Tobacco Smoking Increases the Lung Gene Expression of ACE2, the Receptor of SARS-CoV-2

Guoshuai Cai, PhD[1*]; Yohan Bossé[2]; Feifei Xiao, PhD[3]; Farrah Kheradmand, MD[4]; Christopher I. Amos, PhD[5*]

[1]Department of Environmental Health Sciences, Arnold School of Public Health, University of South Carolina, Columbia, SC 29208, USA

[2]Institut universitaire de cardiologie et de pneumologie de Québec (IUCPQ), Department of Molecular Medicine, Laval University, Quebec City, Canada

[3]Department of Epidemiology and Biostatistics, Arnold School of Public Health, University of South Carolina, Columbia, SC 29208, USA

[4]Department of Medicine, Baylor College of Medicine, & Michael E. DeBakey VA Center, US Department of Veterans Affairs, Houston, TX, 77030, USA

[5]Institute for Clinical and Translational Research and Dan L. Duncan Comprehensive Cancer Center, Department of Medicine, Baylor College of Medicine, Houston, TX 77030, USA

*Corresponding author

Address for Correspondence:

Guoshuai Cai

915 Greene Street

Discovery 517

Columbia, South Carolina 29204

GCAI@mailbox.sc.edu

Phone: 803-777-4120

AJRCCM Articles in Press. Published April 24, 2020 as 10.1164/rccm.202003-0693LE
Copyright © 2020 by the American Thoracic Society

Christopher I. Amos

1 Baylor Plaza

Suite 110.41D

Baylor College of Medicine

Houston, Texas 77030

Phone: 713-798-2102

Word count: 999

Conception and design: GC, CIA; Analysis and interpretation: GC, YB, CIA; Drafting the manuscript for important intellectual content: GC, YB, FX, FK, CIA

Some of the results of these studies have been previously reported in the form of a preprint (medRxiv, 28 February 2020 https://www.medrxiv.org/content/10.1101/2020.02.05.20020107v3)

This article is open access and distributed under the terms of the Creative Commons Attribution Non-Commercial No Derivatives License 4.0 (http://creativecommons.org/licenses/by-nc-nd/4.0/). For commercial usage and reprints please contact Diane Gern (dgern@thoracic.org).

On March 11, 2020, the World Health Organization declared the Coronavirus Disease 2019 (COVID-19) outbreak a pandemic. As of April 20th, 2020, laboratories have confirmed 2,470,410 COVID-19 cases and caused 169,794 deaths in 213 counties, areas, or territories. COVID-19 is caused by a new type of pathogenic coronavirus, severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), which is phylogenetically similar to SARS-CoV, with approximately 80% identity between genomes (1). SARS viruses affect the respiratory tract and cause acute respiratory response through the same cell-entry receptor, angiotensin-converting enzyme 2 (ACE2), which is the only experimentally confirmed SARS-CoV-2 receptor. SARS-CoV-2 infection also employs activation of the spike proteins found on the surface of the virus for cellular entry. The best candidates for priming spike proteins are two host-cell enzymes called Furin and TMPRSS2 (1). In the current severe global emergency, it is imperative to identify potential risk factors for effective prevention and care, such as cigarette smoking, which is a substantial risk factor for various important bacterial and viral infections.

Methods

We evaluated a comprehensive set of transcriptomic datasets to investigate the associations of smoking with *ACE2*, *FURIN* and *TMPRSS2* gene expression in lung tissues. Two datasets were generated using normal lung tissues from patients with Lung Adenocarcinoma: a Caucasian RNA-seq dataset from The Cancer Genome Atlas TCGA (n=48) (2) and an Asian RNA-seq dataset from Gene Expression Omnibus (GEO) GSE40419 (n=74) (3). We included three polyethnic microarray datasets of gene expression of healthy small airway epithelium (SAE) samples (from GSE63127 (n=230) (4), GSE19667 (n=116) (5) and GSE5058 (n=24) (6)) and a large airway epithelium (LAE) samples from GSE7895 (n=104) (7). Also,

AJRCCM Articles in Press. Published April 24, 2020 as 10.1164/rccm.202003-0693LE
Copyright © 2020 by the American Thoracic Society

we analyzed three microarray datasets of samples derived from healthy and chronic obstructive pulmonary disease (COPD) patients, including SAE samples from current smokers (from GSE5058 (6), n=26), bronchial airway epithelium (BAE) samples from current and former smokers (GSE37147 (8), n=238), and lung samples from Caucasian patients (n=438) that underwent lung cancer surgery at the Institut universitaire de cardiologie et de pneumologie de Québec (IUCPQ) (9). RNA-seq datasets were generated with the Illumina HiSeq platform and microarray datasets were generated with Affymetrix arrays. All the data were de-identified and study approval has been stated in original studies. A total of 1286 assay results were evaluated. We considered the Fragments Per kilobase per million mapped reads (FPKM) for RNA-seq data and robust multi-array average (RMA) values for microarray data to represent normalized gene expression. All data were $\log_2$ transformed to improve normality.

Smoking status (never, former and current smokers) was identified based on self-reported smoking history. Association test was performed using a linear model with $\log_2$ *ACE2*, *FURIN*, or *TMPRSS2* gene expression as the dependent variable and smoking status or COPD status as an independent variable. Meta-analysis was performed by pooling effect size and standard error estimated from each study using a random effects model. Age and sex were included as covariates. Although we did not observe significant associations of age and sex with the expression of *ACE2* and *FURIN*, we found a negative correlation between *TMPRSS2* expression and age in some of the datasets. Data management, statistical analyses, and visualizations were performed using R 3.6.1.

A single-cell RNA-seq dataset (GSE131391) (10) was also analyzed. This study profiled bronchial epithelial cells, single ALCAM+ epithelial cells and CD45+

AJRCCM Articles in Press. Published April 24, 2020 as 10.1164/rccm.202003-0693LE
Copyright © 2020 by the American Thoracic Society

white blood cells from six never and six current smokers. Sequencing read counts in single cells were downloaded, and subsequent data analyses, including data normalization, high variable feature selection, data scaling, dimension reduction, and cluster identification, were performed using the Seurat 3.0 package. We used SCANNER for data visualization and cell type identification.

Results

We identified upregulation of pulmonary *ACE2* gene expression in ever-smokers compared to non-smokers in all datasets, irrespective of tissue subset or COPD status (Fig. 1). Meta-analysis showed ever-smoking significantly and substantially increased pulmonary *ACE2* expression by 25% (p-value=1.4×10^{-16}, Fig. 1). Similarly, smoking status defined by never, former and current smokers was also significantly associated with *ACE2* pulmonary expression in meta-analysis (beta=0.14, p-value=2.0×10^{-6}, Fig. 1). The significant smoking effect on *ACE2* pulmonary expression identified in this study may suggest an increased risk for viral binding and entry of SARS-CoV and SARS-CoV-2 in lungs of smokers. *FURIN* was also upregulated by smoking, but to a lower extent compared to *ACE2. TMRPSS2* gene expression in lung was not associated with smoking (Fig. 1).

We also evaluated the effect of COPD on gene expression. Stratifying data by smoking status, we observed a trend (beta=0.08, p-value=0.07) for higher ACE2 levels in COPD patients, but the results were not consistent across datasets (Fig. 1). In the IUCPQ data, *ACE2* expression was upregulated in COPD patients (p=0.0006), but the effect was attenuated after adjustment for smoking status (p=0.03).

We further evaluated the effect of smoking on *ACE2* pulmonary expression in single bronchial epithelial cells from six never and six current smokers. We found smoking remodels cells in the bronchial epithelium with a loss of club cells and

AJRCCM Articles in Press. Published April 24, 2020 as 10.1164/rccm.202003-0693LE
Copyright © 2020 by the American Thoracic Society

extensive hyperplasia of goblet cells. *ACE2* gene is mainly expressed in goblet cells in smokers and club cells in never smokers (Fig. 2). This result is consistent with a very recent study that found the highest ACE2 expression in Alveolar Type II cells (which derive from club cells) and in a transient secretory cell type in subsegmental bronchial branches (11). This may indicate that smokers have a risk of COVID-19 infection complications based on *ACE2* expression profiles, which could contribute to variations in infection susceptibility, disease severity, and treatment outcome.

Despite a significant increase in the prevalence of electronic cigarettes (E-cig), no studies have yet included information about single (E-cig only), or dual users (E-cig and tobacco). The mechanisms underlying tobacco-related upregulation of *ACE2* pulmonary expression are unknown; among many others, the degrees of smoking effects to the infection susceptibility and clinical manifestations are unknown; further mechanistic studies are needed. Despite currently limited knowledge, this study indicates that smoking could be a risk factor for COVID-19 by affecting *ACE2* expression and provides valuable information for identifying and stratifying more susceptible populations.

AJRCCM Articles in Press. Published April 24, 2020 as 10.1164/rccm.202003-0693LE
Copyright © 2020 by the American Thoracic Society

References

1. Walls AC, Park YJ, Tortorici MA, Wall A, McGuire AT, Veesler D. Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* 2020.

2. Cancer Genome Atlas Research N, Weinstein JN, Collisson EA, Mills GB, Shaw KR, Ozenberger BA, Ellrott K, Shmulevich I, Sander C, Stuart JM. The Cancer Genome Atlas Pan-Cancer analysis project. *Nat Genet* 2013; 45: 1113-1120.

3. Seo JS, Ju YS, Lee WC, Shin JY, Lee JK, Bleazard T, Lee J, Jung YJ, Kim JO, Shin JY, Yu SB, Kim J, Lee ER, Kang CH, Park IK, Rhee H, Lee SH, Kim JI, Kang JH, Kim YT. The transcriptional landscape and mutational profile of lung adenocarcinoma. *Genome research* 2012; 22: 2109-2119.

4. Tilley AE, Staudt MR, Salit J, Van de Graaf B, Strulovici-Barel Y, Kaner RJ, Vincent T, Agosto-Perez F, Mezey JG, Raby BA, Crystal RG. Cigarette Smoking Induces Changes in Airway Epithelial Expression of Genes Associated with Monogenic Lung Disorders. *Am J Respir Crit Care Med* 2016; 193: 215-217.

5. Strulovici-Barel Y, Omberg L, O'Mahony M, Gordon C, Hollmann C, Tilley AE, Salit J, Mezey J, Harvey BG, Crystal RG. Threshold of biologic responses of the small airway epithelium to low levels of tobacco smoke. *Am J Respir Crit Care Med* 2010; 182: 1524-1532.

6. Tilley AE, Harvey BG, Heguy A, Hackett NR, Wang R, O'Connor TP, Crystal RG. Down-regulation of the notch pathway in human airway epithelium in association with smoking and chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2009; 179: 457-466.

AJRCCM Articles in Press. Published April 24, 2020 as 10.1164/rccm.202003-0693LE
Copyright © 2020 by the American Thoracic Society

7. Beane J, Sebastiani P, Liu G, Brody JS, Lenburg ME, Spira A. Reversible and permanent effects of tobacco smoke exposure on airway epithelial gene expression. *Genome Biol* 2007; 8: R201.

8. Steiling K, van den Berge M, Hijazi K, Florido R, Campbell J, Liu G, Xiao J, Zhang X, Duclos G, Drizik E, Si H, Perdomo C, Dumont C, Coxson HO, Alekseyev YO, Sin D, Pare P, Hogg JC, McWilliams A, Hiemstra PS, Sterk PJ, Timens W, Chang JT, Sebastiani P, O'Connor GT, Bild AH, Postma DS, Lam S, Spira A, Lenburg ME. A dynamic bronchial airway gene expression signature of chronic obstructive pulmonary disease and lung function impairment. *Am J Respir Crit Care Med* 2013; 187: 933-942.

9. Bosse Y, Postma DS, Sin DD, Lamontagne M, Couture C, Gaudreault N, Joubert P, Wong V, Elliott M, van den Berge M, Brandsma CA, Tribouley C, Malkov V, Tsou JA, Opiteck GJ, Hogg JC, Sandford AJ, Timens W, Pare PD, Laviolette M. Molecular signature of smoking in human lung tissues. *Cancer Res* 2012; 72: 3753-3763.

10. Duclos GE, Teixeira VH, Autissier P, Gesthalter YB, Reinders-Luinge MA, Terrano R, Dumas YM, Liu G, Mazzilli SA, Brandsma CA, van den Berge M, Janes SM, Timens W, Lenburg ME, Spira A, Campbell JD, Beane J. Characterizing smoking-induced transcriptional heterogeneity in the human bronchial epithelium at single-cell resolution. *Sci Adv* 2019; 5: eaaw3413.

11. Lukassen S, Lorenz Chua R, Trefzer T, Kahn NC, Schneider MA, Muley T, Winter H, Meister M, Veith C, Boots AW, Hennig BP, Kreuter M, Conrad C, Eils R. SARS-CoV-2 receptor ACE2 and TMPRSS2 are primarily expressed in bronchial transient secretory cells. *EMBO J* 2020.

AJRCCM Articles in Press. Published April 24, 2020 as 10.1164/rccm.202003-0693LE
Copyright © 2020 by the American Thoracic Society

Figure

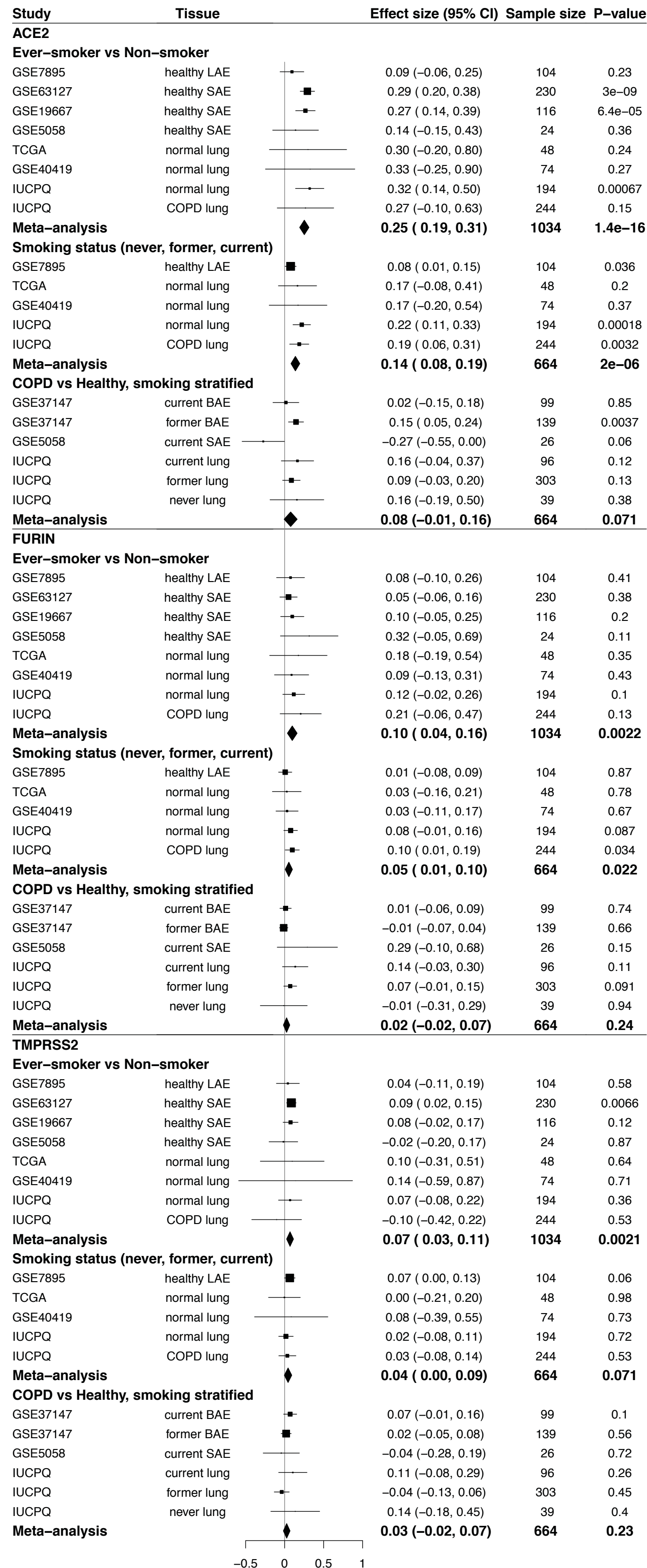
Figure 1. Forest plots for smoking effect on *ACE2*, *FURIN*, and *TMPRSS2* pulmonary gene expression. Non-smoker/never-smoker and ever-smoker including current smoker and former smoker were identified in each original study based on self-reported smoking history. For each gene, the top panel shows the comparison of ever-smoker and non-smoker groups, the middle panel shows the association of *ACE2* gene expression with smoking status (never-, former- and current-smoker), and the bottom panel shows the comparison of COPD and healthy groups, stratified by smoking status. For each study, the estimated effect size and 95% confidence intervals are plotted. The size of square boxes is proportional to weights which were estimated by the standard 'inverse-variance' method for random-effects models in meta-analysis. TCGA, The Cancer Genome Atlas; IUCPQ, "Institut universitaire de cardiologie et de pneumologie de Québec"; SAE, small airway epithelium; LAE, large airway epithelium; BAE, bronchial airway epithelium.

Figure 2. *ACE2* expression in single-cell transcriptomics of bronchial epithelium cells from never and current smokers.

The tSNE plot of single-cell transcriptome profiles from never smokers and current smokers is showed. *ACE2* expression is shown by red stars. Cell types were identified based on gene expression of markers (10). Confirming the original study, we observed smokers showed a remodeled cell composition in bronchial epithelium with a loss of club cells and extensive hyperplasia of goblet cells.

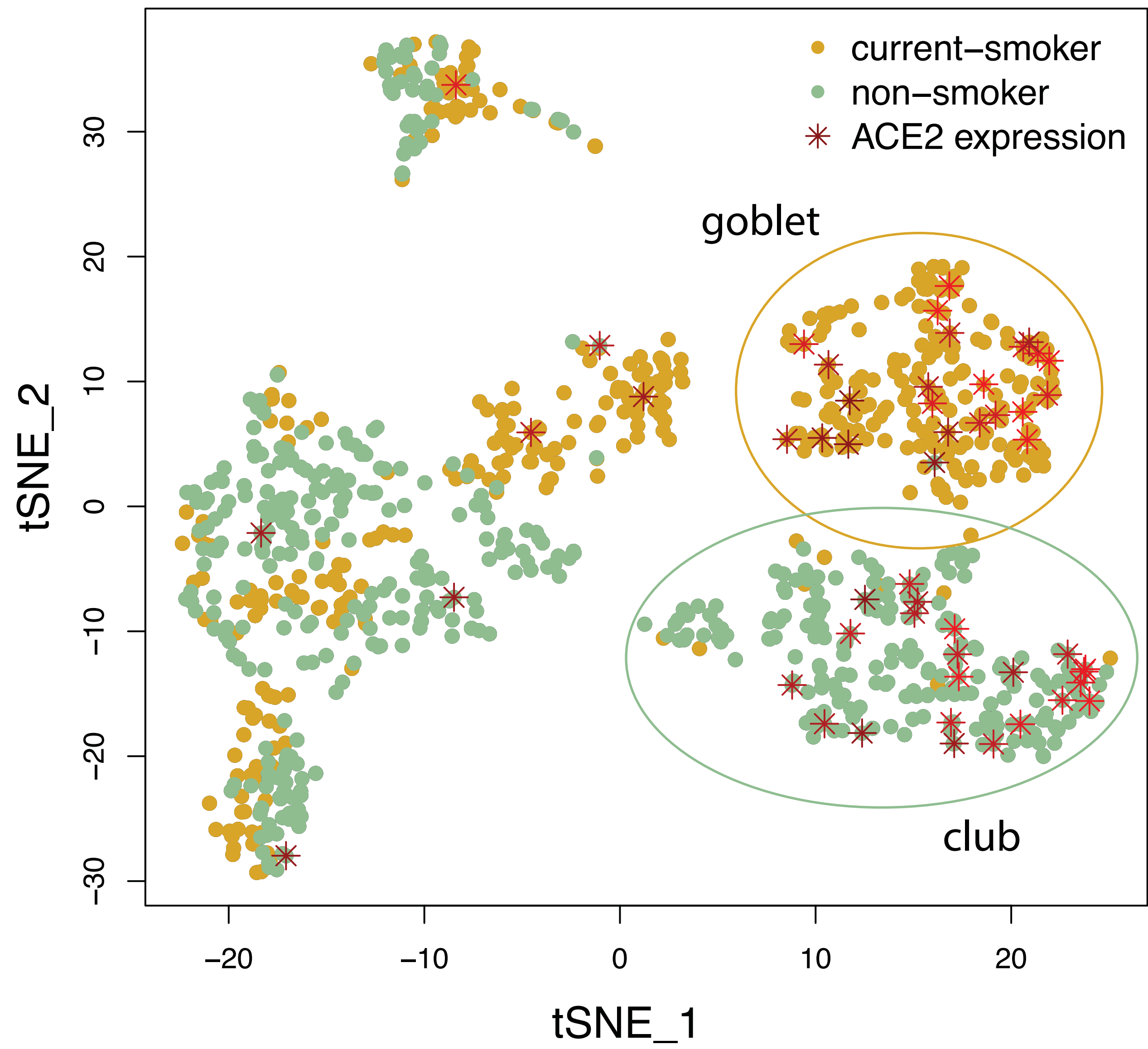
AJRCCM Articles in Press. Published April 24, 2020 as 10.1164/rccm.202003-0693LE
Copyright © 2020 by the American Thoracic Society

Figure 1



Copyright © 2020 by the American Thoracic Society

Figure 2



AJRCCM Articles in Press. Published April 24, 2020 as 10.1164/rccm.202003-0693LE
Copyright © 2020 by the American Thoracic Society