CHEST Online Supplement

Expression Quantitative Trait Methylation Analysis Reveals Methylomic Associations With Gene Expression in Childhood Asthma

Soyeon Kim, PhD; Erick Forno, MD, MPH; Rong Zhang, PhD; Hyun Jung Park, PhD; Zhongli Xu; Qi Yan, PhD; Nadia Boutaoui, PhD; Edna Acosta-Pérez, PhD; Glorisa Canino, PhD; Wei Chen, PhD; and Juan C. Celedón, MD, DrPH

CHEST 2020; 158(5):1841-1856

Online supplements are not copyedited prior to posting and the author(s) take full responsibility for the accuracy of all data.

© 2020 AMERICAN COLLEGE OF CHEST PHYSICIANS. Reproduction of this article is prohibited without written permission from the American College of Chest Physicians. See online for more details. **DOI:** 10.1016/j.chest.2020.05.601

CHEST™ Online Supplement

e-Appendix 1.

Supplementary Methods

Study population

Participants with and without asthma were recruited from households in San Juan (PR) from February 2014 through May 2017, using multistage probability sampling; 638 households had ≥1 eligible subject, and 543 (85.1%) subjects (one per household) agreed to participate. There were no significant differences in age or sex between eligible children who did and did not participate. The study was approved by the institutional review boards of the University of Puerto Rico (San Juan, PR- Protocol #0160713) and the University of Pittsburgh (Pittsburgh, PA- Protocol #PRO14010428). Written parental consent and assent were obtained from participants <18 years old, and consent was obtained from participants ≥18 years old.

The study protocol included questionnaires on respiratory health, measurement of serum allergen-specific IgEs, and collection of nasal epithelial samples for DNA and RNA extraction. Atopy was defined as ≥1 positive IgE (≥0.35 IU/mL) to five common allergens in Puerto Rico: house dust mite (Der p 1), cockroach (Bla g 2), cat dander (Fel d 1), dog dander (Can f 1), and mouse urinary protein (mus m 1). Asthma was defined as a physician's diagnosis plus at least one episode of wheeze in the previous year. Control subjects had neither physician-diagnosed asthma nor wheeze in the previous year.

Normality of methylation and RNA-seq values

The distribution of M values of methylation is closer to a normal distribution than that of beta values. Although M values do not strictly follow a normal distribution and TPM values can be skewed[6], we used multivariable linear regression in our eQTM analysis because of a relatively large sample size (n=455) and given ability to control for possible confounders.

Epigenome-wide association study (EWAS) of atopic asthma

The EWAS conducted by Forno et al[2] included 273 Puerto Rican subjects in EVA-PR (169 with atopic asthma and 104 control subjects without atopy or asthma). After quality controls, 227,836 methylation probes were evaluated in a multivariable logistic regression model, as follows: $logit(p) = \beta_0 + \beta_1 M + \sum \alpha_j Z_j$, where p is the probability of having atopic asthma, M is a methylation value at a probe, Z_j is an adjusted covariate, and $\beta_0, \beta_1,$ and α_j are regression coefficients. Other covariates included in the model were the first five PCs derived from genotypic data, age, gender, methylation batches, and latent factors of methylation -estimated from R package sva[6]. FDR-P values were then calculated based on testing 227,836 methylation probes. Significance was defined as FDR-P <0.01.

Online supplements are not copyedited prior to posting and the author(s) take full responsibility for the accuracy of all data.

Transcriptome-wide association studies (TWAS) of atopic asthma

A TWAS of atopic asthma was recently conducted by Forno et al[8] in 258 Puerto Rican subjects in EVA-PR (157 with atopic asthma and 101 non-atopic non-asthmatic control subjects). In that study, differential gene expression was analyzed based on the raw count table of RNA sequencing data used using the R package DESeq2. Multivariable models of atopic asthma were adjusted for age, gender, RNA batches, RNA cell sorting, and the first five PCs derived from genotypic data. FDR-P were calculated for the 18,311 genes tested. Significance was defined as an FDR-P <0.01.

Analysis stratified by atopic asthma

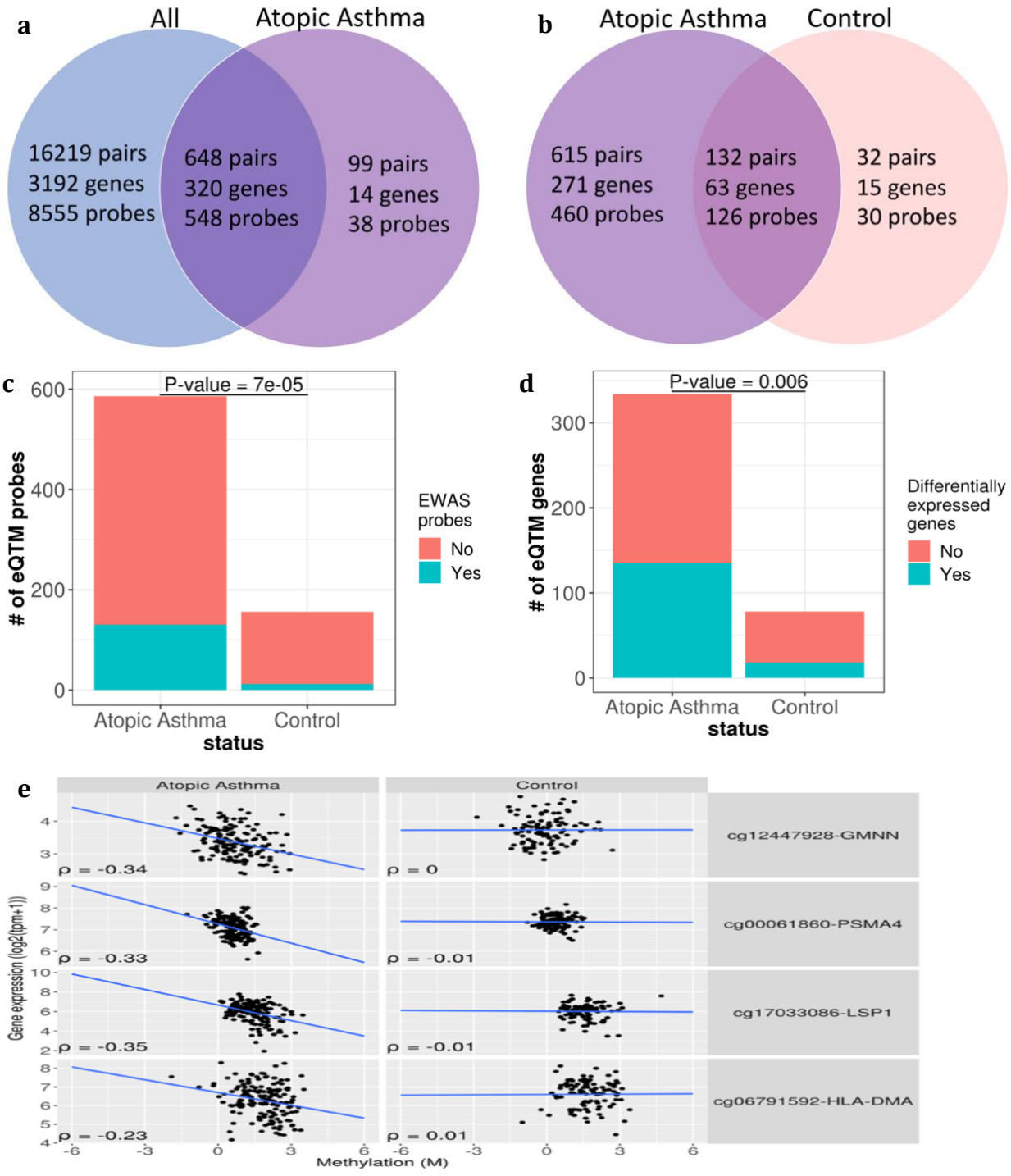
We conducted eQTM analyses separately in subjects with atopic asthma (n=158) and in non-atopic non-asthmatic controls (n=100). Using data from all subjects (n=455), we identified nearly 23 times as many significant associations as in subjects with atopic asthma, most likely due to a larger sample size (e-Figure 1**a**); 86.7% of the associations identified in cases were also identified in all subjects. Whereas only 17.7% of the associations identified in cases were also identified in control subjects, 80.5% of the associations identified in controls were also identified in cases (e-Figure 1**b)**. Such differences can be explained by diverging sample size, as well as by disease status.

Methylation probes identified in the eQTM analysis in subjects with atopic asthma were more likely to be associated with atopic asthma those identified in the eQTM analysis in non-atopic controls (e-Figure 1**c**). Likewise, genes identified in the eQTM analysis in subjects with atopic asthma were more likely to be DEGs in atopic asthma than those identified in the analysis of non-atopic controls (e-Figure 1**d**). Examples of eQTM methylation-gene expression pairs that are significant in the analysis of subjects with atopic asthma but not in that of control subjects are shown in e-Figure 1**e**. The expression of *GMNN* is regulated by a miRNA in a murine model of acute and chronic asthma[1]. *PSMA4* methylation and expression have been associated with SNPs identified in GWAS of COPD[2] [3]. Leukocyte-specific protein 1 (*LSP1*) regulates neutrophil recruitment in acute lung inflammation[4], and HLA-genes have been associated with asthma[5].

Distance of associations between atopic asthma subjects and health control subjects

We also checked the distance between the eQTM CpG-gene expression pairs, separately in subjects with atopic asthma and in non-atopic controls (Table S1). eQTM methylation-gene pairs in atopic-asthma are significantly more likely to be distant than those in control subjects (chi-square p-value: 2.36×10^{-8}). This suggests that methylation regulates gene expression more remotely in cases than in control subjects.

Online supplements are not copyedited prior to posting and the author(s) take full responsibility for the accuracy of all data.



e-Figure 1. eQTM analyses in subjects with atopic asthma only (n=158) and in non-atopic controls (n=100). a Venn diagram of significant eQTMs (FDR < 0.01) in all subjects (n=455) and in subjects with atopic asthma (n=158). Pairs refer to methylation-gene pairs. Probes refer to methylation probes. **b** Venn diagram of significant eQTMs in subjects with atopic asthma and in non-atopic

controls. **c** # of eQTM methylation probes identified in subjects with atopic asthma and in non-atopic controls. The blue bars are the number of eQTM probes that are also EWAS probes. EWAS probes are significantly enriched in eQTM probes in the analysis of subjects with atopic asthma vs. that in control subjects only (P-value = 7x10-5). The Chi-square test was used for the enrichment test. **d** # of eQTM genes identified in subjects with atopic asthma vs. non-atopic control subjects. The blue bars indicate the number of eQTM genes that are also differentially expressed (DE) in atopic asthma. DE genes are significantly enriched in eQTM genes in subjects with atopic asthma compared to control subjects (P-value= 0.006). **e** Examples of the associations between methylation probes and gene expressions that are significant in subjects with atopic asthma but not in non-atopic controls. ρ is Pearson's correlation coefficient. The blue line is the regression line.

References

1. Garbacki N, Di Valentin E, Huynh-Thu VA, et al. MicroRNAs profiling in murine models of acute and chronic asthma: a relationship with mRNAs targets. *PloS one.* 2011;6(1):e16509-e16509.
2. Castaldi PJ, Cho MH, Zhou X, et al. Genetic control of gene expression at novel and established chronic obstructive pulmonary disease loci. *Human molecular genetics.* 2015;24(4):1200-1210.
3. Nedeljkovic I, Carnero-Montoro E, Lahousse L, et al. Understanding the role of the chromosome 15q25.1 in COPD through epigenetics and transcriptomics. *European Journal of Human Genetics.* 2018;26(5):709-722.
4. Le NPK, Channabasappa S, Hossain M, Liu L, Singh B. Leukocyte-specific protein 1 regulates neutrophil recruitment in acute lung inflammation. *American journal of physiology. Lung cellular and molecular physiology.* 2015;309(9):L995-L1008.
5. Kontakioti E, Domvri K, Papakosta D, Daniilidis M. HLA and asthma phenotypes/endotypes: A review. *Human Immunology.* 2014;75(8):930-939.
6. Church BV, Williams HT, Mar JC. Investigating skewness to understand gene expression heterogeneity in large patient cohorts. *BMC Bioinformatics.* 2019;20(24):668.

Online supplements are not copyedited prior to posting and the author(s) take full responsibility for the accuracy of all data.