

# The Effect of Air Pollution on the Transcriptomics of the Immune Response to Respiratory Infection

**Daniel Croft** (✉ [daniel\\_croft@urmc.rochester.edu](mailto:daniel_croft@urmc.rochester.edu))

University of Rochester Medical Center

**David Burton**

University of Rochester

**David Nagel**

University of Rochester Medical Center

**Soumyaroop Bhattacharya**

University of Rochester Medical Center

**Ann Falsey**

University of Rochester Medical Center

**Steve Georas**

University of Rochester Medical Center

**Philip Hopke**

University of Rochester

**Carl Johnston**

University of Rochester Medical Center

**R. Matthew Kottmann**

University of Rochester Medical Center

**Augusto Litonjua**

University of Rochester Medical Center

**Thomas Mariani**

University of Rochester Medical Center

**David Rich**

University of Rochester

**Kelly Thevenet-Morrison**

University of Rochester

**Sally Thurston**

University of Rochester

**Mark Utell**

University of Rochester Medical Center

**Matthew McCall**

University of Rochester

---

## Research Article

**Keywords:** respiratory infection, air pollution, immune response

**Posted Date:** April 23rd, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-440415/v1>

**License:**   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

**The Effect of Air Pollution on the Transcriptomics of the  
Immune Response to Respiratory Infection**

\*Daniel P. Croft<sup>1,2</sup>, David S. Burton<sup>3</sup>, David J. Nagel<sup>1,2</sup>, Soumyaroop Bhattacharya<sup>2,4</sup>, Ann R. Falsey<sup>5</sup>, Steve N. Georas<sup>1,2</sup>, Philip K. Hopke<sup>2,6,7</sup>, Carl J. Johnston<sup>2,4</sup>, R. Matthew Kottmann<sup>1,2</sup>, Augusto A. Litonjua<sup>1,2</sup>, Thomas J. Mariani<sup>2,4,8</sup>, David Q. Rich<sup>1,2,4,6</sup>, Kelly Thevenet-Morrison<sup>6</sup>, Sally W. Thurston<sup>2,3</sup>, Mark J. Utell<sup>1,2</sup>, Matthew N. McCall<sup>2,3</sup>

1. Department of Medicine, Pulmonary and Critical Care Medicine Division. University of Rochester Medical Center, Rochester, New York

2. Environmental Health Science Center, University of Rochester Medical Center, Rochester, New York

3. Department of Biostatistics and Computational Biology, University of Rochester Medical Center, Rochester, New York

4. Department of Pediatrics, University of Rochester Medical Center, Rochester, New York

5. Department of Medicine, Infectious Diseases Division. University of Rochester Medical Center, Rochester, New York

6. Department of Public Health Sciences, University of Rochester Medical Center, Rochester, New York

7. Institute for a Sustainable Environment, and Center for Air Resources Engineering and Science, Clarkson University, Potsdam, New York

8. Department of Biomedical Genetics. University of Rochester Medical Center, Rochester, New York

**\*Corresponding author:**

Daniel P. Croft MD, MPH  
University of Rochester Medical Center  
Pulmonary and Critical Care Division  
601 Elmwood Avenue Box 692  
Rochester, NY 14642  
Phone: 585 275 4161  
Fax: 585 271 1171  
Email: [daniel\\_croft@urmc.rochester.edu](mailto:daniel_croft@urmc.rochester.edu)

## **Abstract**

Combustion related particulate matter air pollution (PM) is associated with an increased risk of respiratory infections in adults. The exact mechanism underlying this association has not been determined. We hypothesized that increased concentrations of combustion related PM would result in dysregulation of the innate immune system. This epidemiological study includes 111 adult patients hospitalized with respiratory infections who underwent transcriptional analysis of their peripheral blood. We examined the association between gene expression at the time of hospitalization and ambient measurements of particulate air pollutants in the 28 days prior to hospitalization. For each pollutant and time lag, gene-specific linear models adjusting for infection type were fit using LIMMA (Linear Models For Microarray Data), and pathway /gene set analyses were performed using the CAMERA (Correlation Adjusted Mean Rank) program. Comparing patients with viral and/or bacterial infection, the expression patterns associated with air pollution exposure differed. Adjusting for the type of infection, increased concentrations of Delta-C (a marker of biomass smoke) and other PM were associated with upregulation of iron homeostasis and protein folding. Increased concentrations of black carbon (BC) were associated with upregulation of viral related pathways and downregulation of pathways related to antigen presentation. The pollutant/pathway associations differed by lag time and by type of infection. This study suggests that the effect of air pollution on the pathogenesis of respiratory infection may be pollutant, timing, and infection specific.

## **Background**

Respiratory infections are a leading cause of adult morbidity and mortality in the United States [1]. Short-term increases in ambient fine particulate air pollution ( $\leq 2.5\mu\text{m}$  in diameter;  $\text{PM}_{2.5}$ ) concentrations have been associated with increased emergency department (ED) visits for influenza and culture negative pneumonia in adults [2]. Studies in the USA have also observed associations between acute increases in air pollution concentrations and increased risks of a healthcare encounter for respiratory viral infection (RVI) or other lower respiratory tract infections (LRTI), including respiratory bacterial infection (RBI), in adults [3-6]. With over 90% of the world breathing unhealthy air [7], and large health care burden of respiratory infections [8], understanding the mechanisms underlying air pollution effects on the innate immune response to respiratory infections is crucial.

The association between the inflammation caused by air pollution and disruption of the lung's innate immune system, including epithelial barrier disruption, macrophage function, and protein/cytokine response, is typically studied via in vitro human cell and in vivo rodent models [9-19]. Although the majority of studies suggest that different air pollution components might predispose patients to RVI or RBI by multiple mechanisms, laboratory-based studies are generally limited to inbred mouse strains and/or only a single pollutant or a single category of pollution (traffic pollution), and do not fully recapitulate complex events occurring in naturally exposed humans. Studying natural exposures also helps characterize pollutant effects that controlled exposure studies may miss. Epidemiologic studies allow examination of health responses to single pollutants as they occur within a pollutant mixture.

By examining the association between multiple pollutants and the transcriptional profiles of the peripheral blood of patients with respiratory infection we can better understand what specific gene pathways may be driving the relationship between specific particulate pollution and different types of infection including RVI and RBI. Studying respiratory infection types in aggregate can give important insight into potential shared mechanisms between the effect of air pollution on respiratory infection in general while studying infections individually may elucidate infections specific mechanisms applicable to only RVI or RBI. Furthermore, studying the potential mechanisms of the effects of air pollution on respiratory infection may help design potential risk reduction strategies (e.g. anti-inflammatory treatment) for patients exposed to air pollution during periods of high rates of respiratory infection.

Although most literature focuses on  $\text{PM}_{2.5}$  and not its constituents, our prior source-specific study in NY State observed associations between combustion related constituents of  $\text{PM}_{2.5}$  and ED visits for influenza [20]. Among the  $\text{PM}_{2.5}$  constituent sources, we hypothesized that ambient combustion sources including traffic related air pollution (black carbon) and biomass burning (Delta-C) would be associated with dysregulation of key gene pathways within the immune response to respiratory infection. In this

study, we focused on the gene expression patterns in patients with RVI as the primary outcome of interest associated with air pollution. We also performed the same analysis in patients with RBI and combined viral and bacterial infection. First, we highlighted the patterns in gene expression across these three infection groups. Next, we analyzed the associations between air pollution concentrations and gene expression in all types of respiratory infection combined. We then analyzed the gene pathways associated with increases in air pollution concentrations in the aggregate group and also visualized patterns of gene expression in key pathways for RVI, RBI and combined infection individually (**Fig. 1**). We report several associations between gene expression and air pollution exposure including several gene pathways relevant to respiratory infections, including immunity, protein folding and iron homeostasis.

## **Results**

### **Patient Characteristics**

The majority of participants were white (76%), female (58%), and older, with median age range of 58-61 years old (**Table 1**). The majority of participants had viral infection (n=66) and fewer patients had bacterial infection (n=20) or coinfection (n=25). The most common viruses included influenza (n=22), respiratory syncytial virus (n=19) and the most common bacterial infection was Streptococcus pneumonia (n=12) (**Supplemental Table S1 online**). Over 90% of patients lived at home and the most common comorbidities included diabetes (34%), COPD (38%) and obesity (39%). Forty percent of participants were active smokers, with the lowest proportion of smokers in the viral infection group (35%) and the largest proportion of smokers in the co-infection group (52%). Inhaled steroid use was most prevalent within the bacterial infection group (65%) and least prevalent in the viral infection group (30%). Home oxygen use was most prevalent in the bacterial infection group (30%) and less common in the viral (15%) and combined infection groups 12%. Air pollution distributions for Rochester, NY during the same season as the hospitalizations are reported in **Table 2**. Although Rochester, NY is a medium-sized city, the average PM<sub>2.5</sub> concentration is well below the EPA 24- hour fine particle standard of 35 µg/m<sup>3</sup>[21].

### **Overview of the highest variance genes, divided by infection**

In the unsupervised analysis, we observed infection specific differences in the expression of 150 high variance genes (**Fig. 2**). Based on the hierarchical clustering of genes, we defined 7 distinct gene clusters. One cluster (#2) showed clear expression differences between infections, as previously shown by Suarez et al. [22]. Of particular interest to our study, increased expression of genes in this cluster corresponded to high pollution levels among the viral infection samples. The expression patterns in patients with bacterial or combined infection did not clearly correspond to high pollution levels.

Within a specific type of respiratory infection, the expression levels were not consistent across all high variance genes. We observed a higher gene expression corresponding to the highest air pollution concentrations in Cluster #2 for viral infection compared with Clusters #3-6. While other gene clusters captured consistent expression patterns across subjects, these differences were not clearly associated with either infection or pollutant levels.

In the heatmap of the 150 most differentially expressed genes (**Fig. 2**), the 38 individual genes within Cluster #2, had the highest gene expression within the viral infection group and included a variety of immune related genes (**Supplemental Table S2 online**). For example, there were multiple interferon induced protein coding genes including IFI44L, OAS1 and MX1. IFI27, IFIT (1-3) and RSAD2 participate in Interferon gamma signaling and the innate immune response. OAS2 also participates in the innate immune response by encoding for the 2-5A synthetase family. The HERC5 gene is upregulated by endothelial cell inflammation. This result suggests exposure to particulate air pollutants in the week prior to hospitalization with RVI may be associated with a more exuberant immune and inflammatory response.

### **Individual gene analysis**

To quantify the association between expression of individual genes and air pollution measurements prior to hospitalization for individuals with a respiratory infection, we performed a LIMMA analysis controlling for sex.

#### ***Differential expression of viral versus bacterial infection (not factoring in air pollutants)***

The top genes in the individual LIMMA analyses that were differentially expressed in peripheral blood samples of patients hospitalized with viral infection vs. bacterial (baseline in the model) infection were ISG15, TIMM10, IFI27, IFI44L and OAS2 (**Supplemental Table S3 online**). These genes were consistently differentially expressed regardless of which pollutant was included in the model. Specifically, in the 0-6 lag period for DC, there was a log<sub>2</sub>-fold increase of 3.3 for ISG15 and 2.2 for TIMM10, two genes participating in antigen presenting cell pathways. IFI 27 and IFI44L, both part of the interferon induced response to RVI, were each observed to have a log<sub>2</sub>-fold increase of around 4. Finally, OAS-2, a part of the innate immune system, was observed to have a 2.5 log<sub>2</sub>-fold increase. With the exception of increased expression of IFI27, no genes showed a statistically significant difference in expression between the co-infection and bacterial infection groups.

### **Air pollution versus individual genes**

We then analyzed the association between air pollution and gene expression in the overall study population including all infection types. The average effect of each 1  $\mu\text{g}/\text{m}^3$  increase in DC in the prior 7 days, regardless of infection type, is a log<sub>2</sub>-fold increase of 3.5 for RNF14 and 2.4 for UBE2F, two genes participating in antigen presenting cell pathways (**Supplemental Table S4 online**). Other notable genes upregulated in association with increased concentrations of DC at the 7 day lag period include a 3.1 log<sub>2</sub>-fold change increase in MAP2K3 (TLR signaling) and 2.9 log<sub>2</sub>-fold change increase in ADORA1 (pro-inflammatory monocyte activation). Two genes involved in hemoglobin synthesis (HMBS) and blood cell size (TMCC2) were also upregulated in association with increased DC concentrations in the seven days prior to hospitalization. No significant changes in genes were associated with increased concentrations of PM<sub>2.5</sub>, BC, UFP or AMP. To further investigate these putative associations between gene expression and air pollution, as well as others not readily apparent from the unsupervised and individual gene analysis, we completed a pathway analysis using the differentially expressed genes group by clusters.

### **Pathways analysis**

With a signal appearing between high concentrations of air pollution and gene expression within Fig. 2, we then identified specific gene pathways that were associated with modelled increases in air pollution concentrations across all infections (**Table 3 and Supplemental Tables S5-S9**). When analyzing the association between air pollutants and gene pathways using the CAMERA method, we observed several broad pathways with implications on innate immunity. Then, given the differences in gene expression across infections, we constructed heatmaps of specific pathways of interest for infection specific changes at multiple pollution lags (**Supplemental Fig. S1-11 Online**).

**Delta-C (DC):** Multiple different gene pathways associated with red blood cell synthesis and iron handling were significantly upregulated in association with increases in DC concentrations at all lag periods (**Table 3 and Supplemental Table S5 online**). Gene pathways associated with endoplasmic reticulum activity (protein folding essential for cytokine production in the immune system) were also upregulated by increased DC concentrations at the 7-13, 14-20, 21-28 lag periods while increased concentrations of DC at the 7-13 and 14-20 lag periods were associated with upregulation of a viral replication pathway.

In the infection specific heatmaps, DC was observed to have different patterns of expression in the iron homeostasis pathway (Protoporphyrinogen IX Biosynthetic Process) when comparing infections at the 0-6 and 21-27 lag times (**Supplemental Fig. S1-6 online**). While the highest gene expression in iron homeostasis was observed in patients with bacterial or viral infection at the highest concentrations of DC,

the combined infection group displayed a mixed to decreased expression. There was no clear pattern of gene expression along the pollution concentration gradient when comparing infection type within an endoplasmic reticulum or viral gene expression pathway.

**Black carbon (BC):** Increased concentrations of BC were associated with upregulation of a combination of anti- viral pathways and pro-viral pathways in the 0-13 day lag periods (**Table 3 and Supplemental Table S6 online**). In the 14-20 day lag period, several immune pathways were significantly down regulated in association with a one-unit increase in BC concentrations. Differences in gene expression of the type 1 interferon pathway were observed between infection types for BC (**Fig. 3**). For BC at the 0-6 lag period, the highest gene expression was found in Cluster #5 of genes for patients with RVI and the lowest expression was observed in this cluster for patients with respiratory bacterial infection (RBI). Patients with RVI appeared to be driving the positive association between BC and the type 1 interferon pathway.

**PM<sub>2.5</sub>:** In the 7-20 day lag periods, multiple ribosomal/ER related pathways were downregulated. Iron binding pathways were upregulated in the 20-27 day lag period and iron homeostasis pathways were also upregulated at all lag times (**Table 3 and Supplemental Table S7 online**).

For PM<sub>2.5</sub> at the 0-6 and 14-20 day lag periods, the highest expression of iron homeostasis genes were observed within patients with bacterial infection (**Supplemental Fig. S7-8 online**). At the 14-20 lag days for PM<sub>2.5</sub>, patients with RVI had the lowest gene expression within the iron homeostasis pathways. In patients with combined infection, low levels of gene expression in iron homeostasis was also observed at the 0-6 lag time for the highest concentrations of PM<sub>2.5</sub>. These findings indicate that the patients with RBI were driving the signal in the combined analysis of all infections.

**Ultrafine particle (UFP):** A one-unit increase in UFP was associated with down regulation of several immune related pathways at the 7-13 day lag period, including neutrophil and myeloid activation and regulation of leukocyte degranulation (**Table 3 and Supplemental Table S8 online**). In the 0-13 day lag period, increased concentrations of UFP were associated with upregulation of both iron binding and heme related iron pathways. Finally, we observed upregulation of ER related pathways in the 14-27 lag time associated with a one unit increase in UFP concentrations.

In the infection specific heatmaps for UFP at the 14-20 and 21-27 lag time, we observed the highest expression of viral gene expression pathway genes in patients with viral infection, but there was no clear pattern within bacterial or combined infection (**Supplemental Fig. S9-10 online**). While the infection specific heatmaps showed similar expression pattern for iron homeostasis across different infections,

there were distinct expression patterns for immune specific (highest expression for viral infection) and protein folding pathways (highest for combined infection) when comparing infections.

***Accumulation mode particles (AMP):*** One unit increases in AMP were associated with upregulation of multiple ER associated pathways in the 0-6 lag period and upregulation of iron homeostasis related pathways at the 21-27 day lag period. An increased concentration of AMP was associated with downregulation of a myeloid leukocyte pathway at the 0-6 lag time (**Table 3 and Supplemental Table S9 online**).

In the infection specific heatmaps for AMP at the 0-6 lag period, we observed the highest level of expression in the mRNA catabolic pathway in patients with bacterial infection and lowest levels in patients with viral infection (**Supplemental Fig. S11 online**).

#### **Summary Comparison of expression patterns across pollutants:**

The pathways that were associated with all pollutants included genes related to the upregulation of endoplasmic reticulum related pathways (except for downregulation seen with PM<sub>2.5</sub>), including ribosomal synthesis (**Table 3**). Iron binding and hemoglobin related pathways were also commonly affected by all pollutants except for BC. Out of all the pollutants, BC and UFP were associated with the greatest number of immune and viral related gene pathways.

The most common discordance in expression was observed between viral and bacterial infection. Combined infection commonly displayed a mixed (indeterminant) pattern. The strongest gene expression differences when comparing between infections occurred with BC at the 0-6 lag period for Type 1 interferon (Figure 3). These differences associated with BC displayed decreased expression in the patients with bacterial infection and higher expression in patients with viral infection. One of the most consistent patterns of discordance across multiple pollutants was observed within an iron homeostasis pathway. The highest levels of pollutants including AMP, DC and UFP corresponded to higher expression of the middle to lower clusters of genes at the 0-6 and 21-27 lag periods for both viral and bacterial infection, but the relationship was inconsistent in the combined infection group. In the infection specific analysis, there were also multiple examples of low gene expression, which appeared to be independent of air pollution levels.

#### **Individual genes within the Interferon pathway.**

Given the contrast between gene expression corresponding to the highest pollutant concentrations between the types of infection within Cluster #5 of the Response to Type 1 Interferon pathway (**Fig. 3**), we examined the 18 individual genes presents within this cluster (**Supplemental Table S10 online**). The

majority of the genes in this cluster were related to the immune response to infection including OAS-1/3, IFI44L, HLA-A and HLA-DRB-1/4. Genes related to ribosomal activity (RPS4Y-1) and heme biosynthesis (ALAS2) were also present in this cluster. This review of individual genes highlights the overlap between infection related pathways (Interferon) and the ribosomal and heme related genes. OAS2 and IFI44L were also found to be expressed in a high level in patients with RVI in the analysis of 150 highest variance genes (**Fig. 2**). The gene expression in the other clusters in the heatmap did not appear to correspond to pollution levels when comparing across infection types. The genes in Cluster #2, for example, contains a variety of genes related to acute phase reaction (ORM1) viral replication (IFIT1) and granulocyte regulation (CD17, CLC). However, Cluster #2 also has genes that are not directly related to immunity including collagen formation (COL9A2) and the urea cycle (ARG1). In summary, the cluster which displays the greatest infection specific differences in gene expression corresponding to high pollutant concentrations (Cluster #5) has a higher proportion of immune related genes than the more diversely populated Cluster #2.

Descriptions of additional sensitivity analysis on participants who are smokers (**Supplemental Table S10 online**) and highlighting individual genes present in the aforementioned significant gene pathways (**Supplemental Tables S11-S15 online**) are provided in the supplemental information.

## **Discussion**

In patients hospitalized with viral, bacterial and combined infection, we observed associations between gene expression and air pollution exposure in the 1 to 4 weeks prior to hospitalization, including gene pathways related to immunity, protein folding and iron homeostasis.

As hypothesized, combustion related air pollutants including DC (marker of wood burning) and BC (marker of traffic pollution) were associated with changes in gene expression of multiple pathways related to the immune response to respiratory infection. All pollutants were associated with genes involved in changes to protein folding, and all but BC were associated with key iron homeostasis pathways. Of all pollutants, BC was associated with upregulation of the largest number of immune specific pathways in the week prior to infection. In the infection specific exploration of selected pathways, we observed similar patterns of gene expression with the overall pathway analysis at the early lag times (0-6 days) but observed different patterns at the later lag times. There appeared to be infection specific patterns of gene expression when comparing patients with RVI, RBI and combined respiratory infection. Overall, this analysis suggests that the mechanistic effects of air pollution on the pathogenesis of respiratory infection may be pollutant, timing, and infection specific.

Discerning the effect of air pollution on the normal immune response to respiratory infection is made more difficult by limitations of our epidemiological design, and by the overlapping effects of both air pollution and respiratory infection on the human immune response.

To date, it is not clear by what mechanism(s) short term air pollution exposures contribute to diagnosed respiratory infection. Independent of air pollution exposure, RVIs can disrupt epithelial barriers [23-25], activate an inflammatory cascade mediated by nuclear factor kappa-light-chain-enhancer of activated B cells (NF- $\kappa$ B), [26] and activate key antiviral proteins, including Type I (e.g. interferon beta [IFN- $\beta$ ]) and Type II (e.g. interferon gamma [IFN- $\gamma$ ]) interferons) [27, 28].

Differentiating RVI from RBI using gene expression is an area of active research [29]. When compared to RVI, RBI has a distinct gene expression signature including the lymphocyte,  $\alpha$ -linoleic acid metabolism and IGF regulation pathways [30]. While another study of narrower scope found that a single gene (IFI27) is able to differentiate influenza from RBI [31], there is no current standard transcriptomic approach to differentiating RVI and RBI. Independent of the pollution effects, we also observed IFI27, among others (IFI44L and OAS2) to be differentially expressed in patients with viral infection compared to those bacterial infection. This consistency between our results and prior studies provides support for the generalizability of our population's transcriptomic data. However, adding complexity to the interpretation, air pollution can also illicit a strong immune response independent of respiratory infection.

Independent of respiratory infection, air pollution exposure is known to broadly lead to immune dysregulation in cell and animal models through pro-inflammatory changes to lung epithelia, dysregulation of cell signaling pathways and direct effects on immune cells including macrophages, dendritic cells and granulocytes [32]. Specifically, in response to diesel exhaust particles, a NF- $\kappa$ B mediated inflammatory cascade is thought to occur within the lung epithelium [10]. This response has the potential to disrupt the tight junctions between epithelial cells, thereby increasing the risk for viral or bacterial penetration and subsequent infection [9, 11]. We did not observe many inflammatory pathways associated with air pollution and only observed two individual inflammatory related genes in our study. ADORA1 participates in the activation of monocytes, which leads to a pro-inflammatory response. HERC5 was another example of a gene upregulated by endothelial inflammation that was differentially expressed between types of infections in our study. The paucity of inflammatory changes may be related to the overall low concentrations of ambient air pollution in the Rochester area compared to other more heavily polluted areas.

In terms of immune effects, decreased levels of IFN- $\gamma$  (important for macrophage activation) were observed in the peripheral blood of mice exposed to traffic pollution in China [12], and in the peripheral blood of humans exposed to diesel exhaust [33]. Further research has also observed dysregulation of the

epithelial cell junction, respiratory microbiome and cytokine response as additional factors increasing pathogenic virulence in the setting of PM exposure [34]. Our analysis suggests that while BC is associated with upregulation of type 1 interferon related pathway in the two weeks prior to infection, there may also be a component of immune suppression associated with exposure to traffic pollution in the later lag periods. Specifically, in the pathway analysis (Table 3), BC was associated with upregulation with numerous immune related pathways at the 0-6 lag period and downregulation of natural killer cell activity and antigen processing in the 14-20 day lag period. BC may have unique health effects when compared to other PM due to its physical shape as a chain aggregate particle. This shape provides a large surface area and concavity between intersecting spheres (most particles are convex/spherical) that improves the ability for BC to serve as a transport vector for other chemicals or possibly infectious organisms into the body [35]. UFP also was associated with upregulation of a viral related pathway at the 0-6 lag day period and suppression of multiple immune related pathways at the 7-13 lag day lag period. Though speculative, these findings suggest that for BC and UFP, an inflection point may exist where the effects of air pollution alone (lag days 13-28) are then comingled with the effect of acute infection during the incubation period in the 7 days preceding infection. In contrast to BC and UFP, PM<sub>2.5</sub> was associated with suppression of multiple immune related pathways at the 0-6 lag day period but had no associations with immune pathways at later lag periods. Though the etiology of the gene suppression from PM<sub>2.5</sub> is not clear, it may suggest that effects on immunity may be pollutant specific in addition to being timing specific. The relative importance of the directionality and timing of these immune changes for specific pollutants in the pathogenesis of respiratory infection deserves further study in prospective manner.

Despite our study population preceding the current COVID-19 pandemic, our participants had a high prevalence of several comorbidities that are risk factors to severe COVID-19 illness including diabetes, COPD, smoking and obesity [36]. A deficiency in the type 1 interferon response has been hypothesized to be a risk factor for a severe clinical course of COVID-19 [37]. Our study observed a difference in type 1 interferon expression when comparing RVI (high expression) and RBI (low expression) in the week prior to hospitalization. We also observed that one cluster of genes within the Type 1 Interferon pathway appeared to drive the association between air pollution and gene expression. Determining the potential gene pathways and individual genes involved in the air pollution/respiratory infection association is a key area of research for the current and future pandemics. A further benefit of improved knowledge of the risk of specific air pollutants and respiratory infections could be the ability to make real time policy changes (e.g. diesel traffic modifications) during a pandemic to reduce pathogen virulence.

In addition to the immune specific pathways, there were two additional general pathways broadly related to protein folding and iron homeostasis, which were associated with changes in air pollution. All

pollutants were associated with three pathways related to the endoplasmic reticulum (ER), an organelle central to protein synthesis (e.g. cytokines or other immune proteins) and transport in the body [38].

**(Table 3)** In a prior in vitro study of PM exposure to bronchial epithelial cells, PM increased stress on the ER (upregulation) that lead to a deleterious unfolded protein response (UPR) [39]. Influenza has also been observed to cause similar dysfunction in the ER [40]. This can impair the function of cells central to innate immunity in the lung including bronchial epithelial cells [41], and also has the potential to lead to dysregulation in the synthesis and transport of other immune related proteins. Aside from AMP, all other pollutant/ER related pathway associated were observed at the 7-27 day lag period, with no association observed in the 0-6 day lag period. This may suggest that the ER (protein folding) related dysregulation precedes the time point of acute infection for infections with incubation periods under 7 days (e.g. Influenza)[42]. While all other pollutants led to upregulation of ER related pathways, PM<sub>2.5</sub> was associated with suppression of the ER pathways. While both upregulation and suppression could lead to dysregulation, upregulation may be more detrimental due to the risk of the UPR.

Iron homeostasis plays an important role in the clinical course of RBI [43] and is increasingly recognized as an important factor in RVI[44] as well. Air pollution is known to induce a relative iron deficiency through dysregulation of iron homeostasis through mechanisms of chelation and/or sequestration [45]. Aside from BC, all other pollutants were observed to upregulate multiple pathways of iron homeostasis **(Table 3)**. The association between multiple pollutants and an upregulation of iron related pathways is consistent with the relative iron deficient state induced by air pollution. Intracellular iron deficiency can lead to an increased oxidative state and inflammation within the host. Whether the observed changes in iron homeostasis are protective (immune priming), deleterious (immunosuppressive) or related to changes in the proportion of blood cells remains unclear and the effect of air pollution on iron homeostasis pathways in viral infection is deserving of further study.

### **Limitations:**

Our study results should be interpreted in light of several limitations. First, our study focused on a small group of severely ill patients requiring hospitalization. Our cohort was also older with multiple comorbidities, potentially limiting the generalizability to younger patients with respiratory infection. Second, the lack of a control population limited the pathway and LIMMA analysis to comparisons between types of infection. Third, this analysis was not able to correct for the blood cell proportions in the peripheral whole blood of our samples. In theory the lack of inclusion of blood cell proportions should only minimally change the magnitude of effect estimates, but would result in a larger standard error and reduced statistical significance. Fourth, the epidemiological design does not allow for causal link between air pollution and immune response to respiratory infection. Fifth, as there was no external validation with

gene sets outside of our study, the generalizability of our findings are reduced. Sixth, there was likely an element of exposure misclassification given central site monitor estimated pollution, which likely reduced the magnitude of the observed effects. Finally, Rochester, NY has a relatively low average concentration of PM air pollution so generalizability to areas of higher pollution may be limited if dose thresholds exist in the pathogenic response to PM.

## **Conclusion**

Overall, this epidemiological study suggests that combustion related pollution, particularly BC, is associated with changes in gene expression within innate immune pathways. Increased concentrations in the majority of pollutants also appear to correspond to changes in expression to protein folding, iron homeostasis. Distinct from other pollutants, PM<sub>2.5</sub> was associated with downregulation of immune and protein folding pathways. The relatively low pollution in the study region may explain the lack of inflammatory changes accompanying the changes in the immune pathways. Future controlled exposure studies informed by broad epidemiological studies exploring the mechanism and timing of changes to the immune response to specific respiratory infections may further explore the relationship between inflammatory and immune responses to particulate air pollution.

## **Methods**

We used data from patients enrolled in the study by Falsey et al.[46] which includes adults over the age of 21 years with symptoms compatible with acute respiratory tract infection admitted through the emergency department at Rochester General Hospital (RGH), Rochester, NY from 2008-2011. All ethical approvals, guidelines and consent were provided in this previous study. Subjects had comprehensive microbiologic testing and cases were adjudicated by specialists as viral alone, bacterial alone, or mixed viral -bacterial infection. From this population, 1-3 ml of peripheral whole blood RNA was collected from 111 patients in Tempus tubes and hybridized using an Illumina Human HT-12 v4 BeadChip kit. Transcripts from the Illumina GenomeStudio based analysis were included if they were present in 10% or more of the samples and if they exhibited a minimum of a 2 fold expression change. In our current study on the association between air pollution and gene expression, we analyzed the data of all 111 patients who had a transcriptional analysis of peripheral blood performed.

### **Air pollution data**

Measurements of particle number concentrations in the size range of 10–500 nm are made continuously and sequentially at a New York State Department of Environmental Conservation (NYS DEC) site in Rochester, NY.[47] From 2004 to the present, measurements have been made at the NYS DEC primary site (latitude 43°09'56" N, longitude 77°33'15" W) on the eastside of Rochester, NY. This sampling site is

close to two major interstates (I-490 and I-590) as well as NY route 96, a major route carrying traffic traveling to and from downtown Rochester. Hourly  $PM_{2.5}$  mass, wind speed and wind direction, ambient temperature and relative humidity are also measured at the above-mentioned site. Size distribution measurements are made using a scanning mobility particle sizer (SMPS, TSI Inc.) system consisting of an electrostatic classifier (TSI model 3071), with an impactor having an orifice size of 0.0457 cm, an 85Kr aerosol neutralizer (TSI model 3077), and a condensation particle counter (CPC; TSI model 3010). The size range bounds are 10.4 nm (lower) and 0.542  $\mu\text{m}$  (upper) leading to measurement of mid-point particle sizes ranging from 11.1 nm to 0.47  $\mu\text{m}$  (32 channels per decade) at a total scan time of 5 minutes per sample. Routine maintenance such as calibrating rates is performed once a week to ensure that the system is functioning properly.  $PM_{2.5}$  is measured with a TEOM (model 1400ab, Thermo Fisher Scientific Inc., USA). Black carbon (BC) was measured with a 2-wavelength aethalometer. Delta-C (DC) is the difference between BC measured at 370 and 880 nm and has been shown to be a marker for biomass burning.[48] Pollution measurements were taken from Nov 1<sup>st</sup>, 2008 to May 31<sup>st</sup>, 2011.

### **Microarray data acquisition and processing**

This dataset, GSE60244 on the Gene Expression Omnibus, contains background corrected, non-normalized whole-blood genome data from microarrays run on the Illumina HumanHT-12 V4.0 expression BeadChip platform. The reported detection p-values were used to infer the mean and variance of the negative control probe intensities in order to perform background correction using the normal exponential convolution model. This method prevents negative values that arise from subtraction-based background correction. Quantile normalization was used after background correction in order to minimize variation between arrays. The probes were matched to gene names using the annotation package IlluminaHumanV4 [49]. Unidentified and non-detected genes were removed.

### **Statistics:**

#### ***Unsupervised analysis:***

After preprocessing, the top 150 high variance genes were selected for initial exploratory analyses. The selected expression values were centered and scaled to have mean zero and variance one. We performed an unsupervised analysis using hierarchical clustering with Euclidean distance and complete linkage on both the genes and samples. The resulting sample dendrogram was qualitatively compared to ambient air pollution levels in the weeks prior to infection diagnosis obtained from a prior study[22]. Finally, the ambient pollution values for each patient were overlaid from the day 0 to day 6 time period preceding diagnosis of infection.

### ***Individual gene analysis***

The Linear Models for MicroArray (LIMMA) package in R [50] was used to test hypotheses about the effect of ambient air pollution levels in the weeks prior to infection diagnosis on patients' gene expression using all 47,231 probes available in the microarray platform.

Each of the five pollutant exposures was tested individually by fitting a separate linear model for each of the four exposure time intervals: 0-6 days, 7-13 days, 14-20 days, and 21-27 days prior to date of diagnosis. This resulted in fitting a total of 20 models of the following form:

$$E[Y_{ij}] = \beta_{0j} + \beta_{1j}I(\text{Viral}_i) + \beta_{2j}I(\text{Coinfection}_i) + \beta_{3j}I(\text{Female}_i) + \beta_{4j}x_i^{\text{pollutant}}$$

Here  $Y_{ij}$  is the gene expression in subject  $i$  for gene probe  $j$ . Since no pollution exposure data was available for the gene expression control group, the bacterial infection group was set as the baseline category; therefore, the difference in expected gene expression between viral and bacterial infection and between coinfection and bacterial infection are  $\beta_{1j}$  and  $\beta_{2j}$ , respectively. The difference in expected gene expression between female and male subjects is  $\beta_{3j}$ . Ambient pollutant levels measured for AMP, PM<sub>2.5</sub>, UFP, DC, and BC in four different time lags are each used separately as the pollution value  $x_i^{\text{pollutant}}$ , and  $\beta_{4j}$  is the change in expected gene expression for a one unit increase in a given pollutant. Standard errors for these coefficients were calculated using the empirical Bayes method [51] central to the LIMMA method. Differential expression was determined by using a false discovery rate (FDR) threshold of 0.1.

### **Pathway analysis**

In order to test the effect of pollutants on gene pathways we used the CAMERA algorithm [52]. CAMERA implements a competitive gene set test comparing each pathway against all other genes not in the pathway. Such tests focus on identifying the most important biological processes relative to all other processes. Gene-wise moderated t-statistics are used as in LIMMA, but here the goal is to determine if the mean of the gene-wise statistics differ between the pathway of interest and all other genes. The key feature of CAMERA is that it accounts for inter-gene correlation in order to better control type I error. In this work, we used a false discovery rate threshold of 0.1.

Pathway definitions came from the Molecular Signatures Database version 7.1 GO: Gene Ontology gene sets listing of 10,192 pathways. The pathways defined here are derived from the Gene Ontology (GO) resource, and they were compiled into R data files which mapped probes to gene symbols, which were subsequently used to define pathway membership. We cross referenced the significant LIMMA results against the significant CAMERA results to limit our scope to only the pathways found significant in

CAMERA which also contained individual genes found to be significant in the LIMMA analysis. From this subset of 448 pathways we chose those known to be associated with infection and clustered those gene sets as in the unsupervised analysis to visually examine trends across patients and expression levels compared with pollutant levels.

## **References**

1. Xu J, Kochanek KD, Murphy SL, Arias E. Mortality in the United States. NCHS Data Brief. 2012;1:8.
2. Croft DP, Zhang WJ, Lin S, Thurston SW, Hopke PK, Masiol M, et al. The Association between Respiratory Infection and Air Pollution in the Setting of Air Quality Policy and Economic Change. *Annals of the American Thoracic Society*. 2019;16(3):321-30.
3. Ciencewicki J, Jaspers I. Air pollution and respiratory viral infection. *Inhal Toxicol*. 2007;19(14):1135-46.
4. Horne BD, Joy EA, Hofmann MG, Gesteland PH, Cannon JB, Lefler JS, et al. Short-Term Elevation of Fine Particulate Matter Air Pollution and Acute Lower Respiratory Infection. *Am J Respir Crit Care Med*. 2018;198(6):759-66.
5. Pirozzi CS, Jones BE, VanDerslice JA, Zhang Y, Paine R, 3rd, Dean NC. Short-Term Air Pollution and Incident Pneumonia. A Case-Crossover Study. *Ann Am Thorac Soc*. 2018;15(4):449-59.
6. Krall JR, Mulholland JA, Russell AG, Balachandran S, Winqvist A, Tolbert PE, et al. Associations between Source-Specific Fine Particulate Matter and Emergency Department Visits for Respiratory Disease in Four U.S. Cities. *Environmental Health Perspectives*. 2017;125(1):97-103.
7. WHO. Ambient air pollution: A global assessment of exposure and burden of disease. 2106 [Available from: <http://www.who.int/phe/publications/air-pollution-global-assessment/en/>].
8. CDC. Stats of the State of New York 2018 [updated 4/10/2018; cited 2019 November 1st].
9. Bayram H, Rusznak C, Khair OA, Sapsford RJ, Abdelaziz MM. Effect of ozone and nitrogen dioxide on the permeability of bronchial epithelial cell cultures of non-asthmatic and asthmatic subjects. *Clin Exp Allergy*. 2002;32(9):1285-92.
10. Weng CM, Lee MJ, He JR, Chao MW, Wang CH, Kuo HP. Diesel exhaust particles up-regulate interleukin-17A expression via ROS/NF-kappaB in airway epithelium. *Biochem Pharmacol*. 2018;151:1-8.
11. Yu XY, Takahashi N, Croxton TL, Spannhake EW. Modulation of bronchial epithelial cell barrier function by in vitro ozone exposure. *Environ Health Perspect*. 1994;102(12):1068-72.
12. Yang J, Chen Y, Yu Z, Ding H, Ma Z. Changes in gene expression in lungs of mice exposed to traffic-related air pollution. *Molecular and cellular probes*. 2018;39:33-40.
13. Sigaud S, Goldsmith CA, Zhou H, Yang Z, Fedulov A, Imrich A, et al. Air pollution particles diminish bacterial clearance in the primed lungs of mice. *Toxicol Appl Pharmacol*. 2007;223(1):1-9.
14. Lee GI, Saravia J, You D, Shrestha B, Jalgama S, Hebert VY, et al. Exposure to combustion generated environmentally persistent free radicals enhances severity of influenza virus infection. *Particle and fibre toxicology*. 2014;11:57.

15. Borcharding JA, Chen H, Caraballo JC, Baltrusaitis J, Pezzulo AA, Zabner J, et al. Coal fly ash impairs airway antimicrobial peptides and increases bacterial growth. *PLoS One*. 2013;8(2):e57673.
16. Harrod KS, Jaramillo RJ, Rosenberger CL, Wang SZ, Berger JA, McDonald JD, et al. Increased susceptibility to RSV infection by exposure to inhaled diesel engine emissions. *Am J Respir Cell Mol Biol*. 2003;28(4):451-63.
17. Jaspers I, Ciencewicz JM, Zhang W, Brighton LE, Carson JL, Beck MA, et al. Diesel exhaust enhances influenza virus infections in respiratory epithelial cells. *Toxicol Sci*. 2005;85(2):990-1002.
18. Bauer RN, Diaz-Sanchez D, Jaspers I. Effects of air pollutants on innate immunity: the role of Toll-like receptors and nucleotide-binding oligomerization domain-like receptors. *J Allergy Clin Immunol*. 2012;129(1):14-24; quiz 5-6.
19. Smyth T, Veazey J, Eliseeva S, Chalupa D, Elder A, Georas SN. Diesel exhaust particle exposure reduces expression of the epithelial tight junction protein Tricellulin. *Particle and fibre toxicology*. 2020;17(1):52.
20. Croft DP, Zhang W, Lin S, Thurston SW, Hopke PK, van Wijngaarden E, et al. Associations between Source-Specific Particulate Matter and Respiratory Infections in New York State Adults. *Environ Sci Technol*. 2020;54(2):975-84.
21. Office USGP. Electronic Code of Federal Regulations 2021 [updated 3/11/2021; cited 2021 March 13]. Available from: [https://www.ecfr.gov/cgi-bin/text-idx?node=pt40.2.50&rgn=div5#se40.2.50\\_17](https://www.ecfr.gov/cgi-bin/text-idx?node=pt40.2.50&rgn=div5#se40.2.50_17).
22. Suarez NM, Bunsow E, Falsey AR, Walsh EE, Mejias A, Ramilo O. Superiority of transcriptional profiling over procalcitonin for distinguishing bacterial from viral lower respiratory tract infections in hospitalized adults. *The Journal of infectious diseases*. 2015;212(2):213-22.
23. Comstock AT, Ganesan S, Chatteraj A, Faris AN, Margolis BL, Hershenson MB, et al. Rhinovirus-Induced Barrier Dysfunction in Polarized Airway Epithelial Cells Is Mediated by NADPH Oxidase 1. *Journal of Virology*. 2011;85(13):6795-808.
24. Rezaee F, DeSando SA, Ivanov AI, Chapman TJ, Knowlden SA, Beck LA, et al. Sustained protein kinase D activation mediates respiratory syncytial virus-induced airway barrier disruption. *J Virol*. 2013;87(20):11088-95.
25. Coyne CB, Shen L, Turner JR, Bergelson JM. Coxsackievirus entry across epithelial tight junctions requires occludin and the small GTPases Rab34 and Rab5. *Cell Host Microbe*. 2007;2(3):181-92.
26. Garcia MA, Gallego P, Campagna M, Gonzalez-Santamaria J, Martinez G, Marcos-Villar L, et al. Activation of NF- $\kappa$ B pathway by virus infection requires Rb expression. *PLoS One*. 2009;4(7):e6422.
27. Li SF, Gong MJ, Zhao FR, Shao JJ, Xie YL, Zhang YG, et al. Type I Interferons: Distinct Biological Activities and Current Applications for Viral Infection. *Cellular Physiology and Biochemistry*. 2018;51(5):2377-96.

28. Platanius LC. Mechanisms of type-I- and type-II-interferon-mediated signalling. *Nat Rev Immunol*. 2005;5(5):375-86.
29. Tsalik EL, Henaó R, Nichols M, Burke T, Ko ER, McClain MT, et al. Host gene expression classifiers diagnose acute respiratory illness etiology. *Science Translational Medicine*. 2016;8(322):322ra11-ra11.
30. Bhattacharya S, Rosenberg AF, Peterson DR, Grzesik K, Baran AM, Ashton JM, et al. Transcriptomic biomarkers to discriminate bacterial from nonbacterial infection in adults hospitalized with respiratory illness. *Scientific reports*. 2017;7(1):1-12.
31. Tang BM, Shojaei M, Parnell GP, Huang S, Nalos M, Teoh S, et al. A novel immune biomarker IFI27 discriminates between influenza and bacteria in patients with suspected respiratory infection. *European Respiratory Journal*. 2017;49(6):1602098.
32. Glencross DA, Ho T-R, Camiña N, Hawrylowicz CM, Pfeffer PE. Air pollution and its effects on the immune system. *Free Radical Biology and Medicine*. 2020;151:56-68.
33. Estrella B, Naumova EN, Cepeda M, Voortman T, Katsikis PD, Drexhage HA. Effects of Air Pollution on Lung Innate Lymphoid Cells: Review of In Vitro and In Vivo Experimental Studies. *Int J Environ Res Public Health*. 2019;16(13).
34. Pompilio A, Di Bonaventura G. Ambient air pollution and respiratory bacterial infections, a troubling association: epidemiology, underlying mechanisms, and future challenges. *Critical Reviews in Microbiology*. 2020;46(5):600-30.
35. Organization WH. Health effects of black carbon: WHO; 2012.
36. Ejaz H, Alsrhani A, Zafar A, Javed H, Junaid K, Abdalla AE, et al. COVID-19 and comorbidities: Deleterious impact on infected patients. *Journal of Infection and Public Health*. 2020;13(12):1833-9.
37. Hadjadj J, Yatim N, Barnabei L, Corneau A, Boussier J, Smith N, et al. Impaired type I interferon activity and inflammatory responses in severe COVID-19 patients. *Science*. 2020;369(6504):718-24.
38. Schwarz DS, Blower MD. The endoplasmic reticulum: structure, function and response to cellular signaling. *Cellular and molecular life sciences : CMLS*. 2016;73(1):79-94.
39. Watterson TL, Hamilton B, Martin R, Coulombe RA, Jr. Urban particulate matter causes ER stress and the unfolded protein response in human lung cells. *Toxicol Sci*. 2009;112(1):111-22.
40. Zhang K, Kaufman RJ. The unfolded protein response. *Neurology*. 2006;66(1 suppl 1):S102.
41. Chen ACH, Burr L, McGuckin MA. Oxidative and endoplasmic reticulum stress in respiratory disease. *Clin Transl Immunology*. 2018;7(6):e1019-e.
42. Lessler J, Reich NG, Brookmeyer R, Perl TM, Nelson KE, Cummings DA. Incubation periods of acute respiratory viral infections: a systematic review. *Lancet Infect Dis*. 2009;9(5):291-300.

43. Gomes AC, Moreira AC, Mesquita G, Gomes MS. Modulation of Iron Metabolism in Response to Infection: Twists for All Tastes. *Pharmaceuticals (Basel)*. 2018;11(3):84.
44. Liu W, Zhang S, Nekhai S, Liu S. Depriving Iron Supply to the Virus Represents a Promising Adjuvant Therapeutic Against Viral Survival. *Curr Clin Microbiol Rep*. 2020:1-7.
45. Ghio AJ, Soukup JM, Dailey LA, Madden MC. Air pollutants disrupt iron homeostasis to impact oxidant generation, biological effects, and tissue injury. *Free Radic Biol Med*. 2020;151:38-55.
46. Falsey AR, Becker KL, Swinburne AJ, Nysten ES, Formica MA, Hennessey PA, et al. Bacterial complications of respiratory tract viral illness: a comprehensive evaluation. *J Infect Dis*. 2013;208(3):432-41.
47. Masiol M, Squizzato S, Chalupa DC, Utell MJ, Rich DQ, Hopke PK. Long-term trends in submicron particle concentrations in a metropolitan area of the northeastern United States. *Science of The Total Environment*. 2018;633:59-70.
48. Wang Y, Hopke PK, Rattigan OV, Chalupa DC, Utell MJ. Multiple-year black carbon measurements and source apportionment using delta-C in Rochester, New York. *Journal of the Air & Waste Management Association*. 2012;62(8):880-7.
49. Dunning M, Lynch A, Eldridge M. illuminaHumanv4. db: Illumina HumanHT12v4 annotation data (chip illuminaHumanv4). R package version. 2015;1(0).
50. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic acids research*. 2015;43(7):e47-e.
51. Law CW, Chen Y, Shi W, Smyth GK. voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol*. 2014;15(2):R29.
52. Wu D, Smyth GK. Camera: a competitive gene set test accounting for inter-gene correlation. *Nucleic acids research*. 2012;40(17):e133-e.

### Acknowledgements:

Research reported in this publication was supported by the Empire Clinical Research Investigator Program of the New York State Department of Health. Support was also provided by the National Institute of Environmental Health Sciences of the National Institutes of Health (NIH) under award numbers K23-ES032459, T32ES007271 and P30ES001247 and from the National Heart, Lung, and Blood Institute of the NIH, T32-HL066988-1.

Author contributions: Study concept and design: DPC, DSB, DN, RMK, SWT, PKH, DQR, SB, SG, TM and MNM. Acquisition, analysis or interpretation of data: All authors. Drafting of the manuscript: DPC, CJJ, DSB, MNM, MJU, PKH, AAL, DQR. Critical review and revision of the manuscript: All authors. Statistical analysis: SWT, DSB, DQR, MNM. Study supervision: DPC, ARF, TM, PKH, DQR, MNM.

### Competing interests

The authors do not have any competing interests to declare.

### **Figure Legends:**

**Figure 1** Study overview [Created with BioRender.com]

**Figure 2.** Heat map of high variance genes and corresponding air pollution concentrations, stratified by type of respiratory infection and divided by participant (Columns) and gene cluster (Rows 1-7)

**Figure 3.** Infection specific heatmap of Type 1 Interferon pathway with corresponding black carbon concentrations divided by participant (Columns) and Cluster (Rows 1-5)

**Table 1 Patient Characteristics**

Variable Name	Descriptions	Overall N(%)	Viral N = 66 (%)	Bacterial N= 20 (%)	Coinfection N= 25 (%)
<b>Age</b>		61.1(26.1)*	59.7(26.7)*	71.6(25.9)*	59.6(17.9)*
<b>Gender</b>	<b>Female</b>	64(57.7)	41(62.1)	10(50)	13(52)
	<b>Male</b>	47(42.3)	25(37.9)	10(50)	12(48)
<b>Race</b>	<b>Black</b>	25(22.5)	11(16.7)	7(35)	7(28)
	<b>Other</b>	2(1.8)	1(1.5)		1(4)
	<b>White</b>	84(75.7)	54(81.8)	13(65)	17(68)
<b>Ethnicity</b>	<b>Hispanic</b>	17(15.3)	10(15.2)	2(10)	5(20)
	<b>Non-Hispanic</b>	94(84.7)	56(84.8)	18(90)	20(80)
<b>Type of Residence</b>	<b>Home</b>	103(92.8)	61(92.4)	19(95)	23(92)
<b>Diabetes</b>	<b>Yes Diabetes</b>	38(34.2)	24(36.4)	6(30)	8(32)
<b>CHF</b>	<b>Yes CHF</b>	19(17.1)	12(18.2)	3(15)	4(16)
<b>COPD</b>	<b>Yes COPD</b>	42(37.8)	22(33.3)	9(45)	11(44)
<b>CRF</b>	<b>Yes CRF</b>	1(0.9)			1(4)
<b>BMI</b>		27.4(13)*	27.2(14.6)*	27.9(11.1)*	27.4(9.8)*
<b>BMI category</b>	<b>Underweight</b>	4(4.6)	1(2)	2(11.8)	1(4.8)
	<b>Normal</b>	31(35.6)	18(36.7)	5(29.4)	8(38.1)
	<b>Overweight</b>	18(20.7)	11(22.4)	3(17.6)	4(19)
	<b>Obese</b>	34(39.1)	19(38.8)	7(41.2)	8(38.1)
<b>Smoking status</b>	<b>Never</b>	30(27)	19(28.8)	6(30)	5(20)
	<b>Active (within 3 months)</b>	44(39.6)	23(34.8)	8(40)	13(52)
	<b>Past</b>	37(33.3)	24(36.4)	6(30)	7(28)
<b>Oral Steroid</b>	<b>Yes</b>	12(10.9)	9(13.8)	1(5)	2(8)
	<b>No</b>	98(89.1)	56(86.2)	19(95)	23(92)
<b>Inhaled Steroid</b>	<b>Yes</b>	47(42.3)	20(30.3)	13(65)	14(56)
	<b>No</b>	64(57.7)	46(69.7)	7(35)	11(44)
<b>Home O2</b>	<b>Yes</b>	19(17.1)	10(15.2)	6(30)	3(12)
<b>Flu vaccine</b>	<b>Yes</b>	66(62.3)	39(61.9)	16(84.2)	11(45.8)
<b>Pneumonia vaccine</b>	<b>Yes</b>	56(58.9)	32(60.4)	13(72.2)	11(45.8)

\*Age and BMI are Median (interquartile range)

**Table 2. Ambient air pollution concentrations for all participants, at multiple lag times, stratified by infection type**

<b>Pollutant</b>	<b>Lag period prior to diagnosis</b>	<b>Overall Mean(Std)</b>	<b>Viral Mean(Std)</b>	<b>Bacterial Mean(Std)</b>	<b>Coinfection Mean(Std)</b>
<b>AMP</b>					
	0 to 6 days	873.22(308.7)	902.39(274.54)	827.48(483.06)	832.53(134.39)
	7 to 13 days	1008.29(323)	984.63(307.87)	1081.82(381.58)	1019.45(327.68)
	14 to 20 days	1142.28(412.83)	1156.35(457.73)	1156.7(422.28)	1095.93(276.88)
	21 to 27 days	1087.91(451.2)	1116.54(472.79)	1013.61(483.29)	1064.06(378.11)
<b>Black Carbon</b>					
	0 to 6 days	0.52(0.21)	0.52(0.18)	0.52(0.31)	0.51(0.16)
	7 to 13 days	0.53(0.16)	0.49(0.13)	0.6(0.19)	0.57(0.18)
	14 to 20 days	0.56(0.18)	0.56(0.17)	0.54(0.24)	0.57(0.15)
	21 to 27 days	0.55(0.21)	0.54(0.2)	0.52(0.21)	0.62(0.22)
<b>Delta C</b>					
	0 to 6 days	0.19(0.11)	0.19(0.11)	0.18(0.11)	0.21(0.09)
	7 to 13 days	0.21(0.09)	0.18(0.08)	0.24(0.1)	0.26(0.1)
	14 to 20 days	0.22(0.1)	0.22(0.09)	0.23(0.13)	0.23(0.09)
	21 to 27 days	0.24(0.13)	0.23(0.14)	0.25(0.12)	0.28(0.13)
<b>PM<sub>2.5</sub></b>					
	0 to 6 days	7.79(2.75)	7.66(2.63)	8.01(3.58)	7.9(2.34)
	7 to 13 days	8.33(2.67)	7.72(2.51)	9.58(2.54)	8.8(2.79)
	14 to 20 day	8.45(3.05)	8.16(3.4)	8.9(3.12)	8.79(1.84)
	21 to 27 days	8.15(3.5)	7.86(3.66)	8.62(3.6)	8.55(2.99)
<b>Ultrafine Particles</b>					
	0 to 6 days	4611.46(1120.64)	4594.73(1026.12)	4890.24(1512.57)	4365.42(909.43)
	7 to 13 days	4835.69(1221.7)	4815.96(1163.35)	5186.72(1646.82)	4619.66(1000.51)
	14 to 20 days	5262.77(1410.88)	5208.64(1473.39)	5703.08(1534.16)	5110.57(1154.33)
	21 to 27 days	5161.78(1595.44)	5276.92(1604.38)	5309.07(1799.33)	4753.53(1425.86)

**Table 3. Pattern of association between gene pathways and increased concentrations of multiple particulate matter pollutants at multiple lag times (red-upregulated, blue-downregulated).**

Gene Pathway Name	BC (lag)	DC (lag)	AMP (lag)	PM <sub>2.5</sub> (lag)	UFP (lag)
Establishment_Of_Protein_Localization_To_Endoplasmic_Reticulum	21-27	7-13	0-6	7-13	14-20
Ribosomal_Subunit	21-27	7-13	0-6	7-13	14-20
Structural_Constituent_Of_Ribosome	21-27	7-13	0-6	7-13	14-20
Specific_Granule_Lumen		21-27	0-6	0-6	0-6
Cotranslational_Protein_Targeting_To_Membrane	21-27	7-13	0-6		14-20
Cytosolic_Large_Ribosomal_Subunit	21-27	7-13	0-6		7-13
Cytosolic_Ribosome	21-27	7-13	0-6		14-20
Ferric_Iron_Binding		0-6	21-27	21-27	0-6
Ferrous_Iron_Binding		0-6	21-27	21-27	0-6
Large_Ribosomal_Subunit	21-27	7-13	0-6		14-20
Nuclear_Transcribed_Mrna_Catabolic_Process_Nonsense_Mediated_Decay	21-27	7-13	0-6		14-20
Protein_Localization_To_Endoplasmic_Reticulum	21-27	7-13	0-6		14-20
Protoporphyrinogen_Ix_Biosynthetic_Process		0-6	21-27	0-6	0-6
Protoporphyrinogen_Ix_Metabolic_Process		0-6	21-27	0-6	0-6
Ribosome	21-27	7-13	0-6		14-20
Translational_Initiation	21-27	7-13	0-6		14-20
Catalytic_Step_2_Spliceosome		0-6	21-27	14-20	
Mitochondrial_Translation		0-6	21-27	7-13	
U2_Type_Catalytic_Step_2_Spliceosome		0-6	21-27	7-13	
U2_Type_Spliceosomal_Complex		0-6	21-27	7-13	
Nuclear_Transcribed_Mrna_Catabolic_Process	21-27		0-6		14-20
Translation_Preinitiation_Complex		14-20	0-6		7-13
Specific_Granule			0-6	0-6	0-6
Small_Ribosomal_Subunit	21-27			7-13	14-20
Cytoplasmic_Translation	21-27	14-20			14-20
Cytosolic_Small_Ribosomal_Subunit	21-27	7-13			14-20
Protein_Targeting_To_Membrane	21-27	7-13			14-20
Viral_Gene_Expression	21-27	7-13			14-20
Cytoplasmic_Side_Of_Rough_Endoplasmic_Reticulum_Membrane			7-13	7-13	
General_Transcription_Initiation_Factor_Activity			21-27	21-27	
Mitochondrial_Gene_Expression			21-27	7-13	
Myeloid_Leukocyte_Mediated_Immunity			0-6		7-13
Positive_Regulation_Of_Double_Strand_Break_Repair_Via_Homologous_Recombination			21-27	21-27	
Secretory_Granule_Membrane			0-6		7-13
Specific_Granule_Membrane			0-6		0-6
Superoxide_Anion_Generation			0-6		7-13
Tertiary_Granule			0-6		7-13
Autolysosome			21-27	21-27	
Oxidoreductase_Activity_Oxidizing_Metal_Ions			21-27	21-27	
Oxidoreductase_Activity_Oxidizing_Metal_Ions_Oxygen_As_Acceptor			21-27	21-27	
Regulation_Of_Ryanodine_Sensitive_Calcium_Release_Channel_Activity	21-27	21-27			
T_Cell_Receptor_Binding	14-20			14-20	
Mda_5_Signaling_Pathway	0-6			0-6	
Polysomal_Ribosome	21-27				14-20
Polysome	21-27				21-27
Cajal_Body		0-6		14-20	
Lewy_Body		0-6			0-6
Porphyrin_Containing_Compound_Metabolic_Process		0-6			0-6
Regulation_Of_Mitochondrial_Electron_Transport_Nadh_To_Ubiquinone		0-6			0-6
Spectrin_Associated_Cytoskeleton		0-6			0-6
Hemoglobin_Metabolic_Process				7-13	0-6
Purine_Deoxyribonucleotide_Binding				0-6	0-6

# Figures

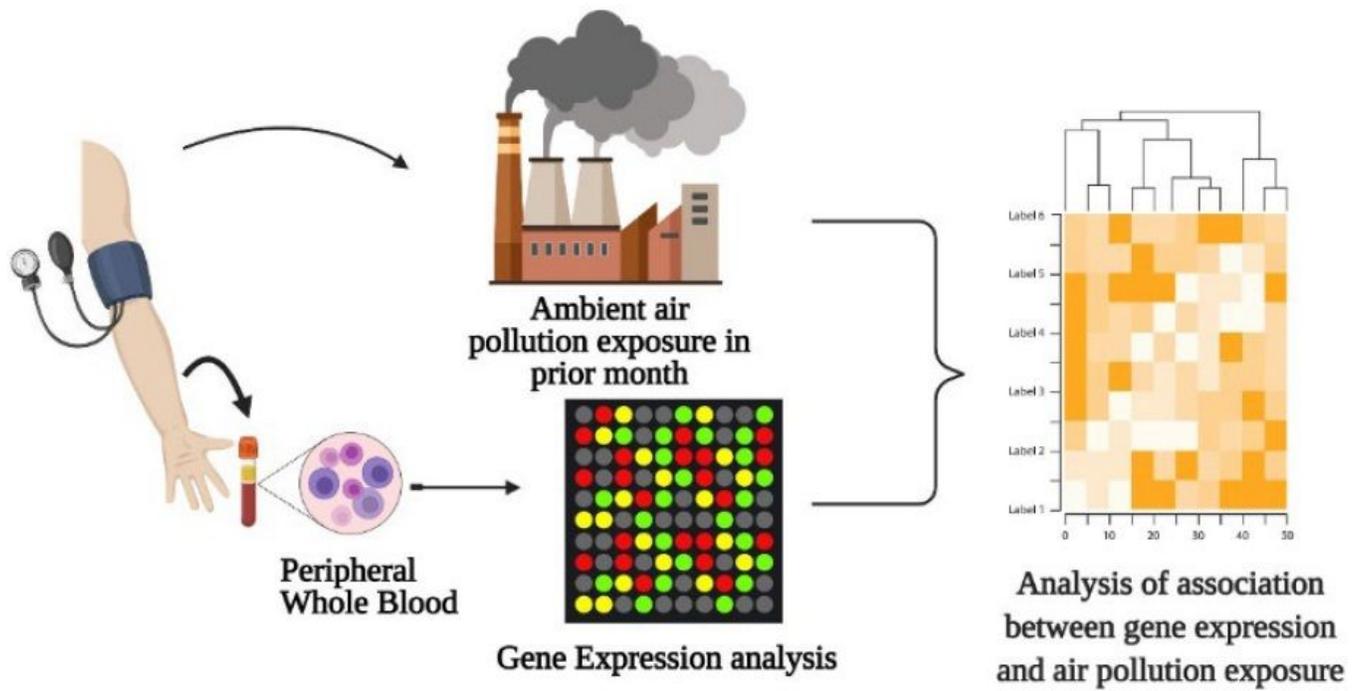


Figure 1

Study overview [Created with BioRender.com]

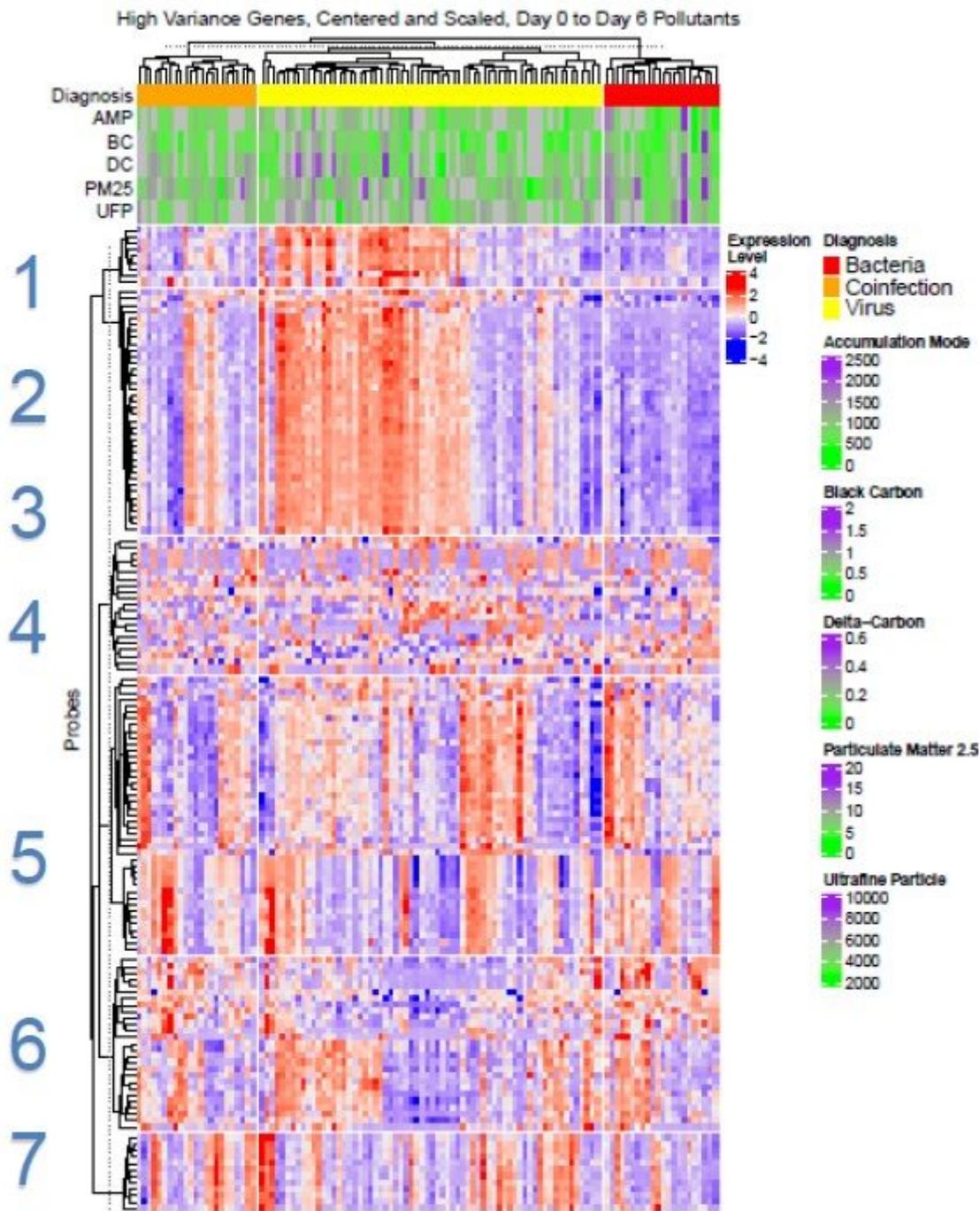
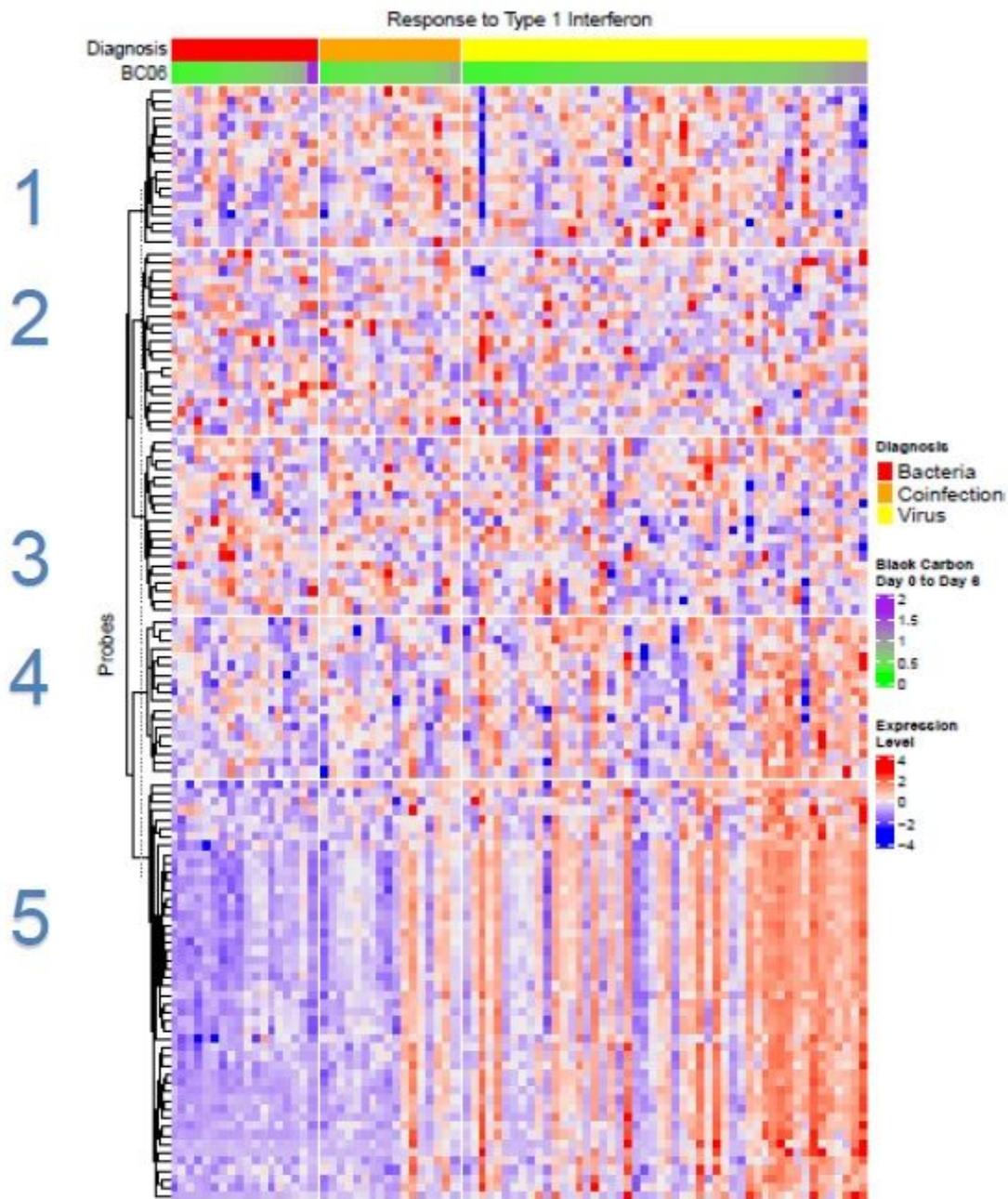


Figure 2

Heat map of high variance genes and corresponding air pollution concentrations at the 0-6 lag days, stratified by type of respiratory infection and divided by participant (Columns) and gene cluster (Rows1-7)



**Figure 3**

Infection specific heatmap of Type 1 Interferon pathway with corresponding black carbon concentrations divided by participant (Columns) and Cluster (Rows 1-5)

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementalFiguresheatmapsfinal.pdf](#)
- [Supplementalinformation4.19finalsubmit.pdf](#)