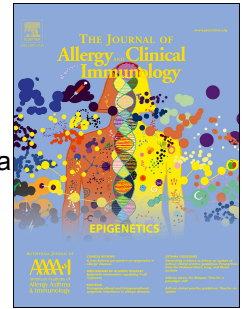


# Journal Pre-proof

Transcriptome-wide and differential expression network analyses of childhood asthma in nasal epithelium

Erick Forno, MD MPH, Rong Zhang, PhD, Yale Jiang, MD, Soyeon Kim, PhD, Qi Yan, PhD, Zhao Ren, PhD, Yueh-Ying Han, PhD, Nadia Boutaoui, PhD, Franziska Rosser, MD MPH, Daniel E. Weeks, PhD, Edna Acosta-Pérez, PhD, Angel Colón-Semidey, MD, María Alvarez, MD, Glorisa Canino, PhD, Wei Chen, PhD, Juan C. Celedón, MD DrPH, FAAAAI



PII: S0091-6749(20)30197-4

DOI: <https://doi.org/10.1016/j.jaci.2020.02.005>

Reference: YMAI 14422

To appear in: *Journal of Allergy and Clinical Immunology*

Received Date: 6 June 2019

Revised Date: 4 February 2020

Accepted Date: 12 February 2020

Please cite this article as: Forno E, Zhang R, Jiang Y, Kim S, Yan Q, Ren Z, Han Y-Y, Boutaoui N, Rosser F, Weeks DE, Acosta-Pérez E, Colón-Semidey A, Alvarez M, Canino G, Chen W, Celedón JC, Transcriptome-wide and differential expression network analyses of childhood asthma in nasal epithelium, *Journal of Allergy and Clinical Immunology* (2020), doi: <https://doi.org/10.1016/j.jaci.2020.02.005>.

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2020 Published by Elsevier Inc. on behalf of the American Academy of Allergy, Asthma & Immunology.

## Transcriptome-wide and differential expression network analyses of childhood asthma in nasal epithelium

Erick Forno, MD MPH<sup>1,2\*</sup>, Rong Zhang, PhD<sup>1,3\*</sup>, Yale Jiang MD<sup>1,4</sup>, Soyeon Kim, PhD<sup>1,2</sup>, Qi Yan, PhD<sup>1,2</sup>, Zhao Ren, PhD<sup>3</sup>, Yueh-Ying Han, PhD<sup>1,2</sup>, Nadia Boutaoui, PhD<sup>1,2</sup>, Franziska Rosser, MD MPH<sup>1,2</sup>, Daniel E. Weeks, PhD<sup>5</sup>, Edna Acosta-Pérez, PhD<sup>6</sup>, Angel Colón-Semidey, MD<sup>7</sup>, María Alvarez, MD<sup>7</sup>, Glorisa Canino, PhD<sup>6</sup>, Wei Chen, PhD<sup>1,2^</sup>, Juan C. Celedón, MD DrPH, FAAAAI<sup>1,2^</sup>

<sup>1</sup>Division of Pulmonary Medicine, Dept. of Pediatrics, UPMC Children's Hospital of Pittsburgh, Pittsburgh, PA. <sup>2</sup>University of Pittsburgh School of Medicine, Pittsburgh, PA. <sup>3</sup>Dept. of Statistics, University of Pittsburgh, PA. <sup>4</sup>School of Medicine, Tsinghua University, Beijing, China. <sup>5</sup>Dept. of Human Genetics, University of Pittsburgh Graduate School of Public Health, Pittsburgh, PA. <sup>6</sup>Behavioral Sciences Research Institute and <sup>7</sup>Dept of Pediatrics, Medical Science Campus, University of Puerto Rico, San Juan, PR.

\*Shared first and ^shared senior authors.

**Corresponding Author:** Juan C. Celedón, MD, DrPH, FAAAAI  
Division of Pediatric Pulmonary Medicine  
UPMC Children's Hospital of Pittsburgh  
4401 Penn Avenue, Pittsburgh, PA 15224  
Phone: 412.692.8429; Fax: 412.692.7636  
Email: [juan.celedon@chp.edu](mailto:juan.celedon@chp.edu)

**Declaration of Funding and Conflicts of Interest:** This study was supported by U.S. National Institutes of Health (NIH) grants HL079966, HL117191, and MD011764 (PI: Celedón JC). Dr Forno's contribution was supported by U.S. NIH grant HL125666. Dr. Celedón has received research materials from Merck and GSK (inhaled steroids) and Pharmavite (vitamin D and placebo tablets), to provide medications free of cost to subjects participating in two NIH-funded studies unrelated to the current work. The other authors have no conflicts of interest or funding to declare.

**Capsule Summary:**

In a transcription-wide association study of nasal epithelium, we identify novel and previously reported susceptibility genes for atopic asthma in children and show that gene co-expression networks differ markedly between children with and without atopic asthma.

**Key words:**

Airway epithelium, TWAS, atopic asthma, childhood asthma, gene network analysis

**Word Count:** 1,286

1 *To the Editor:*

2 Asthma affects ~7 million children in the U.S., where ~11-13% of children report respiratory  
3 or skin allergies. Genetic variants identified in large genome-wide association studies explain  
4 only a modest proportion of asthma risk, suggesting a substantial contribution of environmental  
5 exposures –which could alter the expression of susceptibility genes, ultimately leading to  
6 asthma pathogenesis.

7 Gene expression in nasal airway epithelium is well correlated with that in bronchial  
8 epithelium. Environmental stimuli such as pollutants and allergens could modify the epigenetic  
9 and transcriptomic characteristics of airway epithelium, leading to abnormal immune responses  
10 and asthma. We recently identified >7,000 CpGs associated with atopic asthma in an  
11 epigenome-wide association study of nasal epithelium in Puerto Rican children<sup>1</sup>, and a recent  
12 meta-analysis of airway epithelial gene expression identified 1,273 differentially expressed  
13 genes (DEGs) in asthma.<sup>2</sup> Here, we report a transcriptome-wide study of nasal epithelium and  
14 atopic asthma in Puerto Rican children, with replication in three separate cohorts.

15 Recruitment and procedures for the Epigenetic Variation and Childhood Asthma in Puerto  
16 Ricans (EVAPR) study, a case-control study of asthma in subjects aged 9-20 years, have been  
17 previously described.<sup>1</sup> Detailed methods are found in the **Online Repository**. For this analysis,  
18 cases (n=157) had atopic asthma, defined as a doctor's diagnosis of asthma,  $\geq 1$  episode of  
19 wheezing in the previous year, and atopy ( $\geq 1$  positive IgE to aero-allergens). Controls (n=101)  
20 had neither asthma nor atopy. RNA from nasal specimens was sequenced using 75-bp paired-  
21 end reads at 80M reads/sample, reads were aligned to reference genome (hg19), and  
22 transcripts-per-kilobase-million (TPM) were used as proxy for gene expression. After QC,  
23 18,311 genes were retained for analysis. DEGs were identified by negative binomial regression  
24 in DESeq2 adjusted for age, sex, sequencing plate, sample sorting protocol, and principal  
25 components from genotypic data. Significance was defined as false discovery rate (FDR)-  
26 corrected P-value<0.01. Replication (P-value<0.05 and effect direction concordant with EVAPR)

was attempted in three independent datasets.<sup>3-5</sup> Gene co-expression analysis was performed using SILGGM.<sup>6</sup>

Compared to controls (**Table E1**), cases had higher total IgE (mean 372 IU/mL vs 42 IU/mL). We found 6,058 DEGs (**Figure 1A**; 3,418 over-expressed and 2,640 under-expressed in atopic asthma). Of those, 458 had expression levels  $\geq 1.5$ -fold times higher in cases, and 227 showed expression  $\geq 1.5$ -fold times higher in controls. **Table 1** shows the top 50 DEGs (FDR  $P=2.78 \times 10^{-18}$  to  $2.08 \times 10^{-60}$ ), including *CST1*, *CST2*, *CST4*, *NTRK2*, *CDH26*, *CCL26*, *POSTN*, *TPSAB1*, *CLCA1*, *ALOX15*, *ITLN1*, *CPA3*, and other biologically relevant genes. While most enriched pathways ( $P$ -values=0.048 to  $1.15 \times 10^{-7}$ , see Figure E1) were related to immune system regulation (including antigen presentation, leukocyte extravasation,  $T_H1$  and  $T_H2$  activation, and granulocyte adhesion and diapedesis), a few were related to the biosynthesis of heparan, chondroitin, and dermatan sulfate ( $P$ -values=0.022 to  $7.76 \times 10^{-4}$ ). Upstream regulatory analysis predicted activation of the IL-13 pathway ( $P=6.5 \times 10^{-14}$ ; **Figure E2**); indeed, *IL13* was over-expressed in EVAPR cases (FDR  $P=2.24 \times 10^{-10}$ ). String-db interaction analysis<sup>7</sup> showed significant enrichment, with 179 connections among the top 200 DEGs ( $P=4.88 \times 10^{-15}$ ; **Figure E3**). Literature database analysis yielded 100 publications containing  $\geq 4$  of our top genes (FDR  $P=0.022$  to  $6.17 \times 10^{-7}$ ; **Table E2**); most of them related to asthma, atopy, eosinophilic esophagitis, eosinophils, mast cells, or  $T_H2$  cytokines. We also evaluated the association between gene expression and methylation (**Table E3**): 26 of our top 50 DEGs were also eQTMs (CpG/gene pairs where DNA methylation is inversely associated with gene expression) in which both methylation and expression were associated with atopic asthma in EVAPR –including *CDH26*, *CCL26*, *CPA3*, *MS4A2*, *ALOX15*, *NTRK1* and *NTRK2*.<sup>8</sup>

We replicated our top findings in three datasets (**Table E1**): 40 of the top 50 genes replicated in  $\geq 1$  cohort (combined  $P=4.08 \times 10^{-18}$  to  $2.93 \times 10^{-79}$ , **Table 1**), and 11/50 replicated in all three. Replication rates were higher in pediatric than in adult data: 35/49 (71.4%) of available

top-50 DEGs replicated using data from Giovannini-Chami (GSE19187),<sup>4</sup> as did 20/24 (83.3%) in the pediatric data from Yang (GSE65205);<sup>5</sup> compared to 21/48 (43.8%) from the study in adults by Yang (GSE104472).<sup>3</sup> Of the 6,058 DEGs in EVAPR, 1,318 (21.8%) replicated in  $\geq 1$  dataset: 684 (51.9%) in GSE19187, 747 (56.7%) in GSE65205, and 276 (20.9%) in GSE104472. We also replicated the top results from previous studies on transcriptome-wide gene expression in nasal epithelium and asthma (**Tables E4-E6**).<sup>3-5,9</sup>

Co-expression network analysis of the top 200 DEGs showed marked differences (**Figure E4**), indicating that the way these genes are connected (co-expressed) with each other also differs between atopic asthma and controls. The degree of connectivity (the number of genes with which each DEG is co-expressed) was significantly higher in cases (**Figure E4**; Kolmogorov-Smirnov  $P=5.23 \times 10^{-14}$ ): almost 100% of DEGs connected to  $\leq 2$  others among controls, while among cases  $\sim 50\%$  of DEGs connected to  $\geq 2$  others. We identified 17 gene “hubs” (those connected to  $\geq 4$  other genes in the network) in cases, including *POSTN*, *BRD4*, *DOK1*, *ELOVL5*, *FA2H*, *GCNT3*, and *SEC14L1* (**Figure 1B**, **Table E7**). Of those, 16 were significant in the pediatric cohorts and only 4 in the adult data (**Table E8**). There were no gene “hubs” among controls. Moreover, we found 194 gene pairs whose co-expression was significantly different in cases vs controls (**Table E9**); e.g., *POSTN* and *CTSG* were co-expressed in cases ( $P=4.87 \times 10^{-4}$ ) but not in controls ( $P=0.99$ ), with a significant difference between groups ( $P=0.013$ ). While connectivity distributions can depend on sample size, differences between cases and controls remained significant when we performed sensitivity analyses randomly selecting 101 cases to match the number of controls (**Table E10**).

Finally, we evaluated the ability of DEGs to predict atopic asthma in EVAPR using panels based on the top P-values, largest expression differences, or connectivity (**Figure 1C**;  $n=17$  genes in each panel, to match the number of gene “hubs”): these panels achieved areas under

the curve (AUC) of ~0.89-0.92, sensitivity ~0.80-0.85, specificity ~0.83-0.91, PPV ~0.88-0.93, NPV ~0.74-0.79, and overall accuracy of 82%-86% (**Table E11**). Replication results are shown in **Table E11** and **Figure E5**.

To our knowledge, this is the first study to report both individual-gene and network-level differences in gene expression from nasal epithelium in childhood asthma. We identified biologically relevant DEGs –including genes previously reported in asthma or other atopic diseases– as well as several novel genes. Many top DEGs were eQTMs,<sup>1</sup> with methylation inversely associated with expression and both associated with atopic asthma. As expected, the majority of top genes are related to atopy and immune pathways, but some are related to epithelial barrier processes (e.g., NTRK2 has been implicated in epithelial-mesenchymal transition dysfunction in asthma, while CDH26 regulates airway epithelial cell structure and polarity).<sup>12</sup> Interestingly, enriched pathways also included the biosynthesis of glycosaminoglycans that can participate in fibroblast activation in response to injury/inflammation.

We report a larger number of DEGs than a recent meta-analysis,<sup>2</sup> which pooled data from eight studies (two on pediatric nasal samples and six on adult bronchial samples) that used different platforms. EVAPR participants all met the same inclusion criteria and underwent uniform phenotyping and RNA sequencing protocols. We hypothesize that this allowed us to detect a larger number of significant signals, which nonetheless were statistically robust and largely replicated in independent datasets from different racial/ethnic and environmental backgrounds. We report higher replication rates using data from pediatric cohorts compared to adults, which suggests that many asthma DEGs vary with age.

Beyond the traditional individual-gene approach, we show that transcriptomic networks are quite different in atopic asthma. We identify “hub” genes that are biologically relevant and highly co-expressed with several others and may represent important novel targets for further exploration in functional and experimental studies. Taken together, our findings suggest a

complex interplay between epithelial integrity, immune regulation, and response to injury and inflammation. Finally, we show that a small number of DEGs can achieve high accuracy in identifying subjects with atopic asthma. If replicated in longitudinal studies, our findings may yield novel insights and promising biomarkers.

*Erick Forno, MD MPH\**  
*University of Pittsburgh*

*Rong Zhang, PhD\**  
*University of Pittsburgh*

*Yale Jiang MDc*  
*University of Pittsburgh*

*Soyeon Kim, PhD*  
*University of Pittsburgh*

*Qi Yan, PhD*  
*University of Pittsburgh*

*Zhao Ren, PhD*  
*University of Pittsburgh*

*Yueh-Ying Han, PhD*  
*University of Pittsburgh*

*Nadia Boutaoui, PhD*  
*University of Pittsburgh*

*Franziska Rosser, MD MPH*  
*University of Pittsburgh*

135 *Daniel E. Weeks, PhD*  
136 *University of Pittsburgh*  
137  
138 *Edna Acosta-Pérez, PhD*  
139 *University of Puerto Rico*  
140  
141 *Angel Colón-Semidey, MD*  
142 *University of Puerto Rico*  
143  
144 *María Alvarez, MD*  
145 *University of Puerto Rico*  
146  
147 *Glorisa Canino, PhD*  
148 *University of Puerto Rico*  
149  
150 *Wei Chen, PhD<sup>^</sup>*  
151 *University of Pittsburgh*  
152  
153 *Juan C. Celedón, MD DrPH, FAAAAI<sup>^</sup>*  
154 *University of Pittsburgh*  
155  
156 \*Shared first and <sup>^</sup>shared senior authors  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172

## References

1. Forno E, Wang T, Qi C, Yan Q, Xu CJ, Boutaoui N, et al. DNA methylation in nasal epithelium, atopy, and atopic asthma in children: a genome-wide study. *Lancet Respir Med* 2019;7:336-46.
2. Tsai YH, Parker JS, Yang IV, Kelada SNP. Meta-analysis of airway epithelium gene expression in asthma. *Eur Respir J* 2018;51.
3. Yang IV, Richards A, Davidson EJ, Stevens AD, Kolakowski CA, Martin RJ, et al. The Nasal Methylome: A Key to Understanding Allergic Asthma. *Am J Respir Crit Care Med* 2017;195:829-31.
4. Giovannini-Chami L, Marcet B, Moreilhon C, Chevalier B, Illie MI, Lebrigand K, et al. Distinct epithelial gene expression phenotypes in childhood respiratory allergy. *Eur Respir J* 2012;39:1197-205.
5. Yang IV, Pedersen BS, Liu AH, O'Connor GT, Pillai D, Kattan M, et al. The nasal methylome and childhood atopic asthma. *J Allergy Clin Immunol* 2017;139:1478-88.
6. Zhang R, Ren Z, Chen W. SILGGM: An extensive R package for efficient statistical inference in large-scale gene networks. *PLoS Comput Biol* 2018;14:e1006369.
7. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, et al. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res* 2019;47:D607-D13.
8. Singhanian A, Wallington JC, Smith CG, Horowitz D, Staples KJ, Howarth PH, et al. Multitissue Transcriptomics Delineates the Diversity of Airway T Cell Functions in Asthma. *Am J Respir Cell Mol Biol* 2018;58:261-70.

9. Pandey G, Pandey OP, Rogers AJ, Ahsen ME, Hoffman GE, Raby BA, et al. A Nasal Brush-based Classifier of Asthma Identified by Machine Learning Analysis of Nasal RNA Sequence Data. Sci Rep 2018;8:8826.

## FIGURE LEGEND

**Figure 1** – Transcriptome-wide association study (TWAS) of atopic asthma in nasal epithelium

**Footnote:** **A)** Volcano plot showing differential gene expression in nasal epithelium between cases and controls. *CST1* (P-value  $2.08 \times 10^{-60}$ ) and *CLCA1* ( $\log_2$  fold change 8.14) not shown. **B)** Hub-based sub-network in cases (top) and controls (bottom). Hubs are genes with connected to  $\geq 4$  co-expressed genes. Only hub genes in cases and their direct connections are shown. In controls no genes had  $\geq 4$  connections. **C)** Prediction analysis of using three panels of top genes (by P-value, by  $\log_2$ -fold change, or by “hub” genes). Each  $n=17$  (based on the number of “hub” genes identified) to ensure an equal number of predictors. AUC: Area under the receiver-operating characteristic (ROC) curve.

**Table 1 – Top 50 differentially expressed genes (DEGs) in atopic asthma**

| Gene                      | EVAPR       |                        |                        | GSE19187    |                        | GSE65205    |                        | GSE104472   |                        | Combined P-value <sup>a</sup> |
|---------------------------|-------------|------------------------|------------------------|-------------|------------------------|-------------|------------------------|-------------|------------------------|-------------------------------|
|                           | Fold-change | P-value                | FDR P-value            | Fold-change | P-value                | Fold-change | P-value                | Fold-change | P-value                |                               |
| <b>CST1*</b>              | 79.00       | 1.14×10 <sup>-64</sup> | 2.08×10 <sup>-60</sup> | 145.96      | 1.21×10 <sup>-12</sup> | 10.82       | 1.45×10 <sup>-08</sup> | 140.83      | 6.39×10 <sup>-06</sup> | 2.93×10 <sup>-79</sup>        |
| <b>CLCA1</b>              | 282.24      | 1.40×10 <sup>-47</sup> | 1.28×10 <sup>-43</sup> | 3.23        | 7.96×10 <sup>-03</sup> | n/a         | n/a                    | 13.65       | 5.03×10 <sup>-04</sup> | 3.22×10 <sup>-45</sup>        |
| <b>NTRK2*</b>             | 6.24        | 4.27×10 <sup>-44</sup> | 2.60×10 <sup>-40</sup> | 2.68        | 1.20×10 <sup>-05</sup> | 2.07        | 1.67×10 <sup>-06</sup> | 2.01        | 0.014                  | 2.14×10 <sup>-47</sup>        |
| <b>FETUB</b>              | 43.82       | 7.54×10 <sup>-42</sup> | 3.45×10 <sup>-38</sup> | 4.67        | 8.81×10 <sup>-05</sup> | n/a         | n/a                    | 35.18       | 6.90×10 <sup>-07</sup> | 1.29×10 <sup>-44</sup>        |
| <b>CPA3*</b>              | 21.77       | 3.43×10 <sup>-39</sup> | 1.26×10 <sup>-35</sup> | 6.01        | 2.19×10 <sup>-06</sup> | 3.19        | 1.43×10 <sup>-07</sup> | 6.34        | 4.91×10 <sup>-04</sup> | 5.25×10 <sup>-46</sup>        |
| GATA2                     | 13.14       | 4.45×10 <sup>-38</sup> | 1.36×10 <sup>-34</sup> | 1.12        | 0.14                   | 1.50        | 6.87×10 <sup>-04</sup> | 4.85        | 0.018                  | 3.03×10 <sup>-35</sup>        |
| <b>ITLN1*</b>             | 28.11       | 8.68×10 <sup>-38</sup> | 2.27×10 <sup>-34</sup> | 8.48        | 5.00×10 <sup>-05</sup> | 4.79        | 1.57×10 <sup>-05</sup> | 8.98        | 5.49×10 <sup>-03</sup> | 1.87×10 <sup>-40</sup>        |
| <b>CDH26*</b>             | 3.37        | 8.55×10 <sup>-37</sup> | 1.96×10 <sup>-33</sup> | 3.05        | 3.16×10 <sup>-05</sup> | 2.03        | 3.23×10 <sup>-05</sup> | 2.23        | 1.86×10 <sup>-03</sup> | 6.83×10 <sup>-40</sup>        |
| <b>CCL26*</b>             | 7.16        | 2.15×10 <sup>-35</sup> | 4.38×10 <sup>-32</sup> | 2.00        | 1.32×10 <sup>-03</sup> | 2.24        | 4.47×10 <sup>-04</sup> | 9.74        | 4.48×10 <sup>-04</sup> | 1.67×10 <sup>-36</sup>        |
| <b>CST2</b>               | 29.25       | 4.02×10 <sup>-35</sup> | 7.37×10 <sup>-32</sup> | 4.27        | 2.27×10 <sup>-07</sup> | n/a         | n/a                    | 92.64       | 3.25×10 <sup>-06</sup> | 2.75×10 <sup>-40</sup>        |
| C3orf70                   | 3.18        | 8.30×10 <sup>-35</sup> | 1.38×10 <sup>-31</sup> | 1.24        | 0.018                  | n/a         | n/a                    | 1.32        | 0.34                   | 2.53×10 <sup>-30</sup>        |
| <b>TPSAB1*</b>            | 17.36       | 3.27×10 <sup>-34</sup> | 4.98×10 <sup>-31</sup> | 2.10        | 3.25×10 <sup>-04</sup> | 2.66        | 5.40×10 <sup>-06</sup> | 4.82        | 2.50×10 <sup>-03</sup> | 3.32×10 <sup>-37</sup>        |
| CISH                      | 3.34        | 3.02×10 <sup>-32</sup> | 4.26×10 <sup>-29</sup> | 1.48        | 2.56×10 <sup>-04</sup> | n/a         | n/a                    | 2.00        | 0.26                   | 8.20×10 <sup>-30</sup>        |
| CLC                       | 20.29       | 9.83×10 <sup>-32</sup> | 1.29×10 <sup>-28</sup> | 3.00        | 7.33×10 <sup>-05</sup> | n/a         | n/a                    | 1.80        | 0.32                   | 8.69×10 <sup>-30</sup>        |
| <b>TPSB2</b>              | 15.97       | 7.98×10 <sup>-30</sup> | 9.74×10 <sup>-27</sup> | 2.21        | 2.43×10 <sup>-04</sup> | n/a         | n/a                    | n/a         | n/a                    | 1.64×10 <sup>-28</sup>        |
| <b>ALOX15*</b>            | 2.38        | 3.37×10 <sup>-29</sup> | 3.85×10 <sup>-26</sup> | 1.74        | 2.61×10 <sup>-04</sup> | 2.13        | 1.49×10 <sup>-06</sup> | 1.62        | 0.046                  | 6.86×10 <sup>-32</sup>        |
| <b>MS4A2</b>              | 14.65       | 4.86×10 <sup>-29</sup> | 5.23×10 <sup>-26</sup> | 1.93        | 1.06×10 <sup>-04</sup> | n/a         | n/a                    | 3.82        | 3.65×10 <sup>-03</sup> | 5.52×10 <sup>-29</sup>        |
| TPSG1                     | 32.78       | 5.95×10 <sup>-29</sup> | 6.06×10 <sup>-26</sup> | 0.94        | 0.23                   | 1.08        | 0.42                   | 0.92        | 0.79                   | 1.77×10 <sup>-22</sup>        |
| CEP72                     | 2.19        | 1.26×10 <sup>-28</sup> | 1.22×10 <sup>-25</sup> | 1.27        | 9.47×10 <sup>-03</sup> | 1.38        | 4.74×10 <sup>-04</sup> | 0.87        | 0.75                   | 2.43×10 <sup>-26</sup>        |
| SLC5A5                    | 6.26        | 5.54×10 <sup>-28</sup> | 5.07×10 <sup>-25</sup> | 2.79        | 2.14×10 <sup>-03</sup> | n/a         | n/a                    | 2.38        | 0.21                   | 4.74×10 <sup>-25</sup>        |
| <b>POSTN</b> <sup>†</sup> | 9.10        | 7.69×10 <sup>-28</sup> | 6.70×10 <sup>-25</sup> | 13.93       | 6.82×10 <sup>-07</sup> | n/a         | n/a                    | 5.63        | 5.33×10 <sup>-04</sup> | 7.48×10 <sup>-31</sup>        |
| MYO15A                    | 3.20        | 4.14×10 <sup>-27</sup> | 3.44×10 <sup>-24</sup> | 1.02        | 0.78                   | n/a         | n/a                    | 1.66        | 0.11                   | 4.72×10 <sup>-22</sup>        |
| HS3ST4                    | 5.22        | 5.19×10 <sup>-27</sup> | 4.13×10 <sup>-24</sup> | 0.96        | 0.58                   | n/a         | n/a                    | 3.35        | 8.88×10 <sup>-03</sup> | 3.88×10 <sup>-23</sup>        |
| PRB2                      | 87.44       | 7.05×10 <sup>-27</sup> | 5.38×10 <sup>-24</sup> | 1.23        | 0.13                   | n/a         | n/a                    | 5.06        | 0.010                  | 1.32×10 <sup>-23</sup>        |
| LOC100507472              | 3.80        | 1.07×10 <sup>-26</sup> | 7.87×10 <sup>-24</sup> | n/a         | n/a                    | n/a         | n/a                    | 0.47        | 0.34                   | 1.49×10 <sup>-22</sup>        |
| <b>PCSK6*</b>             | 2.50        | 1.35×10 <sup>-26</sup> | 9.28×10 <sup>-24</sup> | 1.39        | 1.06×10 <sup>-03</sup> | 1.25        | 4.52×10 <sup>-04</sup> | 1.81        | 0.039                  | 1.08×10 <sup>-26</sup>        |
| <b>NTRK1</b>              | 8.60        | 1.37×10 <sup>-26</sup> | 9.28×10 <sup>-24</sup> | 1.11        | 0.036                  | n/a         | n/a                    | 2.54        | 1.70×10 <sup>-03</sup> | 1.15×10 <sup>-24</sup>        |
| WBSCR17                   | 4.27        | 2.54×10 <sup>-26</sup> | 1.66×10 <sup>-23</sup> | 1.40        | 7.63×10 <sup>-03</sup> | n/a         | n/a                    | 2.10        | 0.26                   | 5.87×10 <sup>-23</sup>        |
| CCBL1                     | 2.70        | 2.78×10 <sup>-26</sup> | 1.75×10 <sup>-23</sup> | 1.68        | 2.99×10 <sup>-03</sup> | 1.31        | 2.65×10 <sup>-03</sup> | 1.24        | 0.64                   | 4.16×10 <sup>-24</sup>        |
| ANO1                      | 3.01        | 5.10×10 <sup>-26</sup> | 3.11×10 <sup>-23</sup> | 2.35        | 2.12×10 <sup>-05</sup> | 1.65        | 1.31×10 <sup>-06</sup> | 1.37        | 0.25                   | 1.70×10 <sup>-29</sup>        |
| PTGDR2                    | 4.30        | 7.09×10 <sup>-26</sup> | 4.19×10 <sup>-23</sup> | 0.99        | 0.89                   | n/a         | n/a                    | 1.51        | 0.43                   | 2.29×10 <sup>-20</sup>        |
| ABO                       | 2.52        | 3.05×10 <sup>-25</sup> | 1.74×10 <sup>-22</sup> | 1.18        | 0.078                  | 1.02        | 0.89                   | 1.24        | 0.35                   | 1.18×10 <sup>-19</sup>        |

|                |       |                        |                        |      |                        |      |                        |      |                        |                        |
|----------------|-------|------------------------|------------------------|------|------------------------|------|------------------------|------|------------------------|------------------------|
| <b>CMYA5*</b>  | 2.17  | 4.46×10 <sup>-25</sup> | 2.47×10 <sup>-22</sup> | 2.07 | 6.43×10 <sup>-06</sup> | 1.26 | 9.23×10 <sup>-04</sup> | 1.85 | 7.76×10 <sup>-03</sup> | 7.87×10 <sup>-28</sup> |
| <i>CST4</i>    | 25.58 | 1.89×10 <sup>-24</sup> | 1.02×10 <sup>-21</sup> | 1.03 | 0.66                   | n/a  | n/a                    | n/a  | n/a                    | 3.35×10 <sup>-20</sup> |
| <i>GCSAML</i>  | 8.33  | 2.08×10 <sup>-24</sup> | 1.09×10 <sup>-21</sup> | 1.12 | 0.090                  | n/a  | n/a                    | 0.99 | 0.89                   | 1.17×10 <sup>-19</sup> |
| <i>CDK15</i>   | 9.97  | 1.28×10 <sup>-23</sup> | 6.50×10 <sup>-21</sup> | 1.03 | 0.59                   | n/a  | n/a                    | 1.47 | 0.093                  | 4.53×10 <sup>-19</sup> |
| <i>SLC24A3</i> | 2.64  | 1.32×10 <sup>-23</sup> | 6.52×10 <sup>-21</sup> | 1.06 | 0.45                   | 1.48 | 8.76×10 <sup>-04</sup> | 2.06 | 0.22                   | 1.74×10 <sup>-20</sup> |
| <i>GCNT4</i>   | 2.23  | 1.09×10 <sup>-22</sup> | 5.24×10 <sup>-20</sup> | 2.70 | 7.57×10 <sup>-05</sup> | n/a  | n/a                    | 2.31 | 0.20                   | 1.23×10 <sup>-21</sup> |
| <i>PTCHD4</i>  | 4.05  | 1.40×10 <sup>-22</sup> | 6.58×10 <sup>-20</sup> | 1.11 | 0.28                   | n/a  | n/a                    | 2.27 | 0.17                   | 3.54×10 <sup>-18</sup> |
| <i>SLC7A1</i>  | 1.67  | 1.60×10 <sup>-22</sup> | 7.34×10 <sup>-20</sup> | 1.66 | 1.69×10 <sup>-04</sup> | 1.24 | 2.92×10 <sup>-03</sup> | 0.94 | 0.88                   | 1.14×10 <sup>-21</sup> |
| <i>ST18</i>    | 3.21  | 1.64×10 <sup>-22</sup> | 7.34×10 <sup>-20</sup> | 1.26 | 2.27×10 <sup>-03</sup> | n/a  | n/a                    | 1.24 | 0.26                   | 5.94×10 <sup>-20</sup> |
| <i>SLC45A4</i> | 1.55  | 1.79×10 <sup>-22</sup> | 7.79×10 <sup>-20</sup> | 1.30 | 2.54×10 <sup>-03</sup> | 1.20 | 5.45×10 <sup>-03</sup> | 1.21 | 0.19                   | 6.56×10 <sup>-21</sup> |
| <i>DQX1</i>    | 2.37  | 6.88×10 <sup>-22</sup> | 2.93×10 <sup>-19</sup> | 1.37 | 0.037                  | 1.23 | 0.079                  | 1.47 | 0.074                  | 1.49×10 <sup>-18</sup> |
| <b>GSN*</b>    | 1.64  | 9.39×10 <sup>-22</sup> | 3.86×10 <sup>-19</sup> | 1.53 | 3.44×10 <sup>-05</sup> | 1.53 | 4.74×10 <sup>-06</sup> | 1.96 | 1.35×10 <sup>-03</sup> | 5.38×10 <sup>-27</sup> |
| <i>LGALS12</i> | 13.75 | 9.48×10 <sup>-22</sup> | 3.86×10 <sup>-19</sup> | 1.28 | 0.023                  | n/a  | n/a                    | 1.77 | 0.19                   | 1.99×10 <sup>-18</sup> |
| <i>HRH4</i>    | 3.99  | 2.25×10 <sup>-21</sup> | 8.87×10 <sup>-19</sup> | 1.72 | 0.011                  | n/a  | n/a                    | 1.13 | 0.37                   | 4.08×10 <sup>-18</sup> |
| <i>KCNJ16</i>  | 5.97  | 2.28×10 <sup>-21</sup> | 8.87×10 <sup>-19</sup> | 1.70 | 0.011                  | 1.63 | 5.82×10 <sup>-04</sup> | 1.91 | 0.36                   | 5.90×10 <sup>-20</sup> |
| <i>TAL1</i>    | 7.16  | 2.37×10 <sup>-21</sup> | 9.04×10 <sup>-19</sup> | 0.94 | 0.38                   | 1.16 | 0.060                  | 1.06 | 0.88                   | 3.05×10 <sup>-16</sup> |
| <i>SLC9B2</i>  | 2.70  | 3.50×10 <sup>-21</sup> | 1.31×10 <sup>-18</sup> | 2.72 | 3.66×10 <sup>-04</sup> | n/a  | n/a                    | 3.80 | 0.13                   | 8.51×10 <sup>-20</sup> |
| <i>CTSG</i>    | 16.51 | 7.74×10 <sup>-21</sup> | 2.78×10 <sup>-18</sup> | 1.15 | 0.094                  | 2.38 | 7.16×10 <sup>-08</sup> | 4.47 | 1.44×10 <sup>-03</sup> | 1.35×10 <sup>-24</sup> |

EVAPR analysis adjusted for age, sex, and genotypic principal components. GSE19187, GSE65205, and GSE104472 models adjusted for gender and age.

Replication defined as P-value <0.05 and effect in the same direction as in EVAPR. **Bold**: Replicated in all available cohorts. <sup>a</sup>Fisher combined P-values.

\*Replicated in all three cohorts. <sup>†</sup>Gene also a “hub” gene in the differential gene network analysis. n/a: Not available.

Journal Pre-proof

