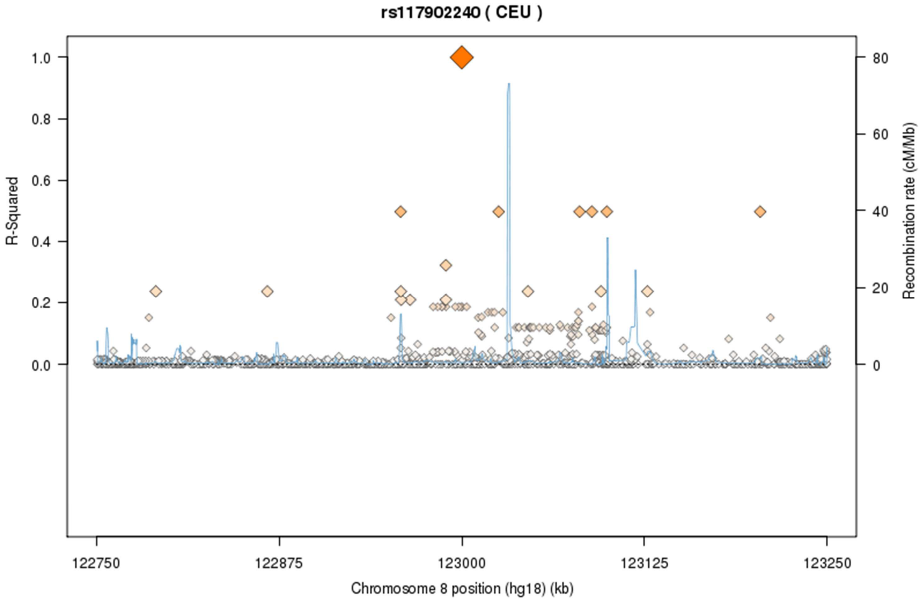
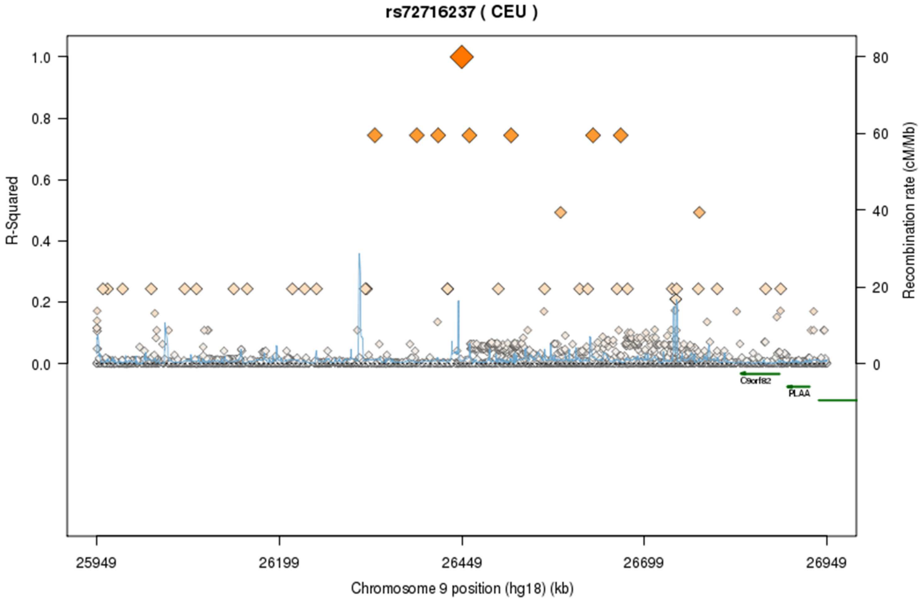
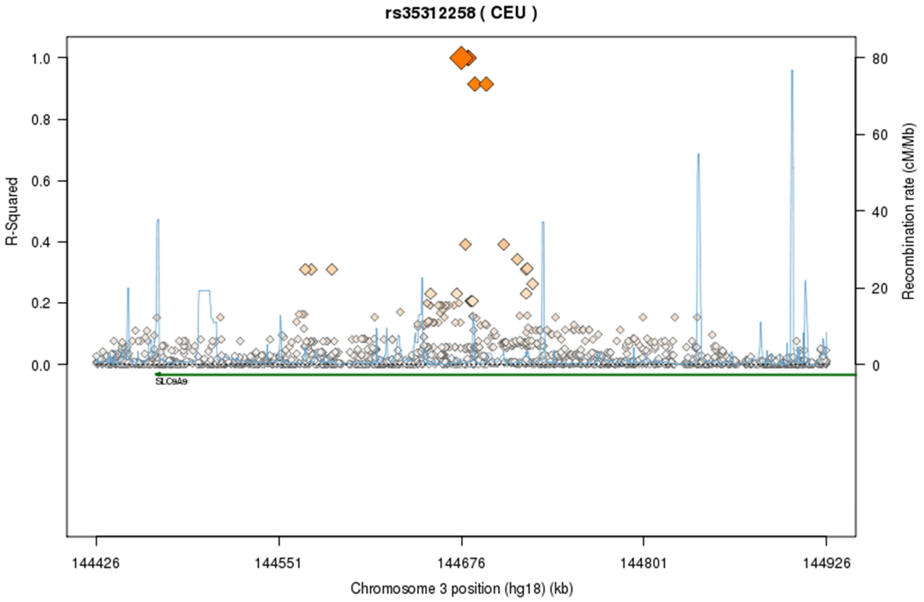


rs117902240 ( CEU )



rs72716237 ( CEU )



rs35312258 ( CEU )

# Genome-wide interaction study of dust mite allergen on lung function in children with asthma

## Online Supplementary Data

Erick Forno, MD MPH[1,2], Joanne Sordillo, MD MPH[3], John Brehm, MD MPH[1,2], Wei Chen, PhD[1,2], Takis Benos, PhD[2], Qi Yan, PhD[1,2], Lydiana Avila, MD[4], Manuel Soto-Quirós, MD[4], Michelle M Cloutier, MD[5], Angel Colón-Semidey, MD[6], Maria Alvarez, MD[6], Edna Acosta-Pérez, PhD[7], Scott T Weiss, MD MS[3], Augusto A Litonjua, MD MPH[3], Glorisa Canino, PhD[7], Juan C Celedón, MD DrPH[1,2]

*[1]Division of Pediatric Pulmonary Medicine, Allergy, and Immunology, Children's Hospital of Pittsburgh of UPMC, Pittsburgh, PA; [2]University of Pittsburgh School of Medicine, Pittsburgh, PA; [3]Channing Division of Network Medicine, Department of Medicine, Harvard Medical School and Brigham and Women's Hospital, Boston, MA; [4]Department of Pediatric Pulmonology, Hospital Nacional de Niños, San José, Costa Rica; [5]Department of Pediatrics, University of Connecticut Health Center, Farmington, CT; [6]Department of Pediatrics and [7]Behavioral Sciences Research Institute, University of Puerto Rico, San Juan, PR.*

**Corresponding author:** Juan C. Celedón, MD, DrPH, FAAAAI
Division of Pulmonary Medicine, Allergy and Immunology
Children's Hospital of Pittsburgh of UPMC
4401 Penn Avenue, Pittsburgh, PA 15224
Phone: 412.692.8429; Fax: 412.692.7636
Email: juan.celedon@chp.edu

## FEV$_1$ – Functional analysis of SNPs rs117902240, rs72716237, and rs35312258

Using data from ENCODE[1] and is-rSNP[2], a software designed to predict *in silico* whether a SNP has a regulatory effect on transcription factor binding sites, we found that while SNP rs117902240 in chromosome 8q24.13 is intergenic, it lies within <1 Kbp from a transcription factor binding island identified in ENCODE using ChIP-seq. Transcription factor C/EBP-beta (CEBPB) binds in this region in human lung fibroblasts (ENCODE accession ENCSR000EFM). CEBPB forms part of the signaling pathway of interleukin (IL)-17, nuclear factor Kappa-B (NF-κB), and interferon (IFN-)γ (see **Figure 3** in the main document)[3 4].

SNP rs72716237, in chromosome 9p21.2, lies right outside (<200bp) a transcription factor binding site for MAFK (ENCODE accession ENCSR000EFH), whose expression can be induced by transcription growth factor (TGF-)β[5]. MAFK regulates expression of NF-κB[6] and NF-E2-related factor 2 (NRF2), implicated in antioxidant pathways[7]. Of note, MAFK and NRF2 may also regulate CFTR gene expression[8].

SNP rs35312258, in chromosome 3q24, is an intronic SNP in gene *SLC9A9* (solute carrier family 9, subfamily A, member 9), a Na+/H+ transmembrane exchanger that has been linked to autism and ADHD[9], and more recently to increased arterial stiffness[10].

## Supplemental Figures – Titles and Legends

### Figure S1 – Regional LD plots of the three top SNPs for $FEV_1$

Linkage disequilibrium (LD) plots generated at the SNP Annotation and Proxy Search (SNAP) website (https://www.broadinstitute.org/mpg/snap/ldplot.php) using data from the 1,000 Genomes Pilot 1.

**Table S1 – Main effects of the three top SNPs on FEV$_1$, in PRGOAL as a whole and stratified by dust mite allergen level.**

| | rs117902240 | rs72716237 | rs35312258 |
|---|---|---|---|
| **Non-stratified** | +0.07 (0.10)<br>p=0.48 | –0.04 (0.09)<br>p=0.64 | +0.03 (0.052)<br>p=0.33 |
| **Low dust mite** | **+0.84 (0.15)**<br>**p=1.2x10$^{-7}$** | *–0.39 (0.11)*<br>*p=5.3x10$^{-4}$* | *–0.14 (0.052)*<br>*p=0.0061* |
| **High dust mite** | *–0.41 (0.14)*<br>*p=0.0032* | **+0.49 (0.15)**<br>**p=0.0011** | **+0.21 (0.051)**<br>**p=1.0x10$^{-4}$** |

Shown are the beta coefficients (standard error) and p-values for the additive, adjusted genotypic models in PRGOAL, for the whole cohort or stratified by low (<2mcg/g, n=227) or high (≥2 mcg/g, n=213) dust mite allergen level. **Positive association in bold**, *negative association in italics*. All models adjusted for age, sex, height, and height$^2$.

**Table S2** – **SNPs in IL-9, IL-10, *TGF-β1*, or P2RY2 associated with FEV1**

| GENE | SNP | A1 | Covariate[1] | BETA | P-VAL |
|---|---|---|---|---|---|
| **IL9** | rs2066758 | G | Genotype main effect | 0.1604 | 0.046 |
| **IL10** | rs4390174 | G | ***Genotype*dust mite*** | 0.082 | 0.039 |
| | rs3024498 | G | Genotype main effect | 0.076 | 0.042 |
| | rs6673928 | A | Genotype main effect | 0.075 | 0.046 |
| **P2RY2** | rs62285907 | G | ***Genotype*dust mite*** | 0.16 | 0.0033 |
| | rs58448575 | A | ***Genotype*dust mite*** | -0.22 | 0.0075 |
| | rs2276763 | A | ***Genotype*dust mite*** | 0.57 | 0.0077 |
| | rs7646193 | C | ***Genotype*dust mite*** | -0.19 | 0.011 |
| | rs58448575 | A | Genotype main effect | -0.19 | 0.031 |
| | rs62285878 | A | ***Genotype*dust mite*** | -0.31 | 0.042 |
| | rs116537786 | A | ***Genotype*dust mite*** | 0.21 | 0.044 |
| | rs16863376 | G | ***Genotype*dust mite*** | -0.29 | 0.047 |
| | rs765281 | G | ***Genotype*dust mite*** | -0.11 | 0.047 |
| **TGFB1** | rs79033882 | G | ***Genotype*dust mite*** | -0.076 | 0.037 |

Models included the main effects of genotype (additive) and dust mite allergen level (dichotomized <2 vs ≥2 mcg/g) and the interaction term (*genotype*dust mite allergen level*), and were additionally adjusted for age, sex, study site (San Juan, PR, or Hartford, CT), height, and height$^2$, and principal components. SNP=single nucleotide polymorphism; A1=minor allele; BETA=beta coefficient; P-VAL=one-sided p-value.

**Table S3** – **GWIS results for dust mite allergen level and $FEV_1$/FVC in PRGOAL**

| SNP | CHR | BP | A1 | A2 | BETA | P-VAL | MAF |
|---|---|---|---|---|---|---|---|
| rs17048089 | 3 | 68923779 | G | A | -0.067 | **4.01E-07** | 0.204 |
| rs115997623 | 4 | 119115917 | A | G | 0.37 | **3.94E-07** | 0.011 |
| rs10021113 | 4 | 159678591 | A | G | 0.36 | **1.51E-07** | 0.013 |
| rs75863350 | 5 | 148423424 | A | G | 0.47 | **2.79E-07** | 0.013 |
| rs78930606 | 6 | 119078652 | C | A | -0.32 | **1.49E-07** | 0.018 |

Models included the main effects of genotype (additive) and dust mite allergen level (dichotomized <2 vs ≥2 mcg/g) and the interaction term (*genotype*dust mite allergen level*), and were additionally adjusted for age, sex, study site (San Juan, PR, or Hartford, CT), height, and height$^2$, and principal components. SNP=single nucleotide polymorphism; CHR=chromosome; BP=base position; A1=minor allele; A2=major allele; BETA=beta coefficient for interaction term; P-VAL=interaction p-value; MAF=minor allele frequency.

**Table S4** – **GWIS replication results for $FEV_1$/FVC in CAMP and the GACRS**

| **$FEV_1$/FVC** | **CAMP Caucasians** | | | **CAMP Hispanics** | | | **GACRS** | | | **COMBINED P-VAL** | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **SNP** | **A1** | **BETA** | **P-VAL** | **A1** | **BETA** | **P-VAL** | **A1** | **BETA** | **P-VAL** | **All cohorts** | **Replication only** |
| rs17048089 | G | 1.35 | 0.20 | G | -1.077 | 0.39 | G | -0.22 | 0.45 | **1.6E-05** | 0.35 |
| rs115997623 | T | 0.64 | 0.44 | T | NA | NA | T | -13.55 | **0.04** | **1.5E-06** | 0.10 |
| rs10021113 | T | NA | NA | T | NA | NA | T | NA | NA | NA | NA |
| rs75863350 | A | NA | NA | A | NA | NA | A | NA | NA | NA | NA |
| rs78930606 | C | 10.99 | **0.037** | C | NA | NA | C | NA | NA | **9.9E-07** | NA |

Models included the main effects of genotype (additive) and dust mite allergen level (dichotomized <2 vs ≥2 mcg/g) and the interaction term (*genotype*dust mite allergen level*), and were additionally adjusted for age, sex, study site, and principal components. SNP=single nucleotide polymorphism; A1=minor allele; BETA=beta coefficient for interaction term; P-VAL=interaction p-value; MAF=minor allele frequency; NA=not available (genotyped or imputed). Combined P-values calculated only if at least two cohorts were significant.

**Table S5 – *Post hoc* power calculations for interaction term based on results from SNP rs117902240**

| **Alpha =** | **0.05** | **0.01** | **0.005** | **0.001** |
|---|---|---|---|---|
| **Power =** | 0.8975 | 0.7427 | 0.6630 | 0.4749 |

Shown are power calculations using different levels of significance. Calculations were performed using Quanto (http://biostats.usc.edu/Quanto.html), given a population mean FEV1 of 1.87 with SD=0.65 (see Table 1, main manuscript), an exposure proportion of 48.4% (Table 1), and risk allele frequency of 0.015 (rs117902240, Table 2), with N=440, and based on the observed main effects of the additive SNP and dust mite level as well as their interaction, using a 2-sided test (null: no significant interaction).

## Online Supplement References

1. ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. Nature 2012;**489**(7414):57-74.
2. Macintyre G, Bailey J, Haviv I, et al. is-rSNP: a novel technique for in silico regulatory SNP detection. Bioinformatics 2010;**26**(18):i524-30.
3. Ruddy MJ, Wong GC, Liu XK, et al. Functional cooperation between interleukin-17 and tumor necrosis factor-alpha is mediated by CCAAT/enhancer-binding protein family members. The Journal of biological chemistry 2004;**279**(4):2559-67.
4. Gaffen SL. Structure and signalling in the IL-17 receptor family. Nat Rev Immunol 2009;**9**(8):556-67.
5. Okita Y, Kamoshida A, Suzuki H, et al. Transforming growth factor-beta induces transcription factors MafK and Bach1 to suppress expression of the heme oxygenase-1 gene. The Journal of biological chemistry 2013;**288**(28):20658-67.
6. Hwang YJ, Lee EW, Song J, et al. MafK positively regulates NF-kappaB activity by enhancing CBP-mediated p65 acetylation. Scientific reports 2013;**3**:3242.
7. Ishikawa T. Genetic polymorphism in the NRF2 gene as a prognosis marker for cancer chemotherapy. Frontiers in genetics 2014;**5**:383.
8. Zhang Z, Leir SH, Harris A. Oxidative stress regulates CFTR gene expression in human airway epithelial cells through a distal antioxidant response element. Am J Respir Cell Mol Biol 2015;**52**(3):387-96.
9. Zayats T, Athanasiu L, Sonderby I, et al. Genome-wide analysis of attention deficit hyperactivity disorder in Norway. PLoS One 2015;**10**(4):e0122501.
10. Levy D, Larson MG, Benjamin EJ, et al. Framingham Heart Study 100K Project: genome-wide associations for blood pressure and arterial stiffness. BMC Med Genet 2007;**8 Suppl 1**:S3.