A genome-wide association study of severe asthma exacerbations in Latino children and adolescents

Qi Yan1,16, Erick Forno1,16, Esther Herrera-Luis2,16, Maria Pino-Yanes2,3, Cancan Qi4,5, Raimon Rios6, Yueh-Ying Han6, Soyeon Kim1, Sam Oh7, Edna Acosta-Pérez8, Rong Zhang1, Donglei Hu7, Celeste Eng7, Scott Huntsman7, Lydiana Avila9, Nadia Boutaoui1, Michelle M. Cloutier10, Manuel E. Soto-Quiros9, Cheng-Jian Xu11,12, Scott T. Weiss13, Jessica Lasky-Su13, Megan R. Kiedrowski14, Camila Figueiredo6, Jennifer Bomberger14, Mauricio L. Barreto15, Glorisa Canino8, Wei Chen1, Gerard H. Koppelman4,5, Esteban G. Burchard7,17 and Juan C. Celedón1,17

@ERSpublications

A novel SNP in FLJ22447 is significantly associated with both severe asthma exacerbations in Latino youth and DNA methylation of a cis-CpG in nasal epithelium. This CpG is linked to nasal epithelial expression of a gene implicated in atopic asthma. https://bit.ly/33D9Joc


ABSTRACT Severe asthma exacerbations are a major cause of school absences and healthcare costs in children, particularly those in high-risk racial/ethnic groups.

To identify susceptibility genes for severe asthma exacerbations in Latino children and adolescents, we conducted a meta-analysis of genome-wide association studies (GWAS) in 4010 Latino youth with asthma in four independent cohorts, including 1693 Puerto Ricans, 1019 Costa Ricans, 640 Mexicans, 256 Brazilians and 402 members of other Latino subgroups. We then conducted methylation quantitative trait locus, expression quantitative trait locus and expression quantitative trait methylation analyses to assess whether the top single nucleotide polymorphism (SNP) in the meta-analysis is linked to DNA methylation and gene expression in nasal (airway) epithelium in separate cohorts of Puerto Rican and Dutch children and adolescents.

In the meta-analysis of GWAS, an SNP in FLJ22447 (rs2253681) was significantly associated with 1.55 increased odds of severe asthma exacerbation (95% CI 1.34–1.79, p=6.3×10−9). This SNP was significantly associated with DNA methylation of a CpG site (cg25024579) at the FLJ22447 locus, which was in turn associated with increased expression of KCNJ2-AS1 in nasal airway epithelium from Puerto Rican children and adolescents (β=0.10, p=2.18×10−7).

SNP rs2253681 was significantly associated with both DNA methylation of a cis-CpG in FLJ22447 and severe asthma exacerbations in Latino youth. This may be partly explained by changes in airway epithelial expression of a gene recently implicated in atopic asthma in Puerto Rican children and adolescents (KCNJ2-AS1).
Introduction
Asthma is the most common chronic respiratory disease among children [1]. In the USA, total costs related to asthma exceed $81 billion per year [2]. Severe asthma exacerbations (SAEs), defined as episodes of disease worsening that require a change in treatment to prevent a serious outcome [3], are a major cause of school or work absences and healthcare costs. Of the ~6.1 million children with asthma in the USA, 2.1% report ≥1 asthma-related hospitalisation and 10.7% report ≥1 asthma-related visit to the emergency department (ED) in the previous year [4]. Despite recent advances, the best predictor of SAEs is having had an SAE in the previous year [5].

Although genome-wide association studies (GWASs) have identified susceptibility loci for asthma [6–8], little is known about genetic determinants of asthma exacerbations, which may be distinct from those for asthma per se. A GWAS in Danish children with asthma (ages 2–6 years) identified cadherin-related family member 3 (CDHHR3), a gene not previously associated with asthma, as a susceptibility locus for recurrent SAEs [9]. In a combined GWAS of two cohorts of non-Hispanic white children with asthma, four intronic single nucleotide polymorphisms (SNPs) in cadherin-associated protein alpha 3 (CTNNA3) were significantly associated with SAEs [10], but the association was not replicated in an independent cohort including 786 children and adults with asthma. Moreover, a GWAS of SAEs among 806 non-Hispanic white children and adults with asthma who were being treated with inhaled corticosteroids found no genome-wide significant results [11].

The burden of asthma varies across racial and ethnic groups in the USA and Latin America. For example, Puerto Rican children have a greater prevalence of and morbidity and mortality from asthma than non-Hispanic white children in the USA [12], and Costa Rican adolescents have a greater burden of asthma than those in other Latin American countries [13]. Moreover, recent evidence suggests that some susceptibility variants for asthma-related outcomes are ethnicity-specific [14, 15]. We hypothesised that there would be susceptibility variants for SAEs that would be more common or exert a greater effect in Latino subgroups at risk for morbidity from asthma. To test this hypothesis, we conducted a meta-analysis of GWAS of SAEs among Latino youth with asthma in four independent studies.

Methods
Please refer to the supplementary material for more details.

Study populations included in the meta-analysis of GWAS of severe exacerbations

Hartford-Puerto Rico study
The Hartford-Puerto Rico (HPR) study is a case–control study of childhood asthma in Puerto Ricans [6]. SAEs were defined as ≥1 hospitalisation for asthma or ≥1 ED/urgent care visit for asthma requiring treatment with systemic corticosteroids in the previous year, or ≥1 course of systemic corticosteroids for asthma in the previous year. After quality control (QC) measures, 554 independent children with asthma (236 of whom had ≥1 SAE in the previous year) were included in the genome-wide association analysis, which was conducted using logistic regression under an additive genetic model, adjusting for age, sex, inhaled steroid use and the first two principal components (PCs), calculated using smartPCA [16].

The Genetics of Asthma in Latino Americans study
The Genetics of Asthma in Latino Americans (GALA II) study is a case–control study of asthma in Latino children and youth [17]. An SAE was defined as having had ≥1 of the following events in the previous

Affiliations: 1Division of Pediatric Pulmonary Medicine, University of Pittsburgh Medical Centre, Children’s Hospital of Pittsburgh, University of Pittsburgh, Pittsburgh, PA, USA. 2Genomics and Health Group, Dept of Biochemistry, Microbiology, Cell Biology and Genetics, Universidad de La Laguna, La Laguna, Spain. 3CIBER de Enfermedades Respiratorias, Instituto de Salud Carlos III, Madrid, Spain. 4Dept of Pediatric Pulmonology and Pediatric Allergy, Beatrix Children’s Hospital, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands. 5GRIAC Research Institute, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands. 6Instituto de Ciências da Saúde, Universidade Federal da Bahia, Salvador, Brazil. 7Dept of Medicine, University of California San Francisco, San Francisco, CA, USA. 8Behavioral Sciences Research Institute, University of Puerto Rico, San Juan, Puerto Rico. 9Dept of Pediatrics, Hospital Nacional de Niños, San José, Costa Rica. 10Dept of Pediatrics, University of Connecticut, Farmington, CT, USA. 11CiIM and TWINCORE, joint ventures between the Hannover Medical School and the Helmholtz Centre for Infection Research, Hannover, Germany. 12Dept of Internal Medicine and Radiobad Center for Infectious Diseases, Radboud University Medical Center, Nijmegen, The Netherlands. 13Channing Division of Network Medicine, Brigham and Women’s Hospital, Harvard Medical School, Boston, MA, USA. 14Dept of Microbiology and Molecular Genetics, University of Pittsburgh, Pittsburgh, PA, USA. 15Instituto de Saúde Coletiva, Federal University of Bahia, Salvador, Brazil. 16Shared first authors. 17Shared senior authors.

Correspondence: Juan C. Celedón, Division of Pulmonary Medicine, UPMC Children’s Hospital of Pittsburgh, 4401 Penn Avenue, Pittsburgh, PA 15224, USA. E-mail: juan.celedon@chp.edu
year: hospitalisations for asthma, ED visits or unscheduled and urgent doctor’s visits because of asthma, or treatment with systemic corticosteroids for asthma. The current analysis focused on 2181 children with asthma (1283 of whom had ≥1 SAE) and self-reported Latino ethnicity (1139 Puerto Rican, 640 Mexican and 402 from other groups). Association testing was conducted using logistic regression under an additive genetic model, adjusting for age, sex, inhaled steroid use, ethnicity and the first two PCs.

The Genetics of Asthma in Costa Rica Study
The Genetics of Asthma in Costa Rica Study (GACRS) is a genetic study of nuclear families of children with asthma in Costa Rica [18, 19]. An SAE was defined as ≥1 ED visit for asthma or ≥1 hospitalisation for asthma in the previous year. After QC, 1019 independent children with asthma (851 with ≥1 SAE) were included in the analysis, which was also conducted using logistic regression under an additive genetic model, adjusting for age, sex, inhaled steroid use and the first two PCs.

The Social Changes, Asthma and Allergy in Latin America study, Bahia, Brazil
The Social Changes, Asthma and Allergy in Latin America (SCAALA) study is a longitudinal study of asthma and allergic diseases of Brazilian children [20]. Children who reported asthma symptoms and had data on SAEs (n=256) were included in this analysis. An SAE was defined as ≥1 ED visit or ≥1 hospitalisation due to wheeze in the previous year. The analysis was conducted using logistic regression under an additive genetic model, with adjustment for age, sex and the first two PCs from genotypic data.

Study populations included in molecular quantitative trait analyses in nasal epithelium
The Epigenetic Variation and Childhood Asthma in Puerto Rico study
Nasal (airway) epithelium can serve as a surrogate marker for DNA methylation and gene expression in bronchial (airway) epithelium [21]. In the Epigenetic Variation and Childhood Asthma in Puerto Rico (EVA-PR) study, whole-genome methylation assays were performed using HumanMethylation450 BeadChips (Illumina, San Diego, CA, USA). After QC, 227 901 CpG probes remained for the analysis of nasal epithelium, and M-values were used in all downstream analyses. RNA sequencing was performed with the Illumina NextSeq 500 platform, with paired-end reads at 75 cycles and 80 million reads per sample. After QC, 16 737 genes were retained for the analysis. The R function sva was used to estimate latent factors that capture unknown data heterogeneity [22]. Of the 543 study participants, 457 had complete genome-wide data for genotypes, methylation and transcriptomics in nasal epithelium (supplementary figure S1).

Prevention and Incidence of Asthma and Mite Allergy study
The Prevention and Incidence of Asthma and Mite Allergy (PIAMA) study is a birth cohort study of children born in the Netherlands in 1996 and 1997 [23, 24]. A total of 479 nasal epithelial samples were hybridised to the Infinium HumanMethylation450 BeadChip arrays. After QC, 455 samples and 436 824 probes remained, and 432 samples had matched genotype data. RNA sequencing was performed with the Illumina HiSeq 2500 platform with paired-end sequencing. After stringent QC, 17 156 genes and 326 samples remained, with 233 samples having matched genotype data [25].

Meta-analysis of GWAS of SAEs
METAL [26] software was used to perform the meta-analysis of GWAS of SAEs, using data from HPR, GALA II, GACRS and SCAALA.

Molecular quantitative trait locus analyses
To estimate the effects of our top SNP on DNA methylation and gene expression, we conducted a methylation quantitative trait locus (mQTL) analysis to test for an association between our top SNP and genome-wide DNA methylation in nasal epithelium, and an expression quantitative trait locus (eQTL) analysis to test for an association between our top SNP and genome-wide gene expression in nasal epithelium. We then conducted an expression quantitative trait methylation (eQTM) analysis to test whether the top CpG site identified in the mQTL analysis was associated with genome-wide gene expression in nasal epithelium. All analyses were adjusted for age, sex, asthma status, atopy status, the top five PCs from genotypic data and latent factors estimated from sva [22]. In addition, RNA sequencing batch and RNA sample sorting protocol (i.e. whole cells or CD326-positive nasal epithelial cells) [21] were adjusted for in the eQTM analyses, and methylation batch was adjusted for in the mQTL and eQTM analyses; both cis and trans effects were considered.

Pathway analysis
A pathway analysis was performed with MAGMA [27], which conducts SNP-wise gene analysis of summary statistics with correction for linkage disequilibrium between variants and genes, to test whether

https://doi.org/10.1183/13993003.02693-2020
sets of genes are jointly associated with a phenotype (i.e. asthma exacerbation) compared to other genes across the genome. Adaptive permutation was used to produce an empirical p-value and false discovery rate (FDR). Gene sets used in the analyses were from Gene Ontology (GO) [28, 29], Kyoto Encyclopedia of Genes and Genomes (KEGG) [30, 31], Reactome [32, 33] and BioCarta pathways.

Results
The characteristics of the 4010 participants in the four studies included in the meta-analysis of GWAS of SAEs are shown in table 1. In total, there were 2509 children with asthma and ≥1 SAE (cases) and 1501 children with asthma but no SAEs (controls). Compared to subjects who participated in the HPR or SCAALA studies, those in GALA II were older and those in GACRS were younger. Compared to subjects in the other studies, those in GACRS were more likely to be male and to have had ≥1 severe asthma exacerbation in the previous year.

Approximately six million genotyped and imputed SNPs with a minor allele frequency (MAF) ≥0.05 were included in the meta-analysis of GWAS of SAEs. In this meta-analysis, one SNP (rs2253681, in FLJ22447 on chromosome 14q23.2) was significantly associated with SAEs (p<5×10^{-8}) (figure 1a). This SNP was genotyped in HPR, GALA II and SCAALA, and imputed in GACRS with imputation quality r^2=0.96. Each copy of the minor allele (A) of SNP rs2253681 was associated with 1.55 times increased odds of SAEs (95% CI 1.34–1.79, p=6.3×10^{-9}, figure 1b). The Q-Q plot (figure 1c) showed no inflation for the results for each of the four individual cohorts or for the pooled analysis (figure 2). There was no significant interaction between SNP rs2253681 and inhaled steroid use on SAEs in the GWAS conducted in the HPR, GALA II or GACRS cohorts (p for interaction≥0.15 in all instances).

We then examined whether SNPs previously associated with asthma in either a multi-ancestry meta-analysis [7] or a meta-analysis from the UK Biobank [34] were associated with SAEs in our analysis (supplementary table S1). None of the previously reported asthma-susceptibility SNPs was significantly associated with SAEs in our meta-analysis of Latino youth. Moreover, no SNPs associated with SAEs or asthma hospitalisations in previous candidate-gene studies (e.g. IL13, IL4RA) or GWAS (e.g. CDHR3, CTNNA3) [9, 10, 18, 35–38] were significantly associated with SAEs in our meta-analysis (meta-p≥0.05 in all instances, supplementary table S2). Although the previously reported SNP rs7216389 in ORMDL3 on chromosome 17q21 was nominally associated with SAEs in HPR (p=1.6×10^{-3}), it did not reach statistical significance in our meta-analysis (meta-p=0.06, supplementary table S2).

To examine whether SNP rs2253681 affects methylation of the FLJ22447 locus in nasal epithelium, we first conducted an mQTL analysis. In this analysis, rs2253681 was a significant cis-acting mQTL in nasal epithelium for cg25024579 in FLJ22447 (β=0.55, p=3.6×10^{-16} and FDR-p=8.1×10^{-11} in EVA-PR; β=0.34, Table 1 Summary of main characteristics of study participants

<table>
<thead>
<tr>
<th>Study site</th>
<th>HPR</th>
<th>GALA II</th>
<th>GACRS</th>
<th>SCAALA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subjects n</td>
<td>554</td>
<td>2181</td>
<td>1019</td>
<td>256</td>
</tr>
<tr>
<td>Age years</td>
<td>10.0±2.7</td>
<td>12.7±3.3</td>
<td>9.2±1.9</td>
<td>7.2±1.9</td>
</tr>
<tr>
<td>Male sex</td>
<td>300 (54.2)</td>
<td>1196 (54.8)</td>
<td>598 (58.7)</td>
<td>139 (54.3)</td>
</tr>
<tr>
<td>Asthma exacerbation</td>
<td>236 (42.6)</td>
<td>1283 (58.6)</td>
<td>851 (83.5)</td>
<td>139 (54.3)</td>
</tr>
<tr>
<td>Inhaled steroid use</td>
<td>187 (33.8)</td>
<td>996 (45.7)</td>
<td>521 (51.1)</td>
<td>Not available</td>
</tr>
<tr>
<td>FEV1 % pred#</td>
<td>86.8±16.0</td>
<td>90.6±16.3</td>
<td>99.1±17.3</td>
<td>Not available</td>
</tr>
<tr>
<td>FEV1/FVC % pred#</td>
<td>91.8±9.8</td>
<td>96.3±8.8</td>
<td>94.5±8.7</td>
<td>Not available</td>
</tr>
<tr>
<td>Study sites</td>
<td>Hartford (CT, USA) and San Juan (Puerto Rico)</td>
<td>Chicago (IL, USA), Bronx (NY, USA), Houston (TX, USA), San Francisco (CA, USA) and Puerto Rico</td>
<td>Costa Rica</td>
<td>Salvador (Bahia), Brazil</td>
</tr>
<tr>
<td>Genotyping platform</td>
<td>Illumina 2.5M</td>
<td>Affymetrix Axiom® LAT1</td>
<td>Illumina Human Omni 2.5v1</td>
<td>Illumina Human Omni Express-12v1_A</td>
</tr>
</tbody>
</table>

Data are presented as n (%) or mean±SD, unless otherwise indicated. HPR: Hartford–Puerto Rico study; GALA II: Genetics of Asthma in Latino America II; GACRS: Genetics of Asthma in Costa Rica Study; SCAALA: Social Changes, Asthma and Allergy in Latin America; FEV1: forced expiratory volume in 1 s; FVC: forced vital capacity. # for comparability, all % predicted values across the three studies were calculated using reference values for Mexican American youth [45].
In addition, among the top 20 mQTL, cg05223396 in SH3PXD2B, cg13127574 in ADSSL1 and cg21115391 in CCDC67 were associated with rs2253681 at \( p < 0.05 \) in PIAMA, in the same direction as in EVA-PR (table 2). Next, we conducted an eQTL analysis of SNP rs2253681 in the FLJ22447 locus and gene expression in nasal epithelium, and found no cis or trans genome-wide significant results (supplementary table S3). Among the top 20 eQTL, SRF was associated with rs2253681 at \( p < 0.05 \) in PIAMA, in the same direction as in EVA-PR (supplementary table S3).

Given the association between SNP rs2253681 and methylation of cg25024579, we then examined the effect of that CpG on gene expression in nasal epithelium by conducting an eQTM analysis (table 3). In this analysis, methylation of cg25024579 was significantly associated in trans with expression of KCNJ2 antisense RNA 1 (KCNJ2-AS1) on chromosome 17q24.3 (\( \beta = 0.10, p = 2.2 \times 10^{-7} \) and FDR\( \text{p} = 3.2 \times 10^{-6} \)) in EVA-PR. Although this was not replicated in PIAMA, the observed association remained significant in the combined analysis of the two cohorts (\( \beta = 0.09, p = 1.4 \times 10^{-6} \)). The top cis-eQTM gene in the current analysis in EVA-PR was protein kinase C-\( \eta \) (PKC\( \eta \)), encoded by PRKCH (\( \beta = 0.07, p = 4.1 \times 10^{-5} \) and
rs2253681

<table>
<thead>
<tr>
<th>Study</th>
<th>Minor allele</th>
<th>MAF</th>
<th>p-value</th>
<th>OR</th>
</tr>
</thead>
<tbody>
<tr>
<td>HPR</td>
<td>A</td>
<td>0.20</td>
<td>1.3×10^{-2}</td>
<td>1.49</td>
</tr>
<tr>
<td>GACRS</td>
<td>A</td>
<td>0.11</td>
<td>0.12</td>
<td>1.38</td>
</tr>
<tr>
<td>GALA II</td>
<td>A</td>
<td>0.16</td>
<td>2.3×10^{-5}</td>
<td>1.54</td>
</tr>
<tr>
<td>SCAALA</td>
<td>A</td>
<td>0.24</td>
<td>3.2×10^{-3}</td>
<td>1.92</td>
</tr>
</tbody>
</table>

Meta 6.3×10^{-9} 1.55

**FIGURE 2** Forest plots of odds ratio and 95% CI for the association with asthma for rs2253681, the most significant single nucleotide polymorphism in the meta-analysis. The heterogeneity measure I²=0 implies no heterogeneity among the odds ratios from the four studies [see supplementary material for details]. HPR: Hartford-Puerto Rico cohort; GACRS: Genetics of Asthma in Costa Rica Study; GALA II: Genetics of Asthma in Latino Americans II; SCAALA: Social Changes, Asthma and Allergy in Latin America; MAF: minor allele frequency.

FDR-adj=1.5×10^{-1}). 

**TABLE 2** Top 20 methylation quantitative trait locus for rs2253681 in EVA-PR nasal epithelial cells and replication results from PIAMA.

<table>
<thead>
<tr>
<th>CpG</th>
<th>Chr</th>
<th>Position</th>
<th>Nearest gene</th>
<th>Distance to gene</th>
<th>Effect</th>
<th>EVA-PR p-value</th>
<th>FDR</th>
<th>PIAMA p-value</th>
<th>Meta p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>cg25024579</td>
<td>14</td>
<td>62 076 297</td>
<td>FLJ22447</td>
<td>0</td>
<td>0.5520</td>
<td>3.55×10^{-16}</td>
<td>8.08×10^{-11}</td>
<td>0.3365</td>
<td>1.16×10^{-11}</td>
</tr>
<tr>
<td>cg02787615</td>
<td>2</td>
<td>99 434 688</td>
<td>KIAA1211L</td>
<td>0</td>
<td>-0.1080</td>
<td>9.68×10^{-07}</td>
<td>1.10×10^{-01}</td>
<td>0.0415</td>
<td>1.73×10^{-02}</td>
</tr>
<tr>
<td>cg05223396</td>
<td>5</td>
<td>171 830 155</td>
<td>SH3PXD2B</td>
<td>0</td>
<td>-0.2175</td>
<td>1.89×10^{-06}</td>
<td>1.44×10^{-01}</td>
<td>-0.0641</td>
<td>2.52×10^{-02}</td>
</tr>
<tr>
<td>cg24718015</td>
<td>17</td>
<td>40 489 721</td>
<td>STAT3</td>
<td>0</td>
<td>-0.1581</td>
<td>5.01×10^{-06}</td>
<td>2.80×10^{-01}</td>
<td>0.0215</td>
<td>5.00×10^{-01}</td>
</tr>
<tr>
<td>cg21785067</td>
<td>1</td>
<td>54 587 182</td>
<td>TCEANC2</td>
<td>8989</td>
<td>0.2305</td>
<td>7.39×10^{-04}</td>
<td>2.80×10^{-01}</td>
<td>0.0826</td>
<td>6.42×10^{-02}</td>
</tr>
<tr>
<td>cg21917349</td>
<td>15</td>
<td>29 213 860</td>
<td>APBA2</td>
<td>0</td>
<td>-0.3427</td>
<td>7.89×10^{-06}</td>
<td>2.80×10^{-01}</td>
<td>-0.0441</td>
<td>2.22×10^{-01}</td>
</tr>
<tr>
<td>cg15904664</td>
<td>14</td>
<td>104 569 653</td>
<td>ASPG</td>
<td>0</td>
<td>-0.2093</td>
<td>9.96×10^{-06}</td>
<td>2.80×10^{-01}</td>
<td>-0.0574</td>
<td>1.24×10^{-01}</td>
</tr>
<tr>
<td>cg11374933</td>
<td>8</td>
<td>9 756 172</td>
<td>MIR124-1</td>
<td>5189</td>
<td>-0.1660</td>
<td>1.18×10^{-05}</td>
<td>2.80×10^{-01}</td>
<td>0.0357</td>
<td>2.55×10^{-01}</td>
</tr>
<tr>
<td>cg15375424</td>
<td>5</td>
<td>131 823 451</td>
<td>IRF1</td>
<td>0</td>
<td>-0.2272</td>
<td>1.37×10^{-05}</td>
<td>2.80×10^{-01}</td>
<td>-0.0623</td>
<td>9.38×10^{-02}</td>
</tr>
<tr>
<td>cg07975634</td>
<td>1</td>
<td>147 232 608</td>
<td>GJA5</td>
<td>0</td>
<td>-0.2236</td>
<td>1.49×10^{-05}</td>
<td>2.80×10^{-01}</td>
<td>0.0396</td>
<td>6.77×10^{-02}</td>
</tr>
<tr>
<td>cg08487909</td>
<td>6</td>
<td>30 904 981</td>
<td>DPCR1</td>
<td>3794</td>
<td>-0.1356</td>
<td>1.59×10^{-05}</td>
<td>2.80×10^{-01}</td>
<td>-0.0364</td>
<td>8.15×10^{-02}</td>
</tr>
<tr>
<td>cg0203736</td>
<td>7</td>
<td>55 022 567</td>
<td>EGRF</td>
<td>64156</td>
<td>-0.2440</td>
<td>1.86×10^{-05}</td>
<td>2.80×10^{-01}</td>
<td>0.0034</td>
<td>9.20×10^{-01}</td>
</tr>
<tr>
<td>cg23696248</td>
<td>19</td>
<td>45 260 501</td>
<td>BCL3</td>
<td>0</td>
<td>-0.3099</td>
<td>2.01×10^{-05}</td>
<td>2.80×10^{-01}</td>
<td>0.0230</td>
<td>7.09×10^{-01}</td>
</tr>
<tr>
<td>cg13127574</td>
<td>14</td>
<td>105 196 523</td>
<td>ADSS1</td>
<td>0</td>
<td>-0.2116</td>
<td>2.11×10^{-05}</td>
<td>2.80×10^{-01}</td>
<td>-0.0445</td>
<td>9.01×10^{-03}</td>
</tr>
<tr>
<td>cg08403064</td>
<td>11</td>
<td>44 068 709</td>
<td>ACCSL</td>
<td>820</td>
<td>-0.2011</td>
<td>2.23×10^{-05}</td>
<td>2.80×10^{-01}</td>
<td>0.0472</td>
<td>1.78×10^{-01}</td>
</tr>
<tr>
<td>cg21115391</td>
<td>11</td>
<td>93 143 810</td>
<td>CDC67</td>
<td>0</td>
<td>-0.2452</td>
<td>2.23×10^{-05}</td>
<td>2.80×10^{-01}</td>
<td>-0.0945</td>
<td>9.51×10^{-03}</td>
</tr>
<tr>
<td>cg01220257</td>
<td>19</td>
<td>45 260 935</td>
<td>BCL3</td>
<td>0</td>
<td>-0.1474</td>
<td>2.30×10^{-05}</td>
<td>2.80×10^{-01}</td>
<td>0.0269</td>
<td>5.78×10^{-01}</td>
</tr>
<tr>
<td>cg12765028</td>
<td>4</td>
<td>13 526 659</td>
<td>LINC01097</td>
<td>1282</td>
<td>0.1566</td>
<td>2.34×10^{-05}</td>
<td>2.80×10^{-01}</td>
<td>-0.0063</td>
<td>7.70×10^{-01}</td>
</tr>
<tr>
<td>cg11136041</td>
<td>1</td>
<td>158 111 350</td>
<td>LOC64628</td>
<td>919</td>
<td>0.2204</td>
<td>2.39×10^{-05}</td>
<td>2.80×10^{-01}</td>
<td>0.0249</td>
<td>4.28×10^{-01}</td>
</tr>
<tr>
<td>cg01753277</td>
<td>15</td>
<td>75 077 691</td>
<td>CSK</td>
<td>0</td>
<td>-0.2291</td>
<td>2.64×10^{-05}</td>
<td>2.80×10^{-01}</td>
<td>-0.0460</td>
<td>1.30×10^{-01}</td>
</tr>
</tbody>
</table>

Chr: chromosome; EVA-PR: Epigenetic Variation and Childhood Asthma in Puerto Rico; PIAMA: Prevention and Incidence of Asthma and Mite Allergy; FDR: false discovery rate. *: FDR is adjusted for the whole genome in EVA-PR.
To further assess the biological relevance of the FLJ22447 locus, we conducted an SNP-wise pathway analysis of SAEs using the meta-analysis results and evaluated public repositories and databases. Although there were no pathways associated with SAEs after adjusting for multiple testing, there were a number of nominally significant pathways that included genes on chromosome 14q23.2, near FLJ22447 (supplementary table S4).

In a sensitivity analysis, we repeated the GWAS of SAEs after imputing genotypes using the 1000 Genomes Ad Mixed American (AMR) reference panel instead of the HaploType Reference Consortium (HRC) r1.1.2016 reference panel (supplementary material), obtaining very similar results (supplementary figure S2 and supplementary table S5).

### Discussion

Our combined analysis including 4010 youths in four study cohorts showed that a novel SNP in FLJ22447 (rs2253681) is significantly associated with SAEs (p=6.3×10^-6) among children and adolescents in Latino subgroups affected with asthma (Puerto Ricans, Costa Ricans, Mexicans and Brazilians) [12]. In an mQTL analysis in nasal epithelium, this SNP was significantly associated with DNA methylation of a CpG site at the FLJ22447 locus (cg25024579), which was in turn significantly associated with the expression of KCNJ2-AS1 in nasal epithelium.

**FLJ22447** codes for a long non-coding RNA, a type of RNA that does not get translated to a protein but may serve pre- and post-transcriptional regulatory functions. **FLJ22447** partially overlaps with **AL355916.3**, which encodes a protein kinase C paralog expressed in epithelial tissues. Of note, an SNP in **AL355916.3** has been associated with forced expiratory volume in 1 s/forced vital capacity and bronchodilator response in adults [34]. **FLJ22447** mediates interleukin (IL)-33 upregulation in activated fibroblasts; together, **FLJ22447** and IL-33 can increase fibroblast expression of α-smooth muscle actin (α-SMA), vimentin and N-cadherin [35], which are associated with myofibroblast differentiation and airway remodelling [36, 37].

**KCNJ2-AS1** also encodes a long non-coding RNA. Expression of **KCNJ2-AS1** (log2(fold-change)=0.34, p=3.9×10^-6, FDR-p=3.1×10^-5) was significantly associated with atopic asthma in our recent transcriptome-wide association study of atopic asthma in Puerto Rican children and adolescents [38]. SNPs rs8066985 and rs312750 near **KCNJ2-AS1** are associated with body mass index and waist-to-hip ratio [39, 40], which are associated with worse asthma outcomes in children and adults. To our knowledge, this is the first study linking **FLJ22447** or **KCNJ2-AS1** to SAEs in children and adolescents.
While not significant genome wide, the top cis-eQTM gene in EVA-PR was PRKCH, which is adjacent to FLJ22447 and most highly expressed in the lungs [41]. Both methylation and transcription of this gene in nasal (airway) epithelium were recently associated with atopic asthma in Puerto Rican children and adolescents [21]. PKCη, encoded by PRKCH, plays a key role in the assembly and maintenance of epithelial tight junctions. PKCη phosphorylates occludin on threonine residues (T403 and T404), and such phosphorylation is required for the assembly and/or maintenance of occludin in epithelial tight junctions that are key to the integrity and function of the airway epithelial barrier [42].

Recent GWAS of exacerbations in non-Hispanic white children identified SNPs in two genes, CDHR3 and CTNNNA3 [9, 10], neither of which replicated in our analysis, despite our having similar or larger sample sizes than those in the original studies. Similarly, a previously reported association between an SNP on chromosome 17q21 (rs7216389) and SAEs was not statistically significant in our meta-analysis (p=0.06). While an association between the 17q21 locus and asthma is widely recognised, whether this locus is linked to severe disease exacerbations among children with asthma is less clear. The initial report in paediatric asthma analysed a cohort of 376 children and reported 80 with recurrent wheezing and 66 with asthma [43]. The authors then analysed the rate of severe exacerbations within the full cohort (57 among the 376 participants), rather than comparing children with asthma and an exacerbation to those with asthma but no exacerbation. A recent meta-analysis reported a significant association between rs7216389 and asthma hospitalisations and ED visits in children, but, while the pooled estimates were significant, only four of the 13 cohorts showed significant results [44]. Although the lack of an association between CDHR3 and SAEs in the current study may be due to differences in the age of participants and outcome definitions across studies, our results for FLJ22447 are novel and further highlight the importance of studying asthma outcomes in children and adolescents of diverse races and ethnicities.

We recognise several study limitations. First, we lack data on viral infections or air pollution, which may interact with genetic or epigenetic mechanisms causing SAEs in children. Second, we cannot assess temporal relationships between DNA methylation or gene expression and SAEs in our cross-sectional analysis. Third, we had insufficient statistical power to detect either uncommon risk alleles or alleles with modest genetic effects on SAEs. Moreover, we had limited statistical power to detect an interaction between our top SNP and inhaled steroid use on SAEs or for our molecular quantitative trait analyses, and lacked a replication cohort for analyses of DNA methylation or gene expression in nasal epithelium and SAEs (because PIAMA lacks adequate data on asthma exacerbations). Fourth, the definitions of SAEs varied across the study cohorts. However, we obtained similar results in a sensitivity analysis for the HPR and GALA II cohorts, in which we re-ran the GWAS after excluding subjects who received systemic corticosteroids but did not report an unscheduled and acute visit for asthma (data not shown).

In summary, we have identified a novel SNP in FLJ22447 that is significantly associated with both SAEs in Latino children and adolescents and DNA methylation of a cis-CpG in FLJ22447 in nasal epithelium. This CpG is, in turn, linked to nasal epithelial expression of a gene recently implicated in atopic asthma in children and adolescents (KCNJ2-AS1). Future longitudinal studies of asthma-omics should assess whether and how genetic variants affect airway epithelial function and SAEs in childhood.

Author contributions: Q. Yan, E. Forno and J.C. Celedón conceived and designed the study. E.G. Burchard, G.H. Koppelman and J.C. Celedón obtained funding. Q. Yan, E. Herrera-Luis, M. Pino-Yanes and R. Rios conducted the primary analysis and interpreted data. C. Qi, Y-Y. Han, S. Kim, S. Oh, E. Acosta-Pérez, R. Zhang, D. Hu, C. Eng, S. Huntsman, L. Avila, N. Boutaoui, M.M. Cloutier, M.E. Soto-Quiros, C-J. Xu, S.T. Weiss, J. Lasky-Su, M.R. Kiedrowski, C. Figueiredo, J. Bomberger, M.L. Barreto, G. Canino, W. Chen and E.G. Burchard participated in data collection and data analysis. E. Herrera-Luis and J.C. Celedón prepared the first draft of the manuscript. All authors reviewed the draft for intellectual content, and approved submission of the final version of the manuscript.

Support statement: This work was supported by grants HL079966, HL117191 and MD011764 from the US National Institutes of Health (NIH) to J.C. Celedón. Q. Yan’s contribution was supported by grant HL138098 from the US NIH. E. Forno’s contribution was supported by grant HL149693 from the US NIH. The GALA II study was supported by US NIH grants to E.G. Burchard: HL088133, HL004646, HL117004, ES015794, ES24844, TRDRP 24RT 0025, MD006902 and GM007346. M. Pino-Yanes was supported by the Ramón y Cajal Program (RVC-2015-17205) and by grant SAF2017-83417R from the Spanish Ministry of Economy, Industry and Competitiveness. The PIAMA study was supported by The Netherlands Organisation for Health Research and Development; The Netherlands Organisation for Scientific Research; the Lung Foundation of the Netherlands (with methylation and gene expression studies supported by AF 4.1.14.001); The Netherlands Ministry of Spatial Planning, Housing and the Environment; and The Netherlands Ministry of Health, Welfare and Sport. C. Qi was supported by a grant from the China Scholarship Council. Funding information for this article has been deposited with the Crossref Funder Registry.

Conflict of interest: Q. Yan reports grants from National Institutes of Health (K01HL138098), during the conduct of the study. E. Forno has nothing to disclose. E. Herrera-Luis reports grants from the Spanish Ministry of Science, Innovation and Universities (PRE2018-083837), during the conduct of the study. M. Pino-Yanes reports grants from Spanish Ministry of Economy, Industry and Competitiveness, the State Research Agency and the European Regional
Development Funds from the European Union (MINECO/AEI/FEDER, UE, SAF2017-83417R), and grants from Spanish Ministry of Economy, Industry and Competitiveness (Ramón y Cajal grant RYC-2015-17205), during the conduct of the study. C. Qi has nothing to disclose. R. Rios has nothing to disclose. Y.-Y. Han has nothing to disclose. S. Kim has nothing to disclose. S. Oh has nothing to disclose. E. Acosta-Pérez has nothing to disclose. R. Zhang has nothing to disclose. D. Hu has nothing to disclose. C. Eng has nothing to disclose. S. Huntsman has nothing to disclose. L. Avila has nothing to disclose. N. Boutoufiq has nothing to disclose. M.M. Cloutier has nothing to disclose. M.E. Soto-Quiros has nothing to disclose. C.-J. Xu has nothing to disclose. J. Laskey-Su has nothing to disclose. M.R. Kiedrowski has nothing to disclose. C. Figueiredo has nothing to disclose. J. Bomberger has nothing to disclose. M.L. Barreto has nothing to disclose. G. Canino has nothing to disclose. W. Chen has nothing to disclose. G.H. Koppelman reports grants from the Lung Foundation of the Netherlands, during the conduct of the study; grants from Ubbo Emmius Foundation, TETRI Foundation, GSK, Vertex, TEVA, the Lung Foundation of the Netherlands and the European Union H2020 Program, outside the submitted work; and has served on advisory boards to GSK and PURE IMS (fees paid to institution). E.G. Burchard reports grants from National Heart, Lung, and Blood Institute, National Institute of Health, General Medical Sciences, National Institute on Minority Health and Health Disparities, Tobacco-Related Disease Research Program, Food and Drug Administration, and the Sandler Family Foundation, during the conduct of the study. J.C. Celeon has received research materials from GSK and Merck (inhaled steroids) and Pharmavite (vitamin D and placebo capsules), in order to provide medications free of cost to participants in NIH-funded studies, unrelated to the current work.

References


