced bronchoconstriction (79). In addition, *CST1* expression was increased in patients with seasonal allergic rhinitis (80). Importantly, these two eQTLs are *trans*-activating and airway specific and were not discovered by previous eQTL studies.

A recent GWAS for childhood asthma and severe exacerbations identified a nonsynonymous coding SNP in the cadherin-related family member 3 gene (*CDHR3*), which is highly expressed in airway epithelium (81). Using our airway epithelium eQTL dataset, we found that *CDHR3* is measured by two probesets and that both probesets have *cis*-eQTLs controlling their expression levels. Specifically, probe 231582_at is influenced by rs75858860 ($P = 1.1 \times 10^{-3}$; $r^2 = 0.09$), and probe 35650_at is influenced by rs2528883 ($P = 8.8 \times 10^{-4}$; $r^2 = 0.10$). It is possible that this gene's effect on asthma and exacerbations is driven by both coding and regulatory SNPs in airway epithelium.

In this study, we identified a number of genes for which the variation in expression was largely explained by airway eQTLs. For example, *GSTT2* expression was strongly controlled by a *cis*-eQTL ($r^2 = 0.72$; $P = 5.8 \times 10^{-30}$) in small airways but not in whole lungs. GSTs are a superfamily of enzymes that catalyze the conjugation of glutathione with electrophilic compounds. They are involved in detoxifying toxic components of tobacco smoke, and the epithelial lining fluid contains 140-fold higher levels of glutathione than plasma, indicating its critical role in protecting airway epithelium from oxidant injury (82, 83). In one study, a missense (Met139Ile) coding SNP in *GSTT2* was associated with non–small cell lung cancer in smokers (84). Other genes for which eQTLs explained most of the variation in expression included Churchill Domain Containing 1 (*CHURC1*) ($r^2 = 0.76$; $P = 7.29 \times 10^{-34}$), Hepatocellular Carcinoma-Associated Antigen 520 (*CHP2*) ($r^2 = 0.75$; $P = 1.03 \times 10^{-32}$), and Family with Sequence Similarity 118, Member A (*FAM118A*) ($r^2 = 0.74$; $P = 8.11 \times 10^{-32}$). Interestingly, the levels of expression of *CHURC1* and *FAM118A* were also identified as being strongly influenced by eQTLs in a study of the genetic control of gene expression in the sputum of patients with COPD (85).

Two additional genes under epithelial eQTL control were *SPINK5* and *SPATS2L*. *SPINK5* encodes for lymphoepithelial Kazal-type–related inhibitor, a serine protease inhibitor (86). The gene is located on 5q31–32, a region repeatedly associated with asthma in linkage studies (87–89). Moreover, SNPs in the *SPINK5* gene showed significant associations with atopy, atopic dermatitis, and asthma (90–92). *SPATS2L* (spermatogenesis associated, serine-rich 2–like) is located on 2q33.1. SNPs near this gene showed the strongest associations in a GWAS of bronchodilator response in individuals with asthma (93). The authors followed up their studies of the association signal with an *in vitro* study of the effect of knocked-down of *SPATS2L* in human airway smooth muscle cells using short interfering RNA. The knock-down led to an increase in the β_2 adrenergic receptor protein expression, suggesting that *SPATS2L* may affect BDR by directly modulating β_2 adrenergic receptor protein expression (93).

The airway epithelium plays an essential role in the pathogenesis of asthma, and the eQTL dataset allows interpretation of asthma GWAS hits in the context of the pathways and networks they alter by empirically linking association loci to the underlying genes. We identified a set of genes whose expression level is under the influence of asthma GWAS SNPs (GABRIEL and NHGRI database) and conducted BioProcess enrichment on this gene set. Results (Figure 3) show strong over-representation of genes in response to IFN-γ, immune response signaling, proliferation of various types of immune cells (T cells, mononuclear cells, lymphocytes, and leukocytes), and IL-4 signaling. Such analyses reveal the functions of the GWAS loci at a higher level and are consistent with current knowledge of asthma pathogenesis. IL-4 is a key cytokine in the development of allergic inflammation, and the activity of IL-4 promotes cellular inflammation in asthma (94). SNPs on IL-4 or its receptor genes are not identified in asthma GWASs with global significance, although they have been reported in some candidate gene studies (94), suggesting that these SNPs have

a small-to-moderate effect size (if any). However, SNPs regulating the expression of other genes of the IL-4 pathway were captured by GWASs, and these SNPs may influence the IL-4 pathway and asthma susceptibility through regulation of gene networks.

In summary, this paper reports the results of a study of small airway epithelial eQTLs that complements previous large-scale studies of lung eQTLs (21) and that can be used to shed light on GWAS findings in pulmonary diseases. With a moderate sample size, we only had statistical power to identify eQTLs, which had substantial effect size (median r^2 = 0.335). Strikingly, a large fraction (48%) of these strong eQTLs was not discovered in the large scale lung tissue study, and are therefore, termed airway eQTLs. Using the results of the largest to date GWASs of asthma as an example, we show how the airway eQTL dataset can be used to identify additional causal genes and pathways underlying asthma. Our results will also serve as a valuable tool to study the pathogenesis of other lung diseases, such as COPD and lung cancer. ■

Author disclosures are available with the text of this article at www.atsjournals.org.

Acknowledgments: The authors thank Dr. Ronald G. Crystal and colleagues for conducting the gene expression and genotype measurement and contributing the results in GEO databases (accession no. GSE5057 and GSE40364).

References

1. Choudhry S, Taub M, Mei R, Rodriguez-Santana J, Rodriguez-Cintron W, Shriver MD, Ziv E, Risch NJ, Burchard EG. Genome-wide screen for asthma in Puerto Ricans: evidence for association with 5q23 region. *Hum Genet* 2008;123:455–468.
2. DeWan AT, Triche EW, Xu X, Hsu LI, Zhao C, Belanger K, Hellenbrand K, Willis-Owen SA, Moffatt M, Cookson WO, *et al*. *PDE11A* associations with asthma: results of a genome-wide association scan. *J Allergy Clin Immunol* 2010;126:871–873.
3. Gudbjartsson DF, Bjornsdottir US, Halapi E, Helgadottir A, Sulem P, Jonsdottir GM, Thorleifsson G, Helgadottir H, Steinthorsdottir V, Stefansson H, *et al*. Sequence variants affecting eosinophil numbers associate with asthma and myocardial infarction. *Nat Genet* 2009;41: 342–347.
4. Hancock DB, Romieu I, Shi M, Sienra-Monge JJ, Wu H, Chiu GY, Li H, del Rio-Navarro BE, Willis-Owen SA, Weiss ST, *et al*. Genome-wide association study implicates chromosome 9q21.31 as a susceptibility locus for asthma in Mexican children. *PLoS Genet* 2009;5:e1000623.
5. Himes BE, Hunninghake GM, Baurley JW, Rafaels NM, Sleiman P, Strachan DP, Wilk JB, Willis-Owen SA, Klanderman B, Lasky-Su J, *et al*. Genome-wide association analysis identifies *PDE4D* as an asthma-susceptibility gene. *Am J Hum Genet* 2009;84:581–593.
6. Hui J, Oka A, James A, Palmer LJ, Musk AW, Beilby J, Inoko H. A genome-wide association scan for asthma in a general Australian population. *Hum Genet* 2008;123:297–306.
7. Kim SH, Cho BY, Park CS, Shin ES, Cho EY, Yang EM, Kim CW, Hong CS, Lee JE, Park HS. Alpha-T-catenin (*CTNNA3*) gene was identified as a risk variant for toluene diisocyanate-induced asthma by genome-wide association analysis. *Clin Exp Allergy* 2009;39:203–212.
8. Mathias RA, Grant AV, Rafaels N, Hand T, Gao L, Vergara C, Tsai YJ, Yang M, Campbell M, Foster C, *et al*. A genome-wide association study on African-ancestry populations for asthma. *J Allergy Clin Immunol* 2010;125:336–346.
9. Moffatt MF, Gut IG, Demenais F, Strachan DP, Bouzigon E, Heath S, von Mutius E, Farrall M, Lathrop M, Cookson WO; GABRIEL Consortium. A large-scale, consortium-based genomewide association study of asthma. *N Engl J Med* 2010;363:1211–1221.
10. Moffatt MF, Kabesch M, Liang L, Dixon AL, Strachan D, Heath S, Depner M, von Berg A, Bufe A, Rietschel E, *et al*. Genetic variants regulating *ORMDL3* expression contribute to the risk of childhood asthma. *Nature* 2007;448:470–473.
11. Ober C, Tan Z, Sun Y, Possick JD, Pan L, Nicolae R, Radford S, Parry RR, Heinzmann A, Deichmann KA, *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.
12. Sleiman PM, Flory J, Imielinski M, Bradfield JP, Annaiah K, Willis-Owen SA, Wang K, Rafaels NM, Michel S, Bonnelykke K, *et al*. Variants of *DENND1B* associated with asthma in children. *N Engl J Med* 2010; 362:36–44.
13. Weidinger S, Gieger C, Rodriguez E, Baurecht H, Mempel M, Klopp N, Gohlke H, Wagenpfeil S, Ollert M, Ring J, *et al*. Genome-wide scan on total serum IgE levels identifies *FCER1A* as novel susceptibility locus. *PLoS Genet* 2008;4:e1000166.
14. Hirota T, Takahashi A, Kubo M, Tsunoda T, Tomita K, Doi S, Fujita K, Miyatake A, Enomoto T, Miyagawa T, *et al*. Genome-wide association study identifies three new susceptibility loci for adult asthma in the Japanese population. *Nat Genet* 2011;43:893–896.
15. Torgerson DG, Ampleford EJ, Chiu GY, Gauderman WJ, Gignoux CR, Graves PE, Himes BE, Levin AM, Mathias RA, Hancock DB, *et al*.; Mexico City Childhood Asthma Study (MCAAS); Children's Health Study (CHS) and HARBORS study; Genetics of Asthma in Latino Americans (GALA) Study, Study of Genes-Environment and Admixture in Latino Americans (GALA2) and Study of African Americans, Asthma, Genes & Environments (SAGE); Childhood Asthma Research and Education (CARE) Network; Childhood Asthma Management Program (CAMP); Study of Asthma Phenotypes and Pharmacogenomic Interactions by Race-Ethnicity (SAPPHIRE); Genetic Research on Asthma in African Diaspora (GRAAD) Study. Meta-analysis of genome-wide association studies of asthma in ethnically diverse North American populations. *Nat Genet* 2011;43:887–892.
16. Noguchi E, Sakamoto H, Hirota T, Ochiai K, Imoto Y, Sakashita M, Kurosaka F, Akasawa A, Yoshihara S, Kanno N, *et al*. Genome-wide association study identifies *HLA-DP* as a susceptibility gene for pediatric asthma in Asian populations. *PLoS Genet* 2011;7:e1002170.
17. Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, McCarthy MI, Ramos EM, Cardon LR, Chakravarti A, *et al*. Finding the missing heritability of complex diseases. *Nature* 2009;461: 747–753.
18. Emilsson V, Thorleifsson G, Zhang B, Leonardson AS, Zink F, Zhu J, Carlson S, Helgason A, Walters GB, Gunnarsdottir S, *et al*. Genetics of gene expression and its effect on disease. *Nature* 2008;452: 423–428.
19. Chen Y, Zhu J, Lum PY, Yang X, Pinto S, MacNeil DJ, Zhang C, Lamb J, Edwards S, Sieberts SK, *et al*. Variations in DNA elucidate molecular networks that cause disease. *Nature* 2008;452:429–435.
20. Teslovich TM, Musunuru K, Smith AV, Edmondson AC, Stylianou IM, Koseki M, Pirruccello JP, Ripatti S, Chasman DI, Willer CJ, *et al*. Biological, clinical and population relevance of 95 loci for blood lipids. *Nature* 2010;466:707–713.
21. Hao K, Bossé Y, Nickle DC, Paré PD, Postma DS, Laviolette M, Sandford A, Hackett TL, Daley D, Hogg JC, *et al*. Lung eQTLs to help reveal the molecular underpinnings of asthma. *PLoS Genet* 2012;8: e1003029.
22. Obeidat M, Miller S, Probert K, Billington CK, Henry AP, Hodge E, Nelson CP, Stewart CE, Swan C, Wain LV, *et al*. GSTCD and INTS12 regulation and expression in the human lung. *PLoS One* 2013;8: e74630.
23. Lamontagne M, Couture C, Postma DS, Timens W, Sin DD, Paré PD, Hogg JC, Nickle D, Laviolette M, Bossé Y. Refining susceptibility loci of chronic obstructive pulmonary disease with lung eqtls. *PLoS One* 2013;8:e70220.
24. Dimas AS, Deutsch S, Stranger BE, Montgomery SB, Borel C, Attar-Cohen H, Ingle C, Beazley C, Gutierrez Arcelus M, Sekowska M, *et al*. Common regulatory variation impacts gene expression in a cell type-dependent manner. *Science* 2009;325: 1246–1250.

25. Tam A, Wadsworth S, Dorscheid D, Man SF, Sin DD. The airway epithelium: more than just a structural barrier. *Ther Adv Respir Dis* 2011;5:255–273.
26. Holgate ST. The sentinel role of the airway epithelium in asthma pathogenesis. *Immunol Rev* 2011;242:205–219.
27. Holgate ST, Roberts G, Arshad HS, Howarth PH, Davies DE. The role of the airway epithelium and its interaction with environmental factors in asthma pathogenesis. *Proc Am Thorac Soc* 2009;6: 655–659.
28. Mercer BA, Lemaître V, Powell CA, D'Armiento J. The epithelial cell in lung health and emphysema pathogenesis. *Curr Respir Med Rev* 2006;2:101–142.
29. Thorley AJ, Tetley TD. Pulmonary epithelium, cigarette smoke, and chronic obstructive pulmonary disease. *Int J Chron Obstruct Pulmon Dis* 2007;2:409–428.
30. Carolan BJ, Heguy A, Harvey BG, Leopold PL, Ferris B, Crystal RG. Up-regulation of expression of the *ubiquitin carboxyl-terminal hydrolase L1* gene in human airway epithelium of cigarette smokers. *Cancer Res* 2006;66:10729–10740.
31. Gao C, Tignor NL, Salit J, Strulovici-Barel Y, Hackett NR, Crystal RG, Mezey JG. HEFT: eQTL analysis of many thousands of expressed genes while simultaneously controlling for hidden factors. *Bioinformatics* 2014;30:369–376.
32. Laurie CC, Doheny KF, Mirel DB, Pugh EW, Bierut LJ, Bhangale T, Boehm F, Caporaso NE, Cornelis MC, Edenberg HJ, *et al*.; GENEVA Investigators. Quality control and quality assurance in genotypic data for genome-wide association studies. *Genet Epidemiol* 2010;34: 591–602.
33. Boyle AP, Hong EL, Hariharan M, Cheng Y, Schaub MA, Kasowski M, Karczewski KJ, Park J, Hitz BC, Weng S, *et al*. Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res* 2012;22:1790–1797.
34. Howie B, Fuchsberger C, Stephens M, Marchini J, Abecasis GR. Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. *Nat Genet* 2012;44:955–959.
35. Schadt EE, Woo S, Hao K. Bayesian method to predict individual SNP genotypes from gene expression data. *Nat Genet* 2012;44:603–608.
36. Kang HM, Sul JH, Service SK, Zaitlen NA, Kong SY, Freimer NB, Sabatti C, Eskin E. Variance component model to account for sample structure in genome-wide association studies. *Nat Genet* 2010;42: 348–354.
37. Di Narzo AF, Cheng H, Lu J, Hao K. Meta-eQTL: a tool set for flexible eQTL meta-analysis. *BMC Bioinformatics* 2014;15:392.
38. Lambert JC, Ibrahim-Verbaas CA, Harold D, Naj AC, Sims R, Bellenguez C, DeStafano AL, Bis JC, Beecham GW, Grenier-Boley B, *et al*.; European Alzheimer's Disease Initiative (EADI); Genetic and Environmental Risk in Alzheimer's Disease; Alzheimer's Disease Genetic Consortium; Cohorts for Heart and Aging Research in Genomic Epidemiology. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat Genet* 2013;45:1452–1458.
39. Morris AP, Voight BF, Teslovich TM, Ferreira T, Segrè AV, Steinthorsdottir V, Strawbridge RJ, Khan H, Grallert H, Mahajan A, *et al*.; Wellcome Trust Case Control Consortium; Meta-Analyses of Glucose and Insulin-related traits Consortium (MAGIC) Investigators; Genetic Investigation of ANthropometric Traits (GIANT) Consortium; Asian Genetic Epidemiology Network–Type 2 Diabetes (AGEN-T2D) Consortium; South Asian Type 2 Diabetes (SAT2D) Consortium; DIAbetes Genetics Replication And Meta-analysis (DIAGRAM) Consortium. Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. *Nat Genet* 2012;44:981–990.
40. Locke AE, Kahali B, Berndt SI, Justice AE, Pers TH, Day FR, Powell C, Vedantam S, Buchkovich ML, Yang J, *et al*.; LifeLines Cohort Study; ADIPOGen Consortium; AGEN-BMI Working Group; CARDIOGRAMplusC4D Consortium; CKDGen Consortium; GLGC; ICBP; MAGIC Investigators; MuTHER Consortium; MIGen Consortium; PAGE Consortium; ReproGen Consortium; GENIE Consortium; International Endogene Consortium. Genetic studies of body mass index yield new insights for obesity biology. *Nature* 2015;518:197–206.
41. Purcell SM, Wray NR, Stone JL, Visscher PM, O'Donovan MC, Sullivan PF, Sklar P; International Schizophrenia Consortium. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* 2009;460:748–752.
42. Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* 2014;511:421–427.
43. Ferreira MA, Matheson MC, Duffy DL, Marks GB, Hui J, Le Souëf P, Danoy P, Baltic S, Nyholt DR, Jenkins M, *et al*.; Australian Asthma Genetics Consortium. Identification of *IL6R* and chromosome 11q13.5 as risk loci for asthma. *Lancet* 2011;378:1006–1014.
44. Schadt EE, Molony C, Chudin E, Hao K, Yang X, Lum PY, Kasarskis A, Zhang B, Wang S, Suver C, *et al*. Mapping the genetic architecture of gene expression in human liver. *PLoS Biol* 2008;6:e107.
45. Genevar - Welcome Trust Sanger Institute. Search tools and software. Date. [accessed 2015 Feb 26]. http://www.sanger.ac.uk/resources/software/genevar/
46. Westra HJ, Peters MJ, Esko T, Yaghootkar H, Schurmann C, Kettunen J, Christiansen MW, Fairfax BP, Schramm K, Powell JE, *et al*. Systematic identification of trans eQTLs as putative drivers of known disease associations. *Nat Genet* 2013;45:1238–1243.
47. Hindorff LA, Sethupathy P, Junkins HA, Ramos EM, Mehta JP, Collins FS, Manolio TA. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc Natl Acad Sci USA* 2009;106:9362–9367.
48. Flutre T, Wen X, Pritchard J, Stephens M. A statistical framework for joint eQTL analysis in multiple tissues. *PLoS Genet* 2013;9: e1003486.
49. Greenawalt DM, Dobrin R, Chudin E, Hatoum IJ, Suver C, Beaulaurier J, Zhang B, Castro V, Zhu J, Sieberts SK, *et al*. A survey of the genetics of stomach, liver, and adipose gene expression from a morbidly obese cohort. *Genome Res* 2011;21:1008–1016.
50. Dehghan A, Dupuis J, Barbalic M, Bis JC, Eiriksdottir G, Lu C, Pellikka N, Wallaschofski H, Kettunen J, Henneman P, *et al*. Meta-analysis of genome-wide association studies in >80 000 subjects identifies multiple loci for C-reactive protein levels. *Circulation* 2011;123:731–738.
51. Imboden M, Bouzigon E, Curjuric I, Ramasamy A, Kumar A, Hancock DB, Wilk JB, Vonk JM, Thun GA, Siroux V, *et al*. Genome-wide association study of lung function decline in adults with and without asthma. *J Allergy Clin Immunol* 2012;129:1218–1228.
52. Hancock DB, Artigas MS, Gharib SA, Henry A, Manichaikul A, Ramasamy A, Loth DW, Imboden M, Koch B, McArdle WL, *et al*. Genome-wide joint meta-analysis of SNP and SNP-by-smoking interaction identifies novel loci for pulmonary function. *PLoS Genet* 2012;8:e1003098.
53. Bossé Y, Hudson TJ. Toward a comprehensive set of asthma susceptibility genes. *Annu Rev Med* 2007;58:171–184.
54. Bernstein BE, Birney E, Dunham I, Green ED, Gunter C, Snyder M; ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature* 2012;489:57–74.
55. RegulomeDB. RegulomeDB has been updated to Version 1.1. 2011 [accessed 2015 Feb 26]. http://regulomedb.org/
56. Wang C, Rose-Zerilli MJ, Koppelman GH, Sandling JK, Holloway JW, Postma DS, Holgate ST, Bours V, Syvänen AC, Dideberg V. Evidence of association between interferon regulatory factor 5 gene polymorphisms and asthma. *Gene* 2012;504:220–225.
57. RegulomeDB. Data supporting chr7:128589426 (rs10954213). 2011 [accessed 2015 Feb 26]. http://regulomedb.org/snp/chr7/128589426
58. Bisgaard H, Bønnelykke K, Sleiman PM, Brasholt M, Chawes B, Kreiner-Møller E, Stage M, Kim C, Tavendale R, Baty F, *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;179:179–185.
59. Ferreira MA, McRae AF, Medland SE, Nyholt DR, Gordon SD, Wright MJ, Henders AK, Madden PA, Visscher PM, Wray NR, *et al.* Association between *ORMDL3, IL1RL1* and a deletion on chromosome 17q21 with asthma risk in Australia. *Eur J Hum Genet* 2011;19:458–464.
60. Galanter J, Choudhry S, Eng C, Nazario S, Rodríguez-Santana JR, Casal J, Torres-Palacios A, Salas J, Chapela R, Watson HG, *et al*. *ORMDL3* gene is associated with asthma in three ethnically diverse populations. *Am J Respir Crit Care Med* 2008;177:1194–1200.

61. Halapi E, Gudbjartsson DF, Jonsdottir GM, Bjornsdottir US, Thorleifsson G, Helgadottir H, Williams C, Koppelman GH, Heinzmann A, Boezen HM, *et al*. A sequence variant on 17q21 is associated with age at onset and severity of asthma. *Eur J Hum Genet* 2010;18:902–908.
62. Madore AM, Tremblay K, Hudson TJ, Laprise C. Replication of an association between 17q21 SNPs and asthma in a French-Canadian familial collection. *Hum Genet* 2008;123:93–95.
63. Sleiman PM, Annaiah K, Imielinski M, Bradfield JP, Kim CE, Frackelton EC, Glessner JT, Eckert AW, Otieno FG, Santa E, *et al*. *ORMDL3* variants associated with asthma susceptibility in North Americans of European ancestry. *J Allergy Clin Immunol* 2008; 122:1225–1227.
64. Tavendale R, Macgregor DF, Mukhopadhyay S, Palmer CN. A polymorphism controlling *ORMDL3* expression is associated with asthma that is poorly controlled by current medications. *J Allergy Clin Immunol* 2008;121:860–863.
65. Wu H, Romieu I, Sienra-Monge JJ, Li H, del Rio-Navarro BE, London SJ. Genetic variation in ORM1-like 3 (*ORMDL3*) and gasdermin-like (*GSDML*) and childhood asthma. *Allergy* 2009;64: 629–635.
66. Duffin R, Leitch AE, Sheldrake TA, Hallett JM, Meyer C, Fox S, Alessandri AL, Martin MC, Brady HJ, Teixeira MM, *et al*. The CDK inhibitor, R-roscovitine, promotes eosinophil apoptosis by down-regulation of Mcl-1. *FEBS Lett* 2009;583:2540–2546.
67. Leitch AE, Haslett C, Rossi AG. Cyclin-dependent kinase inhibitor drugs as potential novel anti-inflammatory and pro-resolution agents. *Br J Pharmacol* 2009;158:1004–1016.
68. Sekigawa T, Tajima A, Hasegawa T, Hasegawa Y, Inoue H, Sano Y, Matsune S, Kurono Y, Inoue I. Gene-expression profiles in human nasal polyp tissues and identification of genetic susceptibility in aspirin-intolerant asthma. *Clin Exp Allergy* 2009;39: 972–981.
69. Daley D, Park JE, He JQ, Yan J, Akhabir L, Stefanowicz D, Becker AB, Chan-Yeung M, Bossé Y, Kozyrskyj AL, *et al*. Associations and interactions of genetic polymorphisms in innate immunity genes with early viral infections and susceptibility to asthma and asthma-related phenotypes. *J Allergy Clin Immunol* 2012;130:1284–1293.
70. Colotta F, Re F, Muzio M, Bertini R, Polentarutti N, Sironi M, Giri JG, Dower SK, Sims JE, Mantovani A. Interleukin-1 type II receptor: a decoy target for IL-1 that is regulated by IL-4. *Science* 1993;261: 472–475.
71. Lang D, Knop J, Wesche H, Raffetseder U, Kurrle R, Boraschi D, Martin MU. The type II IL-1 receptor interacts with the IL-1 receptor accessory protein: a novel mechanism of regulation of IL-1 responsiveness. *J Immunol* 1998;161:6871–6877.
72. Müller B, Peri G, Doni A, Perruchoud AP, Landmann R, Pasqualini F, Mantovani A. High circulating levels of the IL-1 type II decoy receptor in critically ill patients with sepsis: association of high decoy receptor levels with glucocorticoid administration. *J Leukoc Biol* 2002;72: 643–649.
73. Re F, Muzio M, De Rossi M, Polentarutti N, Giri JG, Mantovani A, Colotta F. The type II "receptor" as a decoy target for interleukin 1 in polymorphonuclear leukocytes: characterization of induction by dexamethasone and ligand binding properties of the released decoy receptor. *J Exp Med* 1994;179:739–743.
74. Colotta F, Saccani S, Giri JG, Dower SK, Sims JE, Introna M, Mantovani A. Regulated expression and release of the IL-1 decoy receptor in human mononuclear phagocytes. *J Immunol* 1996;156: 2534–2541.
75. Penton-Rol G, Orlando S, Polentarutti N, Bernasconi S, Muzio M, Introna M, Mantovani A. Bacterial lipopolysaccharide causes rapid shedding, followed by inhibition of mRNA expression, of the IL-1 type II receptor, with concomitant up-regulation of the type I receptor and induction of incompletely spliced transcripts. *J Immunol* 1999; 162:2931–2938.
76. Brown EA, Dare HA, Marsh CB, Wewers MD. The combination of endotoxin and dexamethasone induces type II interleukin 1 receptor (IL-1r II) in monocytes: a comparison to interleukin 1 beta (IL-1 beta) and interleukin 1 receptor antagonist (IL-1ra). *Cytokine* 1996;8: 828–836.
77. Yu PW, Schuler LA, Kehrli M, Mattocks L, Nonnecke BJ, Czuprynski CJ. Effects of dexamethasone treatment on IL-1 receptor mRNA levels in vivo. *J Leukoc Biol* 1997;62:401–404.
78. Lin CH, Nai PL, Bien MY, Yu CC, Chen BC. Thrombin-induced CCAAT/enhancer-binding protein β activation and IL-8/CXCL8 expression via MEKK1, ERK, and p90 ribosomal S6 kinase 1 in lung epithelial cells. *J Immunol* 2014;192:338–348.
79. Hallstrand TS, Lai Y, Henderson WR Jr, Altemeier WA, Gelb MH. Epithelial regulation of eicosanoid production in asthma. *Pulm Pharmacol Ther* 2012;25:432–437.
80. Imoto Y, Tokunaga T, Matsumoto Y, Hamada Y, Ono M, Yamada T, Ito Y, Arinami T, Okano M, Noguchi E, *et al*. Cystatin SN upregulation in patients with seasonal allergic rhinitis. *PLoS One* 2013;8:e67057.
81. Bønnelykke K, Sleiman P, Nielsen K, Kreiner-Møller E, Mercader JM, Belgrave D, den Dekker HT, Husby A, Sevelsted A, Faura-Tellez G, *et al*. A genome-wide association study identifies *CDHR3* as a susceptibility locus for early childhood asthma with severe exacerbations. *Nat Genet* 2014;46:51–55.
82. Cantin AM, North SL, Hubbard RC, Crystal RG. Normal alveolar epithelial lining fluid contains high levels of glutathione. *J Appl Physiol (1985)* 1987;63:152–157.
83. Harju T, Mazur W, Merikallio H, Soini Y, Kinnula VL. Glutathione-S-transferases in lung and sputum specimens, effects of smoking and COPD severity. *Respir Res* 2008;9:80.
84. Zienolddiny S, Campa D, Lind H, Ryberg D, Skaug V, Stangeland LB, Canzian F, Haugen A. A comprehensive analysis of phase I and phase II metabolism gene polymorphisms and risk of non-small cell lung cancer in smokers. *Carcinogenesis* 2008;29:1164–1169.
85. Qiu W, Cho MH, Riley JH, Anderson WH, Singh D, Bakke P, Gulsvik A, Litonjua AA, Lomas DA, Crapo JD, *et al*.; ECLIPSE Investigators. Genetics of sputum gene expression in chronic obstructive pulmonary disease. *PLoS One* 2011;6:e24395.
86. Mägert HJ, Kreutzmann P, Drögemüller K, Ständker L, Adermann K, Walden M, John H, Korting HC, Forssmann WG. The 15-domain serine proteinase inhibitor LEKTI: biochemical properties, genomic organization, and pathophysiological role. *Eur J Med Res* 2002;7:49–56.
87. Postma DS, Bleecker ER, Amelung PJ, Holroyd KJ, Xu J, Panhuysen CI, Meyers DA, Levitt RC. Genetic susceptibility to asthma: bronchial hyperresponsiveness coinherited with a major gene for atopy. *N Engl J Med* 1995;333:894–900.
88. Noguchi E, Shibasaki M, Arinami T, Takeda K, Maki T, Miyamoto T, Kawashima T, Kobayashi K, Hamaguchi H. Evidence for linkage between asthma/atopy in childhood and chromosome 5q31-q33 in a Japanese population. *Am J Respir Crit Care Med* 1997;156: 1390–1393.
89. Yokouchi Y, Nukaga Y, Shibasaki M, Noguchi E, Kimura K, Ito S, Nishihara M, Yamakawa-Kobayashi K, Takeda K, Imoto N, *et al*. Significant evidence for linkage of mite-sensitive childhood asthma to chromosome 5q31-q33 near the interleukin 12 B locus by a genome-wide search in Japanese families. *Genomics* 2000;66: 152–160.
90. Kato A, Fukai K, Oiso N, Hosomi N, Murakami T, Ishii M. Association of *SPINK5* gene polymorphisms with atopic dermatitis in the Japanese population. *Br J Dermatol* 2003;148:665–669.
91. Kabesch M, Carr D, Weiland SK, von Mutius E. Association between polymorphisms in serine protease inhibitor, kazal type 5 and asthma phenotypes in a large German population sample. *Clin Exp Allergy* 2004;34:340–345.
92. Kusunoki T, Okafuji I, Yoshioka T, Saito M, Nishikomori R, Heike T, Sugai M, Shimizu A, Nakahata T. *SPINK5* polymorphism is associated with disease severity and food allergy in children with atopic dermatitis. *J Allergy Clin Immunol* 2005;115:636–638.
93. Himes BE, Jiang X, Hu R, Wu AC, Lasky-Su JA, Klanderman BJ, Ziniti J, Senter-Sylvia J, Lima JJ, Irvin CG, *et al*. Genome-wide association analysis in asthma subjects identifies *SPATS2L* as a novel bronchodilator response gene. *PLoS Genet* 2012;8:e1002824.
94. Steinke JW, Borish L. Th2 cytokines and asthma. Interleukin-4: its role in the pathogenesis of asthma, and targeting it for asthma treatment with interleukin-4 receptor antagonists. *Respir Res* 2001;2: 66–70.