Is Low Alveolar Type II Cell SOD3 in the Lungs of Elderly Linked to the Observed Severity of COVID-19?

Ahmed S Abouhashem†, Kanhaiya Singh1**, Hassan ME Azzazy2, Chandan K Sen1

† These authors contributed equally to this work

1 Indiana Center for Regenerative Medicine and Engineering, Indiana University Health Comprehensive Wound Center, Indiana University School of Medicine, Indianapolis, USA

2 Department of Chemistry, School of Sciences & Engineering, The American University in Cairo, New Cairo, Egypt

*Corresponding Author:

Kanhaiya Singh, PhD
Indiana Center for Regenerative Medicine & Engineering, Department of Surgery, Indiana University School of Medicine, Indianapolis, IN 46202
Tel.: +1 317-278-3411
E-mail: kanh@iu.edu

Abbreviated title: Low SOD3 in the Lungs of Elderly

Word count: 1404

Reference number: 30

Number of color illustrations: online 2 and hardcopy 2

KEYWORDS: COVID19, single cell RNA sequencing, Lung, ROS, SOD3, ATF4
Abstract

Human lungs single cell RNA sequencing data from healthy donors (elderly and young; GEO accession number GSE122960) were analyzed to isolate and specifically study gene expression in alveolar type II cells. Co-localization of ACE2 and TMPRSS2 enables SARS-CoV 2 to enter the cells. Expression of these genes in the alveolar type II cells of elderly and young patients were comparable and therefore do not seem to be responsible for worse outcomes observed in COVID-19 affected elderly. In cells from the elderly, 263 genes were downregulated and 95 upregulated. SOD3 was identified as the top-ranked gene that was most down-regulated in the elderly. Other redox-active genes that were also downregulated in cells from the elderly included ATF4 and M2TA. ATF4, an ER stress sensor that defends lungs via induction of heme oxygenase 1. The study of downstream factors known to be induced by ATF4, according to Ingenuity Pathway Analysis™, identified 24 candidates. Twenty-one of these were significantly downregulated in the cells from the elderly. These downregulated candidates were subjected to enrichment using the Reactome Database identifying that in the elderly, the ability to respond to heme deficiency and the ATF4-dependent ability to respond to endoplasmic reticulum stress is significantly compromised. SOD3-based therapeutic strategies have provided beneficial results in treating lung disorders including fibrosis. The findings of this work propose the hypotheses that lung-specific delivery of SOD3/ATF4 related antioxidants may work in synergy with promising anti-viral drugs such as remdesivir to further improve COVID-19 outcomes in the elderly.
Introduction

The human lung alveolar epithelium is mainly composed of types I/II alveolar cells and macrophages. In contrast with alveolar cells type I, alveolar cells type II are capable of giving rise to both type II and type I alveolar cells (18). Alveolar type II cells serve many critical functions including production of pulmonary surfactant, stabilization of airway epithelial barrier, local immune defense, and airway regeneration following injury. As one of few cells in the human body that co-expresses angiotensin-converting enzyme 2 (ACE2) receptor and the TMPRSS2 protease, required for attachment and cellular entry of COVID-19, alveolar type II cells are readily targeted by coronavirus 2 (SARS-CoV 2) (8). Coronavirus disease 2019 (COVID-19) is an emerging respiratory disease caused by SARS-CoV 2. Most COVID-19+ patients exhibit mild to moderate symptoms, with ~10% developing acute respiratory distress syndrome (ARDS) which is the leading cause of mortality among these patients (9). Histopathologic changes of the lungs related to COVID-19 include diffuse alveolar damage, chronic inflammatory infiltrates and intra-alveolar fibrinous exudates (30). The severity of symptoms and mortality of elderly patients are higher than those of younger patients (14).

Reactive oxygen species (ROS) and reactive nitrogen species (RNS) are known to be produced by cells of the innate immune system and others in response to viral infection (11). ROS/RNS are directly implicated in lung fibrosis and declining lung function (5). Antioxidant enzymes such lecithinized superoxide dismutase (SOD) have proven to be useful in patients suffering from lung fibrosis (25). There is substantial literature demonstrating a causative role of ROS/RNS in the development of lung fibrosis (21). In virus-induced lung disease, antioxidant treatment attenuated lung inflammation and airway hyper-reactivity (3). In this work, we utilized single-cell sequencing data from elderly and younger humans to specifically study alveolar type II cells to identify genes that are most profoundly affected as a function of age. Such effort was intended at developing novel hypothesis to understand why lungs of the elderly are more severely affected in COVID-19.
Results

Human lungs single cell RNA sequencing (scRNA-seq) data from four healthy donors (9,783 cells from old-age group, 57 and 63 years old, and 8,501 cells from young-age group, 22 and 29 years old; GEO accession number GSE122960) were analyzed (20). Analysis of scRNA-seq data generated thirteen clusters visualized using t-Distributed Stochastic Neighbor Embedding (t-SNE). Such clustering was based on a non-linear dimensionality reduction technique for embedding high dimensional data with the objective of visualization in low-dimensional space. As depicted in Figure 1A, these clusters include epithelial cells (alveolar type II cells), macrophages in different states, epithelial cells (alveolar type I cells), monocytes, B cells, T cells/ NK cells, endothelial cells and stem cells. SingleR package in R (1), in conjunction with the Human Primary Cell Atlas (HPCA) dataset, was used to identify alveolar type II cells from the total lung cell population. These cells heavily expressed the characteristic surfactant protein C gene (Sup Fig 1A, B). Genes of this particular alveolar type II cell cluster were analyzed for differential expression as a function of aging (Figure 1B). Co-localization of ACE2 and TMPRSS2 enables SARS-CoV 2 to enter the cells(8). Expression of these genes in the alveolar type II cells of old and young patients were comparable (Figure 1C) and therefore do not seem to be responsible for worse outcomes in COVID-19 affected elderly. Compared to that in alveolar type II cells from younger donors, in cells of older donors 263 genes were differentially downregulated and 95 genes were upregulated (adjusted $p$-value $<0.05$ and 10% log fold change; Figure 1D, Sup Table 1). SOD3 was identified as the top-ranked (by log of fold-change) gene that was most down-regulated in alveolar type II cells (Sup Table 1). This observation piqued our interest in other genes with known redox functions. Other genes with known redox-based functions that were also downregulated in alveolar type II cells of elderly lung donors included Activating Transcription Factor 4 ($ATF4$) and Metallothionein 2A ($M2TA$) (Figure 2A-C). Additional studies to look for SOD3-interacting genes, as predicted by String database (String version 11.0), recognized the following: SOD2, Catalase ($CAT$), Glutathione Peroxidase 1 ($GPX1$), $GPX2$, $GPX3$, $GPX5$, $GPX7$, $GPX8$, Antioxidant 1 Copper Chaperone ($ATOX1$), and ATPase Copper Transporting Alpha ($ATP7A$) (Sup Fig 2A, B). Among these candidates, $GPX1$ was the only gene that was differentially low in expression in cells from...
elderly donors (Sup Fig 2B, Sup Table 1). Viral infection is known to employ ER stress to cause lung fibrosis(12). ATF4, an ER stress sensor that can defend lungs via induction of heme oxygenase 1, was downregulated in alveolar type II cells of the elderly (Figure 2B). Study of downstream factors that are known to be induced by ATF4, according to Ingenuity Pathway Analysis™, identified 24 candidates. Twenty-one of these 24 were significantly downregulated in the alveolar type II cells of the elderly (Figure 2C). The downregulated candidates were subjected to pathway enrichment using the Reactome Database. These analyses identified that in the elderly, the ability to respond to heme deficiency and the ATF4-dependent ability to respond to endoplasmic reticulum stress is significantly compromised (Figure 2D, Table 1).

Discussion

At the time of this reporting, of 13,130 COVID-19 deaths (of reported death from all causes 582,565 on April 19, 2020) in the United States as reported by the Center for Disease Control, 13,001 decedents were of age 35 or above representing 99% of all COVID-19 deaths. Decedents aged 55 years or above account for 91% of all COVID-19 deaths (19). The younger donors of this study were both below 30 while the elderly donors were both above 55 (19). Briefly, the SARS-CoV 2 virion contains four proteins: spike, envelope, membrane, and nucleocapsid; and a single-stranded RNA. The virion binds to ACE2 receptor located on alveolar type II cells making these cells a target for infection. Once SARS-CoV 2 has attached to these receptors, the TMPRSS2 protease cleaves the spike protein to expose a fusion peptide which helps the virus enter the cell (8). This enables the virions to release their RNA into infected cells. Co-expression of ACE2 and TMPRSS2 on type II cells makes them a preferred site for SARS-CoV 2 attachment, entry and replication (16). In late stages of the severe form of COVID-19, cytokine storm has been evident (17). Under these conditions, when the inflammatory system has gone awry in COVID-19 unrelated pathologies, there is ample evidence of oxidative stress (15). Findings of this work, involving an unbiased query of differentially expressed genes specifically in alveolar type II cells of the human lung, point towards specific elements of the antioxidant defense system that are weakened as a function of age. SOD3-deficient mice develop severe lung damage in the presence of normal oxygen tension along with marked inflammatory cell
infiltration and alveolar hemorrhage (6). In the lungs, extracellular SOD3 is primarily expressed in bronchial and alveolar type II epithelial cells, alveolar macrophages, and pulmonary endothelial cells (4). Importantly, SOD3 mRNA expression in alveolar cells correlates with locally secreted enzyme activity defending functional significance of observed changes in gene expression (27). With SOD3 as the top-ranked candidate, this work identifies specific component of the alveolar cell antioxidant defense systems which are weakened in response to aging. Evidence in the current literature shows that SOD administration can decrease the severity of respiratory illness (7). Intravenous SOD administration in rabbit models reversed allergic emphysema (2). SOD-based therapies have also shown encouraging results in managing infectious diseases by improving host immune responses. Melon SOD restored CD4+/CD8+ ratio in cats infected with feline immunodeficiency virus (28). Cu/Zn-SOD significantly inhibited HIV replication (13). Cu/Zn-SOD inhalation protected murine lungs against pulmonary emphysema by decreasing ROS levels and pro-inflammatory cytokines expression (26). A limited-scale prospective double-blinded controlled trial NCT04323228 is about to start recruiting to test the effect of an antioxidant oral nutrition supplement in SARS-CoV 2 positive cases. All of these above-mentioned studies are preliminary at best and may help lay the foundation for a more serious effort testing whether lung-specific delivery of SOD3 related antioxidants may work in synergy with promising anti-virals such as remdesivir (NCT04323761) to further improve COVID-19 outcomes in the elderly.

Innovation

For the first time, single cell RNA sequencing data of the human lungs have been studied as a function of age to reveal that weakening of specific components of the antioxidant defense system of the alveolar type II cells from elderly donors should be tested for a mechanistic connection of COVID-19 severity outcomes.

Notes

Data acquisition. Primary single cell RNA sequencing data were obtained from the GEO database (http://www.ncbi.nlm.nih.gov/geo; accession number GSE122960). Authors of the published study studying lung fibrosis performed single cell RNA sequencing on eight
healthy donor lungs and nine lungs from patients with pulmonary fibrosis (20). Four samples were chosen from the healthy donor lungs in which two of them were 57 and 63 years old and the other two samples were from donors who were 22 and 29 years old. All four donors were females, non-smokers and African American.

**Processing of raw data and quality control.** All analyses were performed using Seurat package (v.3.1.1) in R (v.3.3.5) (22). The initial dataset contained 20,163 cells from four lung samples (2 young and 2 elderly). Gene expression values were log normalized using 10,000 transcripts per cell as scaling factor. Canonical correlational analysis was performed to integrate all four samples to identify the shared cell types using the top 2000 highly variable genes. Cells expressing less than 200 or more than 5000 genes, as detected, were excluded. Further filtering was performed to exclude cells that contain more than 15% of their reads from mitochondrial encoded genes. Cells with total number of counts between 2000 and 25000 were kept for downstream analysis. After the filtration step, 18,284 cells (9,783 from elderly and 8,501 from young donors) were maintained for downstream analysis. Principal component analyses were performed and the first 15 principal components were chosen for clustering.

**Determination of cell-type identity.** To identify cluster identity, SingleR package in R was used to calculate the similarity between each cluster and the Human Primary Cell Atlas (HPCA) dataset(1). In addition, differential gene expression was performed between each cluster and the rest of the cells to identify cluster markers and further identification of the cluster identity using known markers from the literature.

**Differential gene expression analysis** was performed for alveolar type II cells (cluster zero) to compare between young and old-age group using Wilcoxon-Rank Sum test with adjusted p value less than 0.05. To avoid detection of genes only altered in one sample of either group, four additional comparisons were performed (each sample with the 2 samples from the other group). The genes which were not simultaneously upregulated or downregulated and in at least three such comparisons were excluded.

**Ingenuity upstream regulator analysis in Ingenuity Pathway Analysis.** IPA (23,24,29) was used to identify the cascade of upstream transcriptional regulators that can explain the
observed gene expression changes in the dataset. This approach is based on prior knowledge of expected effects between transcriptional regulators and their target genes. Two statistical measures (an overlap p-value and an activation z-score) were computed for each potential transcriptional regulator. The activation z-score was used to infer likely activation states of upstream regulators based on comparison with a model that assigns random regulation directions.

**Pathways enrichment.** Reactome database (10) was used for enrichment of ATF4 downstream targets which were found to be downregulated in alveolar type II cells in the cells from the elderly.

**Authors’ Contributions**


**Author Disclosure Statement**

No competing financial interests exist.

**Supplementary Material**

Supplementary Figure 1

Supplementary Figure 2

Supplementary Table 1
### Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARDS</td>
<td>acute respiratory distress syndrome</td>
</tr>
<tr>
<td>COVID-19</td>
<td>Coronavirus disease 2019</td>
</tr>
<tr>
<td>ROS</td>
<td>reactive oxygen species</td>
</tr>
<tr>
<td>RNS</td>
<td>reactive nitrogen species</td>
</tr>
<tr>
<td>SARS-CoV 2</td>
<td>severe acute respiratory syndrome coronavirus 2</td>
</tr>
<tr>
<td>SOD</td>
<td>superoxide dismutase</td>
</tr>
<tr>
<td>ATF4</td>
<td>activating transcription factor 4</td>
</tr>
</tbody>
</table>
References


Figure Legends:

Figure 1: Identification of thirteen distinct clusters within the lung tissue with unique markers for each. (A) t-distributed stochastic neighbor embedding (t-SNE) projection of the filtered data (18,284 cells; 8501 young; 9783 old). Each cell is represented as dot. (B) tSNE clustering of the epithelial, alveolar type II cells, cells showing clear separation between the cells from young and old group. (C) Expression level of ACE2 (top) and TMPRSS2 (bottom) in alveolar type II cells. (D) Heatmap of the top differentially expressed genes between young and old lung alveolar type II cells (log fold change ± 0.4). Rows represent genes and columns represent cells.
Figure 2. Downregulation of SOD3 and ATF4 in alveolar type II cells of lungs from elderly donors (A) tSNE plots showing SOD3 and (B) ATF4 expression level between young and old age-group in (top) all the 13 clusters and (bottom) in alveolar type II cells. (C) Heatmap of the downstream targets of ATF4. Rows represent genes and columns represent cells. (D) Pathways enrichment for ATF4 downstream targets which are downregulated in elderly. Circles represent genes and diamonds represent the enriched terms.
Table 1. Reactome pathways enrichment for the downregulated genes among ATF4 downstream targets in alveolar type II cells from elderly. Pathways with adjusted p-value <2E-6 are shown.

<table>
<thead>
<tr>
<th>Pathway name</th>
<th>Reactions found/ Total reactions</th>
<th>FDR adjusted p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Response of <em>EIF2AK1</em> (HRI) to heme deficiency</td>
<td>18/ 20</td>
<td>1.41E-14</td>
</tr>
<tr>
<td><em>PERK</em> regulates gene expression</td>
<td>9/ 11</td>
<td>8.72E-12</td>
</tr>
<tr>
<td>Response of <em>EIF2AK4</em> (GCN2) to amino acid deficiency</td>
<td>13/ 16</td>
<td>4.01E-10</td>
</tr>
<tr>
<td><em>ATF4</em> activates genes in response to endoplasmic reticulum stress</td>
<td>7/ 7</td>
<td>9.39E-09</td>
</tr>
<tr>
<td>Cellular responses to external stimuli</td>
<td>33/ 258</td>
<td>1.47E-08</td>
</tr>
<tr>
<td>Unfolded Protein Response (<em>UPR</em>)</td>
<td>10/ 94</td>
<td>8.23E-08</td>
</tr>
<tr>
<td>Cellular responses to stress</td>
<td>27/ 227</td>
<td>1.03E-06</td>
</tr>
</tbody>
</table>

FDR: false discovery rate

*EIF2AK1*: Eukaryotic Translation Initiation Factor 2 Alpha Kinase 4

*HRI*: heme regulated inhibitor

*GCN2*: General Control Nonderepressible 2