Short Report

DNA Methylation Analysis of the COVID-19 host cell receptor, Angiotensin I Converting Enzyme 2 gene (ACE2) in the Respiratory System Reveal Age and Gender Differences

AUTHORS:

1Michael J. Corley and 1Lishomwa C. Ndhlovu*

AFFILIATION:

1Division of Infectious Diseases, Department of Medicine, Weill Cornell Medical College, New York, NY, USA.

*Corresponding Author

Correspondence:

Lishomwa C. Ndhlovu MD, PhD,
Division of Infectious Diseases, Department of Medicine,
Weill Cornell Medical College,
413E 69th St, New York, NY, USA.
Email: lndhlovu@med.cornell.edu

Keywords:

Wuhan 2019-nCoV, ACE2, DNA methylation, epigenetics, profiling, lung tissue, age, gender, COVID-19, coronavirus

CONFLICT OF INTEREST: Authors declare no conflict of interest.
ABSTRACT:

**Background:** Coronavirus disease 2019 (COVID-19) has emerged as a global threat to human health and disease risk increases with advancing age. The regulation of the ACE2 gene that codes for COVID-19 host receptor ACE2 has been shown to be under epigenetic regulation. Here, we examined whether intensive DNA methylation profiling of the ACE2 gene differed by human host tissue and cell type, gender, and age.

**Results:** Accessing four public datasets, we observed unique human cell-type-specific ACE2 DNA methylation patterns. In human lung tissues, gender differences in DNA methylation at 2 sites related to the ACE2 gene were identified. Further, in freshly isolated airway epithelial cells, DNA methylation near the transcription start site of the ACE2 gene associated with biological age.

**Conclusion:** Epigenetic profiling of host tissue may permit discovery of age and gender-related potential risk factors for COVID-19. How perturbations in ACE2 methylation relate to clinical severity across the ages and gender needs to be determined to guide screening tools and potential epigenetic modification targeting to alleviate COVID-19 morbidity in the elderly.
INTRODUCTION:

An outbreak of COVID-2019 emerged in Wuhan, China in December 2019 as a highly contagious coronavirus [1] capable of human infection and transmission [2,3], and possible mortality following infection similar to the SARS coronavirus [4]. As of March 11, 2020, COVID-19 case fatality rate globally is estimated by the WHO at 3.48% based on 109,578 confirmed cases and 3,809 deaths and this infection has evolved into a pandemic [5]. In the United States (US) as of March 16, 2020 a total of 3,487 cases have been reported with total 68 deaths across 49 states including the District of Columbia as a multipronged surveillance and containment strategy has been initiated across the country with a reported mean incubation period of 6.4 days [6]. Recent analysis of epidemiological data suggests that as many as 86% of all COVID-19 infections are undocumented and are contributing to the rapid worldwide increase in cases [7].

Clinical research has described COVID-19 as an acute respiratory tract infection with varied severity of symptoms including fever onset by dry cough. Clinical treatment of patients infect with COVID-19 has evolving management guidelines of the disease [2,3,8,9]. Epidemiology research has documented the spread and fatality rates for the virus [10], virology research has characterized the genomic features and evolution of COVID-2019 [11–14], and vaccine research has started development and testing of candidate vaccines [15,16]. The global spread of the virus and diversity of human host capable of infection highlight the importance of understanding host biological differences related to COVID-2019. There remain numerous key questions on biological factors that contribute to varying host responses and clinical outcomes. Data from other human
pathogenic coronaviruses SARS-CoV and MERS-CoV[17–19] has provided some insight into potential biological factors underlying varied host responses and clinical outcomes. Sex-based differences in susceptibility to SARS-CoV infection have been reported in mice that parallel those observed in patients and this work identified estrogen receptor signaling as critical for protection in females[20]. Initial reports suggest sex and age-related differences in COVID-2019 that warrant further investigation[17,21]. Research is needed to understand specific sex-related biological features that underlie clinical severity and treatment strategies for COVID-2019 in men and women.

Structural analyses have revealed that receptor angiotensin-converting enzyme 2 (ACE2) as a host receptor permitting cell entry and viral infectivity for COVID-2019[22,23]. Notably, ACE2 was also identified as a host receptor for other coronaviruses including SARS-CoV and NL63[24]. ACE2 has been well studied as a central regulator of blood pressure in the renin-angiotensin-aldosterone system[25]. Studies have shown that ACE2 protein and mRNA expression occurs in a variety of human tissues including lung, liver, stomach, ileum, colon, and kidney[26,27]. Recent single cell analyses of normal human tissues have shown that ACE2 is expressed in cells of the respiratory and digestive system suggesting the lung and gut body compartments as routes for COVID-2019 infection, viral replication, and viral shedding[28–30]. Epigenetics research has suggested that the ACE2 gene that codes for ACE2 may be transcriptionally regulated by DNA methylation[31], which is a covalent chemical modification of host DNA. Moreover, ACE2 is located on the X chromosome[32] raising the possibility of gender differences in susceptibility and progression of COVID-
Thus far, no study has examined the epigenetic landscape of \textit{ACE2} and whether it differs by age or gender.

In this study, we accessed four available genome-wide DNA methylation human datasets to examine whether DNA methylation profiling related to the \textit{ACE2} gene differed by host tissue/cell type, sex, and varied by biological age to begin to understand and address the hypothesis of whether epigenetic footprints related to \textit{ACE2} impact susceptibility risk for COVID-19, disease progression, mortality and morbidity.

\textbf{RESULTS AND DISCUSSION:}

We examined evidence for varied DNA methylation patterns related to the \textit{ACE2} gene in human lung, gut, liver, pancreas, brain, and blood by accessing a subset of available raw genome-wide DNA methylation array data\cite{33} (GEO accession: GSE122126). We analyzed the Illumina MethylationEPIC DNA methylation data using the Chip Analysis Methylation Pipeline\cite{34} and observed that DNA methylation levels at loci related to the \textit{ACE2} gene of various human tissue cell types showed that DNA methylation was varied across tissue cell types (\textbf{Fig. 1}). Notably, DNA methylation across three CpGs (cg04013915, cg08559914, cg03536816) assayed for the \textit{ACE2} gene was lowest in lung epithelial cells compared to the other tissue cell types (\textbf{Fig. 1}), suggesting transcription and expression to be highest in the lung/respiratory system compartment which supports emerging single cell RNA-seq analyses of normal human lung/respiratory system datasets\cite{28–30}. Average total DNA methylation at all probes
related to the ACE2 gene was hypermethylated in cortical neurons and leukocytes compared to other cell types examined, suggesting excluded ACE2 transcription and protein expression in these cell types reported previously [26,27].

![Figure 1. Distinct tissue cell type DNA methylation profile related to ACE2. DNA methylation of 11 CpGs surveyed by the Illumina Infinium MethylationEPIC array related to ACE2 gene from lung epithelial, hepatocytes, colon epithelial, pancreatic duct, pancreatic beta, pancreatic acinar, cortical neurons, and leukocytes is shown from GEO accession: GSE122126. DNA methylation level is shown for each CpG as ranging from 0 to 1.0 (methylation normalized beta-value).](image-url)
of DNA methylation at two CpG sites related to the ACE2 gene showed that females were significantly hypomethylated compared to males \((P = 0.0001)\) (Fig. 2). Additionally, DNA methylation trended towards being hyper-methylation in lung disease conditions including smoking and COPD compared to normal lung although not statistically significant due to limited sample size \((P = 0.18)\). Notably, the ACE2 gene is located on the X chromosome raising the possibility of methylation differences due to X chromosome activation [36]. Of note, the genome-wide Illumina HumanMethylation450 array measures DNA methylation in both X chromosomes of the females. These initial gender-associated differences in DNA methylation related to the ACE2 gene in lung tissues warrant further investigation, especially in the context of ACE2 gene and protein expression and COVID-19 severity. To examine the variability in DNA methylation levels related to ACE2 in males and females, we accessed Illumina HumanMethylation450 DNA methylation data from 244 fresh human lung tissues [37]. We observed that DNA methylation at a CpG in the dataset related to the ACE2 gene showed a large degree of variability in both men and women suggesting DNA methylation of ACE2 varies by individual (Fig. 3). Of note, this dataset did not have metadata for age and the differences in DNA methylation related to ACE2 may be reflective of cell type differences in the lung tissues.

These gender-related differences in ACE2 DNA methylation observed in the respiratory system support findings that indicate that Angiotensin II metabolism varies by gender and may relate to hormonal differences or genetic differences in chromosome dosage [38]. Our DNA methylation data contrast a recent preprint study that reanalyzed five bulk transcriptome datasets of normal lung tissue and two single cell transcriptome
dataset reported no significant differences in ACE2 between racial groups, age groups, or gender groups[30]. A major limitation of this study was the sample size and multiple comparisons as well as the focus only on ACE2 transcription. Notably, the ACE2 transcription preprint study did report higher expression in Asian current smokers compared to non-smokers supporting our observations of a trend in differences in ACE2 in lung of male smokers. Smoking dramatically impacts the epigenome and will be a key environmental factor to examine in future epigenetic studies of COVID-19[39,40]. Additional research will need to determine whether protein, gene transcription, or epigenetic landscape of ACE2 is most relevant to COVID-19 infection risk, disease severity, and transmission.
Figure 2. DNA methylation at two CpG loci related to ACE2 gene in human lung tissues of smokers and patients with chronic obstructive pulmonary disease (COPD) by gender from GEO accession: GSM2430978. DNA methylation level is shown for each CpG as ranging from 0 to 1.0 (methylation normalized beta-value).
Next, based on epigenetics work showing DNA methylation sites associated with longitudinal lung function decline and aging in humans[41] and given reports out China highlighting the increased risk of COVID-19 severity in older individuals[3] due to a myriad of potential biological factors, we sought to test our hypothesis that DNA methylation related to the ACE2 gene associated with age in the respiratory system. Therefore, we obtained genome-wide DNA methylation array data from freshly isolated airway epithelial cells of non-asthmatics of varying biological ages and evaluated whether the methylation levels for CpGs related to the ACE2 gene related to biological age. We found that one CpG (cg08559914) near the transcription start site of the ACE2 gene significantly associated with biological age in airway epithelial cells ($r = -0.59, P = 0.001$; **Fig. 4**). This finding was notable given that the Illumina DNA methylation array data analyzed for ACE2 gene did not measure DNA methylation at every CpG but only 7 CpG sites. Whole genome bisulfite sequencing will provide further insight into whether there are additional DNA methylation sites related to the ACE2 gene that dynamically

![Figure 3. DNA methylation variability related to ACE2 gene in human lung tissues men and women from GEO accession: GSE52401. DNA methylation level is shown line at median with minimum and maximum (methylation normalized beta-value).](image-url)
associate with age. These findings suggest and age-related change in the epigenetic regulation of \textit{ACE2} in the respiratory system exist and are relevant to studies of \textit{ACE2} and COVID-19 in elderly. Our findings complement emerging single cell transcriptional profiling data that is suggesting that nasal goblet/secretory and ciliated cells of the respiratory system show the highest \textit{ACE2} transcription[29] and highlight the need for age-related single transcriptomic and epigenetic profiling studies of the respiratory system in COVID-19. Whether this dramatic relationship between \textit{ACE2} DNA methylation and aging is mediated by other aging-related complications or is a major factor contributing to older individuals being more susceptible or younger individuals being less susceptible to severe COVID-19 infection needs to be studied.

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{figure4.png}
\caption{DNA methylation at a CpG loci in the TSS200 region related to \textit{ACE2} gene in airway epithelial cells harvested via bronchial brushing from non-asthmatic females from GEO accession: GSE85566) significantly associates with biological age (r=-0.59, \(P = 0.001\)).}
\end{figure}
We also examined public available data from ENCODE for RNA-Seq from human Lung donors including a 3 year old male, 30 year old female, and fetal samples\cite{42,43}. Interestingly, high levels of transcription were observed in young male lung compared to fetal lung and female lung (Fig. 5). Moreover, we observed in a publicly available ENCODE Hi-C dataset for IMR90 cells\cite{42} used to examine genome-wide chromatin organization that chromatin interaction contacts occurred near the site we observed DNA methylation differences for the \textit{ACE2} gene, suggesting a potential spatiotemporal gene expression program for \textit{ACE2} mediated by DNA methylation. Additional single cell 3D chromatin research in normal respiratory system and during COVID-19 infection will need to examine whether a promoter-enhancer interaction for \textit{ACE2} exist and modulates differential host and respiratory system cell type expression patterns.
Figure 5. Scaled gene view from WashU Epigenome Browser[43] of ACE2 human gene. ACE2 gene track showed in purple from gencodeV29. RNA-Seq tracks from human lung and fetal lung obtained from ENCODE[42] shown in blue. CpG methylation data of the seven 450K array probes over the ACE2 gene shown as blue lines at each CpG. Hi-C genome-wide chromosome organization data from IMR90 cells shown with view of interaction contacts in purple 20kb resolution.
In this study, we observed DNA methylation levels associated with \textit{ACE2} gene varied by tissue type, differed by gender and disease state in lung, and associated with biological age in airway epithelial cells. Notably, the significant relationship between DNA methylation of \textit{ACE2} and biological age suggest age-dependent differences in host response may be mediated by the dysregulation of \textit{ACE2} and increased expression of \textit{ACE2} with aging. While our DNA methylation findings related to the \textit{ACE2} gene are compelling, \textit{ACE2} is dynamically regulated in the body transcriptionally, posttranscriptionally, and posttranslationally[44]. Further work will need to determine whether other pre-existing health conditions such as diabetes and hypertension or medications being taken such as ACE inhibitors and angiotensin II type-I receptor blockers lead the dysregulation of \textit{ACE2} and increased severe outcomes of COVID-19 infection[45]. This notion has recently received attention and controversy and deserves further study[46]. Additionally, the use of ACE inhibitors as therapeutics to prevent COVID-19 viral entry also warrants further investigation[47].

In summary, our work focused on studying available epigenetic DNA methylation data to begin to characterize and focus on the tissue/cell type specific epigenetic landscape of the COVID-19 host receptor gene \textit{ACE2}. Future work should examine varying host epigenetic landscapes during COVID-19 infection to understand whether COVID-2019 antagonizes antigen presentation through epigenetic modulation similar to MERS-CoV [48]. This information will be vital to vaccine development and epigenetic therapeutics for COVID-19. Another compelling question is whether the DNA
methylation profile of ACE2 in COVID-19 viral replication body sites associates with “super spreaders” and enhanced transmissivity.

METHODS

Public dataset acquisition and processing.

Public DNA methylation datasets were downloaded from the Gene Expression Omnibus website. DNA methylation data for human isolated cell populations from tissues was obtained from (GEO: GSE122126). DNA methylation data for airway epithelial cells used to examine the relationship with age was obtained from (GEO: GSE85568). DNA methylation data for lungs tissues of smokers and patients with COPD by gender was obtained from (GEO: GSE92511). DNA methylation data for normal lungs tissues by gender was obtained from (GEO: GSE52401). RAW IDAT files were loaded into ChAMP pipeline, preprocessed, and normalized[34]. We loaded RNA-Seq and Hi-C datasets from the ENCODE portal[49] (https://www.encodeproject.org/) into the WashU Epigenome Browser. Statistical tests were conducted in Graphpad Prism 8 for DNA methylation data related to the ACE2 gene.
CONFLICT OF INTEREST

Authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We would like to acknowledge the authors that deposited raw DNA methylation data for analysis relevant to ACE2. Public access to all formats of data relevant to COVID-19 will aid in overcoming the pandemic.

REFERENCES


45. Fang L, Karakiulakis G, Roth M. Are patients with hypertension and diabetes mellitus at increased risk for COVID-19 infection? Lancet Respir Med [Internet]. 2020; Available from: http://dx.doi.org/10.1016/S2213-2600(20)30116-8

47. Gurwitz D. Angiotensin receptor blockers as tentative SARS-CoV-2 therapeutics. Drug Dev Res [Internet]. 2020; Available from: http://dx.doi.org/10.1002/ddr.21656
