CHEST Online Supplement

# Association of *SERPINE2* With Asthma

*Blanca E. Himes, PhD; Barbara Klanderman, PhD; John Ziniti, BA; Jody Senter-Sylvia, MS; Manuel E. Soto-Quiros, MD, PhD; Lydiana Avila, MD; Juan C. Celedón, MD, DrPH; Christoph Lange, PhD; Thomas J. Mariani, PhD; Jessica Lasky-Su, ScD; Craig P. Hersh, MD, MPH; Benjamin A. Raby, MD, MPH; Edwin K. Silverman, MD, PhD; Scott T. Weiss, MD; and Dawn L. DeMeo, MD, MPH*

 **DOI: 10.1378/chest.10-2973**

**e-Appendix 1.**

## METHODS AND MATERIALS

### Subjects

Childhood Asthma Management Program

Our primary population consisted of 655 non-Hispanic white subjects from the Childhood Asthma Management Program (CAMP), a multi-center clinical trial that followed 1,041 children with asthma for four years and 84% of the original participants for 12 years[1]. Stringent inclusion criteria ensured that participants had mild to moderate asthma. CAMP subjects had increased airway responsiveness, as established by a bronchoprovocation test of up to 12.5mg/dl of methacholine resulting in 20% or greater forced expiratory volume in one second ($FEV_1$) reduction. During the trial, total immunoglobulin E (IgE) was measured from serum specimens using radioimmunoabsorbent assays, and spirometry was performed according to American Thoracic Society recommendations as previously described[1]. For the current study, we focused on phenotype data obtained during the randomization visit to maximize the number of subjects used for association analyses, without having to adjust for effects of trial medications. In addition to asthma diagnosis, we considered intermediate phenotypes that capture components of the Dutch hypothesis: $FEV_1$, the $FEV_1$ to forced vital capacity ratio ($FEV_1$/FVC), airways responsiveness (AHR) quantified as the log-transformed percentage dose of methacholine causing a 20% decrease in $FEV_1$, and log-transformed total serum IgE levels (LogIgE). CAMP participants and their parents provided DNA for family-based genetic studies. Additionally, blood was drawn from a subset of CAMP participants during either the Year 3 or Year 4 CAMPCS/2 follow-up visit at four of the eight CAMP study centers (Baltimore, Boston, Denver, and St. Louis) for gene expression profiling studies of CD4+ T lymphocytes[2].

CAMP/Illumina

In a previous asthma genome-wide association (GWA) study, a subset of 359 CAMP probands of the 655 CAMP probands who had genome-wide genotype data available were genetically matched using the Genetic Matching (GEM) algorithm[3] with 846 publicly available population controls[4]. Because CAMP/Illumina is better powered to measure asthma affection status than CAMP parent-child trios, the CAMP/Illumina results within and near *SERPINE2* were utilized to measure the association of *SERPINE2* variants with asthma affection status. No data is available on the asthma status of the Illumina controls. Hence, these subjects are considered to be "population controls" and can be expected to have asthma at the rate that is prevalent in the general population (~10%). Association of additional phenotypes could not be assessed in CAMP/Illumina due to the lack of pulmonary function and IgE data corresponding to the Illumina controls utilized.

Genetics of Asthma in Costa Rica

This population consisted of 426 probands from the Genetics of Asthma in Costa Rica study, comprised of Costa Rican schoolchildren with asthma and their parents[5,6]. Children had a high probability of having at least six great-grandparents born in the Central Valley of Costa Rica and were included in the study if they had a doctor's diagnosis of asthma and at least two respiratory symptoms or asthma attacks in the year prior to enrollment. Phenotyping of subjects in this study has been described previously[7]. Briefly, total IgE was measured from serum specimens using radioimmunoabsorbent assays and spirometry was performed following American Thoracic Society recommendations. For the current study, $FEV_1$, $FEV_1$/FVC, and LogIgE phenotypes

were defined as they were in CAMP. For asthma affection status and AHR, we selected available phenotypes most similar to those defined in CAMP: asthma affection status was defined as the presence of methacholine responsiveness (PD20 $\leq$8.58 μmoles) or evidence of bronchodilator responsiveness plus the recruitment criteria (i.e., doctor's diagnosis of asthma and at least two respiratory symptoms or asthma attacks in the year prior to enrollment)[6], and AHR was quantified as the log-transformed dose response slope for methacholine[8].

**Variant Selection and Genotyping**

We selected single nucleotide polymorphisms (SNPs) within *SERPINE2* to be genotyped based on two criteria: (1) previously reported associations in COPD cohorts, and (2) capturing maximal variation within the gene based on linkage disequilibrium (LD). For (1), we selected 14 SNPs that were previously associated with $FEV_1$ and/or $FEV_1/FVC$[9]. For (2), we selected 15 SNPs to capture LD across the *SERPINE2* gene with an $r^2$ >0.8 and a minor allele frequency >5%. The selection of LD-tagging SNPs was performed with TAGGER[10] using HapMap Phase 2 European American (CEU) data[11]. The data for 29 SNPs selected for genotyping in 655 CAMP trios were supplemented by GWA data for 10 additional SNPs that were available for a cohort of 403 CAMP probands and their parents. Because the candidate gene data were collected for a larger set of trios than the GWA data, the candidate gene results were used whenever possible. Genotyping was performed using four methods. Most CAMP *SERPINE2* SNPs selected for this study were genotyped using either a SEQUENOM MassARRAY matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometer (Sequenom, San Diego, CA) or an Illumina BeadStation 500G (Illumina, Inc., San Diego, CA) using the GoldenGate assay. Remaining CAMP SNPs and all Costa Rica SNPs were genotyped using Taqman real-

time PCR with an ABI Prism 7900 machine (Applied Biosystems, Foster City, CA). The CAMP GWA data was obtained with Illumina's HumanHap550 Genotyping BeadChip (Illumina, Inc., San Diego, CA) as detailed previously[4]. iCAP SNPs were genotyped using the Illumina BeadStation 500G with a GoldenGate assay. Genotyping pass rates for individual SNPs were >98%. All SNPs were in Hardy-Weinberg Equilibrium among parents or controls of each population (p-values >0.05).

**Peripheral Blood CD4+ T Lymphocyte Expression Profiling**

Peripheral CD4+ T lymphocytes were isolated from blood samples of CAMP participants as described previously[2]. Genome-wide gene expression profiles were generated with Illumina HumanRef8 v2 BeadChip arrays (Illumina Inc., San Diego, CA) using 100ng of CD4+ total RNA from each sample and the Illumina BeadStation 500G according to protocol. Resultant microarray data have been deposited in NCBI's Gene Expression Omnibus[12] and are accessible through GEO Series accession number GSE22324 (http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE22324).

**Statistical Analysis**

Family-based association tests were calculated under an additive inheritance model using Golden Helix PBAT version 6.4.0 in CAMP and Costa Rica parent-child trios to measures associations with asthma affection status, LogIgE, $FEV_1$, $FEV_1$/FVC and AHR[13]. No covariates were included in the analysis of asthma affection status or LogIgE. Analyses of $FEV_1$, $FEV_1$/FVC and AHR included statistical adjustment for age, gender, and height. LD among SNPs was inferred from $r^2$ measures in family data using Haploview[14]. The Cochran-Armitage trend test as

implemented in PLINK was used to measure association of asthma affection status in CAMP/Illumina and iCAP[4]. All other statistical analyses were performed in R[15]. Joint evidence for association across populations was measured by combining p-values using the Liptak method[16]. In combining p-values, hypothesis tests in replication populations had one-sided alternatives (based on the direction of association in CAMP or CAMP/Illumina) so that SNPs with association tests in opposite directions would not produce inappropriately small p-values. The nominal significance threshold, after correcting for multiple comparisons, was equal to 2.9E-04 (= 0.05/175, where 175 was the number of measurements made: (39 SNPs x 4 phenotypes) + (19 SNPs x 1 phenotype)). Population-based expression quantitative trait loci (eQTL) analyses of quantile-normalized *SERPINE2* transcript (Illumina probe ILMN_1655595) abundance were conducted using generalized least squares models with the *nlme* R package, adjusting for age, sex and four principle components derived from EIGENSTRAT[17] analysis of genome-wide SNP data (see[2] for details). For this study, analyses were restricted to SNPs mapping to within 50kb of *SERPINE2* (n=32), and associations were declared significant at a False Discovery Rate of 0.05 in the context of a genome-wide screen for cis-acting (i.e. within 50kb) eQTL, corresponding to a p-value of 0.0008.

**References**


1. The Childhood Asthma Management Program (CAMP): design, rationale, and methods. Childhood Asthma Management Program Research Group. *Control Clin Trials.* 1999; 20(1):91-120.
2. Murphy A, Chu JH, Xu M, Carey VJ, Lazarus R, Liu A, et al. Mapping of numerous disease-associated expression polymorphisms in primary peripheral blood CD4+ lymphocytes. *Hum Mol Genet.* 2010; 19(23):4745-4757.
3. Luca D, Ringquist S, Klei L, Lee AB, Gieger C, Wichmann HE, et al. On the use of general control samples for genome-wide association studies: genetic matching highlights causal variants. *Am J Hum Genet.* 2008; 82(2):453-463.
4. Himes BE, Hunninghake GM, Baurley JW, Rafaels NM, Sleiman P, Strachan DP, et al. Genome-wide association analysis identifies PDE4D as an asthma-susceptibility gene. *Am J Hum Genet.* 2009; 84(5):581-593.
5. Hunninghake GM, Soto-Quiros ME, Avila L, Ly NP, Liang C, Sylvia JS, et al. Sensitization to Ascaris lumbricoides and severity of childhood asthma in Costa Rica. *J Allergy Clin Immunol.* 2007; 119(3):654-661.
6. Hunninghake GM, Soto-Quiros ME, Avila L, Su J, Murphy A, Demeo DL, et al. Polymorphisms in IL13, total IgE, eosinophilia, and asthma exacerbations in childhood. *J Allergy Clin Immunol.* 2007; 120(1):84-90.
7. Hersh CP, Soto-Quiros ME, Avila L, Lake SL, Liang C, Fournier E, et al. Genome-wide linkage analysis of pulmonary function in families of children with asthma in Costa Rica. *Thorax.* 2007; 62(3):224-230.
8. O'Connor GT, Sparrow D, Weiss ST. A prospective longitudinal study of methacholine airway responsiveness as a predictor of pulmonary-function decline: the Normative Aging Study. *Am J Respir Crit Care Med.* 1995; 152(1):87-92.
9. Demeo DL, Mariani TJ, Lange C, Srisuma S, Litonjua AA, Celedon JC, et al. The SERPINE2 gene is associated with chronic obstructive pulmonary disease. *Am J Hum Genet.* 2006; 78(2):253-264.
10. de Bakker PI, Yelensky R, Pe'er I, Gabriel SB, Daly MJ, Altshuler D. Efficiency and power in genetic association studies. *Nat Genet.* 2005; 37(11):1217-1223.
11. The International HapMap Consortium. The International HapMap Project. *Nature.* 2003; 426(6968):789-796.
12. Edgar R, Domrachev M, Lash AE. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.* 2002; 30(1):207-210.
13. Lange C, DeMeo D, Silverman EK, Weiss ST, Laird NM. PBAT: tools for family-based association studies. *Am J Hum Genet.* 2004; 74(2):367-369.
14. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics.* 2005; 21(2):263-265.
15. *R Development Core Team. R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing, 2008.
16. Liptak T. On the combination of independent tests. *Magyar Tud Akad Mat Kutato Int Kozl.* 1958; 3(171-197.
17. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet.* 2006; 38(8):904-909.

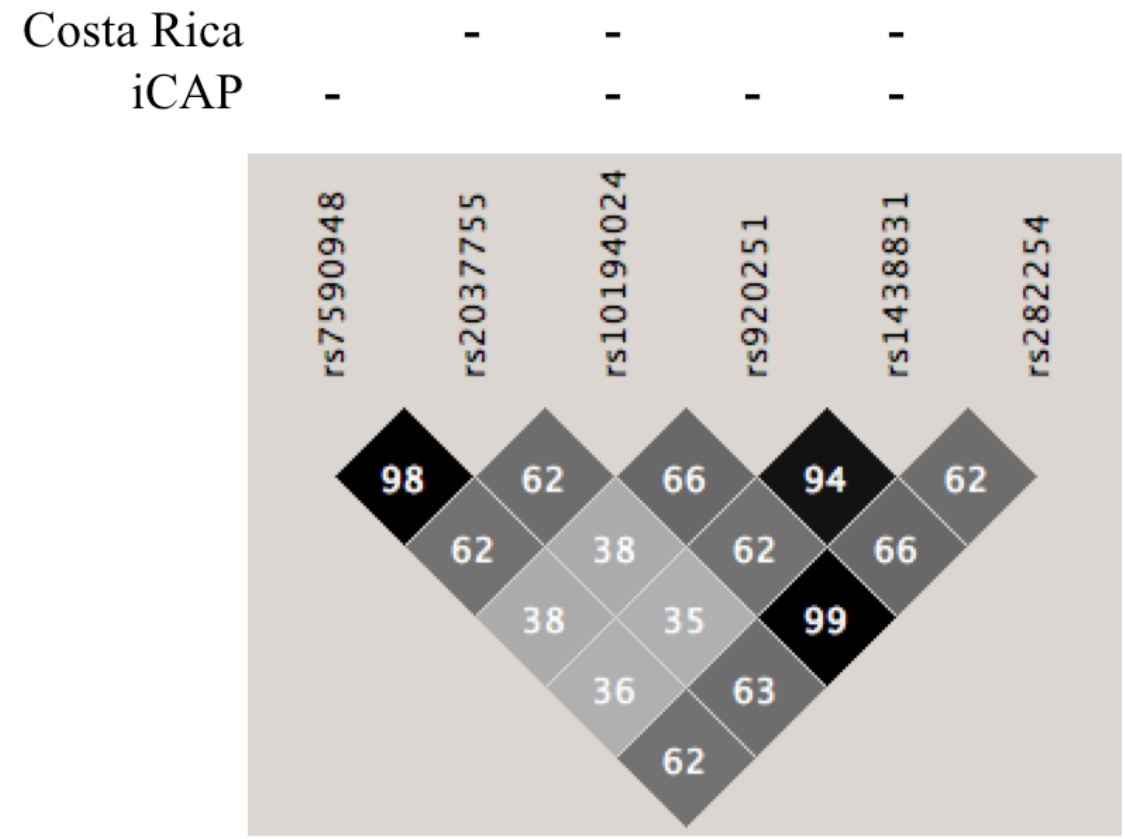
**e-Table 1.** SNPs genotyped in CAMP parent-child trios for this *SERPINE2* candidate gene study.

| SNP | CHR | BP | Allele | MAF | Parent MAF |
|---|---|---|---|---|---|
| rs6734100 | 2 | 224550239 | G | 0.14 | 0.12 |
| rs7597833 | 2 | 224550394 | T | 0.49 | 0.49 |
| rs729631 | 2 | 224553163 | C | 0.18 | 0.17 |
| rs10164837 | 2 | 224555512 | C | 0.30 | 0.29 |
| rs975278 | 2 | 224555951 | T | 0.18 | 0.17 |
| rs7605945 | 2 | 224556050 | C | 0.21 | 0.20 |
| rs4674835 | 2 | 224562037 | A | 0.31 | 0.31 |
| rs7608941 | 2 | 224562312 | C | 0.24 | 0.23 |
| rs6712954 | 2 | 224564894 | A | 0.07 | 0.07 |
| rs10191694 | 2 | 224565510 | A | 0.34 | 0.34 |
| rs6738983 | 2 | 224566994 | T | 0.32 | 0.33 |
| rs6715768 | 2 | 224570735 | A | 0.19 | 0.20 |
| rs3795879 | 2 | 224571065 | T | 0.19 | 0.20 |
| rs6747096 | 2 | 224571086 | G | 0.19 | 0.20 |
| rs1866153 | 2 | 224572796 | T | 0.15 | 0.15 |
| rs3795877 | 2 | 224574421 | G | 0.19 | 0.20 |
| rs7581619 | 2 | 224574631 | A | 0.11 | 0.10 |
| rs2118409 | 2 | 224577887 | G | 0.26 | 0.26 |
| rs3948261 | 2 | 224578136 | C | 0.26 | 0.26 |
| rs1530020 | 2 | 224579614 | C | 0.26 | 0.26 |
| rs7590948 | 2 | 224581934 | G | 0.44 | 0.45 |
| rs6436459 | 2 | 224588580 | T | 0.24 | 0.24 |
| rs13022548 | 2 | 224592334 | C | 0.16 | 0.16 |
| rs4368321 | 2 | 224595584 | A | 0.18 | 0.19 |
| rs10194024 | 2 | 224599084 | C | 0.43 | 0.43 |
| rs7562213 | 2 | 224606671 | A | 0.33 | 0.34 |
| rs840088 | 2 | 224608088 | T | 0.34 | 0.35 |
| rs7579646 | 2 | 224608940 | A | 0.19 | 0.19 |
| rs1438831 | 2 | 224614559 | T | 0.32 | 0.34 |

**e-Table 2.** SNPs genotyped (X) in each population or with CAMP eQTL data available.

| SNP | CHR | BP | CAMP Trio | CAMP/ Illumina | Costa Rica | iCAP | CAMP eQTL |
|---|---|---|---|---|---|---|---|
| rs6754561 | 2 | 224547940 | | X | | | X |
| rs6734100 | 2 | 224550239 | X | | X | X | |
| rs7597833 | 2 | 224550394 | X | X | X | X | X |
| rs729631 | 2 | 224553163 | X | | | X | |
| rs10164837 | 2 | 224555512 | X | | X | X | |
| rs975278 | 2 | 224555951 | X | X | | X | X |
| rs7605945 | 2 | 224556050 | X | | | X | |
| rs4674835 | 2 | 224562037 | X | | | X | |
| rs7608941 | 2 | 224562312 | X | | | X | |
| rs6712954 | 2 | 224564894 | X | | X | X | |
| rs10191694 | 2 | 224565510 | X | X | | X | X |
| rs6738983 | 2 | 224566994 | X | X | X | X | X |
| rs13396978 | 2 | 224568460 | | X | | | X |
| rs11695803 | 2 | 224569405 | | X | | | X |
| rs6715768 | 2 | 224570735 | X | | | X | |
| rs3795879 | 2 | 224571065 | X | | X | | |
| rs6747096 | 2 | 224571086 | X | | X | X | |
| rs1866153 | 2 | 224572796 | X | | | X | |
| rs3795877 | 2 | 224574421 | X | X | X | X | X |
| rs7581619 | 2 | 224574631 | X | | | | |
| rs2118409 | 2 | 224577887 | X | | | X | |
| rs3948261 | 2 | 224578136 | X | | | X | |
| rs1530020 | 2 | 224579614 | X | | | X | |
| rs7590948 | 2 | 224581934 | X | X | | X | |
| rs7563931 | 2 | 224585805 | | X | | | X |
| rs6436459 | 2 | 224588580 | X | X | | X | X |
| rs2037755 | 2 | 224591549 | | X | | | X |
| rs13022548 | 2 | 224592334 | X | | | X | |
| rs2083121 | 2 | 224593246 | | | | | |
| rs4368321 | 2 | 224595584 | X | | | X | |
| rs13008520 | 2 | 224597627 | | X | | | X |
| rs10194024 | 2 | 224599084 | X | X | | X | X |
| rs920251 | 2 | 224601189 | | X | | X | X |
| rs6719480 | 2 | 224601673 | | X | | | X |
| rs7562213 | 2 | 224606671 | X | | | X | |
| rs840088 | 2 | 224608088 | X | | | X | |
| rs7579646 | 2 | 224608940 | X | | | X | |
| rs1438831 | 2 | 224614559 | X | X | | X | X |
| rs282254 | 2 | 224615591 | | X | | | X |

**e-Figure 1.** LD among 8 SNPs that were genotyped in both CAMP and Costa Rica. LD was calculated using corresponding trio data and is expressed as $r^2$ values.



**e- Figure 2.** LD among SNPs that were associated with asthma at a nominally significant level (i.e. $p<0.05$) in CAMP/Illumina. LD was calculated using CAMP trio data and is expressed as $r^2$ values. Dashes denote SNPs genotyped in Costa Rica and/or iCAP. At least one SNP from each pair of SNPs that is in LD at $r^2>.94$ was genotyped in Costa Rica or iCAP.