Pharmacogenomics: novel loci identification via integrating gene differential analysis and eQTL analysis

Weiliang Qiu1, Angela J Rogers3, Amy Damask4, Benjamin A Raby1,2, Barbara J. Klanderman5, Qing Ling Duan1, Shiva Tyagi3, Simin Niu1, Christopher Anderson6, Ellen Cahir-McFarland8, Thomas J. Mariani7, Vincent Carey1 and Kelan G Tantisira1,2,*

1Channing Division of Network Medicine and 2Pulmonary Division, Department of Medicine, Brigham and Women’s Hospital and Harvard Medical School, Boston MA 02115, USA, 3Division of Pulmonary and Critical Care Medicine, Stanford University Medical Center, Stanford CA 94305–5236, USA, 4Novartis Institutes for BioMedical Research, Cambridge, MA 02139, USA, 5Partners Center for Personalized Genetic Medicine, Boston, MA 02115, USA, 6Division of Immunology, and 7Division of Neonatology and Program in Pediatric Molecular and Personalized Medicine, Department of Pediatrics, University of Rochester, Rochester, NY 14642, USA and 8Translational Sciences, Virology/PML, Biogen Idec, Weston, MA 02493, USA

Received October 10, 2013; Revised April 1, 2014; Accepted April 22, 2014

Nearly one-half of asthmatic patients do not respond to the most commonly prescribed controller therapy, inhaled corticosteroids (ICS). We conducted an expression quantitative trait loci (eQTL) analysis using >300 expression microarrays (from 117 lymphoblastoid cell lines) in corticosteroid (dexamethasone) treated and untreated cells derived from asthmatic subjects in the Childhood Asthma Management Program (CAMP) clinical trial. We then tested the associations of eQTL with longitudinal change in airway responsiveness to methacholine (LnPC20) on ICS. We identified 2484 cis-eQTL affecting 767 genes following dexamethasone treatment. A significant over-representation of LnPC20-associated cis-eQTL [190 single-nucleotide polymorphisms (SNPs)] among differentially expressed genes (odds ratio = 1.76, 95% confidence interval: 1.35–2.29) was noted in CAMP Caucasians. Forty-six of these 190 clinical associations were replicated in CAMP African Americans, including seven SNPs near six genes meeting criteria for genome-wide significance (P < 2 × 10⁻⁷). Notably, the majority of genome-wide findings would not have been uncovered via analysis of untreated samples. These results indicate that identifying eQTL after relevant environmental perturbation enables identification of true pharmacogenetic variants.

INTRODUCTION

Asthma is the most common chronic lung disease of childhood, affecting >3 million US children and leading to >200 000 hospitalizations per year (1). For long-term control of asthma, inhaled corticosteroids (ICS) are the mainstay of asthma therapy (2). However, clinical response to ICS therapy is widely variable among individuals (3) and is associated with important side effects.

Candidate genes studies have implicated multiple pharmacogenetic variants with replicated and/or functional association with ICS response in asthma (4–6). Genome-wide association studies (GWAS) have recently been applied to ICS response in asthma, identifying replicable associations in the T (7) and GLCCI1 (8) genes. One method to further enrich GWAS results is to study expression quantitative trait loci (eQTL) as an intermediate phenotype to identify functional candidates. In asthma, an eQTL approach was used to identify the highly replicated ORMNL3-GSDMB locus (9). For gene-environment studies, such as pharmacogenetics, several groups have studied eQTL in immortalized cells after exposure to environmental perturbations including radiation (10) and therapeutics such as

*To whom correspondence should be addressed. Tel: +1 6175250863; Fax: +1 6175250958; Email: kelan.tantisira@channing.harvard.edu

© The Author 2014. Published by Oxford University Press. All rights reserved. For Permissions, please email: journals.permissions@oup.com
corticocorticoids (11,12). These studies have consistently reported that cis-acting eQTL in association with environmentally mediated expression change are rare. Nonetheless, we previously reported ‘drug-specific’ eQTL associated with both TNC expression change following cellular corticosteroid stimulation and clinical response to ICS in asthma (12).

Recently, we identified a functional variant in GLCCI1 associated with reduced ICS response in four asthmatic cohorts that explains ~6% of variation in ICS lung function response (8). Notably, this promoter variant encoded for a cis-eQTL for GLCCI1 expression that was not drug-specific—i.e. the cis-signal in cellular models was noted in both the baseline and the corticosteroid-treated samples. However, GLCCI1 was strongly differentially expressed in these samples. Moreover, dexamethasone-stimulated GLCCI1 expression in CAMP lymphoblastoid cell lines (LCLs) was associated with a robust response to ICS [odds ratio, 3.22; 95% confidence interval (CI), 1.41–7.38], a finding consistent with altered expression as the functional basis for the observed genotypic associations. Thus, assuming differential expression marks the subset of genes truly responsive to treatment, post-treatment cis-eQTL within differentially expressed genes are likely to affect clinical treatment response, even if they do not fully mediate the differential expression response. We therefore hypothesized that differential expression in conjunction with cis-eQTL in corticosteroid-treated cells derived from asthmatics would harbor important clinical ICS response variants.

RESULTS

We conducted an expression eQTL analysis using 117 LCLs in paired fashion [corticosteroid (dexamethasone) treated and untreated cells], which were derived from 117 asthmatic subjects in the Childhood Asthma Management Program (CAMP) clinical trial. We then tested the associations of eQTL with longitudinal change in airway responsiveness to methacholine (LnPC20) on ICS.

Among 21 175 gene probes (tagging 17 370 unique genes), we detected 5846 gene probes (tagging 5437 unique genes) differentially expressed between dexamethasone-treated LCLs and sham-treated LCLs at false discovery rate (FDR) <5%. We performed a pathway analysis using DAVID to further evaluate the differentially expressed genes (13,14) and identified 54 biologic processes that are significantly over-represented (with FDR-adjusted $P$ value <0.01) in response to dexamethasone, including cell apoptosis, DNA replication and repair, immune response and leukocyte and lymphocyte activation (Supplementary Material, Table S3).

Focusing on dexamethasone-treated LCLs, we identified 2484 significant cis-eQTL [2376 single-nucleotide polymorphisms (SNPs) affecting 767 genes] (Supplementary Material, Table S4). Transcript variance explained at these loci was large (median = 0.16; range: 0.07–0.86). Genes harboring cis-eQTL identified in dexamethasone-treated cell lines were more likely to be differentially expressed by dexamethasone (OR = 2.7, 95% CI: 2.3–3.1).

To further check the robustness of our cis-eQTL results, we did two additional cis-eQTL analyses based on gene expression from dexamethasone-treated cell lines by further adjusting for a RNA batch variable and for the top 10 principal components from gene expression data, respectively. These two further adjustments did not significantly change our cis-eQTL results.

We tested these 2484 cis-eQTL for association with LnPC20 in 172 CAMP Caucasians randomized to ICS therapy. As shown in Table 1, 279 (12%) cis-eQTL are significantly associated with LnPC20. There is a greater degree of enrichment in the dexamethasone-treated cell lines and sham-treated cell lines) in this group of 279 cis-eQTL that were associated with LnPC20, with 190 associated versus 89 not associated with a differentially expressed gene (14 versus 9%) with an OR of 1.76 (1.35–2.30). To validate the importance of these loci, we tested the 190 cis-eQTL SNPs mapped to genes that are differentially expressed between dexamethasone-treated cell lines and sham-treated cell lines (Supplementary Material, Table S5) for association with LnPC20 in the CAMP African Americans (AAs) randomized to ICS therapy. We replicated the association in 46 (24%) of the SNPs (Supplementary Material, Table S7). Notably, seven variants near six genes met criteria for genome-wide significance when we applied Bonferroni $P$-value correction for the Fisher combined $P$-values for these seven variants (affecting SLFN5, NAPRTI, GATC2, CDN1A, THEM177 and SLC2A9, Table 2, with $P < 2 \times 10^{-7}$ for the 240 000 SNP tested).

To evaluate the cis-eQTL for treatment-responsive effects, we tested the interaction between the 2484 cis-eQTL and treatment assignment (inhaled corticosteroid or placebo) on LnPC20 in 581 CAMP Caucasians. We identified 266 cis-eQTL associated with a treatment-responsive change in airways responsiveness, of which 176 mapped to differentially expressed genes (Supplementary Material, Table S6), with an OR of 1.67 (1.28–2.20), (Table 1). We tested the 176 treatment-responsive eQTL for

Table 1. ORs for cis-eQTL association and SNP-treatment interaction with ICS response analyzed under dexamethasone treatment condition

<table>
<thead>
<tr>
<th>cis-eQTL association</th>
<th>Differentially expressed</th>
<th>Not differentially expressed</th>
<th>SNP-treatment interaction</th>
<th>Differentially expressed</th>
<th>Not differentially expressed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Associated with airways responsiveness**</td>
<td>190</td>
<td>89</td>
<td>Associated with treatment-responsive differential airways responsiveness**</td>
<td>176</td>
<td>86</td>
</tr>
<tr>
<td>Not associated with airways responsiveness*</td>
<td>1149</td>
<td>948</td>
<td>Not associated with treatment-responsive airways responsiveness*</td>
<td>1163</td>
<td>951</td>
</tr>
<tr>
<td>OR [95% CI]</td>
<td>1.76 [1.35, 2.30]</td>
<td>OR [95% CI]</td>
<td>1.67 [1.28, 2.20]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Airways responsiveness as measured by LnPC20 while on ICS medications over the 4-year-CAMP clinical trial.
**FDR-adjusted P-values < 0.05.
replication in the CAMP AAs and replicated the interaction in 39 loci (Supplementary Material, Table S8). Seven loci affected expression of five genes (SPATA20, ACOT4, BRWD1, ALG8 and NAPRT1) with a treatment interaction P-value that met genome-wide significance (Table 3). Figure 1 A and B shows the treatment-responsive effects of SNP rs12981009 on ACOT4 gene expression in both populations. As shown, among asthmatics taking placebo, the minor allele is associated with lower LnPC20 (risk), while among those assigned to ICS, the minor allele is associated with higher LnPC20 (protective) in both populations. For the SNPs that did not replicate in the AAs, we checked the distribution of the difference of minor allele frequencies (MAFs) between Caucasians and AAs with the distribution of the difference for SNPs that replicated in the AAs; we did not find a significant difference between the two distributions (data not shown).

**DISCUSSION**

In this work, we identified multiple novel genome-wide significant pharmacogenetic loci using a drug-response eQTL model. We found that twice as many clinically significant post-treatment cis-eQTL were located within genes differentially expressed between dexamethasone-treated cells and sham-treated cells in both our main effects and gene-environment analyses. Furthermore, our results demonstrated generalizability, with nearly one-quarter of ICS response eQTL identified in Caucasians also associated in AAs. Combined, 13 unique variants met criteria for genome-wide significance for the clinical association (Tables 2 and 3). Notably, only one SNP (rs2272053) from the ICS specific treatment analysis was also associated (P-value = 0.002) in the 409 subjects randomized to non-steroid therapy. These results indicate that focusing on SNPs that impact gene expression after exposure to a pertinent therapeutic stimulus enables identification of clinically relevant pharmacogenetic loci.

While none of our genome-wide significant eQTL genes has been previously studied in asthma, several are potentially compelling pharmacogenetic loci. Acyl-CoA-thiotransferase 4 (ACOT4) (Fig. 1) is a primary regulator of lipid oxidation in human peroxisomes. Peroxisome proliferator-activated receptor gamma (PPARγ) ligands demonstrate anti-inflammatory effects over and above those of corticosteroids by inhibiting cell growth and inducing apoptosis in human airways smooth muscle cells (15). Nicotinamide phosphoribosyltransferase 1 (NAPRT1) catalyzes the conversion of nicotinic acid (NA) to NA mononucleotide (NaMN) thereby increasing cellular nicotinamide adenine dinucleotide (NAD+) levels and preventing oxidative stress of the cells. Importantly, NA also results in increased cysteinyl leukotriene production (16); cysteinyl leukotrienes augment smooth muscle constriction and airway edema in asthma and are a recognized therapeutic target (17). Therefore, variation in the NA pathway has direct relevance to asthma pathogenesis and treatment response. Figure 2 illustrates the network connecting the five genes in Table 3 and their functionally similar genes from the literature using GeneMANIA (18). By applying DAVID functional annotation, we note that the genes shown in Table 2. Seven genome-wide significant dexamethasone cis-eQTL associated with LnPC20 change during ICS therapy

<table>
<thead>
<tr>
<th>eQTL SNP</th>
<th>Gene</th>
<th>chr</th>
<th>eQTL P-value</th>
<th>Differential expression P-value</th>
<th>Caucasian beta</th>
<th>Caucasian P-value*</th>
<th>AA beta</th>
<th>AA P-value*</th>
<th>Fisher’s combined P-value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs267182</td>
<td>SLFN5</td>
<td>17</td>
<td>1.88E−07</td>
<td>2.94E−07</td>
<td>0.31</td>
<td>2.57E−06</td>
<td>0.62</td>
<td>6.51E−06</td>
<td>4.33E−10</td>
</tr>
<tr>
<td>rs3793371</td>
<td>NAPRT1</td>
<td>8</td>
<td>1.25E−03</td>
<td>2.84E−03</td>
<td>−0.66</td>
<td>1.98E−06</td>
<td>−2.34</td>
<td>4.09E−05</td>
<td>1.97E−09</td>
</tr>
<tr>
<td>rs265125</td>
<td>GAPTC2</td>
<td>1</td>
<td>1.26E−06</td>
<td>2.21E−06</td>
<td>0.40</td>
<td>2.79E−06</td>
<td>0.51</td>
<td>4.27E−04</td>
<td>2.56E−08</td>
</tr>
<tr>
<td>rs4713999</td>
<td>CDKNIA</td>
<td>6</td>
<td>3.68E−08</td>
<td>1.51E−10</td>
<td>−0.40</td>
<td>8.22E−08</td>
<td>−0.34</td>
<td>2.31E−02</td>
<td>4.01E−08</td>
</tr>
<tr>
<td>rs2272053</td>
<td>TMEM177</td>
<td>2</td>
<td>1.38E−07</td>
<td>7.85E−08</td>
<td>0.26</td>
<td>2.59E−03</td>
<td>1.43</td>
<td>2.03E−06</td>
<td>1.06E−07</td>
</tr>
<tr>
<td>rs10939605</td>
<td>SLCA29</td>
<td>4</td>
<td>4.42E−05</td>
<td>1.06E−02</td>
<td>−0.27</td>
<td>1.28E−04</td>
<td>−0.80</td>
<td>4.62E−05</td>
<td>1.18E−07</td>
</tr>
<tr>
<td>rs6449090</td>
<td>SLCA29</td>
<td>4</td>
<td>1.04E−04</td>
<td>1.06E−02</td>
<td>−0.26</td>
<td>2.20E−04</td>
<td>−0.80</td>
<td>4.62E−05</td>
<td>1.98E−07</td>
</tr>
</tbody>
</table>

*Fisher’s combined P-value combines P-values for Caucasian and AA analyses.

Table 3. Seven genome-wide significant dexamethasone cis-eQTL associated with treatment-responsive change in airways responsiveness

<table>
<thead>
<tr>
<th>eQTL SNP</th>
<th>Gene</th>
<th>chr</th>
<th>eQTL P-value</th>
<th>Differential expression P-value</th>
<th>Caucasian beta</th>
<th>Caucasian P-value*</th>
<th>AA Beta</th>
<th>AA P-value*</th>
<th>Fisher’s combined P-value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs6506666</td>
<td>SPATA20</td>
<td>17</td>
<td>1.49E−04</td>
<td>1.34E−03</td>
<td>−0.45</td>
<td>2.05E−07</td>
<td>−1.48</td>
<td>1.78E−06</td>
<td>1.08E−11</td>
</tr>
<tr>
<td>rs132891009</td>
<td>ACOT4</td>
<td>14</td>
<td>2.23E−05</td>
<td>9.19E−14</td>
<td>−0.28</td>
<td>1.23E−03</td>
<td>−1.05</td>
<td>2.30E−07</td>
<td>6.50E−09</td>
</tr>
<tr>
<td>rs2037925</td>
<td>BRWD1</td>
<td>21</td>
<td>1.36E−05</td>
<td>5.32E−05</td>
<td>−0.42</td>
<td>3.02E−07</td>
<td>−0.48</td>
<td>3.91E−03</td>
<td>2.54E−08</td>
</tr>
<tr>
<td>rs2836987</td>
<td>BRWD1</td>
<td>21</td>
<td>1.10E−05</td>
<td>5.32E−05</td>
<td>−0.42</td>
<td>2.37E−07</td>
<td>−0.48</td>
<td>5.25E−03</td>
<td>2.67E−08</td>
</tr>
<tr>
<td>rs1380657</td>
<td>SPATA20</td>
<td>17</td>
<td>8.21E−05</td>
<td>1.34E−03</td>
<td>−0.42</td>
<td>1.30E−06</td>
<td>−1.25</td>
<td>1.63E−03</td>
<td>4.46E−08</td>
</tr>
<tr>
<td>rs1144764</td>
<td>ALG8</td>
<td>11</td>
<td>3.95E−12</td>
<td>2.02E−06</td>
<td>−0.28</td>
<td>4.81E−04</td>
<td>−0.82</td>
<td>5.71E−06</td>
<td>5.69E−08</td>
</tr>
<tr>
<td>rs3793371</td>
<td>NAPRT1</td>
<td>8</td>
<td>1.25E−05</td>
<td>2.84E−03</td>
<td>−0.73</td>
<td>1.98E−05</td>
<td>−2.29</td>
<td>2.39E−04</td>
<td>9.54E−08</td>
</tr>
</tbody>
</table>

*Fisher’s combined P-value combines P-values for Caucasian and AA analyses.

*P-values for Caucasians are two-sided, while those in AA are one-sided.
We identified few log-difference cis-eQTL (Supplementary Material, Table S9), that is, cis-eQTL associated with dexamethasone-induced changes in gene expression versus sham. Although two of the four (rs10505755, affecting C12orf59; rs962817 affecting ANKRD15) were associated with LnPC20 response in the 172 CAMP Caucasians, this finding was not replicated in CAMP AAs.

The rarity of cis-eQTL meeting stringent criteria in differential response to treatment is consistent with the prior studies of environmental perturbation eQTL (10–12). For instance, a recent study by Mangravite et al. (19), only six such differential cis-eQTL in LCLs treated with statins were detected. Similarly, Maranville et al. (11), noted just nine cis-eQTL in a study of dexamethasone-induced gene expression changes in the 172 CAMP Caucasians, this finding was not replicated in CAMP AAs. The rarity of cis-eQTL meeting stringent criteria in differential response to treatment is consistent with the prior studies of environmental perturbation eQTL (10–12). For instance, a recent study by Mangravite et al. (19), only six such differential cis-eQTL in LCLs treated with statins were detected. Similarly, Maranville et al. (11), noted just nine cis-eQTL in a study of dexamethasone-induced gene expression changes in the 172 CAMP Caucasians, this finding was not replicated in CAMP AAs.

Of the 74 dex-eQTL, 26 (35%) were also nominal eQTL for the difference in expression (log-difference cis-eQTL with P, 0.05), but only at P-values less stringent than the FDR used in the log-difference analysis (Fig. 3 and Supplementary Material, Table S10). Thus, clinically relevant log-difference (i.e. treatment response) cis-eQTL are likely to be missed by focusing on genome-wide thresholds for identification. Despite this, many sham-eQTL were indistinguishable from dex-eQTL in terms of their genotypic effect on gene expression. We reiterate that the sham-eQTL overlapping with our clinically relevant dex-eQTL all mapped to differentially expressed genes. Therefore, it is likely that in these genes the treatment effect on expression presents an additive interaction between the SNP and corticosteroid treatment that influences ICS response (Supplementary Material, Fig. S1).

To assess consistency of our cis-eQTL results with other datasets, we applied the web tool GTEx (Genotype-Tissue

Figure 3 are enriched for metabolic biological processes (Supplementary Material, Table S11).

While we focus on the SNPs associated (i.e. FDR-adjusted P-value < 0.05) to gene expressions from dexamethasone-treated cell lines (denoted as dex-eQTL) rather than SNPs associated (i.e. FDR-adjusted P-value < 0.05) to gene expressions from sham-treated samples (denoted as sham-eQTL) in this manuscript, we emphasize that among the 190 dex-eQTL that were associated with LnPC20 and were associated to genes that were differentially expressed between dexamethasone-treated and sham-treated cell lines, only 116 (60%) SNPs were also sham-eQTL. (The P-values and FDR adjusted P-values of eQTL tests for the other 40% (74) SNPs were shown in Supplementary Material, Table S10). Therefore, focusing only on eQTL from untreated (basal) samples would have missed many clinically relevant eQTL, including the majority of the genome-wide significant associations (Tables 2, 3 and Fig. 3).
Expression) eQTL browser (20). We found that 326 (13.1%) of the 2484 dex-eQTL we detected were previously published in the literature based on Lymphoblastoid tissues and that 269 (10.8%) of the 2484 cis-eQTL have also been detected by other researchers in the brain tissue. Unfortunately, lung tissue data are not yet available in GTEx.

Several potential limitations to our analysis exist. First, our clinical sample sizes were relatively small. However, the longitudinal nature of our trial increased statistical power. Additionally, we were able to identify novel replicable airways responsiveness loci in both our main effects and interaction analyses despite the sample size. Secondly, we tested immortalized B-lymphocytes rather than primary B-cells from trial participants. However, with regard to LCLs, Ding et al. (21) recently reported that 70% of cis-eQTL in LCLs is shared with skin, suggesting that our LCL analysis may also have large overlap with primary cells. Moreover, B-lymphocytes are crucial inflammatory mediators in asthma that produce IgE in response to cytokines and T-cell interactions to propagate allergic asthma (22–25). In turn, IgE levels are very highly correlated with PC20, our primary outcome phenotype (26). The fact that the results from differential expression and cis-eQTL analysis in LCL showed enrichment of LnPC20-association SNP further highlights the biological relevance of our findings.

In summary, cis-eQTL derived from steroid-treated cell lines, in conjunction with differential expression, identified multiple novel, genome-wide loci in asthma that would not have been noted via the analysis of basal expression alone. These replicated eQTL represent true pharmacogenetic loci in asthma. It is likely that our method of utilizing both eQTL data and gene expression response to environmental perturbation to identify SNPs associated with clinically important outcomes can be used.

Figure 2. Networks constructed by geneMANIA based on the five genes in Table 3 show these five genes are connected based on the literature. The black circles indicate the five genes we detected, while the gray circles are genes in the literature. The edges indicate that there exist evidences from literatures linking the genes connected by the edges.
more broadly to identify novel pharmacogenetic and other gene-environment loci.

MATERIALS AND METHODS

Population

The CAMP Genetics Ancillary Study was approved by each study center’s internal review board, and informed consent/assent was obtained from all participants and their parents. The CAMP clinical trial enrolled 1041 children aged 5–12 years (27,28) with mild-to-moderate asthma, who were randomly assigned to treatment with budesonide (a corticosteroid), nedocromil, or placebo and followed for a mean of 4.8 years. Because there was no difference in treatment effects of nedocromil or placebo, those groups are analyzed jointly in this work. We focus in this analysis on a subset of 581 Caucasian (172 randomized to ICS) and 97 AA (32 randomized to ICS) CAMP subjects for whom genome-wide genotyping (GWAS) is available. The descriptive statistics of these subjects are summarized in Supplementary Material, Table S1.

Airways responsiveness, as measured by the provocative concentration of methacholine causing a 20% decrement in FEV₁ (PC₂₀) values were obtained at 0, 8, 20, 32 and 44 months. Airway responsiveness was determined by methacholine testing with the Wright nebulizer technique using methods as previously described (29,30). PC₂₀ was log transformed (LnPC₂₀) for all analyses.

Gene expression arrays

LCLs derived from 151 CAMP Caucasian participants were utilized. Each cell line culture was split into two equal parts: one part was treated with 10⁻⁶ M dexamethasone (a corticosteroid); the other part was sham treated. After 6 h, expression levels were measured using the Illumina HumanRef8 v2 BeadChip (Illumina, San Diego, CA). Each pair of dexamethasone-treated and sham-treated arrays was in the same batch. The resulting gene expression profile data include expression levels for 22184 gene probes and 2 arrays for each individual. There were 464 arrays in total. In data pre-process stage, we removed 35 technically failed arrays, 65 arrays from subjects with missing clinical variables, 104 arrays without genotype data, 2 arrays without information about the principal components of genotype data, 1 array with the ratio of the 95th percentile of expression level to the 5th percentile, and 31 replicated arrays. For each of the two types (dexamethasone-treated and sham treated) of arrays, we kept only the array with the largest median expression level for subjects with replicated arrays. We also removed 49 probes with bad chromosome annotation and 960 probes in X or Y chromosome. We then did vst transformation and quantile normalization for all the remaining 234
arrays together (117 dexamethasone-treated arrays and 117 sham-treated arrays) to reduce the effects of technical noises and to make the distribution of expression level for each array closer to normal distribution. Data pre-processing steps are summarized in the Supplementary Material, Table S2. After filtering, we evaluated 21 175 gene probes (17 370 unique genes). We limited cis-eQTL analysis to 117 Caucasian individuals who were randomized to ICS treatment in the clinical trial and for whom GWAS genotyping data was available.

SNP genotyping

SNP genotyping was performed using Illumina’s Human-Hap550v3 Genotyping BeadChip (Illumina, Inc., San Diego, CA). Detailed QC methods have been published previously (31). Briefly, we removed markers for low clustering scores, for flanking sequences that failed to map to a unique position on the HG17 reference genome sequence, monomorphic markers, those with five or more Mendel errors, missing in >5% of subjects, minor allele frequency (MAF) <5%, and mapping to the X or Y chromosome. Among the remaining 431 484 SNPs, the average completion rate was >99%. We assessed genotype reproducibility by plating four subjects on each of 14 genotyping plates. All of these replicates had at least 99.8% concordance. We used Eigenstrat software (32) to adjust for four principal components of population substructure in all analyses.

Statistical analysis

As noted above, each of the 117 CAMP cell lines was split into two equal parts; one exposed to dexamethasone, the other to sham treatment (placebo). We first assessed the effect of dexamethasone treatment on gene expression levels by testing if $b_1 = 0$ for linear regression $\log 2(Dex/Sham) = b_1 + b_2age + b_3gender + b_44\text{PCs} + \text{error term}$.

We then performed cis-eQTL analyses to detect for SNP effects on gene expression, by testing a general linear model in an additive model (i.e. testing if $b_1 = 0$ for log2(Dex) = $b_0 + b_1SNP + b_2age + b_3gender + b_44\text{PCs} + \text{error term}$) focusing on log2 expression following dexamethasone treatment. We used this model to test for cis-eQTL (SNPs within 50 kb of a gene tested with the 21 175 gene probes).

We next evaluated the associations of the cis-eQTL with LnPC20 over time in 172 CAMP Caucasian subjects treated with ICS using generalized estimating equations (GEE). In the GEE model, we adjusted for potential confounding factors: gender and longitudinal measurements of age and height (i.e. treating age and height as time-dependant covariates given known change in PC20 with age).

In addition to our main effects model, we performed a SNP-treatment group interaction analysis on the entire cohort of 581 subjects for the eQTL, including the 172 subjects randomized to ICS and 409 subjects randomized to non-steroid therapy in the GEE model.

Given our main hypothesis, we specifically assessed whether cis-eQTL associated with LnPC20 are also more likely to affect genes that are differentially expressed after dexamethasone treatment. Finally, we tested the generalizability of the impact of the cis-eQTL on PC20 by testing whether they also affect methacholine responsiveness in the CAMP AA subjects.

For all analyses except replication tests, we used two-sided tests and adjusted the $P$-values using Benjamini and Hochberg’s method to control multiple testing (33). For clinical association tests, if multiple cis-eQTL SNP were in perfect LD, we tested only one of them for association with LnPC20. A test is claimed as significant if the FDR-adjusted $P$-value < 0.05. For replication, we used one-sided test and un-adjusted $P$-value with a $P$-value threshold of <0.05. The statistical software R was used to perform all analyses. Bioconductor (www.bioconductor.org, last accessed date for www.bioconductor.org is 28 April 2014) packages limma, GGtools, and gee were used to perform gene expression differential analysis, eQTL analyses and GEE analyses, respectively (34,35).

SUPPLEMENTARY MATERIAL

Supplementary Material is available at HMG online.

ACKNOWLEDGEMENTS

We thank all subjects for their ongoing participation in this study. We acknowledge the CAMP investigators and research team, supported by the National Heart, Lung, and Blood Institute (NHLBI), for collection of CAMP Genetic Ancillary Study data.

Conflict of Interest statement. All authors have no conflict of interests to report. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

FUNDING

This work was supported by National Institutes of Health (R01 HL092197, K23 HG003983, R01 NR13391) and U01 HL065899).

REFERENCES


